LOW-DOSE NEOANTIGEN VACCINE THERAPY

Abstract

Disclosed herein are compositions that include antigen-encoding nucleic acid sequences and/or antigen peptides. Also disclosed are nucleotides, cells, and methods associated with the compositions including their use as vaccines, including vectors and methods for a heterologous prime/boost vaccination strategy.

Inventors: **Jooss; Karin (San Diego, CA), Chauder; Brian (DeRuyter, NY), Ferguson; Andrew (Hingham, MA)**

Applicant: **Gritstone bio, Inc.** (Emeryville, CA)

Family ID: **88102186**

Appl. No.: **18/891398**

Filed: **September 20, 2024**

Related U.S. Application Data

parent US continuation PCT/US2023/064791 20230321 PENDING child US 18891398
us-provisional-application US 63384567 20221121
us-provisional-application US 63322143 20220321

Publication Classification

Int. Cl.: **A61K39/00** (20060101); **C12N15/86** (20060101)

U.S. Cl.:

CPC **A61K39/001164** (20180801); **C12N15/86** (20130101); A61K2039/53 (20130101); A61K2039/545 (20130101); A61K2039/55555 (20130101); C12N2770/36143 (20130101)

Background/Summary

CROSS REFERENCE TO RELATED APPLICATIONS [0001] This application is a continuation of International Patent Application No. PCT/US2023/064791 filed Mar. 21, 2023, which claims the benefit of U.S. Provisional Application No. 63/322,143 filed Mar. 21, 2022, and U.S. Provisional Application No. 63/384,567 filed Nov. 21, 2022, each of which is hereby incorporated in their entirety by reference for all purposes.

SEQUENCE LISTING
[0002] The instant application contains a Sequence Listing which has been submitted electronically in .XML format and is hereby incorporated by reference in its entirety. Said. XML copy, created on Feb. 2, 2023, is named GSO-105WO and is 6,031,351 bytes in size.
BACKGROUND

[0003] Therapeutic vaccines based on tumor-specific antigens hold great promise as a next-generation of personalized cancer immunotherapy..sup.1-3 For example, cancers with a high mutational burden, such as non-small cell lung cancer (NSCLC) and melanoma, are particularly attractive targets of such therapy given the relatively greater likelihood of neoantigen generation..sup.4,5 Early evidence shows that neoantigen-based vaccination can elicit T-cell responses.sup.6 and that neoantigen targeted cell-therapy can cause tumor regression under certain circumstances in selected patients..sup.7

[0004] One question for neoantigen vaccine design is which of the many coding mutations present in subject tumors can generate the "best" therapeutic neoantigens, e.g., antigens that can elicit anti-tumor immunity and cause tumor regression.

[0005] Initial methods have been proposed incorporating mutation-based analysis using next-generation sequencing, RNA gene expression, and prediction of MHC binding affinity of candidate neoantigen peptides.sup.8. However, these proposed methods can fail to model the entirety of the epitope generation process, which contains many steps (e.g., TAP transport, proteasomal cleavage, and/or TCR recognition) in addition to gene expression and MHC binding.sup.9. Consequently, existing methods are likely to suffer from reduced low positive predictive value (PPV).

[0006] Indeed, analyses of peptides presented by tumor cells performed by multiple groups have shown that <5% of peptides that are predicted to be presented using gene expression and MHC binding affinity can be found on the tumor surface MHC.sup.10,11. This low correlation between binding prediction and MHC presentation was further reinforced by recent observations of the lack of predictive accuracy improvement of binding-restricted neoantigens for checkpoint inhibitor response over the number of mutations alone..sup.12

[0007] This low positive predictive value (PPV) of existing methods for predicting presentation presents a problem for neoantigen-based vaccine design. If vaccines are designed using predictions with a low PPV, most patients are unlikely to receive a therapeutic neoantigen and fewer still are likely to receive more than one (even assuming all presented peptides are immunogenic). Thus, neoantigen vaccination with current methods is unlikely to succeed in a substantial number of subjects having tumors.

[0008] Additionally, previous approaches generated candidate neoantigens using only cis-acting mutations, and largely neglected to consider additional sources of neo-ORFs, including mutations in splicing factors, which occur in multiple tumor types and lead to aberrant splicing of many genes.sup.13, and mutations that create or remove protease cleavage sites.

[0009] Finally, standard approaches to tumor genome and transcriptome analysis can miss somatic mutations that give rise to candidate neoantigens due to suboptimal conditions in library construction, exome and transcriptome capture, sequencing, or data analysis. Likewise, standard tumor analysis approaches can inadvertently promote sequence artifacts or germline polymorphisms as neoantigens, leading to inefficient use of vaccine capacity or auto-immunity risk, respectively.

[0010] In addition to the challenges of current neoantigen prediction methods certain challenges also exist with the available vector systems that can be used for neoantigen delivery in humans, many of which are derived from humans. For example, many humans have pre-existing immunity to human viruses as a result of previous natural exposure, and this immunity can be a major obstacle to the use of recombinant human viruses for neoantigen delivery for cancer treatment.

[0011] In addition, targeting antigens that are shared among patients with cancer hold great promise as a vaccine strategy, including targeting both neoantigens with a mutation as well as tumor antigens without a mutation (e.g., tumors antigens that are improperly expressed). The challenges with shared antigen vaccine strategies include at least those discussed above.

SUMMARY

[0012] Disclosed herein is a composition for delivery of a self-amplifying alphavirus-based expression system, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises: (A) the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, wherein the one or more vectors comprises: (a) an RNA alphavirus backbone, wherein the RNA alphavirus backbone comprises: (i) at least one promoter nucleotide sequence, and (ii) at least one polyadenylation (poly(A)) sequence; and (b) a cassette, wherein the cassette comprises: (i) at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; (ii) optionally, a second promoter nucleotide sequence operably linked to the at least one antigen-encoding nucleic acid sequence; and (iii) optionally, at least one second poly(A) sequence, wherein the second poly(A) sequence is a native poly(A) sequence or an exogenous poly(A) sequence to the alphavirus, and (B) a lipid-nanoparticle (LNP), wherein the LNP encapsulates the self-amplifying alphavirus-based expression system, and wherein the composition comprises a therapeutically effective amount comprising 30 µg or less of each of the one or more vectors.

[0013] Also disclosed herein is a composition for delivery of a self-amplifying alphavirus-based expression system, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises: (A) the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system

comprises one or more vectors, wherein the one or more vectors comprises: (a) an RNA alphavirus backbone, wherein the RNA alphavirus backbone comprises the nucleic acid sequence set forth in SEQ ID NO:6, wherein the RNA alphavirus backbone sequence comprises a 26S promoter nucleotide sequence and a poly(A) sequence, wherein the 26S promoter sequence is endogenous to the RNA alphavirus backbone, and wherein the poly(A) sequence is endogenous to the RNA alphavirus backbone; and (b) a cassette integrated between the 26S promoter nucleotide sequence and the poly(A) sequence, wherein the cassette is operably linked to the 26S promoter nucleotide sequence, and wherein the cassette comprises at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; and (B) a lipid-nanoparticle (LNP), wherein the LNP encapsulates the self-amplifying alphavirus-based expression system, and wherein the composition comprises a therapeutically effective amount comprising 30 µg or less of each of the one or more vectors.

[0014] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises a therapeutically effective amount comprising 30 µg or less in total for all of the one or more vectors combined. In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises 30 µg of each of the one or more vectors. In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises 30 µg in total for all of the one or more vectors combined of each of the one or more vectors. In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises 30 µg in total for all of the one or more vectors combined, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises only a single distinct vector.

[0015] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises only a single distinct vector.

[0016] In some aspects, the therapeutically effective amount is capable of stimulating a T cell response following administration to a subject. In some aspects, the T cell response is assessed by monitoring ex vivo stimulation with the encoded epitope in PBMCs from the subject. In some aspects, monitoring ex vivo stimulation comprises an ELISpot assay, optionally wherein the ELISpot assay quantifies IFN-gamma production.

[0017] In some aspects, the therapeutically effective amount does not stimulate an inhibitory IFNα response and/or reduces the inhibitory IFNα response following administration to a subject.

[0018] In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is between 10-40 to 1. In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is between 16-32 to 1. In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is about 24 to 1. In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is 24 to 1.

[0019] In some aspects, the therapeutically effective amount is at least 1 µg, at least 3 µg, at least 10 µg, about 30 µg or less, or is 30 µg. In some aspects, the therapeutically effective amount is at least 1 µg. In some aspects, the therapeutically effective amount is at least 3 µg. In some aspects, the therapeutically effective amount is at least 10 µg. In some aspects, the therapeutically effective amount is about 30 µg or less. In some aspects, the therapeutically effective amount is 30 µg.

[0020] In some aspects, the one or more vectors is at a concentration of 1 mg/mL.

[0021] In some aspects, an ordered sequence of each element of the cassette in the composition for delivery of the self-amplifying alphavirus-based expression system is described in the formula, from 5′ to 3′, comprising $P_a\text{-}(L5_b\text{-}N_c\text{-}L3_d)_X\text{-}(G5_e\text{-}U_f)_Y\text{-}G3_g$ wherein P comprises the second promoter nucleotide sequence, where a=0 or 1, N comprises one of the epitope-encoding nucleic acid sequences, wherein the epitope-encoding nucleic acid sequence comprises an MHC class I epitope-encoding nucleic acid sequence, where c=1, L5 comprises the 5′ linker sequence, where b=0 or 1, L3 comprises the 3′ linker sequence, where d=0 or 1, G5 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where e=0 or 1, G3 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where g=0 or 1, U comprises one of the at least one MHC class II epitope-encoding nucleic acid sequence, where f=1, X=1 to 400, where for each X the corresponding Ne is an MHC class I epitope-encoding nucleic acid sequence, and Y=0, 1, or 2, where for each Y the corresponding Ur is an MHC class II epitope-encoding nucleic acid sequence. In some aspects, for each X the corresponding Ne is a distinct MHC class I epitope-encoding nucleic acid sequence. In some aspects, for each Y the corresponding Ur is a distinct MHC class II epitope-encoding nucleic acid sequence. In some aspects, wherein a=0, b=1, d=1, c=1, g=1, h=1, X=at least 2, Y=2, the at least one promoter nucleotide sequence is a single 26S promoter nucleotide sequence provided by the RNA alphavirus backbone, the at least one polyadenylation poly(A) sequence is a poly(A) sequence of at least 100 consecutive A nucleotides (SEQ ID NO: 29358) provided by the RNA alphavirus backbone, the cassette is integrated between the 26S promoter nucleotide sequence and the poly(A) sequence, wherein the cassette is operably linked to the 26S promoter nucleotide sequence and the poly(A) sequence, each N encodes a MHC class I epitope 7-15 amino acids in length, L5 is a native 5′ linker sequence that encodes a native N-terminal amino acid sequence of the MHC I epitope, and wherein the 5′ linker sequence encodes a peptide that is at least 3 amino acids in length, L3 is a native 3′ linker sequence that encodes a native C-terminal amino acid sequence of the MHC I epitope, and wherein the 3′ linker sequence encodes a peptide that

is at least 3 amino acids in length, U is each of a PADRE class II sequence and a Tetanus toxoid MHC class II sequence, the RNA alphavirus backbone is the sequence set forth in SEQ ID NO:6, and each of the MHC class I epitope-encoding nucleic acid sequences encodes a polypeptide that is between 13 and 25 amino acids in length.

[0022] In some aspects, the LNP comprises a lipid selected from the group consisting of: an ionizable amino lipid, a phosphatidylcholine, cholesterol, a PEG-based coat lipid, or a combination thereof. In some aspects, the LNP comprises an ionizable amino lipid, a phosphatidylcholine, cholesterol, and a PEG-based coat lipid. In some aspects, the ionizable amino lipids comprise MC3-like (dilinoleylmethyl-4-dimethylaminobutyrate) molecules. In some aspects, the LNP-encapsulated expression system has a diameter of about 100 nm.

[0023] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system is formulated for intramuscular (IM), intradermal (ID), subcutaneous (SC), or intravenous (IV) administration. In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system is formulated for intramuscular (IM) administration.

[0024] In some aspects, the cassette is integrated between the at least one promoter nucleotide sequence and the at least one poly(A) sequence. In some aspects, the at least one promoter nucleotide sequence is operably linked to the cassette.

[0025] In some aspects, the one or more vectors comprise one or more +-stranded RNA vectors. In some aspects, the one or more +-stranded RNA vectors comprise a 5′ 7-methylguanosine (m7g) cap. In some aspects, the one or more +-stranded RNA vectors are produced by in vitro transcription. In some aspects, the one or more vectors are self-amplifying within a mammalian cell. In some aspects, the RNA alphavirus backbone comprises at least one nucleotide sequence of an Aura virus, a Fort Morgan virus, a Venezuelan equine encephalitis virus, a Ross River virus, a Semliki Forest virus, a Sindbis virus, or a Mayaro virus. In some aspects, the RNA alphavirus backbone comprises at least one nucleotide sequence of a Venezuelan equine encephalitis virus. In some aspects, the RNA alphavirus backbone comprises at least sequences for nonstructural protein-mediated amplification, a 26S promoter sequence, a poly(A) sequence, a nonstructural protein 1 (nsP1) gene, a nsP2 gene, a nsP3 gene, and a nsP4 gene encoded by the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus. In some aspects, the RNA alphavirus backbone comprises at least sequences for nonstructural protein-mediated amplification, a 26S promoter sequence, and a poly(A) sequence encoded by the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus. In some aspects, sequences for nonstructural protein-mediated amplification are selected from the group consisting of: an alphavirus 5′ UTR, a 51-nt CSE, a 24-nt CSE, a 26S subgenomic promoter sequence, a 19-nt CSE, an alphavirus 3′ UTR, or combinations thereof. In some aspects, the RNA alphavirus backbone does not encode structural virion proteins capsid, E2 and E1. In some aspects, the cassette is inserted in place of structural virion proteins within the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus. In some aspects, the Venezuelan equine encephalitis virus comprises the sequence of SEQ ID NO:3 or SEQ ID NO:5. In some aspects, the Venezuelan equine encephalitis virus comprises the sequence of SEQ ID NO:3 or SEQ ID NO:5 further comprising a deletion between base pair 7544 and 11176. In some aspects, the RNA alphavirus backbone comprises the sequence set forth in SEQ ID NO:6 or SEQ ID NO:7. In some aspects, the cassette is inserted at position 7544 to replace the deletion between base pairs 7544 and 11176 as set forth in the sequence of SEQ ID NO:3 or SEQ ID NO:5. In some aspects, the insertion of the cassette provides for transcription of a polycistronic RNA comprising the nsP1-4 genes and the at least one nucleic acid sequence, wherein the nsP1-4 genes and the at least one nucleic acid sequence are in separate open reading frames.

[0026] In some aspects, the at least one promoter nucleotide sequence is the native 26S promoter nucleotide sequence encoded by the RNA alphavirus backbone. In some aspects, the at least one promoter nucleotide sequence is an exogenous RNA promoter. In some aspects, the second promoter nucleotide sequence is a 26S promoter nucleotide sequence. In some aspects, the second promoter nucleotide sequence comprises multiple 26S promoter nucleotide sequences, wherein each 26S promoter nucleotide sequence provides for transcription of one or more of the separate open reading frames.

[0027] In some aspects, one or more of the cassettes are at least 100, 200, 300, 400, 500, 600, 700, 800, or 900 nucleotides in length. In some aspects, one or more of the cassettes are at least 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, or 10000 nucleotides in length. In some aspects, the one or more vectors are capable of driving expression of a cassette that is at least 3500 nucleotides in length. In some aspects, the one or more vectors are capable of driving expression of a cassette that is at least 6000 nucleotides in length.

[0028] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises two or more antigen-encoding nucleic acid sequences. In some aspects, each antigen-encoding nucleic acid sequence is linked directly to one another. In some aspects, each antigen-encoding nucleic acid sequence is linked to a distinct antigen-encoding nucleic acid sequence with a nucleic acid sequence encoding a linker. In some aspects, the linker links two epitope-encoding nucleic acid sequences or an epitope-encoding nucleic acid sequence to an MHC class II epitope-encoding nucleic acid sequence. In some aspects, the linker is selected from the group consisting of: (1) consecutive glycine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (2) consecutive alanine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (3) two arginine residues (RR); (4) alanine, alanine, tyrosine (AAY); (5) a consensus

sequence at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues in length that is processed efficiently by a mammalian proteasome; and (6) one or more native sequences flanking the antigen derived from the cognate protein of origin and that is at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 2-20 amino acid residues in length. In some aspects, the linker links two MHC class II epitope-encoding nucleic acid sequences or an MHC class II sequence to an epitope-encoding nucleic acid sequence. In some aspects, the linker comprises the sequence GPGPG (SEQ ID NO: 56).

[0029] In some aspects, the antigen-encoding nucleic acid sequences is linked, operably or directly, to a separate or contiguous sequence that enhances the expression, stability, cell trafficking, processing and presentation, and/or immunogenicity of the antigen-encoding nucleic acid sequence. In some aspects, the separate or contiguous sequence comprises at least one of: a ubiquitin sequence, a ubiquitin sequence modified to increase proteasome targeting (e.g., the ubiquitin sequence contains a Gly to Ala substitution at position 76), an immunoglobulin signal sequence (e.g., IgK), a major histocompatibility class I sequence, lysosomal-associated membrane protein (LAMP)-1, human dendritic cell lysosomal-associated membrane protein, and a major histocompatibility class II sequence; optionally wherein the ubiquitin sequence modified to increase proteasome targeting is A76.

[0030] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-400 antigen-encoding nucleic acid sequences and wherein at least two of the antigen-encoding nucleic acid sequences encode epitope sequences or portions thereof that are presented by MHC class I on a cell surface. In some aspects, the MHC class I epitopes are presented by MHC class I on the tumor cell surface.

[0031] In some aspects, the epitope-encoding nucleic acid sequences comprises at least one MHC class I epitope-encoding nucleic acid sequence, and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence between 8 and 35 amino acids in length, optionally 9-17, 9-25, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35 amino acids in length.

[0032] In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present. In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present and comprises at least one MHC class II epitope-encoding nucleic acid sequence that comprises at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises an MHC class II epitope-encoding nucleic acid sequence and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence that is 12-20, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 20-40 amino acids in length. In some aspects, the epitope-encoding nucleic acid sequences comprises an MHC class II epitope-encoding nucleic acid sequence, wherein the at least one MHC class II epitope-encoding nucleic acid sequence is present, and wherein the at least one MHC class II epitope-encoding nucleic acid sequence comprises at least one universal MHC class II epitope-encoding nucleic acid sequence, optionally wherein the at least one universal sequence comprises at least one of Tetanus toxoid and PADRE.

[0033] In some aspects, the at least one promoter nucleotide sequence or the second promoter nucleotide sequence is inducible. In some aspects, the at least one promoter nucleotide sequence or the second promoter nucleotide sequence is non-inducible.

[0034] In some aspects, the at least one poly(A) sequence comprises a poly(A) sequence native to the alphavirus. In some aspects, the at least one poly(A) sequence comprises a poly(A) sequence exogenous to the alphavirus. In some aspects, the at least one poly(A) sequence is operably linked to at least one of the at least one nucleic acid sequences. In some aspects, the at least one poly(A) sequence is at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, or at least 90 consecutive A nucleotides (SEQ ID NO: 29361). In some aspects, the at least one poly(A) sequence is at least 100 consecutive A nucleotides (SEQ ID NO: 29358).

[0035] In some aspects, the epitope-encoding nucleic acid sequence comprises a MHC class I epitope-encoding nucleic acid sequence, and wherein the MHC class I epitope-encoding nucleic acid sequence is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the MHC class I epitope-encoding nucleic acid sequence. In some aspects, each of the MHC class I epitope-encoding nucleic acid sequences is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor,

wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the at least 20 MHC class I epitope-encoding nucleic acid sequences. In some aspects, a number of the set of selected epitopes is 2-20. In some aspects, the presentation model represents dependence between: (a) presence of a pair of a particular one of the MHC alleles and a particular amino acid at a particular position of a peptide sequence; and (b) likelihood of presentation on the tumor cell surface, by the particular one of the MHC alleles of the pair, of such a peptide sequence comprising the particular amino acid at the particular position. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being presented on the tumor cell surface relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of inducing a tumor-specific immune response in the subject relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of being presented to naïve T cells by professional antigen presenting cells (APCs) relative to unselected epitopes based on the presentation model, optionally wherein the APC is a dendritic cell (DC). In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being subject to inhibition via central or peripheral tolerance relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being capable of inducing an autoimmune response to normal tissue in the subject relative to unselected epitopes based on the presentation model. In some aspects, exome or transcriptome nucleotide sequencing data is obtained by performing sequencing on the tumor tissue. In some aspects, the sequencing is next generation sequencing (NGS) or any massively parallel sequencing approach.

[0036] Also disclosed herein is a composition for delivery of a chimpanzee adenovirus (ChAdV)-based expression system, wherein the composition for delivery of the ChAdV-based expression system comprises: the ChAdV-based expression system, wherein the ChAdV-based expression system comprises a viral particle comprising a ChAdV vector, wherein the ChAdV vector comprises: (a) a ChAdV backbone, wherein the ChAdV backbone comprises: (i) at least one promoter nucleotide sequence, and (ii) at least one polyadenylation (poly(A)) sequence; and (b) a cassette, wherein the cassette comprises: (i) at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; and wherein the cassette is operably linked to the at least one promoter nucleotide sequence and the at least one poly(A) sequence, and wherein the composition comprises 1×10.sup.12 or less of the viral particles.

[0037] Also disclosed herein is a composition for delivery of a ChAdV-based expression system, wherein the composition for delivery of the ChAdV-based expression system comprises: the ChAdV-based expression system, wherein the ChAdV-based expression system comprises a viral particle comprising a ChAdV vector, wherein the ChAdV vector comprises: (a) a ChAdV backbone, wherein the ChAdV backbone comprises: (i) a modified ChAdV68 sequence comprising at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1, wherein the nucleotides 2 to 36,518 lack: (1) nucleotides 577 to 3403 of the sequence shown in SEQ ID NO: 1 corresponding to an E1 deletion and (2) nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO: I corresponding to an E3 deletion; and optionally lacks (3) nucleotides 34,916 to 35,642 of the sequence shown in SEQ ID NO: 1 corresponding to a partial E4 deletion; (ii) a CMV promoter nucleotide sequence; and (iii) an SV40 polyadenylation (poly(A)) sequence; and (b) a cassette, wherein the cassette comprises: (i) at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; and wherein the cassette is inserted within the E1 deletion and the cassette is operably linked to the CMV promoter nucleotide sequence and the SV40 poly(A) sequence, and wherein the composition comprises 1×10.sup.12 or less of the viral particles.

[0038] In some aspects, the composition for delivery of the ChAdV-based expression system comprises 3×10.sup.11 or less of the viral particles. In some aspects, the composition for delivery of the ChAdV-based expression system comprises at least 1×10.sup.11 of the viral particles. In some aspects, the composition for delivery of the ChAdV-based expression system comprises between 1×10.sup.11 and 1×10.sup.12, between 3×10.sup.11 and 1×10.sup.12, or between 1×10.sup.11 and 3×10.sup.11 of the viral particles. In some aspects, the composition for delivery of the ChAdV-based expression system comprises 1×10.sup.11, **3**×10.sup.11, or 1×10.sup.12 of the viral particles.

[0039] In some aspects, the viral particles are at a concentration of at 5×10.sup.11 vp/mL.

[0040] In some aspects, the epitope-encoding nucleic acid sequence encodes an epitope known or suspected to be presented by MHC class I on a surface of a cell, optionally wherein the surface of the cell is a tumor cell surface or an infected cell surface, and optionally wherein the cell is a subject's cell. In some aspects, the cell is a tumor cell selected

from the group consisting of: lung cancer, melanoma, breast cancer, ovarian cancer, prostate cancer, kidney cancer, gastric cancer, colon cancer, testicular cancer, head and neck cancer, pancreatic cancer, brain cancer, B-cell lymphoma, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, T cell lymphocytic leukemia, non-small cell lung cancer, and small cell lung cancer, or wherein the cell is an infected cell selected from the group consisting of: a pathogen infected cell, a virally infected cell, a bacterially infected cell, an fungally infected cell, and a parasitically infected cell. In some aspects, the virally infected cell is an HIV infected cell.

[0041] In some aspects, an ordered sequence of each element of the cassette in the composition for delivery of the ChAdV-based expression system is described in the formula, from 5′ to 3′, comprising P.sub.a-(L5.sub.b-N.sub.c-L3.sub.d) x-(G5.sub.e-U.sub.f).sub.Y-G3.sub.g wherein P comprises the at least one promoter sequence operably linked to at least one of the at least one antigen-encoding nucleic acid sequences, where a=1, N comprises one of the epitope-encoding nucleic acid sequences, wherein the epitope-encoding nucleic acid sequence comprises an MHC class I epitope-encoding nucleic acid sequence, where c=1, L5 comprises the 5′ linker sequence, where b=0 or 1, L3 comprises the 3′ linker sequence, where d=0 or 1, G5 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where e=0 or 1, G3 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where g=0 or 1, U comprises one of the at least one MHC class II epitope-encoding nucleic acid sequence, where f=1, X=1 to 400, where for each X the corresponding Ne is an MHC class I epitope-encoding nucleic acid sequence, and Y=0, 1, or 2, where for each Y the corresponding Ur is an MHC class II epitope-encoding nucleic acid sequence. In some aspects, for each X the corresponding Ne is a distinct MHC class I epitope-encoding nucleic acid sequence. In some aspects, for each Y the corresponding Ur is a distinct MHC class II epitope-encoding nucleic acid sequence. In some aspects, b=1, d=1, c=1, g=1, h=1, X=at least 2, Y=2, P is a CMV promoter sequence, each N encodes a MHC class I epitope 7-15 amino acids in length, L5 is a native 5′ linker sequence that encodes a native N-terminal amino acid sequence of the MHC I epitope, and wherein the 5′ linker sequence encodes a peptide that is at least 3 amino acids in length, L3 is a native 3′ linker sequence that encodes a native C-terminal amino acid sequence of the MHC I epitope, and wherein the 3′ linker sequence encodes a peptide that is at least 3 amino acids in length, U is each of a PADRE class II sequence and a Tetanus toxoid MHC class II sequence, the ChAdV vector comprises a modified ChAdV68 sequence comprising the sequence of SEQ ID NO: 1 with an E1 (nt 577 to 3403) deletion and an E3 (nt 27,125-31,825) deletion and the neoantigen cassette is inserted within the E1 deletion, and each of the MHC class I antigen-encoding nucleic acid sequences encodes a polypeptide that is 25 amino acids in length.

[0042] In some aspects, the composition for delivery of the ChAdV-based expression system is formulated for intramuscular (IM), intradermal (ID), subcutaneous (SC), or intravenous (IV) administration. In some aspects, the composition for delivery of the ChAdV-based expression system is formulated for intramuscular (IM) administration.

[0043] In some aspects, the cassette is integrated between the at least one promoter nucleotide sequence and the at least one poly(A) sequence. In some aspects, the at least one promoter nucleotide sequence is operably linked to the cassette.

[0044] In some aspects, the ChAdV backbone comprises a ChAdV68 vector backbone. In some aspects, the ChAdV68 vector backbone comprises the sequence set forth in SEQ ID NO:1. In some aspects, the ChAdV68 vector backbone comprises a functional deletion in at least one gene selected from the group consisting of an adenovirus EIA, EIB, E2A, E2B, E3, L1, L2, L3, L4, and L5 gene with reference to a ChAdV68 genome or with reference to the sequence shown in SEQ ID NO: 1, optionally wherein the adenoviral backbone or modified ChAdV68 sequence is fully deleted or functionally deleted in: (1) ElA and EIB; or (2) E1A, EIB, and E3 with reference to the adenovirus genome or with reference to the sequence shown in SEQ ID NO:1, optionally wherein the E1 gene is functionally deleted through an E1 deletion of at least nucleotides 577 to 3403 with reference to the sequence shown in SEQ ID NO: 1 and optionally wherein the E3 gene is functionally deleted through an E3 deletion of at least nucleotides 27,125 to 31,825 with reference to the sequence shown in SEQ ID NO: 1. In some aspects, the ChAdV68 vector backbone comprises one or more genes or regulatory sequences with reference to a ChAdV68 genome or with reference to the sequence shown in SEQ ID NO: 1, optionally wherein the one or more genes or regulatory sequences are selected from the group consisting of the chimpanzee adenovirus inverted terminal repeat (ITR), E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4, and L5 genes. In some aspects, the ChAdV68 vector backbone comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO: 1, wherein the nucleotides 2 to 36,518 lack: (1) nucleotides 577 to 3403 of the sequence shown in SEQ ID NO: 1 corresponding to an E1 deletion and (2) nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO:1 corresponding to an E3 deletion. In some aspects, the ChAdV68 vector backbone comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1, wherein the nucleotides 2 to 36,518 lack: (1) nucleotides 577 to 3403 of the sequence shown in SEQ ID NO:1 corresponding to an E1 deletion; (2) nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO: 1 corresponding to an E3 deletion; and (3) nucleotides 34,916 to 35,642 of the sequence shown in SEQ ID NO: 1 corresponding to a partial E4 deletion; optionally wherein the antigen cassette is inserted within the E1 deletion.

[0045] In some aspects, the cassette is inserted in the ChAdV backbone at the E1 region, E3 region, and/or any deleted AdV region that allows incorporation of the cassette. In some aspects, the ChAdV backbone is generated from one of a first generation, a second generation, or a helper-dependent adenoviral vector.

[0046] In some aspects, the at least one promoter nucleotide sequence is selected from the group consisting of: a CMV,

a SV40, an EF-1, a RSV, a PGK, a HSA, a MCK, and a EBV promoter sequence. In some aspects, the at least one promoter nucleotide sequence is a CMV promoter sequence.

[0047] In some aspects, at least one of the epitope-encoding nucleic acid sequences encodes an epitope that, when expressed and translated, is capable of being presented by MHC class I on a cell of a subject. In some aspects, at least one of the epitope-encoding nucleic acid sequences encodes an epitope that, when expressed and translated, is capable of being presented by MHC class II on a cell of a subject.

[0048] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises two or more antigen-encoding nucleic acid sequences. In some aspects, each antigen-encoding nucleic acid sequence is linked directly to one another. In some aspects, each antigen-encoding nucleic acid sequence is linked to a distinct antigen-encoding nucleic acid sequence with a nucleic acid sequence encoding a linker. In some aspects, the linker links two MHC class I epitope-encoding nucleic acid sequences or an MHC class I epitope-encoding nucleic acid sequence to an MHC class II epitope-encoding nucleic acid sequence. In some aspects, the linker is selected from the group consisting of: (1) consecutive glycine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (2) consecutive alanine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (3) two arginine residues (RR); (4) alanine, alanine, tyrosine (AAY); (5) a consensus sequence at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues in length that is processed efficiently by a mammalian proteasome; and (6) one or more native sequences flanking the antigen derived from the cognate protein of origin and that is at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 2-20 amino acid residues in length. In some aspects, the linker links two MHC class II epitope-encoding nucleic acid sequences or an MHC class II sequence to an MHC class I epitope-encoding nucleic acid sequence. In some aspects, the linker comprises the sequence GPGPG (SEQ ID NO: 56).

[0049] In some aspects, the antigen-encoding nucleic acid sequences is linked, operably or directly, to a separate or contiguous sequence that enhances the expression, stability, cell trafficking, processing and presentation, and/or immunogenicity of the antigen-encoding nucleic acid sequence. In some aspects, the separate or contiguous sequence comprises at least one of: a ubiquitin sequence, a ubiquitin sequence modified to increase proteasome targeting (e.g., the ubiquitin sequence contains a Gly to Ala substitution at position 76), an immunoglobulin signal sequence (e.g., IgK), a major histocompatibility class I sequence, lysosomal-associated membrane protein (LAMP)-1, human dendritic cell lysosomal-associated membrane protein, and a major histocompatibility class II sequence; optionally wherein the ubiquitin sequence modified to increase proteasome targeting is A76.

[0050] In some aspects, the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have increased binding affinity to its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have increased binding stability to its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have an increased likelihood of presentation on its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence. In some aspects, the at least one alteration comprises a point mutation, a frameshift mutation, a non-frameshift mutation, a deletion mutation, an insertion mutation, a splice variant, a genomic rearrangement, or a proteasome-generated spliced antigen. In some aspects, the epitope-encoding nucleic acid sequence encodes an epitope known or suspected to be expressed in a subject known or suspected to have cancer. In some aspects, the cancer is a solid tumor. In some aspects, the cancer is selected from the group consisting of: MSS-CRC, NSCLC, and PDA. In some aspects, the cancer is selected from the group consisting of: lung cancer, melanoma, breast cancer, ovarian cancer, prostate cancer, kidney cancer, gastric cancer, colon cancer, testicular cancer, head and neck cancer, pancreatic cancer, bladder cancer, brain cancer, B-cell lymphoma, acute myelogenous leukemia, adult acute lymphoblastic leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, T cell lymphocytic leukemia, non-small cell lung cancer, and small cell lung cancer.

[0051] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-400 antigen-encoding nucleic acid sequences and wherein at least two of the antigen-encoding nucleic acid sequences encode epitope sequences or portions thereof that are presented by MHC class I on a cell surface. In some aspects, at least two of the MHC class I epitopes are presented by MHC class I on the tumor cell surface. In some aspects, the epitope-encoding nucleic acid sequences comprises at least one MHC class I epitope-encoding nucleic acid sequence, and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence between 8 and 35 amino acids in length, optionally 9-17, 9-25, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35 amino acids in length.

[0052] In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present. In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present and comprises at least one MHC class II epitope-encoding nucleic acid sequence that comprises at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises an MHC class II epitope-encoding nucleic acid sequence and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence that is 12-20, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 20-40 amino acids in length. In some aspects, the epitope-encoding nucleic acid sequences comprises an MHC class II epitope-encoding nucleic acid sequence, wherein the at least one MHC class II epitope-encoding nucleic acid sequence is present, and wherein the at least one MHC class II epitope-encoding nucleic acid sequence comprises at least one universal MHC class II epitope-encoding nucleic acid sequence, optionally wherein the at least one universal sequence comprises at least one of Tetanus toxoid and PADRE.

[0053] In some aspects, the at least one promoter nucleotide sequence is inducible. In some aspects, the at least one promoter nucleotide sequence is non-inducible.

[0054] In some aspects, the at least one poly(A) sequence comprises a Bovine Growth Hormone (BGH) SV40 polyA sequence. In some aspects, the at least one poly(A) sequence is at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, or at least 90 consecutive A nucleotides (SEQ ID NO: 29361). In some aspects, the at least one poly(A) sequence is at least 100 consecutive A nucleotides (SEQ ID NO: 29358).

[0055] In some aspects, the cassette further comprises at least one of: an intron sequence, a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) sequence, an internal ribosome entry sequence (IRES) sequence, a nucleotide sequence encoding a 2A self cleaving peptide sequence, a nucleotide sequence encoding a Furin cleavage site, or a sequence in the 5′ or 3′ non-coding region known to enhance the nuclear export, stability, or translation efficiency of mRNA that is operably linked to at least one of the at least one antigen-encoding nucleic acid sequences. In some aspects, the cassette further comprises a reporter gene, including but not limited to, green fluorescent protein (GFP), a GFP variant, secreted alkaline phosphatase, luciferase, a luciferase variant, or a detectable peptide or epitope. In some aspects, the detectable peptide or epitope is selected from the group consisting of an HA tag, a Flag tag, a His-tag, or a V5 tag.

[0056] In some aspects, the one or more vectors further comprises one or more nucleic acid sequences encoding at least one immune modulator. In some aspects, the immune modulator is an anti-CTLA4 antibody or an antigen-binding fragment thereof, an anti-PD-1 antibody or an antigen-binding fragment thereof, an anti-PD-L1 antibody or an antigen-binding fragment thereof, an anti-4-1BB antibody or an antigen-binding fragment thereof, or an anti-OX-40 antibody or an antigen-binding fragment thereof. In some aspects, the antibody or antigen-binding fragment thereof is a Fab fragment, a Fab′ fragment, a single chain Fv (scFv), a single domain antibody (sdAb) either as single specific or multiple specificities linked together (e.g., camelid antibody domains), or full-length single-chain antibody (e.g., full-length IgG with heavy and light chains linked by a flexible linker). In some aspects, the heavy and light chain sequences of the antibody are a contiguous sequence separated by either a self-cleaving sequence such as 2A or IRES; or the heavy and light chain sequences of the antibody are linked by a flexible linker such as consecutive glycine residues. In some aspects, the immune modulator is a cytokine. In some aspects, the cytokine is at least one of IL-2, IL-7, IL-12, IL-15, or IL-21 or variants thereof of each.

[0057] In some aspects, the epitope-encoding nucleic acid sequence comprises a MHC class I epitope-encoding nucleic acid sequence, and wherein the MHC class I epitope-encoding nucleic acid sequence is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the MHC class I epitope-encoding nucleic acid sequence. In some aspects, each of the MHC class I epitope-encoding nucleic acid sequences is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the at least 20 MHC class I epitope-encoding nucleic acid sequences. In some aspects, a number of the set of selected epitopes is 2-20. In some aspects, the presentation model represents dependence between: (a) presence of a pair of a particular one of the MHC alleles and a particular amino acid at a particular position of a peptide sequence; and

[0058] (b) likelihood of presentation on the tumor cell surface, by the particular one of the MHC alleles of the pair, of such a peptide sequence comprising the particular amino acid at the particular position. In some aspects, selecting the

set of selected epitopes comprises selecting epitopes that have an increased likelihood of being presented on the tumor cell surface relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of inducing a tumor-specific immune response in the subject relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of being presented to naïve T cells by professional antigen presenting cells (APCs) relative to unselected epitopes based on the presentation model, optionally wherein the APC is a dendritic cell (DC). In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being subject to inhibition via central or peripheral tolerance relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being capable of inducing an autoimmune response to normal tissue in the subject relative to unselected epitopes based on the presentation model. In some aspects, exome or transcriptome nucleotide sequencing data is obtained by performing sequencing on the tumor tissue. In some aspects, the sequencing is next generation sequencing (NGS) or any massively parallel sequencing approach.

[0059] In some aspects, the cassette comprises junctional epitope sequences formed by adjacent sequences in the cassette. In some aspects, at least one or each junctional epitope sequence has an affinity of greater than 500 nM for MHC. In some aspects, each junctional epitope sequence is non-self. In some aspects, the cassette does not encode a non-therapeutic MHC class I or class II epitope nucleic acid sequence comprising a translated, wild-type nucleic acid sequence, wherein the non-therapeutic epitope is predicted to be displayed on an MHC allele of the subject. In some aspects, the non-therapeutic predicted MHC class I or class II epitope sequence is a junctional epitope sequence formed by adjacent sequences in the cassette. In some aspects, the prediction is based on presentation likelihoods generated by inputting sequences of the non-therapeutic epitopes into a presentation model. In some aspects, an order of the antigen-encoding nucleic acid sequences in the cassette is determined by a series of steps comprising: (a) generating a set of candidate cassette sequences corresponding to different orders of the antigen-encoding nucleic acid sequences; (b) determining, for each candidate cassette sequence, a presentation score based on presentation of non-therapeutic epitopes in the candidate cassette sequence; and (c) selecting a candidate cassette sequence associated with a presentation score below a predetermined threshold as the cassette sequence for a vaccine.

[0060] In some aspects, the composition for delivery of the ChAdV-based expression system is formulated in a pharmaceutical composition comprising a pharmaceutically acceptable carrier.

[0061] In some aspects, one or more of the epitope-encoding nucleic acid sequences are derived from a tumor of a subject. In some aspects, each of the epitope-encoding nucleic acid sequences are derived from a tumor of a subject. In some aspects, one or more of the epitope-encoding nucleic acid sequences are not derived from a tumor of a subject. In some aspects, each of the epitope-encoding nucleic acid sequences are not derived from a tumor of a subject. In some aspects, the epitope-encoding nucleic acid sequence comprises an epitope selected from the group consisting of SEQ ID NO: 57-29,357 and SEQ ID NO: 29,512-29,519. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 tumor-specific MHC class I antigen-encoding nucleic acid sequences linearly linked to each other, comprising: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 14,954; 19,848; and 19,850, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; and 19,863, and (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974, wherein each of the tumor-specific MHC class I antigen-encoding nucleic acid sequences comprises a class I epitope encoding nucleic acid sequence, optionally wherein each MHC I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 57-29,357 and SEQ ID NO: 29,512-29,519.

[0062] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 20 tumor-specific MHC class I antigen-encoding nucleic acid sequences linearly linked to each other, comprising: (A) a KRAS_G12A MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 14,954; 19,848; and 19,850, (C) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; 19,863, and (D) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974, (E) a KRAS_G13D MHC class I epitope encoding nucleic acid sequence, (F) a KRAS_Q61K MHC class I epitope encoding nucleic acid sequence, (G) a TP53_R249M MHC class I epitope encoding nucleic acid sequence, (H) a CTNNB1_S45P MHC class I epitope encoding nucleic acid sequence, (I) a CTNNB1_S45F MHC class I epitope encoding nucleic acid sequence, (J) a ERBB2_Y772_A775dup MHC class I

epitope encoding nucleic acid sequence, (K) a KRAS_Q61R MHC class I epitope encoding nucleic acid sequence, (L) a CTNNB1_T41A MHC class I epitope encoding nucleic acid sequence, (M) a TP53_K132N MHC class I epitope encoding nucleic acid sequence, (N) a KRAS_Q61L MHC class I epitope encoding nucleic acid sequence, (O) a TP53_R213L MHC class I epitope encoding nucleic acid sequence, (P) a BRAF_G466V MHC class I epitope encoding nucleic acid sequence, (Q) a KRAS_Q61H MHC class I epitope encoding nucleic acid sequence, (R) a CTNNB1_S37F MHC class I epitope encoding nucleic acid sequence, (S) a TP53_S127Y MHC class I epitope encoding nucleic acid sequence, (T) a TP53_K132E MHC class I epitope encoding nucleic acid sequence, and (U) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence.

[0063] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, D) a KRAS Q61H MHC class I epitope encoding nucleic acid sequence, or combinations thereof. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, and (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, and (D) a KRAS Q61H MHC class I epitope encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope comprising the sequence of SEQ ID NO: 14,954; 19,848; 19,850, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; 19,863, and (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope comprising the sequence of SEQ ID NO: 14,954; 19,848; 19,850, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; 19,863, (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974, and (D) a KRAS Q61H MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_Q61H MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 14,409.

[0064] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises: (A) a TP53_R213L MHC class I epitope encoding nucleic acid sequence, (B) a TP53_S127Y MHC class I epitope encoding nucleic acid sequence, (C) a TP53_R249M MHC class I epitope encoding nucleic acid sequence, or combinations thereof. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a TP53_R213L MHC class I epitope encoding nucleic acid sequence, (B) a TP53_S127Y MHC class I epitope encoding nucleic acid sequence, and (C) a TP53_R249M MHC class I epitope encoding nucleic acid sequence. In some aspects, the TP53_R213L MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 29,519.

[0065] Also provided for herein is a kit comprising any of the compositions for delivery of the ChAdV-based expression system described herein, and instructions for use.

[0066] Also disclosed herein is a method for stimulating an immune response in a subject, the method comprising administering to the subject a composition for delivery of a self-amplifying alphavirus-based expression system and administering to the subject a composition for delivery of a chimpanzee adenovirus (ChAdV)-based expression system, and wherein either: a. the composition for delivery of the ChAdV-based expression system comprises the ChAdV-based expression system, wherein the ChAdV-based expression system comprises a viral particle comprising a ChAdV vector, and wherein the composition comprises 1×10.sup.12 or less of the viral particles, b, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, and wherein the composition comprises a therapeutically effective amount comprising 30 µg or less of each of the one or more vectors, or c. the composition for delivery of the ChAdV-based expression system comprises the ChAdV-based expression system, wherein the ChAdV-based expression system comprises a viral particle comprising a ChAdV vector, and wherein the composition comprises 1×10.sup.12 or less of the viral particles and wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, and

wherein the composition comprises a therapeutically effective amount comprising 30 µg or less of each of the one or more vectors.

[0067] In some aspects, the composition for delivery of the ChAdV-based expression system is administered as a priming dose and the composition for delivery of the self-amplifying alphavirus-based expression system is administered as one or more boosting doses. In some aspects, the priming dose is administered on day 1 and the one or more boosting doses are administered every 4 weeks (Q4W) following the priming dose. In some aspects, the one or more boosting doses are administered every 4 weeks for a time period. In some aspects, the time period is the first 6 months following the priming dose. In some aspects, one or more additional boosting doses are administered at a second interval following the time period. In some aspects, the second interval is every 3 months. In some aspects, two or more boosting doses are administered. In some aspects, 1, 2, 3, 4, 5, 6, 7, or 8 boosting doses are administered.

[0068] In some aspects, the composition for delivery of the ChAdV-based expression system is administered intramuscularly (IM), intradermally (ID), subcutaneously (SC), or intravenously (IV). In some aspects, the composition for delivery of the ChAdV-based expression system is administered (IM). In some aspects, the IM administration is administered at separate injection sites. In some aspects, the separate injection sites are in opposing deltoid muscles. In some aspects, the separate injection sites are in gluteus or rectus femoris sites on each side.

[0069] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system is administered intramuscularly (IM), intradermally (ID), subcutaneously (SC), or intravenously (IV). In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system is administered (IM). In some aspects, the IM administration is administered at separate injection sites. In some aspects, the separate injection sites are in opposing deltoid muscles. In some aspects, the separate injection sites are in gluteus or rectus femoris sites on each side. In some aspects, the injection site of the one or more boosting doses is as close as possible to the injection site of the priming dose.

[0070] In some aspects, the method further comprises determining or having determined the HLA-haplotype of the subject.

[0071] In some aspects, the method further comprises administering nivolumab. In some aspects, nivolumab is administered as an intravenous (IV) infusion. In some aspects, nivolumab is administered at a dose of 480 mg. In some aspects, nivolumab is administered on day 1. In some aspects, nivolumab is on administered day 1 and administered every 4 weeks (Q4W) following the priming dose. In some aspects, nivolumab is on administered on the same day as the priming dose or on the same day as the one or more boosting doses. In some aspects, nivolumab is formulated in solution at 10 mg/mL.

[0072] In some aspects, the method further comprises administering ipilimumab. In some aspects, ipilimumab is administered an intravenous (IV) infusion. In some aspects, ipilimumab is administered subcutaneously (SC). In some aspects, the SC administration is injected proximally (within ˜2 cm) to one or more of the priming dose injection site or the one or more boosting dose injection sites. In some aspects, the SC administration is administered as 4 separate injections or administered as 6 separate injections. In some aspects, ipilimumab is administered at a dose of 30 mg. In some aspects, ipilimumab is administered on day 1. In some aspects, ipilimumab is on administered day 1 and administered every 4 weeks (Q4W) following the priming dose. In some aspects, ipilimumab is on administered on the same day as the priming dose or on the same day as the one or more boosting doses. In some aspects, ipilimumab is formulated in solution at 5 mg/mL.

[0073] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises: (A) the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, wherein the one or more vectors comprises: (a) an RNA alphavirus backbone, wherein the RNA alphavirus backbone comprises: (i) at least one promoter nucleotide sequence, and (ii) at least one polyadenylation (poly(A)) sequence; and (b) a cassette, wherein the cassette comprises: (i) at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; (ii) optionally, a second promoter nucleotide sequence operably linked to the at least one antigen-encoding nucleic acid sequence; and (iii) optionally, at least one second poly(A) sequence, wherein the second poly(A) sequence is a native poly(A) sequence or an exogenous poly(A) sequence to the alphavirus, and (B) a lipid-nanoparticle (LNP), wherein the LNP encapsulates the self-amplifying alphavirus-based expression system.

[0074] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises, (A) the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, wherein the one or more vectors comprises: (a) an RNA alphavirus backbone, wherein the RNA alphavirus backbone comprises the nucleic acid sequence set forth in SEQ ID NO:6, wherein the RNA alphavirus backbone sequence comprises a 26S promoter nucleotide sequence and a poly(A) sequence, wherein the 26S promoter sequence is endogenous to the RNA alphavirus backbone, and wherein the poly(A) sequence is endogenous to the RNA alphavirus backbone; and (b) a cassette integrated between the 26S promoter nucleotide sequence and the poly(A) sequence, wherein the cassette is operably linked to the 26S promoter nucleotide sequence, and wherein the cassette comprises at least one antigen-encoding nucleic acid sequence

comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; and (B) a lipid-nanoparticle (LNP), wherein the LNP encapsulates the self-amplifying alphavirus-based expression system.

[0075] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises a therapeutically effective amount comprising 30 µg or less in total for all of the one or more vectors combined. In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises 30 µg of each of the one or more vectors. In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises 30 µg in total for all of the one or more vectors combined of each of the one or more vectors. In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises 30 µg in total for all of the one or more vectors combined, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises only a single distinct vector.

[0076] In some aspects, the composition for delivery of the self-amplifying alphavirus-based expression system comprises only a single distinct vector.

[0077] In some aspects, the therapeutically effective amount is capable of stimulating a T cell response following administration to a subject. In some aspects, the T cell response is assessed by monitoring ex vivo stimulation with the encoded epitope in PBMCs from the subject. In some aspects, monitoring ex vivo stimulation comprises an ELISpot assay, optionally wherein the ELISpot assay quantifies IFN-gamma production.

[0078] In some aspects, the therapeutically effective amount does not stimulate an inhibitory IFNα response and/or reduces the inhibitory IFNα response following administration to a subject.

[0079] In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is between 10-40 to 1. In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is between 16-32 to 1. In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is about 24 to 1. In some aspects, the weight to weight ratio of the LNP to total weight of the one or more vectors is 24 to 1.

[0080] In some aspects, the one or more vectors is at a concentration of 1 mg/mL.

[0081] In some aspects, an ordered sequence of each element of the cassette in the composition for delivery of the self-amplifying alphavirus-based expression system is described in the formula, from 5′ to 3′, comprising $P_a\text{-}(L5_b\text{-}N_c\text{-}L3_d)_X\text{-}(G5_e\text{-}U_f)_Y\text{-}G3_g$ wherein P comprises the second promoter nucleotide sequence, where a=0 or 1, N comprises one of the epitope-encoding nucleic acid sequences, wherein the epitope-encoding nucleic acid sequence comprises an MHC class I epitope-encoding nucleic acid sequence, where c=1, L5 comprises the 5′ linker sequence, where b=0 or 1, L3 comprises the 3′ linker sequence, where d=0 or 1, G5 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where e=0 or 1, G3 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where g=0 or 1, U comprises one of the at least one MHC class II epitope-encoding nucleic acid sequence, where f=1, X=1 to 400, where for each X the corresponding Ne is an MHC class I epitope-encoding nucleic acid sequence, and Y=0, 1, or 2, where for each Y the corresponding Ur is an MHC class II epitope-encoding nucleic acid sequence. In some aspects, for each X the corresponding Ne is a distinct MHC class I epitope-encoding nucleic acid sequence. In some aspects, for each Y the corresponding Ur is a distinct MHC class II epitope-encoding nucleic acid sequence. In some aspects, a=0, b=1, d=1, c=1, g=1, h=1, X=at least 2, Y=2, the at least one promoter nucleotide sequence is a single 26S promoter nucleotide sequence provided by the RNA alphavirus backbone, the at least one polyadenylation poly(A) sequence is a poly(A) sequence of at least 100 consecutive A nucleotides (SEQ ID NO: 29358) provided by the RNA alphavirus backbone, the cassette is integrated between the 26S promoter nucleotide sequence and the poly(A) sequence, wherein the cassette is operably linked to the 26S promoter nucleotide sequence and the poly(A) sequence, each N encodes a MHC class I epitope 7-15 amino acids in length, L5 is a native 5′ linker sequence that encodes a native N-terminal amino acid sequence of the MHC I epitope, and wherein the 5′ linker sequence encodes a peptide that is at least 3 amino acids in length, L3 is a native 3′ linker sequence that encodes a native C-terminal amino acid sequence of the MHC I epitope, and wherein the 3′ linker sequence encodes a peptide that is at least 3 amino acids in length, U is each of a PADRE class II sequence and a Tetanus toxoid MHC class II sequence, the RNA alphavirus backbone is the sequence set forth in SEQ ID NO: 6, and each of the MHC class I epitope-encoding nucleic acid sequences encodes a polypeptide that is between 13 and 25 amino acids in length.

[0082] In some aspects, the LNP comprises a lipid selected from the group consisting of: an ionizable amino lipid, a phosphatidylcholine, cholesterol, a PEG-based coat lipid, or a combination thereof. In some aspects, the LNP comprises an ionizable amino lipid, a phosphatidylcholine, cholesterol, and a PEG-based coat lipid. In some aspects, the ionizable amino lipids comprise MC3-like (dilinoleylmethyl-4-dimethylaminobutyrate) molecules. In some aspects, the LNP-encapsulated expression system has a diameter of about 100 nm.

[0083] In some aspects, the cassette is integrated between the at least one promoter nucleotide sequence and the at least one poly(A) sequence.

[0084] In some aspects, the at least one promoter nucleotide sequence is operably linked to the cassette.

[0085] In some aspects, the one or more vectors comprise one or more +-stranded RNA vectors. In some aspects, the one or more +-stranded RNA vectors comprise a 5′ 7-methylguanosine (m7g) cap. In some aspects, the one or more +-

stranded RNA vectors are produced by in vitro transcription. In some aspects, the one or more vectors are self-amplifying within a mammalian cell. In some aspects, the RNA alphavirus backbone comprises at least one nucleotide sequence of an Aura virus, a Fort Morgan virus, a Venezuelan equine encephalitis virus, a Ross River virus, a Semliki Forest virus, a Sindbis virus, or a Mayaro virus. In some aspects, the RNA alphavirus backbone comprises at least one nucleotide sequence of a Venezuelan equine encephalitis virus. In some aspects, the RNA alphavirus backbone comprises at least sequences for nonstructural protein-mediated amplification, a 26S promoter sequence, a poly(A) sequence, a nonstructural protein 1 (nsP1) gene, a nsP2 gene, a nsP3 gene, and a nsP4 gene encoded by the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus. In some aspects, the RNA alphavirus backbone comprises at least sequences for nonstructural protein-mediated amplification, a 26S promoter sequence, and a poly(A) sequence encoded by the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus. In some aspects, sequences for nonstructural protein-mediated amplification are selected from the group consisting of: an alphavirus 5′ UTR, a 51-nt CSE, a 24-nt CSE, a 26S subgenomic promoter sequence, a 19-nt CSE, an alphavirus 3′ UTR, or combinations thereof. In some aspects, the RNA alphavirus backbone does not encode structural virion proteins capsid, E2 and E1. In some aspects, the cassette is inserted in place of structural virion proteins within the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus. In some aspects, the Venezuelan equine encephalitis virus comprises the sequence of SEQ ID NO:3 or SEQ ID NO:5. In some aspects, the Venezuelan equine encephalitis virus comprises the sequence of SEQ ID NO:3 or SEQ ID NO:5 further comprising a deletion between base pair 7544 and 11176. In some aspects, the RNA alphavirus backbone comprises the sequence set forth in SEQ ID NO:6 or SEQ ID NO:7. In some aspects, the cassette is inserted at position 7544 to replace the deletion between base pairs 7544 and 11176 as set forth in the sequence of SEQ ID NO:3 or SEQ ID NO:5. In some aspects, the insertion of the cassette provides for transcription of a polycistronic RNA comprising the nsP1-4 genes and the at least one nucleic acid sequence, wherein the nsP1-4 genes and the at least one nucleic acid sequence are in separate open reading frames.

[0086] In some aspects, the at least one promoter nucleotide sequence is the native 26S promoter nucleotide sequence encoded by the RNA alphavirus backbone. In some aspects, the at least one promoter nucleotide sequence is an exogenous RNA promoter. In some aspects, the second promoter nucleotide sequence is a 26S promoter nucleotide sequence. In some aspects, the second promoter nucleotide sequence comprises multiple 26S promoter nucleotide sequences, wherein each 26S promoter nucleotide sequence provides for transcription of one or more of the separate open reading frames.

[0087] In some aspects, one or more of the cassettes are at least 100, 200, 300, 400, 500, 600, 700, 800, or 900 nucleotides in length. In some aspects, one or more of the cassettes are at least 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, or 10000 nucleotides in length. In some aspects, the one or more vectors are capable of driving expression of a cassette that is at least 3500 nucleotides in length. In some aspects, the one or more vectors are capable of driving expression of a cassette that is at least 6000 nucleotides in length.

[0088] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises two or more antigen-encoding nucleic acid sequences. In some aspects, each antigen-encoding nucleic acid sequence is linked directly to one another. In some aspects, each antigen-encoding nucleic acid sequence is linked to a distinct antigen-encoding nucleic acid sequence with a nucleic acid sequence encoding a linker. In some aspects, the linker links two MHC class I epitope-encoding nucleic acid sequences or an MHC class I epitope-encoding nucleic acid sequence to an MHC class II epitope-encoding nucleic acid sequence. In some aspects, the linker is selected from the group consisting of: (1) consecutive glycine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (2) consecutive alanine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (3) two arginine residues (RR); (4) alanine, alanine, tyrosine (AAY); (5) a consensus sequence at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues in length that is processed efficiently by a mammalian proteasome; and (6) one or more native sequences flanking the antigen derived from the cognate protein of origin and that is at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 2-20 amino acid residues in length. In some aspects, the linker links two MHC class II epitope-encoding nucleic acid sequences or an MHC class II sequence to an MHC class I epitope-encoding nucleic acid sequence. In some aspects, the linker comprises the sequence GPGPG (SEQ ID NO: 56).

[0089] In some aspects, the antigen-encoding nucleic acid sequences is linked, operably or directly, to a separate or contiguous sequence that enhances the expression, stability, cell trafficking, processing and presentation, and/or immunogenicity of the antigen-encoding nucleic acid sequence. In some aspects, the separate or contiguous sequence comprises at least one of: a ubiquitin sequence, a ubiquitin sequence modified to increase proteasome targeting (e.g., the ubiquitin sequence contains a Gly to Ala substitution at position 76), an immunoglobulin signal sequence (e.g., IgK), a major histocompatibility class I sequence, lysosomal-associated membrane protein (LAMP)-1, human dendritic cell lysosomal-associated membrane protein, and a major histocompatibility class II sequence; optionally wherein the ubiquitin sequence modified to increase proteasome targeting is A76.

[0090] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence

encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-400 antigen-encoding nucleic acid sequences and wherein at least two of the antigen-encoding nucleic acid sequences encode epitope sequences or portions thereof that are presented by MHC class I on a cell surface. In some aspects, at least two of the MHC class I epitopes are presented by MHC class I on the tumor cell surface.

[0091] In some aspects, the epitope-encoding nucleic acid sequences comprises at least one MHC class I epitope-encoding nucleic acid sequence, and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence between 8 and 35 amino acids in length, optionally 9-17, 9-25, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35 amino acids in length.

[0092] In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present. In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present and comprises at least one MHC class II epitope-encoding nucleic acid sequence that comprises at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises an MHC class II epitope-encoding nucleic acid sequence and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence that is 12-20, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 20-40 amino acids in length. In some aspects, the epitope-encoding nucleic acid sequences comprises an MHC class II epitope-encoding nucleic acid sequence, wherein the at least one MHC class II epitope-encoding nucleic acid sequence is present, and wherein the at least one MHC class II epitope-encoding nucleic acid sequence comprises at least one universal MHC class II epitope-encoding nucleic acid sequence, optionally wherein the at least one universal sequence comprises at least one of Tetanus toxoid and PADRE.

[0093] In some aspects, the at least one promoter nucleotide sequence or the second promoter nucleotide sequence is inducible. In some aspects, the at least one promoter nucleotide sequence or the second promoter nucleotide sequence is non-inducible. In some aspects, the at least one poly(A) sequence comprises a poly(A) sequence native to the alphavirus. In some aspects, the at least In some aspects, the at least one poly(A) sequence is operably linked to at least one of the at least one nucleic acid sequences. In some aspects, the at least one poly(A) sequence is at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, or at least 90 consecutive A nucleotides (SEQ ID NO: 29361). In some aspects, the at least one poly(A) sequence is at least 100 consecutive A nucleotides (SEQ ID NO: 29358).

[0094] In some aspects, the epitope-encoding nucleic acid sequence comprises a MHC class I epitope-encoding nucleic acid sequence, and wherein the MHC class I epitope-encoding nucleic acid sequence is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the MHC class I epitope-encoding nucleic acid sequence. In some aspects, each of the MHC class I epitope-encoding nucleic acid sequences is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the at least 20 MHC class I epitope-encoding nucleic acid sequences. In some aspects, a number of the set of selected epitopes is 2-20. In some aspects, the presentation model represents dependence between: (a) presence of a pair of a particular one of the MHC alleles and a particular amino acid at a particular position of a peptide sequence; and (b) likelihood of presentation on the tumor cell surface, by the particular one of the MHC alleles of the pair, of such a peptide sequence comprising the particular amino acid at the particular position. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being presented on the tumor cell surface relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of inducing a tumor-specific immune response in the subject relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of being presented to naïve T cells by professional antigen presenting cells

(APCs) relative to unselected epitopes based on the presentation model, optionally wherein the APC is a dendritic cell (DC). In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being subject to inhibition via central or peripheral tolerance relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being capable of inducing an autoimmune response to normal tissue in the subject relative to unselected epitopes based on the presentation model. In some aspects, exome or transcriptome nucleotide sequencing data is obtained by performing sequencing on the tumor tissue. In some aspects, the sequencing is next generation sequencing (NGS) or any massively parallel sequencing approach.

[0095] In some aspects, the ChAdV vector comprises: (a) a ChAdV backbone, wherein the ChAdV backbone comprises: (i) at least one promoter nucleotide sequence, and (ii) at least one polyadenylation (poly(A)) sequence; and (b) a cassette, wherein the cassette comprises: (i) at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; and wherein the cassette is operably linked to the at least one promoter nucleotide sequence and the at least one poly(A) sequence.

[0096] In some aspects, the ChAdV vector comprises: (a) a ChAdV backbone, wherein the ChAdV backbone comprises: (i) a modified ChAdV68 sequence comprising at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1, wherein the nucleotides 2 to 36,518 lack: (1) nucleotides 577 to 3403 of the sequence shown in SEQ ID NO: 1 corresponding to an E1 deletion and (2) nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO:1 corresponding to an E3 deletion; and optionally lacks (3) nucleotides 34,916 to 35,642 of the sequence shown in SEQ ID NO: 1 corresponding to a partial E4 deletion; (ii) a CMV promoter nucleotide sequence; and (iii) an SV40 polyadenylation (poly(A)) sequence; and (b) a cassette, wherein the cassette comprises: (i) at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; and wherein the cassette is inserted within the E1 deletion and the cassette is operably linked to the CMV promoter nucleotide sequence and the SV40 poly(A) sequence.

[0097] In some aspects, the composition for delivery of the ChAdV-based expression system comprises 3×10^{11} or less of the viral particles. In some aspects, the composition for delivery of the ChAdV-based expression system comprises at least 1×10^{11} of the viral particles. In some aspects, the composition for delivery of the ChAdV-based expression system comprises between 1×10^{11} and 1×10^{12}, between 3×10^{11} and 1×10^{12}, or between 1×10^{11} and 3×10^{11} of the viral particles. In some aspects, the composition for delivery of the ChAdV-based expression system comprises 1×10^{11}, 3×10^{11}, or 1×10^{12} of the viral particles. In some aspects, the viral particles are at a concentration of at 5×10^{11} vp/mL.

[0098] In some aspects, the epitope-encoding nucleic acid sequence encodes an epitope known or suspected to be presented by MHC class I on a surface of a cell, optionally wherein the surface of the cell is a tumor cell surface or an infected cell surface, and optionally wherein the cell is the subject's cell. In some aspects, the cell is a tumor cell selected from the group consisting of: lung cancer, melanoma, breast cancer, ovarian cancer, prostate cancer, kidney cancer, gastric cancer, colon cancer, testicular cancer, head and neck cancer, pancreatic cancer, brain cancer, B-cell lymphoma, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, T cell lymphocytic leukemia, non-small cell lung cancer, and small cell lung cancer, or wherein the cell is an infected cell selected from the group consisting of: a pathogen infected cell, a virally infected cell, a bacterially infected cell, an fungally infected cell, and a parasitically infected cell. In some aspects, the virally infected cell is an HIV infected cell.

[0099] In some aspects, an ordered sequence of each element of the cassette in the composition for delivery of the ChAdV-based expression system is described in the formula, from 5′ to 3′, comprising $P_a-(L5_b-N_c-L3_d)_X-(G5_e-U_f)_Y-G3_g$ wherein P comprises the at least one promoter sequence operably linked to at least one of the at least one antigen-encoding nucleic acid sequences, where a=1, N comprises one of the epitope-encoding nucleic acid sequences, wherein the epitope-encoding nucleic acid sequence comprises an MHC class I epitope-encoding nucleic acid sequence, where c=1, L5 comprises the 5′ linker sequence, where b=0 or 1, L3 comprises the 3′ linker sequence, where d=0 or 1, G5 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where e=0 or 1, G3 comprises one of the at least one nucleic acid sequences encoding a GPGPG amino acid linker (SEQ ID NO: 56), where g=0 or 1, U comprises one of the at least one MHC class II epitope-encoding nucleic acid sequence, where f=1, X=1 to 400, where for each X the corresponding Ne is an MHC class I epitope-encoding nucleic acid sequence, and Y=0, 1, or 2, where for each Y the corresponding Ur is an MHC class II epitope-encoding nucleic acid sequence. In some aspects, for each X the corresponding Ne is a distinct MHC class I epitope-encoding nucleic acid sequence. In some aspects, for each Y the corresponding Ur is a distinct MHC class II epitope-encoding nucleic acid sequence. In some aspects, b=1, d=1, e=1, g=1, h=1, X=at least 2, Y=2, P is a CMV promoter sequence, each N encodes a MHC class I epitope 7-15 amino acids in length, L5 is a native 5′ linker sequence that encodes a native N-terminal amino acid sequence of the MHC I epitope, and wherein the 5′ linker sequence encodes a peptide that is at least 3 amino acids in length, L3 is a native 3′ linker sequence that encodes a

native C-terminal amino acid sequence of the MHC I epitope, and wherein the 3′ linker sequence encodes a peptide that is at least 3 amino acids in length, U is each of a PADRE class II sequence and a Tetanus toxoid MHC class II sequence, the ChAdV vector comprises a modified ChAdV68 sequence comprising the sequence of SEQ ID NO:1 with an E1 (nt 577 to 3403) deletion and an E3 (nt 27,125-31,825) deletion and the neoantigen cassette is inserted within the E1 deletion, and each of the MHC class I antigen-encoding nucleic acid sequences encodes a polypeptide that is 25 amino acids in length.

[0100] In some aspects, the cassette is integrated between the at least one promoter nucleotide sequence and the at least one poly(A) sequence. In some aspects, the at least one promoter nucleotide sequence is operably linked to the cassette.

[0101] In some aspects, the ChAdV backbone comprises a ChAdV68 vector backbone. In some aspects, the ChAdV68 vector backbone comprises the sequence set forth in SEQ ID NO:1. In some aspects, the ChAdV68 vector backbone comprises a functional deletion in at least one gene selected from the group consisting of an adenovirus E1A, E1B, E2A, E2B, E3, L1, L2, L3, L4, and L5 gene with reference to a ChAdV68 genome or with reference to the sequence shown in SEQ ID NO: 1, optionally wherein the adenoviral backbone or modified ChAdV68 sequence is fully deleted or functionally deleted in: (1) E1A and E1B; or (2) E1A, E1B, and E3 with reference to the adenovirus genome or with reference to the sequence shown in SEQ ID NO:1, optionally wherein the E1 gene is functionally deleted through an E1 deletion of at least nucleotides 577 to 3403 with reference to the sequence shown in SEQ ID NO: 1 and optionally wherein the E3 gene is functionally deleted through an E3 deletion of at least nucleotides 27,125 to 31,825 with reference to the sequence shown in SEQ ID NO: 1. In some aspects, the ChAdV68 vector backbone comprises one or more genes or regulatory sequences with reference to a ChAdV68 genome or with reference to the sequence shown in SEQ ID NO:1, optionally wherein the one or more genes or regulatory sequences are selected from the group consisting of the chimpanzee adenovirus inverted terminal repeat (ITR), E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4, and L5 genes. In some aspects, the ChAdV68 vector backbone comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1, wherein the nucleotides 2 to 36,518 lack: (1) nucleotides 577 to 3403 of the sequence shown in SEQ ID NO: 1 corresponding to an E1 deletion and (2) nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO:1 corresponding to an E3 deletion. In some aspects, the ChAdV68 vector backbone comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO: 1, wherein the nucleotides 2 to 36,518 lack: (1) nucleotides 577 to 3403 of the sequence shown in SEQ ID NO:1 corresponding to an E1 deletion; (2) nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO: 1 corresponding to an E3 deletion; and (3) nucleotides 34,916 to 35,642 of the sequence shown in SEQ ID NO: 1 corresponding to a partial E4 deletion; optionally wherein the antigen cassette is inserted within the E1 deletion. In some aspects, the cassette is inserted in the ChAdV backbone at the E1 region, E3 region, and/or any deleted AdV region that allows incorporation of the cassette. In some aspects, the ChAdV backbone is generated from one of a first generation, a second generation, or a helper-dependent adenoviral vector.

[0102] In some aspects, the at least one promoter nucleotide sequence is selected from the group consisting of: a CMV, a SV40, an EF-1, a RSV, a PGK, a HSA, a MCK, and a EBV promoter sequence. In some aspects, the at least one promoter nucleotide sequence is a CMV promoter sequence.

[0103] In some aspects, at least one of the epitope-encoding nucleic acid sequences encodes an epitope that, when expressed and translated, is capable of being presented by MHC class I on a cell of the subject. In some aspects, at least one of the epitope-encoding nucleic acid sequences encodes an epitope that, when expressed and translated, is capable of being presented by MHC class II on a cell of the subject.

[0104] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises two or more antigen-encoding nucleic acid sequences. In some aspects, each antigen-encoding nucleic acid sequence is linked directly to one another.

[0105] In some aspects, each antigen-encoding nucleic acid sequence is linked to a distinct antigen-encoding nucleic acid sequence with a nucleic acid sequence encoding a linker. In some aspects, the linker links two MHC class I epitope-encoding nucleic acid sequences or an MHC class I epitope-encoding nucleic acid sequence to an MHC class II epitope-encoding nucleic acid sequence. In some aspects, the linker is selected from the group consisting of: (1) consecutive glycine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (2) consecutive alanine residues, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues in length; (3) two arginine residues (RR); (4) alanine, alanine, tyrosine (AAY); (5) a consensus sequence at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues in length that is processed efficiently by a mammalian proteasome; and (6) one or more native sequences flanking the antigen derived from the cognate protein of origin and that is at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 2-20 amino acid residues in length. In some aspects, the linker links two MHC class II epitope-encoding nucleic acid sequences or an MHC class II sequence to an MHC class I epitope-encoding nucleic acid sequence. In some aspects, the linker comprises the sequence GPGPG (SEQ ID NO: 56).

[0106] In some aspects, the antigen-encoding nucleic acid sequences is linked, operably or directly, to a separate or contiguous sequence that enhances the expression, stability, cell trafficking, processing and presentation, and/or immunogenicity of the antigen-encoding nucleic acid sequence. In some aspects, the separate or contiguous sequence comprises at least one of: a ubiquitin sequence, a ubiquitin sequence modified to increase proteasome targeting (e.g., the ubiquitin sequence contains a Gly to Ala substitution at position 76), an immunoglobulin signal sequence (e.g.,

IgK), a major histocompatibility class I sequence, lysosomal-associated membrane protein (LAMP)-1, human dendritic cell lysosomal-associated membrane protein, and a major histocompatibility class II sequence; optionally wherein the ubiquitin sequence modified to increase proteasome targeting is A76.

[0107] In some aspects, the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have increased binding affinity to its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have increased binding stability to its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have an increased likelihood of presentation on its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence. In some aspects, the at least one alteration comprises a point mutation, a frameshift mutation, a non-frameshift mutation, a deletion mutation, an insertion mutation, a splice variant, a genomic rearrangement, or a proteasome-generated spliced antigen.

[0108] In some aspects, the epitope-encoding nucleic acid sequence encodes an epitope known or suspected to be expressed in the subject known or suspected to have cancer. In some aspects, the cancer comprises a solid tumor. In some aspects, the cancer is selected from the group consisting of: microsatellite stable-colorectal cancer (MSS-CRC), non-small cell lung cancer (NSCLC), pancreatic ductal adenocarcinoma (PDA), and gastroesophageal adenocarcinoma (GEA). In some aspects, the cancer is selected from the group consisting of: lung cancer, melanoma, breast cancer, ovarian cancer, prostate cancer, kidney cancer, gastric cancer, colon cancer, testicular cancer, head and neck cancer, pancreatic cancer, bladder cancer, brain cancer, B-cell lymphoma, acute myelogenous leukemia, adult acute lymphoblastic leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, T cell lymphocytic leukemia, non-small cell lung cancer, and small cell lung cancer.

[0109] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences, optionally wherein each antigen-encoding nucleic acid sequence encodes a distinct antigen-encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 2-400 antigen-encoding nucleic acid sequences and wherein at least two of the antigen-encoding nucleic acid sequences encode epitope sequences or portions thereof that are presented by MHC class I on a cell surface. In some aspects, at least two of the MHC class I epitopes are presented by MHC class I on the tumor cell surface.

[0110] In some aspects, the epitope-encoding nucleic acid sequences comprises at least one MHC class I epitope-encoding nucleic acid sequence, and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence between 8 and 35 amino acids in length, optionally 9-17, 9-25, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35 amino acids in length.

[0111] In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present. In some aspects, the at least one MHC class II epitope-encoding nucleic acid sequence is present and comprises at least one MHC class II epitope-encoding nucleic acid sequence that comprises at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence. In some aspects, the epitope-encoding nucleic acid sequence comprises an MHC class II epitope-encoding nucleic acid sequence and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence that is 12-20, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 20-40 amino acids in length. In some aspects, the epitope-encoding nucleic acid sequences comprises an MHC class II epitope-encoding nucleic acid sequence, wherein the at least one MHC class II epitope-encoding nucleic acid sequence is present, and wherein the at least one MHC class II epitope-encoding nucleic acid sequence comprises at least one universal MHC class II epitope-encoding nucleic acid sequence, optionally wherein the at least one universal sequence comprises at least one of Tetanus toxoid and PADRE.

[0112] In some aspects, the at least one promoter nucleotide sequence is inducible. In some aspects, wherein the at least one promoter nucleotide sequence is non-inducible. In some aspects, the at least one poly(A) sequence comprises a Bovine Growth Hormone (BGH) SV40 polyA sequence. In some aspects, the at least one poly(A) sequence is at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, or at least 90 consecutive A nucleotides (SEQ ID NO: 29361). In some aspects, the at least one poly(A) sequence is at least 100 consecutive A nucleotides (SEQ ID NO: 29358).

[0113] In some aspects, the cassette further comprises at least one of: an intron sequence, a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) sequence, an internal ribosome entry sequence (IRES) sequence, a nucleotide sequence encoding a 2A self cleaving peptide sequence, a nucleotide sequence encoding a Furin cleavage site, or a sequence in the 5′ or 3′ non-coding region known to enhance the nuclear export, stability, or translation efficiency of mRNA that is operably linked to at least one of the at least one antigen-encoding nucleic acid sequences.

In some aspects, the cassette further comprises a reporter gene, including but not limited to, green fluorescent protein (GFP), a GFP variant, secreted alkaline phosphatase, luciferase, a luciferase variant, or a detectable peptide or epitope. In some aspects, the detectable peptide or epitope is selected from the group consisting of an HA tag, a Flag tag, a His-tag, or a V5 tag.

[0114] In some aspects, the one or more vectors further comprises one or more nucleic acid sequences encoding at least one immune modulator. In some aspects, the immune modulator is an anti-CTLA4 antibody or an antigen-binding fragment thereof, an anti-PD-1 antibody or an antigen-binding fragment thereof, an anti-PD-L1 antibody or an antigen-binding fragment thereof, an anti-4-1BB antibody or an antigen-binding fragment thereof, or an anti-OX-40 antibody or an antigen-binding fragment thereof. In some aspects, the antibody or antigen-binding fragment thereof is a Fab fragment, a Fab′ fragment, a single chain Fv (scFv), a single domain antibody (sdAb) either as single specific or multiple specificities linked together (e.g., camelid antibody domains), or full-length single-chain antibody (e.g., full-length IgG with heavy and light chains linked by a flexible linker). In some aspects, the heavy and light chain sequences of the antibody are a contiguous sequence separated by either a self-cleaving sequence such as 2A or IRES; or the heavy and light chain sequences of the antibody are linked by a flexible linker such as consecutive glycine residues. In some aspects, the immune modulator is a cytokine. In some aspects, the cytokine is at least one of IL-2, IL-7, IL-12, IL-15, or IL-21 or variants thereof of each.

[0115] In some aspects, the epitope-encoding nucleic acid sequence comprises a MHC class I epitope-encoding nucleic acid sequence, and wherein the MHC class I epitope-encoding nucleic acid sequence is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the MHC class I epitope-encoding nucleic acid sequence. In some aspects, each of the MHC class I epitope-encoding nucleic acid sequences is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the at least 20 MHC class I epitope-encoding nucleic acid sequences. In some aspects, a number of the set of selected epitopes is 2-20. In some aspects, the presentation model represents dependence between: (a) presence of a pair of a particular one of the MHC alleles and a particular amino acid at a particular position of a peptide sequence; and (b) likelihood of presentation on the tumor cell surface, by the particular one of the MHC alleles of the pair, of such a peptide sequence comprising the particular amino acid at the particular position. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being presented on the tumor cell surface relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of inducing a tumor-specific immune response in the subject relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of being presented to naïve T cells by professional antigen presenting cells (APCs) relative to unselected epitopes based on the presentation model, optionally wherein the APC is a dendritic cell (DC). In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being subject to inhibition via central or peripheral tolerance relative to unselected epitopes based on the presentation model. In some aspects, selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being capable of inducing an autoimmune response to normal tissue in the subject relative to unselected epitopes based on the presentation model. In some aspects, exome or transcriptome nucleotide sequencing data is obtained by performing sequencing on the tumor tissue. In some aspects, the sequencing is next generation sequencing (NGS) or any massively parallel sequencing approach.

[0116] In some aspects, the cassette comprises junctional epitope sequences formed by adjacent sequences in the cassette. In some aspects, at least one or each junctional epitope sequence has an affinity of greater than 500 nM for MHC. In some aspects, each junctional epitope sequence is non-self.

[0117] In some aspects, the cassette does not encode a non-therapeutic MHC class I or class II epitope nucleic acid sequence comprising a translated, wild-type nucleic acid sequence, wherein the non-therapeutic epitope is predicted to be displayed on an MHC allele of the subject. In some aspects, the non-therapeutic predicted MHC class I or class II epitope sequence is a junctional epitope sequence formed by adjacent sequences in the cassette. In some aspects, the prediction is based on presentation likelihoods generated by inputting sequences of the non-therapeutic epitopes into a presentation model. In some aspects, an order of the antigen-encoding nucleic acid sequences in the cassette is

determined by a series of steps comprising: (a) generating a set of candidate cassette sequences corresponding to different orders of the antigen-encoding nucleic acid sequences; (b) determining, for each candidate cassette sequence, a presentation score based on presentation of non-therapeutic epitopes in the candidate cassette sequence; and (c) selecting a candidate cassette sequence associated with a presentation score below a predetermined threshold as the cassette sequence for a vaccine.

[0118] In some aspects, the composition for delivery of the ChAdV-based expression system is formulated in a pharmaceutical composition comprising a pharmaceutically acceptable carrier.

[0119] In some aspects, one or more of the epitope-encoding nucleic acid sequences are derived from a tumor of the subject. In some aspects, each of the epitope-encoding nucleic acid sequences are derived from a tumor of the subject. In some aspects, one or more of the epitope-encoding nucleic acid sequences are not derived from a tumor of the subject. In some aspects, each of the epitope-encoding nucleic acid sequences are not derived from a tumor of the subject. In some aspects, the epitope-encoding nucleic acid sequence comprises an epitope selected from the group consisting of SEQ ID NO: 57-29,357 and SEQ ID NO: 29,512-29,519. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 tumor-specific MHC class I antigen-encoding nucleic acid sequences linearly linked to each other, comprising: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 14,954; 19,848; and 19,850, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; 19,863, and (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974, wherein each of the tumor-specific MHC class I antigen-encoding nucleic acid sequences comprises a class I epitope encoding nucleic acid sequence, optionally wherein each MHC I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 57-29,357 and SEQ ID NO: 29,512-29,519. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least at least 20 tumor-specific MHC class I antigen-encoding nucleic acid sequences linearly linked to each other, comprising: (A) a KRAS_G12A MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 14,954; 19,848; and 19,850, (C) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; 19,863, and (D) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974, (E) a KRAS_G13D MHC class I epitope encoding nucleic acid sequence, (F) a KRAS_Q61K MHC class I epitope encoding nucleic acid sequence, (G) a TP53_R249M MHC class I epitope encoding nucleic acid sequence, (H) a CTNNB1_S45P MHC class I epitope encoding nucleic acid sequence, (I) a CTNNB1_S45F MHC class I epitope encoding nucleic acid sequence, (J) a ERBB2_Y772_A775dup MHC class I epitope encoding nucleic acid sequence, (K) a KRAS_Q61R MHC class I epitope encoding nucleic acid sequence, (L) a CTNNB1_T41A MHC class I epitope encoding nucleic acid sequence, (M) a TP53_K132N MHC class I epitope encoding nucleic acid sequence, (N) a KRAS_Q61L MHC class I epitope encoding nucleic acid sequence, (O) a TP53_R213L MHC class I epitope encoding nucleic acid sequence, (P) a BRAF_G466V MHC class I epitope encoding nucleic acid sequence, (Q) a KRAS_Q61H MHC class I epitope encoding nucleic acid sequence, (R) a CTNNB1_S37F MHC class I epitope encoding nucleic acid sequence, (S) a TP53_S127Y MHC class I epitope encoding nucleic acid sequence, (T) a TP53_K132E MHC class I epitope encoding nucleic acid sequence, and (U) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence.

[0120] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, D) a KRAS Q61H MHC class I epitope encoding nucleic acid sequence, or combinations thereof. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, and (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, and (D) a KRAS Q61H MHC class I epitope encoding nucleic acid sequence. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope comprising the sequence of SEQ ID NO: 14,954; 19,848; 19,850, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope

encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; 19,863, and (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12C MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope comprising the sequence of SEQ ID NO: 14,954; 19,848; 19,850, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12D MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,749; 19,865; 19,863, (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_G12V MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 19,976; 19,779; 11,495; and 19,974, and (D) a KRAS Q61H MHC class I epitope encoding nucleic acid sequence, wherein the KRAS_Q61H MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 14,409.

[0121] In some aspects, the at least one antigen-encoding nucleic acid sequence comprises: (A) a TP53_R213L MHC class I epitope encoding nucleic acid sequence, (B) a TP53_S127Y MHC class I epitope encoding nucleic acid sequence, (C) a TP53_R249M MHC class I epitope encoding nucleic acid sequence, or combinations thereof. In some aspects, the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a TP53_R213L MHC class I epitope encoding nucleic acid sequence, (B) a TP53_S127Y MHC class I epitope encoding nucleic acid sequence, and (C) a TP53_R249M MHC class I epitope encoding nucleic acid sequence. In some aspects, the TP53_R213L MHC class I epitope encoding nucleic acid sequence encodes a MHC class I epitope selected from the group consisting of SEQ ID NO: 29,519.

[0122] In some aspects, the cassette of the composition for delivery of the ChAdV-based expression system is identical to the cassette of the composition for delivery of the self-amplifying alphavirus-based expression system. In some aspects, the cassette of the composition for delivery of the ChAdV-based expression system is different from the cassette of the composition for delivery of the self-amplifying alphavirus-based expression system.

[0123] In some aspects, stimulating the immune response comprises cliciting a cytotoxic T lymphocyte response to at least one of the one or more antigens. In some aspects, stimulating the immune response comprises a reduction in a tumor of the subject. In some aspects, the reduction is at least a 5%, at least a 10%, at least a 15%, at least a 20%, at least a 25%, at least a 30%, at least a 35%, at least a 40%, at least a 45%, at least a 50%, at least a 55%, at least a 60%, at least a 65%, at least a 70%, at least a 75%, at least a 80%, at least a 85%, at least a 90%, or at least a 95% reduction. In some aspects, the reduction is at least a 15% reduction. In some aspects, the reduction is at least a 20% reduction.

[0124] In some aspects, stimulating the immune response comprises stabilization of a tumor of the subject. In some aspects, stimulating the immune response comprises ameliorating a disease of the subject. In some aspects, ameliorating the disease comprises a complete response (CR), a partial response (PR), or a stable disease (SD).

[0125] In some aspects, the method further comprises administering one or more immune modulators. In some aspects, the one or more immune modulators are administered before, concurrently with, or after administration of any of the above compositions or pharmaceutical compositions. In some aspects, the one or more immune modulators are selected from the group consisting of: an anti-CTLA4 antibody or an antigen-binding fragment thereof, an anti-PD-1 antibody or an antigen-binding fragment thereof, an anti-PD-L1 antibody or an antigen-binding fragment thereof, an anti-4-1BB antibody or an antigen-binding fragment thereof, or an anti-OX-40 antibody or an antigen-binding fragment thereof. In some aspects, the anti-CTLA4 antibody is Ipilimumab. In some aspects, the anti-PD-1 is Nivolumab. In some aspects, the one or more immune modulators is administered intravenously (IV), intramuscularly (IM), intradermally (ID), or subcutaneously (SC). In some aspects, the subcutaneous administration is near the site of the composition or pharmaceutical composition administration or in close proximity to one or more vector or composition draining lymph nodes.

[0126] In some aspects, at least one of the one or more immune modulators is Ipilimumab. In some aspects, the Ipilimumab is administered subcutaneously (SC). In some aspects, the subcutaneous administration is proximal to a draining lymph node of the administration site of the self-amplifying alphavirus-based expression system or the composition for delivery of the ChAdV-based expression system. In some aspects, the Ipilimumab is administered at a dose of 30 mg. In some aspects, the dose of 30 mg is administered as four separate doses. In some aspects, at least one of the one or more immune modulators is Nivolumab. In some aspects, the Nivolumab is administered intravenously (IV). In some aspects, the Nivolumab is administered at a dose of 480 mg. In some aspects, the one or more immune modulators is each of Ipilimumab and Nivolumab. In some aspects, the Ipilimumab modulator is administered subcutaneously (SC) and wherein the Nivolumab modulator is administered intravenously (IV). In some aspects, the one or more immune modulators are administered concurrently with each administration of the self-amplifying alphavirus-based expression system or the composition for delivery of the ChAdV-based expression system.

[0127] In some aspects, the self-replicating alphavirus-based expression system is administered as a maintenance therapy and/or an adjuvant therapy. In some aspects, the maintenance therapy comprises a combination therapy with chemotherapy, immune checkpoint inhibitor therapy, radiation therapy, or combinations thereof. In some aspects, the

maintenance therapy comprises a chemotherapy comprising fluoropyrimidine and/or bevacizumab. In some aspects, the adjuvant therapy comprises a combination therapy with chemotherapy, immune checkpoint inhibitor therapy, radiation therapy, or combinations thereof, optionally wherein the chemotherapy comprises fluoropyrimidine and/or bevacizumab. In some aspects, the combination therapy comprises fluoropyrimidine and/or bevacizumab. In some aspects, the combination therapy comprises fluoropyrimidine and bevacizumab.

[0128] In some aspects, the combination therapy comprises fluoropyrimidine, bevacizumab, and an immune checkpoint inhibitor therapy. In some aspects, the immune checkpoint inhibitor comprises (1) an anti-PD-1 antibody or an antigen-binding fragment thereof, (2) an anti-PD-L1 antibody or an antigen-binding fragment thereof, and/or (3) an anti-CTLA-4 antibody or an antigen-binding fragment thereof. In some aspects, the immune checkpoint inhibitor therapy comprises administration of an anti-CTLA-4 antibody or an antigen-binding fragment thereof only with the priming dose and the first boosting dose. In some aspects, the anti-CTLA-4 antibody comprises ipilimumab. In some aspects, the ipilimumab is administered at a dose of 30 mg subcutaneously. In some aspects, the immune checkpoint inhibitor therapy comprises administration of an anti-PD-L1 antibody or an antigen-binding fragment thereof every 4 weeks (Q4W). In some aspects, the anti-PD-L1 antibody comprises atezolizumab or nivolumab. In some aspects, the atezolizumab is administered at a dose of 1680 mg intravenously or the nivolumab is administered at a dose of 480 mg intravenously.

[0129] In some aspects, the immune checkpoint inhibitor therapy comprises at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 administrations, optionally wherein the administration of the anti-PD-L1 antibody or an antigen-binding fragment thereof comprises at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 administrations. In some aspects, the immune checkpoint inhibitor therapy comprises at least 13 administrations, optionally wherein the administration of the anti-PD-L1 antibody or an antigen-binding fragment thereof comprises at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 administrations.

[0130] In some aspects, the subject has previously undergone surgery to remove a tumor and/or cancerous tissue, chemotherapy, immunotherapy (e.g., immune checkpoint inhibitor therapy), radiation therapy, or combinations thereof. In some aspects, the prior chemotherapy comprises oxaliplatin, fluoropyrimidine, irinotecan, and/or bevacizumab. In some aspects, the prior chemotherapy comprises oxaliplatin, fluoropyrimidine, and bevacizumab. In some aspects, the prior chemotherapy comprises oxaliplatin, fluoropyrimidine, irinotecan, and bevacizumab. In some aspects, the prior chemotherapy comprises FOLFOX. In some aspects, the prior chemotherapy comprises FOLFOXIRI. In some aspects, the prior chemotherapy further comprises bevacizumab. In some aspects, the prior chemotherapy was administered for up to 24 weeks prior to administration of the maintenance therapy.

[0131] In some aspects, the subject has colorectal cancer (CRC). In some aspects, the CRC is classified as Stage IV, microsatellite-stable, and BRAFwt. In some aspects, the CRC is classified as Stage II or III. In some aspects, the subject is classified as ctDNA positive.

[0132] The method of any one of the above claims, wherein two or more boosting doses are administered.

[0133] In some aspects, 1, 2, 3, 4, 5, 6, 7, or 8 boosting doses are administered.

[0134] In some aspects, the ChAdV-based expression system is further administered as a boosting dose. In some aspects, the ChAdV-based boosting dose is only administered as a single boosting dose. In some aspects, the ChAdV-based expression system is administered as the boosting dose on or about day 140 after the priming dose of the ChAdV-based expression system. In some aspects, the ChAdV-based expression system is administered as the boosting dose on or about week 20 after the priming dose of the ChAdV-based expression system. In some aspects, the ChAdV-based expression system is administered as the boosting dose on or about month 5 after the priming dose of the ChAdV-based expression system. In some aspects, the ChAdV-based expression system is administered as the boosting dose on or after day 140 after the priming dose of the ChAdV-based expression system. In some aspects, the ChAdV-based expression system is administered as the boosting dose on or after week 20 after the priming dose of the ChAdV-based expression system. In some aspects, the ChAdV-based expression system is administered as the boosting dose on or after month 5 after the priming dose of the ChAdV-based expression system.

[0135] In some aspects, the self-replicating alphavirus-based expression system is administered as at least two boosting doses. In some aspects, the self-replicating alphavirus-based expression system is administered as at least two or more boosting doses at least 28 days apart. In some aspects, the self-replicating alphavirus-based expression system is administered as at least two or more boosting doses at least 4 weeks (Q4W) apart. In some aspects, the self-replicating alphavirus-based expression system is administered as at least two or more boosting doses at least one month apart. In some aspects, the self-replicating alphavirus-based expression system is administered as at least two or more boosting doses at least 56 days apart.

[0136] In some aspects, the self-replicating alphavirus-based expression system is administered as at least two or more boosting doses at least 8 weeks (Q8W) apart.

[0137] In some aspects, the self-replicating alphavirus-based expression system is administered as at least two or more boosting doses at least 2 months apart.

[0138] In some aspects, the self-replicating alphavirus-based expression system is administered as at least two boosting doses on or about days 28 and 84 after the priming dose of the ChAdV-based expression system.

[0139] In some aspects, the self-replicating alphavirus-based expression system is administered as at least two boosting

doses on or about weeks 4 and 12 after the priming dose of the ChAdV-based expression system.

[0140] In some aspects, the self-replicating alphavirus-based expression system is administered as at least two boosting doses on or about months 1 and 3 after the priming dose of the ChAdV-based expression system.

[0141] In some aspects, the self-replicating alphavirus-based expression system is administered as at least four boosting doses. In some aspects, the self-replicating alphavirus-based expression system is administered on or about days 28, 84, 224, and 308 relative to the priming dose of the ChAdV-based expression system. In some aspects, the self-replicating alphavirus-based expression system is administered on or about weeks 4, 12, 32, and 44 relative to the priming dose of the ChAdV-based expression system. In some aspects, the self-replicating alphavirus-based expression system is administered on or about months 1, 3, 8, and 11 relative to the priming dose of the ChAdV-based expression system.

Description

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0142] These and other features, aspects, and advantages of the present invention will become better understood with regard to the following description, and accompanying drawings, where:

[0143] FIG. **1** shows serum IFNa2a levels assessed by ELISA at 8 hours following a first samRNA dose at the indicated dosages in non-human primates (NHPs).

[0144] FIG. **2** shows serum IFNa2a levels assessed by ELISA at 24 hours following a first samRNA dose at the indicated dosages in humans.

[0145] FIG. **3** shows ex-vivo IFNγ ELISpot following overnight stimulation with patient-specific peptide pool of predicted CD8 epitopes. PBMCs assessed at peak timepoint post samRNA administration in each patient. Shown are dosing regimens and associated T cell responses for subjects G1 and G2, who both received 30 μg doses, and subjects G10 and G11, who both received 300 μg doses.

[0146] FIG. **4** shows ex-vivo IFNγ ELISpot following overnight stimulation with patient-specific peptide pool of predicted CD8 epitopes. PBMCs assessed at peak timepoint post samRNA administration in each patient. Shown are T cell responses for subjects who received 30 μg, 100 μg, and/or 300 μg doses.

DETAILED DESCRIPTION

I. Definitions

[0147] In general, terms used in the claims and the specification are intended to be construed as having the plain meaning understood by a person of ordinary skill in the art. Certain terms are defined below to provide additional clarity. In case of conflict between the plain meaning and the provided definitions, the provided definitions are to be used.

[0148] As used herein the term "antigen" is a substance that stimulates an immune response. An antigen can be a neoantigen. An antigen can be a "shared antigen" that is an antigen found among a specific population, e.g., a specific population of cancer patients.

[0149] As used herein the term "neoantigen" is an antigen that has at least one alteration that makes it distinct from the corresponding wild-type antigen, e.g., via mutation in a tumor cell or post-translational modification specific to a tumor cell. A neoantigen can include a polypeptide sequence or a nucleotide sequence. A mutation can include a frameshift or non-frameshift indel, missense or nonsense substitution, splice site alteration, genomic rearrangement or gene fusion, or any genomic or expression alteration giving rise to a neoORF. A mutations can also include a splice variant. Post-translational modifications specific to a tumor cell can include aberrant phosphorylation. Post-translational modifications specific to a tumor cell can also include a proteasome-generated spliced antigen. See Liepe et al., A large fraction of HLA class I ligands are proteasome-generated spliced peptides; Science. 2016 Oct. 21; 354 (6310): 354-358. Exemplary shared neoantigens are shown in Table A and in the AACR GENIE Results (SEQ ID NO: 10,755-29,357); corresponding HLA allele(s) for each antigen are also shown. Such shared neoantigens are useful for inducing an immune response in a subject via administration. The subject can be identified for administration through the use of various diagnostic methods, e.g., patient selection methods described further below.

[0150] As used herein the term "tumor antigen" is an antigen present in a subject's tumor cell or tissue but not in the subject's corresponding normal cell or tissue, or derived from a polypeptide known to or have been found to have altered expression in a tumor cell or cancerous tissue in comparison to a normal cell or tissue.

[0151] As used herein the term "antigen-based vaccine" is a vaccine composition based on one or more antigens, e.g., a plurality of antigens. The vaccines can be nucleotide-based (e.g., virally based, RNA based, or DNA based), protein-based (e.g., peptide based), or a combination thereof.

[0152] As used herein the term "candidate antigen" is a mutation or other aberration giving rise to a sequence that may represent an antigen.

[0153] As used herein the term "coding region" is the portion(s) of a gene that encode protein.

[0154] As used herein the term "coding mutation" is a mutation occurring in a coding region.

[0155] As used herein the term "ORF" means open reading frame.

[0156] As used herein the term "NEO-ORF" is a tumor-specific ORF arising from a mutation or other aberration such as splicing.

[0157] As used herein the term "missense mutation" is a mutation causing a substitution from one amino acid to another.

[0158] As used herein the term "nonsense mutation" is a mutation causing a substitution from an amino acid to a stop codon or causing removal of a canonical start codon.

[0159] As used herein the term "frameshift mutation" is a mutation causing a change in the frame of the protein.

[0160] As used herein the term "indel" is an insertion or deletion of one or more nucleic acids.

[0161] As used herein, the term percent "identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that have a specified percentage of nucleotides or amino acid residues that are the same, when compared and aligned for maximum correspondence, as measured using one of the sequence comparison algorithms described below (e.g., BLASTP and BLASTN or other algorithms available to persons of skill) or by visual inspection. Depending on the application, the percent "identity" can exist over a region of the sequence being compared, e.g., over a functional domain, or, alternatively, exist over the full length of the two sequences to be compared.

[0162] For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters. Alternatively, sequence similarity or dissimilarity can be established by the combined presence or absence of particular nucleotides, or, for translated sequences, amino acids at selected sequence positions (e.g., sequence motifs).

[0163] Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by visual inspection (see generally Ausubel et al., infra).

[0164] One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul et al., J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information.

[0165] As used herein the term "non-stop or read-through" is a mutation causing the removal of the natural stop codon.

[0166] As used herein the term "epitope" is the specific portion of an antigen typically bound by an antibody or T cell receptor.

[0167] As used herein the term "immunogenic" is the ability to stimulate an immune response, e.g., via T cells, B cells, or both.

[0168] As used herein the term "HLA binding affinity" "MHC binding affinity" means affinity of binding between a specific antigen and a specific MHC allele.

[0169] As used herein the term "bait" is a nucleic acid probe used to enrich a specific sequence of DNA or RNA from a sample.

[0170] As used herein the term "variant" is a difference between a subject's nucleic acids and the reference human genome used as a control.

[0171] As used herein the term "variant call" is an algorithmic determination of the presence of a variant, typically from sequencing.

[0172] As used herein the term "polymorphism" is a germline variant, i.e., a variant found in all DNA-bearing cells of an individual.

[0173] As used herein the term "somatic variant" is a variant arising in non-germline cells of an individual.

[0174] As used herein the term "allele" is a version of a gene or a version of a genetic sequence or a version of a protein.

[0175] As used herein the term "HLA type" is the complement of HLA gene alleles.

[0176] As used herein the term "nonsense-mediated decay" or "NMD" is a degradation of an mRNA by a cell due to a premature stop codon.

[0177] As used herein the term "truncal mutation" is a mutation originating early in the development of a tumor and present in a substantial portion of the tumor's cells.

[0178] As used herein the term "subclonal mutation" is a mutation originating later in the development of a tumor and present in only a subset of the tumor's cells.

[0179] As used herein the term "exome" is a subset of the genome that codes for proteins. An exome can be the collective exons of a genome.

[0180] As used herein the term "logistic regression" is a regression model for binary data from statistics where the logit of the probability that the dependent variable is equal to one is modeled as a linear function of the dependent variables.

[0181] As used herein the term "neural network" is a machine learning model for classification or regression consisting of multiple layers of linear transformations followed by element-wise nonlinearities typically trained via stochastic

gradient descent and back-propagation.

[0182] As used herein the term "proteome" is the set of all proteins expressed and/or translated by a cell, group of cells, or individual.

[0183] As used herein the term "peptidome" is the set of all peptides presented by MHC-I or MHC-II on the cell surface. The peptidome may refer to a property of a cell or a collection of cells (e.g., the tumor peptidome, meaning the union of the peptidomes of all cells that comprise the tumor).

[0184] As used herein the term "ELISpot" means Enzyme-linked immunosorbent spot assay [0185] which is a common method for monitoring immune responses in humans and animals.

[0186] As used herein the term "dextramers" is a dextran-based peptide-MHC multimers used for antigen-specific T-cell staining in flow cytometry.

[0187] As used herein the term "tolerance or immune tolerance" is a state of immune non-responsiveness to one or more antigens, e.g. self-antigens.

[0188] As used herein the term "central tolerance" is a tolerance affected in the thymus, either by deleting self-reactive T-cell clones or by promoting self-reactive T-cell clones to differentiate into immunosuppressive regulatory T-cells (Tregs).

[0189] As used herein the term "peripheral tolerance" is a tolerance affected in the periphery by downregulating or anergizing self-reactive T-cells that survive central tolerance or promoting these T cells to differentiate into Tregs.

[0190] The term "sample" can include a single cell or multiple cells or fragments of cells or an aliquot of body fluid, taken from a subject, by means including venipuncture, excretion, ejaculation, massage, biopsy, needle aspirate, lavage sample, scraping, surgical incision, or intervention or other means known in the art.

[0191] The term "subject" encompasses a cell, tissue, or organism, human or non-human, whether in vivo, ex vivo, or in vitro, male or female. The term subject is inclusive of mammals including humans.

[0192] The term "mammal" encompasses both humans and non-humans and includes but is not limited to humans, non-human primates, canines, felines, murines, bovines, equines, and porcines.

[0193] The term "clinical factor" refers to a measure of a condition of a subject, e.g., disease activity or severity. "Clinical factor" encompasses all markers of a subject's health status, including non-sample markers, and/or other characteristics of a subject, such as, without limitation, age and gender. A clinical factor can be a score, a value, or a set of values that can be obtained from evaluation of a sample (or population of samples) from a subject or a subject under a determined condition. A clinical factor can also be predicted by markers and/or other parameters such as gene expression surrogates. Clinical factors can include tumor type, tumor sub-type, and smoking history.

[0194] The term "antigen-encoding nucleic acid sequences derived from a tumor" refers to nucleic acid sequences obtained from the tumor, e.g. via RT-PCR; or sequence data obtained by sequencing the tumor and then synthesizing the nucleic acid sequences using the sequencing data, e.g., via various synthetic or PCR-based methods known in the art. Derived sequences can include nucleic acid sequence variants, such as sequence-optimized nucleic acid sequence variants (e.g., codon-optimized and/or otherwise optimized for expression), that encode the same polypeptide sequence as the corresponding native nucleic acid sequence obtained from a tumor.

[0195] The term "alphavirus backbone" refers to minimal sequence(s) of an alphavirus that allow for self-replication of the viral genome. Minimal sequences can include conserved sequences for nonstructural protein-mediated amplification, a nonstructural protein 1 (nsP1) gene, a nsP2 gene, a nsP3 gene, a nsP4 gene, and a polyA sequence, as well as sequences for expression of subgenomic viral RNA including a subgenomic (e.g., a 26S) promoter element.

[0196] The term "sequences for nonstructural protein-mediated amplification" includes alphavirus conserved sequence elements (CSE) well known to those in the art. CSEs include, but are not limited to, an alphavirus 5′ UTR, a 51-nt CSE, a 24-nt CSE, a subgenomic promoter sequence (e.g., a 26S subgenomic promoter sequence), a 19-nt CSE, and an alphavirus 3′ UTR.

[0197] The term "sequences for nonstructural protein-mediated amplification" includes alphavirus conserved sequence elements (CSE) well known to those in the art. CSEs include, but are not limited to, an alphavirus 5′ UTR, a 51-nt CSE, a 24-nt CSE, or other 26S subgenomic promoter sequence, a 19-nt CSE, and an alphavirus 3′ UTR.

[0198] The term "RNA polymerase" includes polymerases that catalyze the production of RNA polynucleotides from a DNA template. RNA polymerases include, but are not limited to, bacteriophage derived polymerases including T3, T7, and SP6.

[0199] The term "lipid" includes hydrophobic and/or amphiphilic molecules. Lipids can be cationic, anionic, or neutral. Lipids can be synthetic or naturally derived, and in some instances biodegradable. Lipids can include cholesterol, phospholipids, lipid conjugates including, but not limited to, polyethyleneglycol (PEG) conjugates (PEGylated lipids), waxes, oils, glycerides, fats, and fat-soluble vitamins. Lipids can also include dilinoleylmethyl-4-dimethylaminobutyrate (MC3) and MC3-like molecules.

[0200] The term "lipid nanoparticle" or "LNP" includes vesicle like structures formed using a lipid containing membrane surrounding an aqueous interior, also referred to as liposomes. Lipid nanoparticles includes lipid-based compositions with a solid lipid core stabilized by a surfactant. The core lipids can be fatty acids, acylglycerols, waxes, and mixtures of these surfactants. Biological membrane lipids such as phospholipids, sphingomyelins, bile salts (sodium taurocholate), and sterols (cholesterol) can be utilized as stabilizers. Lipid nanoparticles can be formed using

defined ratios of different lipid molecules, including, but not limited to, defined ratios of one or more cationic, anionic, or neutral lipids. Lipid nanoparticles can encapsulate molecules within an outer-membrane shell and subsequently can be contacted with target cells to deliver the encapsulated molecules to the host cell cytosol. Lipid nanoparticles can be modified or functionalized with non-lipid molecules, including on their surface. Lipid nanoparticles can be single-layered (unilamellar) or multi-layered (multilamellar). Lipid nanoparticles can be complexed with nucleic acid. Unilamellar lipid nanoparticles can be complexed with nucleic acid, wherein the nucleic acid is in the aqueous interior. Multilamellar lipid nanoparticles can be complexed with nucleic acid, wherein the nucleic acid is in the aqueous interior, or to form or sandwiched between.

[0201] The term "pharmaceutically effective amount" is an amount of a vaccine component (such as a peptide, engineered vector, and/or adjuvant) that is effective in a route of administration to provide a cell with sufficient levels of protein, protein expression, and/or cell-signaling activity (e.g., adjuvant-mediated activation) to provide a vaccinal benefit, i.e., some measurable level of immunity.

[0202] Abbreviations: MHC: major histocompatibility complex; HLA: human leukocyte antigen, or the human MHC gene locus; NGS: next-generation sequencing; PPV: positive predictive value; TSNA: tumor-specific neoantigen; FFPE: formalin-fixed, paraffin-embedded; NMD: nonsense-mediated decay; NSCLC: non-small-cell lung cancer; DC: dendritic cell.

[0203] It should be noted that, as used in the specification and the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise.

[0204] Unless specifically stated or otherwise apparent from context, as used herein the term "about" is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from context, all numerical values provided herein are modified by the term about.

[0205] Any terms not directly defined herein shall be understood to have the meanings commonly associated with them as understood within the art of the invention. Certain terms are discussed herein to provide additional guidance to the practitioner in describing the compositions, devices, methods and the like of aspects of the invention, and how to make or use them. It will be appreciated that the same thing may be said in more than one way. Consequently, alternative language and synonyms may be used for any one or more of the terms discussed herein. No significance is to be placed upon whether or not a term is elaborated or discussed herein. Some synonyms or substitutable methods, materials and the like are provided. Recital of one or a few synonyms or equivalents does not exclude use of other synonyms or equivalents, unless it is explicitly stated. Use of examples, including examples of terms, is for illustrative purposes only and does not limit the scope and meaning of the aspects of the invention herein.

[0206] All references, issued patents and patent applications cited within the body of the specification are hereby incorporated by reference in their entirety, for all purposes.

II. Antigen Identification

[0207] Research methods for NGS analysis of tumor and normal exome and transcriptomes have been described and applied in the antigen identification space. 6.14.15 Certain optimizations for greater sensitivity and specificity for antigen identification in the clinical setting can be considered. These optimizations can be grouped into two areas, those related to laboratory processes and those related to the NGS data analysis. The research methods described can also be applied to identification of antigens in other settings, such as identification of identifying antigens from an infectious disease organism, an infection in a subject, or an infected cell of a subject. Examples of optimizations are known to those skilled in the art, for example the methods described in more detail in U.S. Pat. No. 10,055,540, US Application Pub. No. US20200010849A1, U.S. application Ser. No. 16/606,577, and international patent application publications WO2020181240A1, WO/2018/195357 and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes.

[0208] Methods for identifying shared antigens (e.g., neoantigens) include identifying antigens from a tumor of a subject that are likely to be presented on the cell surface of the tumor or immune cells, including professional antigen presenting cells such as dendritic cells, and/or are likely to be immunogenic. As an example, one such method may comprise the steps of: obtaining at least one of exome, transcriptome or whole genome tumor nucleotide sequencing and/or expression data from the tumor cell of the subject, wherein the tumor nucleotide sequencing and/or expression data is used to obtain data representing peptide sequences of each of a set of antigens (e.g., in the case of neoantigens wherein the peptide sequence of each neoantigen comprises at least one alteration that makes it distinct from the corresponding wild-type peptide sequence or in cases of shared antigens without a mutation where peptides are derived from any polypeptide known to or have been found to have altered expression in a tumor cell or cancerous tissue in comparison to a normal cell or tissue); inputting the peptide sequence of each antigen into one or more presentation models to generate a set of numerical likelihoods that each of the antigens is presented by one or more MHC alleles on the tumor cell surface of the tumor cell of the subject or cells present in the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and selecting a subset of the set of antigens based on the set of numerical likelihoods to generate a set of selected antigens.

III. Identification of Tumor Specific Mutations in Neoantigens

[0209] Also disclosed herein are methods for the identification of certain mutations (e.g., the variants or alleles that are

present in cancer cells). In particular, these mutations can be present in the genome, transcriptome, proteome, or exome of cancer cells of a subject having cancer but not in normal tissue from the subject. Specific methods for identifying neoantigens, including shared neoantigens, that are specific to tumors are known to those skilled in the art, for example the methods described in more detail in U.S. Pat. No. 10,055,540, US Application Pub. No. US20200010849A1, and international patent application publications WO/2018/195357 and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes. Examples of shared neoantigens that are specific to tumors are described in more detail in international patent application publication WO2019226941A1, herein incorporated by reference in its entirety, for all purposes.

[0210] Genetic mutations in tumors can be considered useful for the immunological targeting of tumors if they lead to changes in the amino acid sequence of a protein exclusively in the tumor. Useful mutations include: (1) non-synonymous mutations leading to different amino acids in the protein; (2) read-through mutations in which a stop codon is modified or deleted, leading to translation of a longer protein with a novel tumor-specific sequence at the C-terminus; (3) splice site mutations that lead to the inclusion of an intron in the mature mRNA and thus a unique tumor-specific protein sequence; (4) chromosomal rearrangements that give rise to a chimeric protein with tumor-specific sequences at the junction of 2 proteins (i.e., gene fusion); (5) frameshift mutations or deletions that lead to a new open reading frame with a novel tumor-specific protein sequence. Mutations can also include one or more of non-frameshift indel, missense or nonsense substitution, splice site alteration, genomic rearrangement or gene fusion, or any genomic or expression alteration giving rise to a neoORF.

[0211] Peptides with mutations or mutated polypeptides arising from for example, splice-site, frameshift, readthrough, or gene fusion mutations in tumor cells can be identified by sequencing DNA, RNA or protein in tumor versus normal cells.

[0212] Also mutations can include previously identified tumor specific mutations. Known tumor mutations can be found at the Catalogue of Somatic Mutations in Cancer (COSMIC) databasc.

[0213] A variety of methods are available for detecting the presence of a particular mutation or allele in an individual's DNA or RNA. Advancements in this field have provided accurate, easy, and inexpensive large-scale SNP genotyping. For example, several techniques have been described including dynamic allele-specific hybridization (DASH), microplate array diagonal gel electrophoresis (MADGE), pyrosequencing, oligonucleotide-specific ligation, the TaqMan system as well as various DNA "chip" technologies such as the Affymetrix SNP chips. These methods utilize amplification of a target genetic region, typically by PCR. Still other methods, based on the generation of small signal molecules by invasive cleavage followed by mass spectrometry or immobilized padlock probes and rolling-circle amplification. Several of the methods known in the art for detecting specific mutations are summarized below.

[0214] PCR based detection means can include multiplex amplification of a plurality of markers simultaneously. For example, it is well known in the art to select PCR primers to generate PCR products that do not overlap in size and can be analyzed simultaneously. Alternatively, it is possible to amplify different markers with primers that are differentially labeled and thus can each be differentially detected. Of course, hybridization based detection means allow the differential detection of multiple PCR products in a sample. Other techniques are known in the art to allow multiplex analyses of a plurality of markers.

[0215] Several methods have been developed to facilitate analysis of single nucleotide polymorphisms in genomic DNA or cellular RNA. For example, a single base polymorphism can be detected by using a specialized exonuclease-resistant nucleotide, as disclosed, e.g., in Mundy, C. R. (U.S. Pat. No. 4,656,127). According to the method, a primer complementary to the allelic sequence immediately 3′ to the polymorphic site is permitted to hybridize to a target molecule obtained from a particular animal or human. If the polymorphic site on the target molecule contains a nucleotide that is complementary to the particular exonuclease-resistant nucleotide derivative present, then that derivative will be incorporated onto the end of the hybridized primer. Such incorporation renders the primer resistant to exonuclease, and thereby permits its detection. Since the identity of the exonuclease-resistant derivative of the sample is known, a finding that the primer has become resistant to exonucleases reveals that the nucleotide(s) present in the polymorphic site of the target molecule is complementary to that of the nucleotide derivative used in the reaction. This method has the advantage that it does not require the determination of large amounts of extraneous sequence data.

[0216] A solution-based method can be used for determining the identity of a nucleotide of a polymorphic site. Cohen, D. et al. (French Patent 2,650,840; PCT Appln. No. WO91/02087). As in the Mundy method of U.S. Pat. No. 4,656,127, a primer is employed that is complementary to allelic sequences immediately 3′ to a polymorphic site. The method determines the identity of the nucleotide of that site using labeled dideoxynucleotide derivatives, which, if complementary to the nucleotide of the polymorphic site will become incorporated onto the terminus of the primer.

[0217] An alternative method, known as Genetic Bit Analysis or GBA is described by Goelet, P. et al. (PCT Appln. No. 92/15712). The method of Goclet, P. et al. uses mixtures of labeled terminators and a primer that is complementary to the sequence 3′ to a polymorphic site. The labeled terminator that is incorporated is thus determined by, and complementary to, the nucleotide present in the polymorphic site of the target molecule being evaluated. In contrast to the method of Cohen et al. (French Patent 2,650,840; PCT Appln. No. WO91/02087) the method of Goclet, P. et al. can be a heterogeneous phase assay, in which the primer or the target molecule is immobilized to a solid phase.

[0218] Several primer-guided nucleotide incorporation procedures for assaying polymorphic sites in DNA have been

described (Komher, J. S. et al., Nucl. Acids. Res. 17:7779-7784 (1989); Sokolov, B. P., Nucl. Acids Res. 18:3671 (1990); Syvanen, A.-C., et al., Genomics 8:684-692 (1990); Kuppuswamy, M. N. et al., Proc. Natl. Acad. Sci. (U.S.A.) 88:1143-1147 (1991); Prezant, T. R. et al., Hum. Mutat. 1:159-164 (1992); Ugozzoli, L. et al., GATA 9:107-112 (1992); Nyren, P. et al., Anal. Biochem. 208:171-175 (1993)). These methods differ from GBA in that they utilize incorporation of labeled deoxynucleotides to discriminate between bases at a polymorphic site. In such a format, since the signal is proportional to the number of deoxynucleotides incorporated, polymorphisms that occur in runs of the same nucleotide can result in signals that are proportional to the length of the run (Syvanen, A.-C., et al., Amer. J. Hum. Genet. 52:46-59 (1993)).

[0219] A number of initiatives obtain sequence information directly from millions of individual molecules of DNA or RNA in parallel. Real-time single molecule sequencing-by-synthesis technologies rely on the detection of fluorescent nucleotides as they are incorporated into a nascent strand of DNA that is complementary to the template being sequenced. In one method, oligonucleotides 30-50 bases in length are covalently anchored at the 5′ end to glass cover slips. These anchored strands perform two functions. First, they act as capture sites for the target template strands if the templates are configured with capture tails complementary to the surface-bound oligonucleotides. They also act as primers for the template directed primer extension that forms the basis of the sequence reading. The capture primers function as a fixed position site for sequence determination using multiple cycles of synthesis, detection, and chemical cleavage of the dye-linker to remove the dye. Each cycle includes adding the polymerase/labeled nucleotide mixture, rinsing, imaging and cleavage of dye. In an alternative method, polymerase is modified with a fluorescent donor molecule and immobilized on a glass slide, while each nucleotide is color-coded with an acceptor fluorescent moiety attached to a gamma-phosphate. The system detects the interaction between a fluorescently-tagged polymerase and a fluorescently modified nucleotide as the nucleotide becomes incorporated into the de novo chain. Other sequencing-by-synthesis technologies also exist.

[0220] Any suitable sequencing-by-synthesis platform can be used to identify mutations. As described above, four major sequencing-by-synthesis platforms are currently available: the Genome Sequencers from Roche/454 Life Sciences, the 1G Analyzer from Illumina/Solexa, the SOLID system from Applied BioSystems, and the Heliscope system from Helicos Biosciences. Sequencing-by-synthesis platforms have also been described by Pacific BioSciences and VisiGen Biotechnologies. In some embodiments, a plurality of nucleic acid molecules being sequenced is bound to a support (e.g., solid support). To immobilize the nucleic acid on a support, a capture sequence/universal priming site can be added at the 3′ and/or 5′ end of the template. The nucleic acids can be bound to the support by hybridizing the capture sequence to a complementary sequence covalently attached to the support. The capture sequence (also referred to as a universal capture sequence) is a nucleic acid sequence complementary to a sequence attached to a support that may dually serve as a universal primer.

[0221] As an alternative to a capture sequence, a member of a coupling pair (such as, e.g., antibody/antigen, receptor/ligand, or the avidin-biotin pair as described in, e.g., US Patent Application No. 2006/0252077) can be linked to each fragment to be captured on a surface coated with a respective second member of that coupling pair.

[0222] Subsequent to the capture, the sequence can be analyzed, for example, by single molecule detection/sequencing, e.g., as described in the Examples and in U.S. Pat. No. 7,283,337, including template-dependent sequencing-by-synthesis. In sequencing-by-synthesis, the surface-bound molecule is exposed to a plurality of labeled nucleotide triphosphates in the presence of polymerase. The sequence of the template is determined by the order of labeled nucleotides incorporated into the 3′ end of the growing chain. This can be done in real time or can be done in a step-and-repeat mode. For real-time analysis, different optical labels to each nucleotide can be incorporated and multiple lasers can be utilized for stimulation of incorporated nucleotides.

[0223] Sequencing can also include other massively parallel sequencing or next generation sequencing (NGS) techniques and platforms. Additional examples of massively parallel sequencing techniques and platforms are the Illumina HiSeq or MiSeq, Thermo PGM or Proton, the Pac Bio RS II or Sequel, Qiagen's Gene Reader, and the Oxford Nanopore MinION. Additional similar current massively parallel sequencing technologies can be used, as well as future generations of these technologies.

[0224] Any cell type or tissue can be utilized to obtain nucleic acid samples for use in methods described herein. For example, a DNA or RNA sample can be obtained from a tumor or a bodily fluid, e.g., blood, obtained by known techniques (e.g. venipuncture) or saliva. Alternatively, nucleic acid tests can be performed on dry samples (e.g. hair or skin). In addition, a sample can be obtained for sequencing from a tumor and another sample can be obtained from normal tissue for sequencing where the normal tissue is of the same tissue type as the tumor. A sample can be obtained for sequencing from a tumor and another sample can be obtained from normal tissue for sequencing where the normal tissue is of a distinct tissue type relative to the tumor.

[0225] Tumors can include one or more of lung cancer, melanoma, breast cancer, ovarian cancer, prostate cancer, kidney cancer, gastric cancer, colon cancer, testicular cancer, head and neck cancer, pancreatic cancer, brain cancer, B-cell lymphoma, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, and T cell lymphocytic leukemia, non-small cell lung cancer, and small cell lung cancer.

[0226] Alternatively, protein mass spectrometry can be used to identify or validate the presence of mutated peptides bound to MHC proteins on tumor cells. Peptides can be acid-eluted from tumor cells or from HLA molecules that are

immunoprecipitated from tumor, and then identified using mass spectrometry.

IV. Antigens

[0227] Antigens can include nucleotides or polypeptides. For example, an antigen can be an RNA sequence that encodes for a polypeptide sequence. Antigens useful in vaccines can therefore include nucleotide sequences or polypeptide sequences. Shared neoantigens are shown in Table A (see SEQ ID NO:10,755-21,015) and in the AACR GENIE results (see SEQ ID NO: 21,016-29,357). Shared antigens are shown in Table 1.2 (scc SEQ ID NO:57-10,754).

[0228] Disclosed herein are isolated peptides that comprise tumor specific mutations identified by the methods disclosed herein, peptides that comprise known tumor specific mutations, and mutant polypeptides or fragments thereof identified by methods disclosed herein. Neoantigen peptides can be described in the context of their coding sequence where a neoantigen includes the nucleotide sequence (e.g., DNA or RNA) that codes for the related polypeptide sequence.

[0229] Also disclosed herein are peptides derived from any polypeptide known to or have been found to have altered expression in a tumor cell or cancerous tissue in comparison to a normal cell or tissue, for example any polypeptide known to or have been found to be aberrantly expressed in a tumor cell or cancerous tissue in comparison to a normal cell or tissue. Suitable polypeptides from which the antigenic peptides can be derived can be found for example in the COSMIC database. COSMIC curates comprehensive information on somatic mutations in human cancer. The peptide contains the tumor specific mutation.

[0230] Tumor antigens (e.g., shared tumor antigens and tumor neoantigens) can include, but are not limited to, those described in U.S. application Ser. No. 17/058,128, herein incorporated by reference for all purposes. Antigen peptides can be described in the context of their coding sequence where an antigen includes the nucleotide sequence (e.g., DNA or RNA) that codes for the related polypeptide sequence. Antigens can be selected that are predicted to be presented on the cell surface of a cell, such as a tumor cell or an immune cell, including professional antigen presenting cells such as dendritic cells. Antigens can be selected that are predicted to be immunogenic.

[0231] One or more polypeptides encoded by an antigen nucleotide sequence can comprise at least one of: a binding affinity with MHC with an IC50 value of less than 1000 nM, for MHC Class I peptides a length of 8-15, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids, presence of sequence motifs within or near the peptide promoting proteasome cleavage, and presence or sequence motifs promoting TAP transport. For MHC Class II peptides a length 6-30, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 amino acids, presence of sequence motifs within or near the peptide promoting cleavage by extracellular or lysosomal proteases (e.g., cathepsins) or HLA-DM catalyzed HLA binding.

[0232] One or more antigens can be presented on the surface of a tumor.

[0233] One or more antigens can be immunogenic in a subject having a tumor, e.g., capable of stimulating a T cell response and/or a B cell response in the subject. One or more antigens can be capable of stimulating a B cell response, such as the production of antibodies that recognize the one or more antigens (e.g., antibodies that recognize a tumor). Antibodies can recognize linear polypeptide sequences or recognize secondary and tertiary structures. Accordingly, B cell antigens can include linear polypeptide sequences or polypeptides having secondary and tertiary structures, including, but not limited to, full-length proteins, protein subunits, protein domains, or any polypeptide sequence known or predicted to have secondary and tertiary structures. Antigens capable of stimulating a B cell response to a tumor can be an antigen found on the surface of tumor cell. Antigens capable of eliciting a B cell response to a tumor can be an intracellular neoantigen expressed in a tumor.

[0234] One or more antigens can include a combination of antigens capable of stimulating a T cell response (e.g., peptides including predicted T cell epitope sequences) and distinct antigens capable of stimulating a B cell response (e.g., full-length proteins, protein subunits, protein domains).

[0235] One or more antigens that stimulate an autoimmune response in a subject can be excluded from consideration in the context of vaccine generation for a subject.

[0236] The size of at least one antigenic peptide molecule (e.g., an epitope sequence) can comprise, but is not limited to, about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36, about 37, about 38, about 39, about 40, about 41, about 42, about 43, about 44, about 45, about 46, about 47, about 48, about 49, about 50, about 60, about 70, about 80, about 90, about 100, about 110, about 120 or greater amino molecule residues, and any range derivable therein. In specific embodiments the antigenic peptide molecules are equal to or less than 50 amino acids.

[0237] Antigenic peptides and polypeptides can be: for MHC Class I 15 residues or less in length and usually consist of between about 8 and about 11 residues, particularly 9 or 10 residues; for MHC Class II, 6-30 residues, inclusive.

[0238] If desirable, a longer peptide can be designed in several ways. In one case, when presentation likelihoods of peptides on HLA alleles are predicted or known, a longer peptide could consist of either: (1) individual presented peptides with an extensions of 2-5 amino acids toward the N- and C-terminus of each corresponding gene product; (2) a concatenation of some or all of the presented peptides with extended sequences for each. In another case, when sequencing reveals a long (>10 residues) neocpitope sequence present in the tumor (e.g. due to a frameshift, read-through or intron inclusion that leads to a novel peptide sequence), a longer peptide would consist of: (3) the entire

stretch of novel tumor-specific amino acids—thus bypassing the need for computational or in vitro test-based selection of the strongest HLA-presented shorter peptide. In both cases, use of a longer peptide allows endogenous processing by patient cells and may lead to more effective antigen presentation and stimulation of T cell responses. Longer peptides can also include a full-length protein, a protein subunit, a protein domain, and combinations thereof of a peptide, such as those expressed in a tumor. Longer peptides (e.g., full-length protein, protein subunit, or protein domain) and combinations thereof can be included to stimulate a B cell response.

[0239] Antigenic peptides and polypeptides can be presented on an HLA protein. In some aspects antigenic peptides and polypeptides are presented on an HLA protein with greater affinity than a wild-type peptide. In some aspects, an antigenic peptide or polypeptide can have an IC50 of at least less than 5000 nM, at least less than 1000 nM, at least less than 500 nM, at least less than 250 nM, at least less than 200 nM, at least less than 150 nM, at least less than 100 nM, at least less than 50 nM or less.

[0240] In some aspects, antigenic peptides and polypeptides do not induce an autoimmune response and/or invoke immunological tolerance when administered to a subject.

[0241] Also provided are compositions comprising at least two or more antigenic peptides. In some embodiments the composition contains at least two distinct peptides. At least two distinct peptides can be derived from the same polypeptide. By distinct polypeptides is meant that the peptide vary by length, amino acid sequence, or both. A peptide can include a tumor-specific mutation. Tumor-specific peptides can be derived from any polypeptide known to or have been found to contain a tumor specific mutation or peptides derived from any polypeptide known to or have been found to have altered expression in a tumor cell or cancerous tissue in comparison to a normal cell or tissue, for example any polypeptide known to or have been found to be aberrantly expressed in a tumor cell or cancerous tissue in comparison to a normal cell or tissue. The peptides can be derived from any polypeptide known to or suspected to be associated with an infectious disease organism, or peptides derived from any polypeptide known to or have been found to have altered expression in an infected cell in comparison to a normal cell or tissue (e.g., an infectious disease polynucleotide or polypeptide, including infectious disease polynucleotides or polypeptides with expression restricted to a host cell). Suitable polypeptides from which the antigenic peptides can be derived can be found for example in the COSMIC database or the AACR Genomics Evidence Neoplasia Information Exchange (GENIE) database. COSMIC curates comprehensive information on somatic mutations in human cancer. AACR GENIE aggregates and links clinical-grade cancer genomic data with clinical outcomes from tens of thousands of cancer patients. In some aspects the tumor specific mutation is a driver mutation for a particular cancer type.

[0242] Antigenic peptides and polypeptides having a desired activity or property can be modified to provide certain desired attributes, e.g., improved pharmacological characteristics, while increasing or at least retaining substantially all of the biological activity of the unmodified peptide to bind the desired MHC molecule and activate the appropriate T cell. For instance, antigenic peptide and polypeptides can be subject to various changes, such as substitutions, either conservative or non-conservative, where such changes might provide for certain advantages in their use, such as improved MHC binding, stability or presentation. By conservative substitutions is meant replacing an amino acid residue with another which is biologically and/or chemically similar, e.g., one hydrophobic residue for another, or one polar residue for another. The substitutions include combinations such as Gly, Ala; Val, Ile, Leu, Met; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr. The effect of single amino acid substitutions may also be probed using D-amino acids. Such modifications can be made using well known peptide synthesis procedures, as described in e.g., Merrifield, Science 232:341-347 (1986), Barany & Merrifield, The Peptides, Gross & Meienhofer, eds. (N.Y., Academic Press), pp. 1-284 (1979); and Stewart & Young, Solid Phase Peptide Synthesis, (Rockford, Ill., Pierce), 2d Ed. (1984).

[0243] Modifications of peptides and polypeptides with various amino acid mimetics or unnatural amino acids can be particularly useful in increasing the stability of the peptide and polypeptide in vivo. Stability can be assayed in a number of ways. For instance, peptidases and various biological media, such as human plasma and serum, have been used to test stability. Sec, e.g., Verhoef et al., Eur. J. Drug Metab Pharmacokin. 11:291-302 (1986). Half-life of the peptides can be conveniently determined using a 25% human serum (v/v) assay. The protocol is generally as follows. Pooled human serum (Type AB, non-heat inactivated) is delipidated by centrifugation before use. The serum is then diluted to 25% with RPMI tissue culture media and used to test peptide stability. At predetermined time intervals a small amount of reaction solution is removed and added to either 6% aqueous trichloracetic acid or ethanol. The cloudy reaction sample is cooled (4 degrees C.) for 15 minutes and then spun to pellet the precipitated serum proteins. The presence of the peptides is then determined by reversed-phase HPLC using stability-specific chromatography conditions.

[0244] The peptides and polypeptides can be modified to provide desired attributes other than improved serum half-life. For instance, the ability of the peptides to stimulate CTL activity can be enhanced by linkage to a sequence which contains at least one epitope that is capable of stimulating a T helper cell response. Immunogenic peptides/T helper conjugates can be linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, e.g., Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus can be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more

usually three to six residues. Alternatively, the peptide can be linked to the T helper peptide without a spacer.

[0245] An antigenic peptide can be linked to the T helper peptide cither directly or via a spacer either at the amino or carboxy terminus of the peptide. The amino terminus of either the antigenic peptide or the T helper peptide can be acylated. Exemplary T helper peptides include tetanus toxoid 830-843, influenza 307-319, malaria circumsporozoite 382-398 and 378-389.

[0246] Proteins or peptides can be made by any technique known to those of skill in the art, including the expression of proteins, polypeptides or peptides through standard molecular biological techniques, the isolation of proteins or peptides from natural sources, or the chemical synthesis of proteins or peptides. The nucleotide and protein, polypeptide and peptide sequences corresponding to various genes have been previously disclosed, and can be found at computerized databases known to those of ordinary skill in the art. One such database is the National Center for Biotechnology Information's Genbank and GenPept databases located at the National Institutes of Health website. The coding regions for known genes can be amplified and/or expressed using the techniques disclosed herein or as would be known to those of ordinary skill in the art. Alternatively, various commercial preparations of proteins, polypeptides and peptides are known to those of skill in the art.

[0247] In a further aspect an antigen includes a nucleic acid (e.g. polynucleotide) that encodes an antigenic peptide or portion thereof. The polynucleotide can be, e.g., DNA, cDNA, PNA, CNA, RNA (e.g., mRNA), either single- and/or double-stranded, or native or stabilized forms of polynucleotides, such as, e.g., polynucleotides with a phosphorothioate backbone, or combinations thereof and it may or may not contain introns. A polynucleotide sequence encoding an antigen can be sequence-optimized to improve expression, such as through improving transcription, translation, post-transcriptional processing, and/or RNA stability. For example, polynucleotide sequence encoding an antigen can be codon-optimized. "Codon-optimization" herein refers to replacing infrequently used codons, with respect to codon bias of a given organism, with frequently used synonymous codons. Polynucleotide sequences can be optimized to improve post-transcriptional processing, for example optimized to reduce unintended splicing, such as through removal of splicing motifs (e.g., canonical and/or cryptic/non-canonical splice donor, branch, and/or acceptor sequences) and/or introduction of exogenous splicing motifs (e.g., splice donor, branch, and/or acceptor sequences) to bias favored splicing events. Exogenous intron sequences include, but are not limited to, those derived from SV40 (e.g., an SV40 mini-intron) and derived from immunoglobulins (e.g., human β-globin gene). Exogenous intron sequences can be incorporated between a promoter/enhancer sequence and the antigen(s) sequence. Exogenous intron sequences for use in expression vectors are described in more detail in Callendret et al. (Virology. 2007 Jul. 5; 363 (2): 288-302), herein incorporated by reference for all purposes. Polynucleotide sequences can be optimized to improve transcript stability, for example through removal of RNA instability motifs (e.g., AU-rich elements and 3′ UTR motifs) and/or repetitive nucleotide sequences. Polynucleotide sequences can be optimized to improve accurate transcription, for example through removal of cryptic transcriptional initiators and/or terminators. Polynucleotide sequences can be optimized to improve translation and translational accuracy, for example through removal of cryptic AUG start codons, premature polyA sequences, and/or secondary structure motifs. Polynucleotide sequences can be optimized to improve nuclear export of transcripts, such as through addition of a Constitutive Transport Element (CTE), RNA Transport Element (RTE), or Woodchuck Posttranscriptional Regulatory Element (WPRE). Nuclear export signals for use in expression vectors are described in more detail in Callendret et al. (Virology. 2007 Jul. 5; 363 (2): 288-302), herein incorporated by reference for all purposes. Polynucleotide sequences can be optimized with respect to GC content, for example to reflect the average GC content of a given organism. Sequence optimization can balance one or more sequence properties, such as transcription, translation, post-transcriptional processing, and/or RNA stability. Sequence optimization can generate an optimal sequence balancing each of transcription, translation, post-transcriptional processing, and RNA stability. Sequence optimization algorithms are known to those of skill in the art, such as GeneArt (Thermo Fisher), Codon Optimization Tool (IDT), Cool Tool (University of Singapore), SGI-DNA (La Jolla California). One or more regions of an antigen-encoding protein can be sequence-optimized separately.

[0248] A still further aspect provides an expression vector capable of expressing a polypeptide or portion thereof. Expression vectors for different cell types are well known in the art and can be selected without undue experimentation. Generally, DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, DNA can be linked to the appropriate transcriptional and translational regulatory control nucleotide sequences recognized by the desired host, although such controls are generally available in the expression vector. The vector is then introduced into the host through standard techniques. Guidance can be found e.g. in Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

V. Vaccine Compositions

[0249] Also disclosed herein is an immunogenic composition, e.g., a vaccine composition, capable of raising a specific immune response, e.g., a tumor-specific immune response. Vaccine compositions typically comprise one or a plurality of antigens, e.g., selected using a method described herein or as set forth in Table A, Table 1.2, or AACR GENIE Results. Vaccine compositions can also be referred to as vaccines.

[0250] A vaccine can contain between 1 and 30 peptides, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 different peptides, 6, 7, 8, 9, 10 11, 12, 13, or 14 different peptides, or 12, 13 or

14 different peptides. Peptides can include post-translational modifications. A vaccine can contain between 1 and 100 or more nucleotide sequences, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more different nucleotide sequences, 6, 7, 8, 9, 10 11, 12, 13, or 14 different nucleotide sequences, or 12, 13 or 14 different nucleotide sequences. A vaccine can contain between 1 and 30 antigen sequences, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more different antigen sequences, 6, 7, 8, 9, 10 11, 12, 13, or 14 different antigen sequences, or 12, 13 or 14 different antigen sequences.

[0251] A vaccine can contain between 1 and 30 antigen-encoding nucleic acid sequences, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more different antigen-encoding nucleic acid sequences, 6, 7, 8, 9, 10 11, 12, 13, or 14 different antigen-encoding nucleic acid sequences, or 12, 13 or 14 different antigen-encoding nucleic acid sequences. Antigen-encoding nucleic acid sequences can refer to the antigen encoding portion of an "antigen cassette." Features of an antigen cassette are described in greater detail herein. An antigen-encoding nucleic acid sequence can contain one or more epitope-encoding nucleic acid sequences (e.g., an antigen-encoding nucleic acid sequence encoding concatenated T cell epitopes).

[0252] A vaccine can contain between 1 and 30 distinct epitope-encoding nucleic acid sequences, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more distinct epitope-encoding nucleic acid sequences, 6, 7, 8, 9, 10 11, 12, 13, or 14 distinct epitope-encoding nucleic acid sequences, or 12, 13 or 14 distinct epitope-encoding nucleic acid sequences. Epitope-encoding nucleic acid sequences can refer to sequences for individual epitope sequences, such as each of the T cell epitopes in an antigen-encoding nucleic acid sequence encoding concatenated T cell epitopes.

[0253] A vaccine can contain at least two iterations of an epitope-encoding nucleic acid sequence. A used herein, an "iteration" (or interchangeably a "repeat") refers to two or more identical nucleic acid epitope-encoding nucleic acid sequences (inclusive of the optional 5′ linker sequence and/or the optional 3′ linker sequences described herein) within an antigen-encoding nucleic acid sequence. In one example, the antigen-encoding nucleic acid sequence portion of a cassette encodes at least two iterations of an epitope-encoding nucleic acid sequence. In further non-limiting examples, the antigen-encoding nucleic acid sequence portion of a cassette encodes more than one distinct epitope, and at least one of the distinct epitopes is encoded by at least two iterations of the nucleic acid sequence encoding the distinct epitope (i.e., at least two distinct epitope-encoding nucleic acid sequences). In illustrative non-limiting examples, an antigen-encoding nucleic acid sequence encodes epitopes A, B, and C encoded by epitope-encoding nucleic acid sequences epitope-encoding sequence A (E.sub.A), epitope-encoding sequence B (E.sub.B), and epitope-encoding sequence C (E.sub.C), and exemplary antigen-encoding nucleic acid sequences having iterations of at least one of the distinct epitopes are illustrated by, but is not limited to, the formulas below:

Iteration of One Distinct Epitope (Iteration of Epitope A):

[0254] E.sub.A-E.sub.B-E.sub.C-E.sub.A; or [0255] E.sub.A-E.sub.A-E.sub.B-E.sub.C

Iteration of Multiple Distinct Epitopes (Iterations of Epitopes a, B, and C):

[0256] E.sub.A-E.sub.B-E.sub.C-E.sub.A-E.sub.B-E.sub.C; or [0257] E.sub.A-E.sub.A-E.sub.B-E.sub.B-E.sub.C-E.sub.C

Multiple Iterations of Multiple Distinct Epitopes (Iterations of Epitopes a, B, and C):

[0258] E.sub.A-E.sub.B-E.sub.C-E.sub.A-E.sub.B-E.sub.C-E.sub.A-E.sub.B-E.sub.C; or [0259] E.sub.A-E.sub.A-E.sub.A-E.sub.B-E.sub.B-E.sub.B-E.sub.C-E.sub.C-E.sub.C

[0260] The above examples are not limiting and the antigen-encoding nucleic acid sequences having iterations of at least one of the distinct epitopes can encode each of the distinct epitopes in any order or frequency. For example, the order and frequency can be a random arrangement of the distinct epitopes, e.g., in an example with epitopes A, B, and C, by the formula E.sub.A-E.sub.B-E.sub.C-E.sub.C-E.sub.A-E.sub.B-E.sub.A-E.sub.C-E.sub.A-E.sub.C-E.sub.C-E.sub.B.

[0261] Also provided for herein is an antigen-encoding cassette, the antigen-encoding cassette having at least one antigen-encoding nucleic acid sequence described, from 5′ to 3′, by the formula:

(E.sub.x-(E.sup.N.sub.n).sub.y).sub.z [0262] where E represents a nucleotide sequence including a distinct epitope-encoding nucleic acid sequences, [0263] n represents the number of separate distinct epitope-encoding nucleic acid sequences and is any integer including 0, [0264] E.sup.N represents a nucleotide sequence comprising the separate distinct epitope-encoding nucleic acid sequence for each corresponding n, [0265] for each iteration of z: x=0 or 1, y=0

or 1 for each n, and at least one of x or y=1, and z=2 or greater, wherein the antigen-encoding nucleic acid sequence comprises at least two iterations of E, a given E^N, or a combination thereof.

[0266] Each E or E^N can independently comprise any epitope-encoding nucleic acid sequence described herein (e.g., a peptide encoding an infectious disease T cell epitope and/or a neoantigen epitope). For example, Each E or E^N can independently comprises a nucleotide sequence described, from 5′ to 3′, by the formula (L5$_{b}$-N$_{c}$-L3$_{d}$), where N comprises the distinct epitope-encoding nucleic acid sequence associated with each E or E^N, where c=1, L5 comprises a 5′ linker sequence, where b=0 or 1, and L3 comprises a 3′ linker sequence, where d=0 or 1. Epitopes and linkers that can be used are further described herein.

[0267] Iterations of an epitope-encoding nucleic acid sequences (inclusive of optional 5′ linker sequence and/or the optional 3′ linker sequences) can be linearly linked directly to one another (e.g., E_A-E_A-. . . as illustrated above). Iterations of an epitope-encoding nucleic acid sequences can be separated by one or more additional nucleotides sequences. In general, iterations of an epitope-encoding nucleic acid sequences can be separated by any size nucleotide sequence applicable for the compositions described herein. In one example, iterations of an epitope-encoding nucleic acid sequences can be separated by a separate distinct epitope-encoding nucleic acid sequence (e.g., E_A-E_B-E_C-E_A . . . , as illustrated above). In examples where iterations are separated by a single separate distinct epitope-encoding nucleic acid sequence, and each epitope-encoding nucleic acid sequences (inclusive of optional 5′ linker sequence and/or the optional 3′ linker sequences) encodes a peptide 25 amino acids in length, the iterations can be separated by 75 nucleotides, such as in antigen-encoding nucleic acid represented by E_A-E_B-E_A . . . , E_A is separated by 75 nucleotides. In an illustrative example, an antigen-encoding nucleic acid having the sequence VTNTEMFVTAPDNLGYMYEVQWPGQTQPQIANCSVYDFFVWLHYYSVRDTVTNTEMF VTAPDNLGYMYEVQWPGQTQPQIANCSVYDFFVWLHYYSVRDT (SEQ ID NO:85) encoding iterations of 25mer antigens Trp1 (VTNTEMFVTAPDNLGYMYEVQWPGQ [SEQ ID NO: 86]) and Trp2 (TQPQIANCSVYDFFVWLHYYSVRDT [SEQ ID NO:87]), the iterations of Trp1 are separated by the 25mer Trp2 and thus the repeats of the Trp1 epitope-encoding nucleic acid sequences are separated the 75 nucleotide Trp2 epitope-encoding nucleic acid sequence. In examples where iterations are separated by 2, 3, 4, 5, 6, 7, 8, or 9 separate distinct epitope-encoding nucleic acid sequence, and each epitope-encoding nucleic acid sequences (inclusive of optional 5′ linker sequence and/or the optional 3′ linker sequences) encodes a peptide 25 amino acids in length, the iterations can be separated by 150, 225, 300, 375, 450, 525, 600, or 675 nucleotides, respectively.

[0268] In some instances, an antigen or epitope in a cassette encoding additional antigens and/or epitopes may be an immunodominant epitope relative to the others encoded.

[0269] Immunodominance, in general, is the skewing of an immune response towards only one or a few specific immunogenic peptides. Immunodominance can be assessed as part of an immune monitoring protocol. For example, immunodominance can be assessed through evaluating T cell and/or B cell responses to the encoded antigens.

[0270] In some instances, it may be desired to avoid vaccine compositions containing immunodominant epitope. For example, it may be desired to avoid designing a vaccine cassette encoding an immunodominant epitope. Without wishing to be bound by theory, administering and/or encoding an immunodominant epitope together with additional epitope may reduce the immune response to the additional epitopes, including potentially ultimately reducing vaccine efficacy against the additional epitopes. As an illustrative non-limiting example, vaccine compositions including TP53-associated neoepitopes may have the immune response, e.g., a T cell response, skewed towards the TP53-associated neoepitope negatively impacting the immune response to other antigens or epitopes in the vaccine composition.

[0271] In one embodiment, different peptides and/or polypeptides or nucleotide sequences encoding them are selected so that the peptides and/or polypeptides capable of associating with different MHC molecules, such as different MHC class I molecules and/or different MHC class II molecules. In some aspects, one vaccine composition comprises coding sequence for peptides and/or polypeptides capable of associating with the most frequently occurring MHC class I molecules and/or different MHC class II molecules. Hence, vaccine compositions can comprise different fragments capable of associating with at least 2 preferred, at least 3 preferred, or at least 4 preferred MHC class I molecules and/or different MHC class II molecules.

[0272] The vaccine composition can be capable of stimulating a specific cytotoxic T-cell response and/or a specific helper T-cell response. The vaccine composition can be capable of stimulating a specific cytotoxic T-cell response and a specific helper T-cell response.

[0273] The vaccine composition can be capable of stimulating a specific B-cell response (e.g., an antibody response).

[0274] The vaccine composition can be capable of stimulating a specific cytotoxic T-cell response, a specific helper T-cell response, and/or a specific B-cell response. The vaccine composition can be capable of stimulating a specific cytotoxic T-cell response and a specific B-cell response. The vaccine composition can be capable of stimulating a specific helper T-cell response and a specific B-cell response. The vaccine composition can be capable of stimulating a specific cytotoxic T-cell response, a specific helper T-cell response, and a specific B-cell response.

[0275] A vaccine composition can further comprise an adjuvant and/or a carrier. Examples of useful adjuvants and carriers are given herein below. A composition can be associated with a carrier such as e.g. a protein or an antigen-presenting cell such as e.g. a dendritic cell (DC) capable of presenting the peptide to a T-cell.

[0276] Adjuvants are any substance whose admixture into a vaccine composition increases or otherwise modifies the

immune response to an antigen. Carriers can be scaffold structures, for example a polypeptide or a polysaccharide, to which an antigen, is capable of being associated. Optionally, adjuvants are conjugated covalently or non-covalently.

[0277] The ability of an adjuvant to increase an immune response to an antigen is typically manifested by a significant or substantial increase in an immune-mediated reaction, or reduction in disease symptoms. For example, an increase in humoral immunity is typically manifested by a significant increase in the titer of antibodies raised to the antigen, and an increase in T-cell activity is typically manifested in increased cell proliferation, or cellular cytotoxicity, or cytokine secretion. An adjuvant may also alter an immune response, for example, by changing a primarily humoral or Th response into a primarily cellular, or Th response.

[0278] Suitable adjuvants include, but are not limited to 1018 ISS, alum, aluminum salts, Amplivax, AS15, BCG, CP-870,893, CpG7909, CyaA, dSLIM, GM-CSF, IC30, IC31, Imiquimod, ImuFact IMP321, IS Patch, ISS, ISCOMATRIX, JuvImmunc, Lipo Vac, MF59, monophosphoryl lipid A, Montanide IMS 1312, Montanide ISA 206, Montanide ISA 50V, Montanide ISA-51, OK-432, OM-174, OM-197-MP-E.sub.C, ONTAK, PepTel vector system, PLG microparticles, resiquimod, SRL172, Virosomes and other Virus-like particles, YF-17D, VEGF trap, R848, beta-glucan, Pam3Cys, Aquila's QS21 stimulon (Aquila Biotech, Worcester, Mass., USA) which is derived from saponin, mycobacterial extracts and synthetic bacterial cell wall mimics, and other proprietary adjuvants such as Ribi's Detox. Quil or Superfos. Adjuvants such as incomplete Freund's or GM-CSF are useful. Several immunological adjuvants (e.g., MF59) specific for dendritic cells and their preparation have been described previously (Dupuis M, ct al., Cell Immunol. 1998; 186 (1): 18-27; Allison A C; Dev Biol Stand. 1998; 92:3-11). Also cytokines can be used. Several cytokines have been directly linked to influencing dendritic cell migration to lymphoid tissues (e.g., TNF-alpha), accelerating the maturation of dendritic cells into efficient antigen-presenting cells for T-lymphocytes (e.g., GM-CSF, IL-1 and IL-4) (U.S. Pat. No. 5,849,589, specifically incorporated herein by reference in its entirety) and acting as immunoadjuvants (e.g., IL-12) (Gabrilovich D I, et al., J Immunother Emphasis Tumor Immunol. 1996 (6): 414-418).

[0279] CpG immunostimulatory oligonucleotides have also been reported to enhance the effects of adjuvants in a vaccine setting. Other TLR binding molecules such as RNA binding TLR 7, TLR 8 and/or TLR 9 may also be used.

[0280] Other examples of useful adjuvants include, but are not limited to, chemically modified CpGs (e.g. CpR, Idera), Poly(I:C) (e.g. polyi:CI2U), non-CpG bacterial DNA or RNA as well as immunoactive small molecules and antibodies such as cyclophosphamide, sunitinib, bevacizumab, celebrex, NCX-4016, sildenafil, tadalafil, vardenafil, sorafinib, XL-999, CP-547632, pazopanib, ZD2171, AZD2171, ipilimumab, tremclimumab, and SC58175, which may act therapeutically and/or as an adjuvant. The amounts and concentrations of adjuvants and additives can readily be determined by the skilled artisan without undue experimentation. Additional adjuvants include colony-stimulating factors, such as Granulocyte Macrophage Colony Stimulating Factor (GM-CSF, sargramostim).

[0281] A vaccine composition can comprise more than one different adjuvant. Furthermore, a therapeutic composition can comprise any adjuvant substance including any of the above or combinations thereof. It is also contemplated that a vaccine and an adjuvant can be administered together or separately in any appropriate sequence.

[0282] A carrier (or excipient) can be present independently of an adjuvant. The function of a carrier can for example be to increase the molecular weight of in particular mutant to increase activity or immunogenicity, to confer stability, to increase the biological activity, or to increase serum half-life. Furthermore, a carrier can aid presenting peptides to T-cells. A carrier can be any suitable carrier known to the person skilled in the art, for example a protein or an antigen presenting cell. A carrier protein could be but is not limited to keyhole limpet hemocyanin, serum proteins such as transferrin, bovine serum albumin, human serum albumin, thyroglobulin or ovalbumin, immunoglobulins, or hormones, such as insulin or palmitic acid. For immunization of humans, the carrier is generally a physiologically acceptable carrier acceptable to humans and safe. However, tetanus toxoid and/or diphtheria toxoid are suitable carriers. Alternatively, the carrier can be dextrans for example sepharose.

[0283] Cytotoxic T-cells (CTLs) recognize an antigen in the form of a peptide bound to an MHC molecule rather than the intact foreign antigen itself. The MHC molecule itself is located at the cell surface of an antigen presenting cell. Thus, an activation of CTLs is possible if a trimeric complex of peptide antigen, MHC molecule, and APC is present. Correspondingly, it may enhance the immune response if not only the peptide is used for activation of CTLs, but if additionally APCs with the respective MHC molecule are added. Therefore, in some embodiments a vaccine composition additionally contains at least one antigen presenting cell.

[0284] Antigens can also be included in viral vector-based vaccine platforms, such as vaccinia, fowlpox, self-replicating alphavirus, marabavirus, adenovirus (See, e.g., Tatsis et al., Adenoviruses, *Molecular Therapy* (2004) 10, 616-629), or lentivirus, including but not limited to second, third or hybrid second/third generation lentivirus and recombinant lentivirus of any generation designed to target specific cell types or receptors (See, e.g., Hu et al., Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases, *Immunol. Rev.* (2011) 239 (1): 45-61, Sakuma et al., Lentiviral vectors: basic to translational, *Biochem J.* (2012) 443 (3): 603-18, Cooper et al., Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter, *Nucl. Acids Res.* (2015) 43 (1): 682-690, Zufferey et al., Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery, *J. Virol.* (1998) 72 (12): 9873-9880). Dependent on the packaging capacity of the above mentioned viral vector-based vaccine platforms, this approach can deliver one or more nucleotide sequences that encode one or more antigen peptides. The sequences may be flanked by non-mutated sequences, may be separated by

linkers or may be preceded with one or more sequences targeting a subcellular compartment (See, e.g., Gros et al., Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients, *Nat Med*. (2016) 22 (4): 433-8, Stronen et al., Targeting of cancer neoantigens with donor-derived T cell receptor repertoires, *Science*. (2016) 352 (6291): 1337-41, Lu et al., Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions, *Clin Cancer Res*. (2014) 20 (13): 3401-10). Upon introduction into a host, infected cells express the antigens, and thereby stimulate a host immune (e.g., CTL) response against the peptide(s). Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Pat. No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (Nature 351:456-460 (1991)). A wide variety of other vaccine vectors useful for therapeutic administration or immunization of antigens, e.g., *Salmonella typhi* vectors, and the like will be apparent to those skilled in the art from the description herein.

V.A. Antigen Cassette

[0285] The methods employed for the selection of one or more antigens, the cloning and construction of an "antigen cassette" and its insertion into a viral vector are within the skill in the art given the teachings provided herein. By "antigen cassette" or "cassette" is meant the combination of a selected antigen or plurality of antigens (e.g., antigen-encoding nucleic acid sequences) and the other regulatory elements necessary to transcribe the antigen(s) and express the transcribed product. The selected antigen or plurality of antigens can refer to distinct epitope sequences, e.g., an antigen-encoding nucleic acid sequence in the cassette can encode an epitope-encoding nucleic acid sequence (or plurality of epitope-encoding nucleic acid sequences) such that the epitopes are transcribed and expressed. An antigen or plurality of antigens can be operatively linked to regulatory components in a manner which permits transcription. Such components include conventional regulatory elements that can drive expression of the antigen(s) in a cell transfected with the viral vector. Thus the antigen cassette can also contain a selected promoter which is linked to the antigen(s) and located, with other, optional regulatory elements, within the selected viral sequences of the recombinant vector. Cassettes can include one or more neoantigens shown in Table A and/or AACR GENIE Results, and/or one or more antigens shown in Table 1.2.

[0286] A cassette can have one or more antigen-encoding nucleic acid sequences, such as a cassette containing multiple antigen-encoding nucleic acid sequences each independently operably linked to separate promoters and/or linked together using other multicistonic systems, such as 2A ribosome skipping sequence elements (e.g., E2A, P2A, F2A, or T2A sequences) or Internal Ribosome Entry Site (IRES) sequence elements. A linker can also have a cleavage site, such as a TEV or furin cleavage site. Linkers with cleavage sites can be used in combination with other elements, such as those in a multicistronic system. In a non-limiting illustrative example, a furin protease cleavage site can be used in conjunction with a 2A ribosome skipping sequence element such that the furin protease cleavage site is configured to facilitate removal of the 2A sequence following translation. In a cassette containing more than one antigen-encoding nucleic acid sequences, each antigen-encoding nucleic acid sequence can contain one or more epitope-encoding nucleic acid sequences (e.g., an antigen-encoding nucleic acid sequence encoding concatenated T cell epitopes).

[0287] Useful promoters can be constitutive promoters or regulated (inducible) promoters, which will enable control of the amount of antigen(s) to be expressed. For example, a desirable promoter is that of the cytomegalovirus immediate early promoter/enhancer [see, e.g., Boshart et al, Cell, 41:521-530 (1985)]. Another desirable promoter includes the Rous sarcoma virus LTR promoter/enhancer. Still another promoter/enhancer sequence is the chicken cytoplasmic beta-actin promoter [T. A. Kost et al, Nucl. Acids Res., 11 (23): 8287 (1983)]. Other suitable or desirable promoters can be selected by one of skill in the art.

[0288] The antigen cassette can also include nucleic acid sequences heterologous to the viral vector sequences including sequences providing signals for efficient polyadenylation of the transcript (poly(A), poly-A or pA) and introns with functional splice donor and acceptor sites. A common poly-A sequence which is employed in the exemplary vectors of this invention is that derived from the papovavirus SV-40. The poly-A sequence generally can be inserted in the cassette following the antigen-based sequences and before the viral vector sequences. A common intron sequence can also be derived from SV-40, and is referred to as the SV-40 T intron sequence. An antigen cassette can also contain such an intron, located between the promoter/enhancer sequence and the antigen(s). Selection of these and other common vector elements are conventional [see, e.g., Sambrook et al, "Molecular Cloning. A Laboratory Manual.", 2d edit., Cold Spring Harbor Laboratory, New York (1989) and references cited therein] and many such sequences are available from commercial and industrial sources as well as from Genbank.

[0289] An antigen cassette can have one or more antigens. For example, a given cassette can include 1-10, 1-20, 1-30, 10-20, 15-25, 15-20, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more antigens. Antigens can be linked directly to one another. Antigens can also be linked to one another with linkers. Antigens can be in any orientation relative to one another including N to C or C to N.

[0290] As described elsewhere herein, the antigen cassette can be located in the site of any selected deletion in a viral vector, such as the deleted structural proteins of a VEE backbone or the site of the E1 gene region deletion or E3 gene region deletion of a ChAd-based vector, among others which may be selected.

[0291] The antigen cassette can be described using the following formula to describe the ordered sequence of each element, from 5′ to 3′:

(P.sub.a-(L5.sub.b-N.sub.c-L3.sub.d).sub.X).sub.Z-(P2.sub.h-(G5.sub.e-U.sub.f).sub.Y).sub.W-G3.sub.g [0292] wherein P and P2 comprise promoter nucleotide sequences, N comprises an MHC class I epitope encoding nucleic acid sequence, L5 comprises a 5′ linker sequence, L3 comprises a 3′ linker sequence, G5 comprises a nucleic acid sequences encoding an amino acid linker, G3 comprises one of the at least one nucleic acid sequences encoding an amino acid linker, U comprises an MHC class II antigen-encoding nucleic acid sequence, where for each X the corresponding Ne is an epitope encoding nucleic acid sequence, where for each Y the corresponding Uf is a MHC class II epitope-encoding nucleic acid sequence (e.g., universal MHC class II epitope-encoding nucleic acid sequence). A universal sequence can comprise at least one of Tetanus toxoid and PADRE. A universal sequence can comprise a Tetanus toxoid peptide. A universal sequence can comprise a PADRE peptide. A universal sequence can comprise a Tetanus toxoid and PADRE peptides. The composition and ordered sequence can be further defined by selecting the number of elements present, for example where a=0 or 1, where b=0 or 1, where c=1, where d=0 or 1, where e=0 or 1, where f=1, where g=0 or 1, where h=0 or 1, X=1 to 400, Y=0, 1, 2, 3, 4 or 5, Z=1 to 400, and W=0, 1, 2, 3, 4 or 5.

[0293] In one example, elements present include where a=0, b=1, d=1, c=1, g=1, h=0, X=10, Y=2, Z=1, and W=1, describing where no additional promoter is present (e.g. only the promoter nucleotide sequence provided by a vector backbone, such as an RNA alphavirus backbone is present), 10 MHC class I epitopes are present, a 5′ linker is present for each N, a 3′ linker is present for each N, 2 MHC class II epitopes are present, a linker is present linking the two MHC class II epitopes, a linker is present linking the 5′ end of the two MHC class II epitopes to the 3′ linker of the final MHC class I epitope, and a linker is present linking the 3′ end of the two MHC class II epitopes to the to a vector backbone (e.g., an RNA alphavirus backbone). Examples of linking the 3′ end of the antigen cassette to a vector backbone (e.g., an RNA alphavirus backbone) include linking directly to the 3′ UTR elements provided by the vector backbone, such as a 3′ 19-nt CSE. Examples of linking the 5′ end of the antigen cassette to a vector backbone (e.g., an RNA alphavirus backbone) include linking directly to a promoter or 5′ UTR element of the vector backbone, such as a subgenomic promoter sequence (e.g., a 26S subgenomic promoter sequence), an alphavirus 5′ UTR, a 51-nt CSE, or a 24-nt CSE.

[0294] Other examples include: where a=1 describing where a promoter other than the promoter nucleotide sequence provided by a vector backbone (e.g., an RNA alphavirus backbone) is present; where a=1 and Z is greater than 1 where multiple promoters other than the promoter nucleotide sequence provided by the vector backbone are present each driving expression of 1 or more distinct MHC class I epitope encoding nucleic acid sequences; where h =1 describing where a separate promoter is present to drive expression of the MHC class II epitope-encoding nucleic acid sequences; and where g=0 describing the MHC class II epitope-encoding nucleic acid sequence, if present, is directly linked to a vector backbone (e.g., an RNA alphavirus backbone).

[0295] Other examples include where each MHC class I epitope that is present can have a 5′ linker, a 3′ linker, neither, or both. In examples where more than one MHC class I epitope is present in the same antigen cassette, some MHC class I epitopes may have both a 5′ linker and a 3′ linker, while other MHC class I epitopes may have either a 5′ linker, a 3′ linker, or neither. In other examples where more than one MHC class I epitope is present in the same antigen cassette, some MHC class I epitopes may have either a 5′ linker or a 3′ linker, while other MHC class I epitopes may have either a 5′ linker, a 3′ linker, or neither.

[0296] In examples where more than one MHC class II epitope is present in the same antigen cassette, some MHC class II epitopes may have both a 5′ linker and a 3′ linker, while other MHC class II epitopes may have either a 5′ linker, a 3′ linker, or neither. In other examples where more than one MHC class II epitope is present in the same antigen cassette, some MHC class II epitopes may have either a 5′ linker or a 3′ linker, while other MHC class II epitopes may have either a 5′ linker, a 3′ linker, or neither.

[0297] Other examples include where each antigen that is present can have a 5′ linker, a 3′ linker, neither, or both. In examples where more than one antigen is present in the same antigen cassette, some antigens may have both a 5′ linker and a 3′ linker, while other antigens may have either a 5′ linker, a 3′ linker, or neither. In other examples where more than one antigen is present in the same antigen cassette, some antigens may have either a 5′ linker or a 3′ linker, while other antigens may have either a 5′ linker, a 3′ linker, or neither.

[0298] The promoter nucleotide sequences P and/or P2 can be the same as a promoter nucleotide sequence provided by a vector backbone, such as an RNA alphavirus backbone. For example, the promoter sequence provided by the vector backbone, Pn and P2, can each comprise a subgenomic promoter sequence (e.g., a 26S subgenomic promoter sequence) or a CMV promoter. The promoter nucleotide sequences P and/or P2 can be different from the promoter nucleotide sequence provided by a vector backbone (e.g., an RNA alphavirus backbone), as well as can be different from each other.

[0299] The 5′ linker L5 can be a native sequence or a non-natural sequence. Non-natural sequence include, but are not limited to, AAY, RR, and DPP. The 3′ linker L3 can also be a native sequence or a non-natural sequence. Additionally, L5 and L3 can both be native sequences, both be non-natural sequences, or one can be native and the other non-natural. For each X, the amino acid linkers can be 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more amino acids in length. For each X, the amino acid linkers can be

also be at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, or at least 30 amino acids in length.

[0300] The amino acid linker G5, for each Y, can be 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more amino acids in length. For each Y, the amino acid linkers can be also be at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, or at least 30 amino acids in length.

[0301] The amino acid linker G3 can be 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more amino acids in length. G3 can be also be at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, or at least 30 amino acids in length.

[0302] For each X, each N can encode a MHC class I epitope, a MHC class II epitope, an epitope/antigen capable of stimulating a B cell response, or a combination thereof. For each X, each N can encode a combination of a MHC class I epitope, a MHC class II epitope, and an epitope/antigen capable of stimulating a B cell response. For each X, each N can encode a combination of a MHC class I epitope and a MHC class II epitope. For each X, each N can encode a combination of a MHC class I epitope and an epitope/antigen capable of stimulating a B cell response. For each X, each N can encode a combination of a MHC class II epitope and an epitope/antigen capable of stimulating a B cell response. For each X, each N can encode a MHC class II epitope. For each X, each N can encode an epitope/antigen capable of stimulating a B cell response. For each X, each N can encode a MHC class I epitope 7-15 amino acids in length. For each X, each N can also encodes a MHC class I epitope 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 amino acids in length. For each X, each N can also encodes a MHC class I epitope at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, or at least 30 amino acids in length.

[0303] The cassette encoding the one or more antigens can be 700 nucleotides or less. The cassette encoding the one or more antigens can be 700 nucleotides or less and encode 2 distinct epitope-encoding nucleic acid sequences (e.g., encode 2 distinct infectious disease or tumor derived nucleic acid sequences encoding an immunogenic polypeptide). The cassette encoding the one or more antigens can be 700 nucleotides or less and encode at least 2 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be 700 nucleotides or less and encode 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be 700 nucleotides or less and encode at least 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be 700 nucleotides or less and include 1-10, 1-5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more antigens.

[0304] The cassette encoding the one or more antigens can be between 375-700 nucleotides in length. The cassette encoding the one or more antigens can be between 375-700 nucleotides in length and encode 2 distinct epitope-encoding nucleic acid sequences (e.g., encode 2 distinct infectious disease or tumor derived nucleic acid sequences encoding an immunogenic polypeptide). The cassette encoding the one or more antigens can be between 375-700 nucleotides in length and encode at least 2 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be between 375-700 nucleotides in length and encode 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens be between 375-700 nucleotides in length and encode at least 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be between 375-700 nucleotides in length and include 1-10, 1-5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more antigens.

[0305] The cassette encoding the one or more antigens can be 600, 500, 400, 300, 200, or 100 nucleotides in length or less. The cassette encoding the one or more antigens can be 600, 500, 400, 300, 200, or 100 nucleotides in length or less and encode 2 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be 600, 500, 400, 300, 200, or 100 nucleotides in length or less and encode at least 2 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be 600, 500, 400, 300, 200, or 100 nucleotides in length or less and encode 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be 600, 500, 400, 300, 200, or 100 nucleotides in length or less and encode at least 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be 600, 500, 400, 300, 200, or 100 nucleotides in length or less and include 1-10, 1-5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more antigens.

[0306] The cassette encoding the one or more antigens can be between 375-600, between 375-500, or between 375-400 nucleotides in length. The cassette encoding the one or more antigens can be between 375-600, between 375-500, or between 375-400 nucleotides in length and encode 2 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be between 375-600, between 375-500, or between 375-400 nucleotides in

length and encode at least 2 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be between 375-600, between 375-500, or between 375-400 nucleotides in length and encode 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be between 375-600, between 375-500, or between 375-400 nucleotides in length and encode at least 3 distinct epitope-encoding nucleic acid sequences. The cassette encoding the one or more antigens can be between 375-600, between 375-500, or between 375-400 nucleotides in length and include 1-10, 1-5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more antigens.

V.B. Immune Modulators

[0307] Vectors described herein, such as C68 vectors described herein or alphavirus vectors described herein, can comprise a nucleic acid which encodes at least one antigen and the same or a separate vector can comprise a nucleic acid which encodes at least one immune modulator. An immune modulator can include a binding molecule (e.g., an antibody such as an scFv) which binds to and blocks the activity of an immune checkpoint molecule. An immune modulator can include a cytokine, such as IL-2, IL-7, IL-12 (including IL-12 p35, p40, p70, and/or p70-fusion constructs), IL-15, or IL-21. An immune modulator can include a modified cytokine (e.g., pegIL-2). Vectors can comprise an antigen cassette and one or more nucleic acid molecules encoding an immune modulator.

[0308] Illustrative immune checkpoint molecules that can be targeted for blocking or inhibition include, but are not limited to, CTLA-4, 4-1BB (CD137), 4-1BBL (CD137L), PDL1, PDL2, PD1, B7-H3, B7-H4, BTLA, HVEM, TIM3, GAL9, LAG3, TIM3, B7H3, B7H4, VISTA, KIR, 2B4 (belongs to the CD2 family of molecules and is expressed on all NK, □□, and memory CD8+ (□□) T cells), CD160 (also referred to as BY55), and CGEN-15049. Immune checkpoint inhibitors include antibodies, or antigen binding fragments thereof, or other binding proteins, that bind to and block or inhibit the activity of one or more of CTLA-4, PDL1, PDL2, PD1, B7-H3, B7-H4, BTLA, HVEM, TIM3, GAL9, LAG3, TIM3, B7H3, B7H4, VISTA, KIR, 2B4, CD160, and CGEN-15049. Illustrative immune checkpoint inhibitors include Tremelimumab (CTLA-4 blocking antibody), anti-OX40, PD-L1 monoclonal Antibody (Anti-B7-H1; MEDI4736), ipilimumab, MK-3475 (PD-1 blocker), Nivolumab (anti-PD1 antibody), Cemiplimab (anti-PD1 antibody), CT-011 (anti-PD1 antibody), BY55 monoclonal antibody, AMP224 (anti-PDL1 antibody), BMS-936559 (anti-PDL1 antibody), MPDL3280A/Atczolizumab (anti-PDL1 antibody), MSB0010718C (anti-PDL1 antibody) and Yervoy/ipilimumab (anti-CTLA-4 checkpoint inhibitor). Antibody-encoding sequences can be engineered into vectors such as C68 using ordinary skill in the art. An exemplary method is described in Fang et al., Stable antibody expression at therapeutic levels using the 2A peptide. Nat Biotechnol. 2005 May; 23 (5): 584-90. Epub 2005 Apr. 17; herein incorporated by reference for all purposes.

V.C. Additional Considerations for Vaccine Design and Manufacture

V.C.1. Determination of a Set of Peptides that Cover All Tumor Subclones

[0309] Truncal peptides, meaning those presented by all or most tumor subclones, can be prioritized for inclusion into a vaccine..sup.53 Optionally, if there are no truncal peptides predicted to be presented and immunogenic with high probability, or if the number of truncal peptides predicted to be presented and immunogenic with high probability is small enough that additional non-truncal peptides can be included in the vaccine, then further peptides can be prioritized by estimating the number and identity of tumor subclones and choosing peptides so as to maximize the number of tumor subclones covered by the vaccine..sup.54

V.C.2. Antigen Prioritization

[0310] After all of the above antigen filters are applied, more candidate antigens may still be available for vaccine inclusion than the vaccine technology can support. Additionally, uncertainty about various aspects of the antigen analysis may remain and tradeoffs may exist between different properties of candidate vaccine antigens. Thus, in place of predetermined filters at each step of the selection process, an integrated multi-dimensional model can be considered that places candidate antigens in a space with at least the following axes and optimizes selection using an integrative approach. [0311] 1. Risk of auto-immunity or tolerance (risk of germlinc) (lower risk of auto-immunity is typically preferred) [0312] 2. Probability of sequencing artifact (lower probability of artifact is typically preferred) [0313] 3. Probability of immunogenicity (higher probability of immunogenicity is typically preferred) [0314] 4. Probability of presentation (higher probability of presentation is typically preferred) [0315] 5. Gene expression (higher expression is typically preferred) [0316] 6. Coverage of HLA genes (larger number of HLA molecules involved in the presentation of a set of antigens may lower the probability that a tumor will escape immune attack via downregulation or mutation of HLA molecules) [0317] 7. Coverage of HLA classes (covering both HLA-I and HLA-II may increase the probability of therapeutic response and decrease the probability of tumor escape)

[0318] Additionally, optionally, antigens can be deprioritized (e.g., excluded) from the vaccination if they are predicted to be presented by HLA alleles lost or inactivated in either all or part of the patient's tumor. HLA allele loss can occur by either somatic mutation, loss of heterozygosity, or homozygous deletion of the locus. Methods for detection of HLA allele somatic mutation are well known in the art, e.g. (Shukla et al., 2015). Methods for detection of somatic LOH and homozygous deletion (including for HLA locus) are likewise well described. (Carter et al., 2012; McGranahan et al., 2017; Van Loo et al., 2010). Antigens can also be deprioritized if mass-spectrometry data indicates a predicted antigen is not presented by a predicted HLA allele.

V.D. Alphavirus

V.D.1. Alphavirus Biology

[0319] Alphaviruses are members of the family Togaviridae, and are positive-sense single stranded RNA viruses. Members are typically classified as either Old World, such as Sindbis, Ross River, Mayaro, Chikungunya, and Semliki Forest viruses, or New World, such as eastern equine encephalitis, Aura, Fort Morgan, or Venezuelan equine encephalitis virus and its derivative strain TC-83 (Strauss Microbial Review 1994). A natural alphavirus genome is typically around 12 kb in length, the first two-thirds of which contain genes encoding non-structural proteins (nsPs) that form RNA replication complexes for self-replication of the viral genome, and the last third of which contains a subgenomic expression cassette encoding structural proteins for virion production (Frolov RNA 2001).

[0320] A model lifecycle of an alphavirus involves several distinct steps (Strauss Microbial Review 1994, Jose Future Microbiol 2009). Following virus attachment to a host cell, the virion fuses with membranes within endocytic compartments resulting in the eventual release of genomic RNA into the cytosol. The genomic RNA, which is in a plus-strand orientation and comprises a 5′ methylguanylate cap and 3′ polyA tail, is translated to produce non-structural proteins nsP1-4 that form the replication complex. Early in infection, the plus-strand is then replicated by the complex into a minus-stand template. In the current model, the replication complex is further processed as infection progresses, with the resulting processed complex switching to transcription of the minus-strand into both full-length positive-strand genomic RNA, as well as the 26S subgenomic positive-strand RNA containing the structural genes. Several conserved sequence elements (CSEs) of alphavirus have been identified to potentially play a role in the various RNA replication steps including; a complement of the 5′ UTR in the replication of plus-strand RNAs from a minus-strand template, a 51-nt CSE in the replication of minus-strand synthesis from the genomic template, a 24-nt CSE in the junction region between the nsPs and the 26S RNA in the transcription of the subgenomic RNA from the minus-strand, and a 3′ 19-nt CSE in minus-strand synthesis from the plus-strand template.

[0321] Following the replication of the various RNA species, virus particles are then typically assembled in the natural lifecycle of the virus. The 26S RNA is translated and the resulting proteins further processed to produce the structural proteins including capsid protein, glycoproteins E1 and E2, and two small polypeptides E3 and 6K (Strauss 1994). Encapsidation of viral RNA occurs, with capsid proteins normally specific for only genomic RNA being packaged, followed by virion assembly and budding at the membrane surface.

V.D.2. Alphavirus as a Delivery Vector

[0322] Alphaviruses (including alphavirus sequences, features, and other elements) can be used to generate alphavirus-based delivery vectors (also be referred to as alphavirus vectors, alphavirus viral vectors, alphavirus vaccine vectors, self-replicating RNA (srRNA) vectors, self-amplifying mRNA (SAM) vectors, or samRNA vectors). Alphaviruses have previously been engineered for use as expression vector systems (Pushko 1997, Rheme 2004). Alphaviruses offer several advantages, particularly in a vaccine setting where heterologous antigen expression can be desired. Due to its ability to self-replicate in the host cytosol, alphavirus vectors are generally able to produce high copy numbers of the expression cassette within a cell resulting in a high level of heterologous antigen production. Additionally, the vectors are generally transient, resulting in improved biosafety as well as reduced induction of immunological tolerance to the vector. The public, in general, also lacks pre-existing immunity to alphavirus vectors as compared to other standard viral vectors, such as human adenovirus. Alphavirus based vectors also generally result in cytotoxic responses to infected cells. Cytotoxicity, to a certain degree, can be important in a vaccine setting to properly stimulate an immune response to the heterologous antigen expressed. However, the degree of desired cytotoxicity can be a balancing act, and thus several attenuated alphaviruses have been developed, including the TC-83 strain of VEE. Thus, an example of an antigen expression vector described herein can utilize an alphavirus backbone that allows for a high level of antigen expression, stimulates a robust immune response to antigen, does not stimulate an immune response to the vector itself, and can be used in a safe manner. Furthermore, the antigen expression cassette can be designed to stimulate different levels of an immune response through optimization of which alphavirus sequences the vector uses, including, but not limited to, sequences derived from VEE or its attenuated derivative TC-83.

[0323] Several expression vector design strategies have been engineered using alphavirus sequences (Pushko 1997). In one strategy, a alphavirus vector design includes inserting a second copy of the 26S promoter sequence elements downstream of the structural protein genes, followed by a heterologous gene (Frolov 1993). Thus, in addition to the natural non-structural and structural proteins, an additional subgenomic RNA is produced that expresses the heterologous protein. In this system, all the elements for production of infectious virions are present and, therefore, repeated rounds of infection of the expression vector in non-infected cells can occur.

[0324] Another expression vector design makes use of helper virus systems (Pushko 1997). In this strategy, the structural proteins are replaced by a heterologous gene. Thus, following self-replication of viral RNA mediated by still intact non-structural genes, the 26S subgenomic RNA provides for expression of the heterologous protein. Traditionally, additional vectors that expresses the structural proteins are then supplied in trans, such as by co-transfection of a cell line, to produce infectious virus. A system is described in detail in U.S. Pat. No. 8,093,021, which is herein incorporated by reference in its entirety, for all purposes. The helper vector system provides the benefit of limiting the possibility of forming infectious particles and, therefore, improves biosafety. In addition, the helper vector system reduces the total vector length, potentially improving the replication and expression efficiency. Thus, an example of an antigen expression vector described herein can utilize an alphavirus backbone wherein the structural proteins are replaced by an antigen cassette, the resulting vector both reducing biosafety concerns, while at the same

time promoting efficient expression due to the reduction in overall expression vector size.

V.D.3. Alphavirus Production In Vitro

[0325] Alphavirus delivery vectors are generally positive-sense RNA polynucleotides. A convenient technique well-known in the art for RNA production is in vitro transcription IVT. In this technique, a DNA template of the desired vector is first produced by techniques well-known to those in the art, including standard molecular biology techniques such as cloning, restriction digestion, ligation, gene synthesis (e.g., chemical and/or enzymatic synthesis), and polymerase chain reaction (PCR). The DNA template contains a RNA polymerase promoter at the 5′ end of the sequence desired to be transcribed into RNA. Promoters include, but are not limited to, bacteriophage polymerase promoters such as T3, T7, or SP6. The DNA template is then incubated with the appropriate RNA polymerase enzyme, buffer agents, and nucleotides (NTPs). The resulting RNA polynucleotide can optionally be further modified including, but limited to, addition of a 5′ cap structure such as 7-methylguanosine or a related structure, and optionally modifying the 3′ end to include a polyadenylate (polyA) tail. The RNA can then be purified using techniques well-known in the field, such as phenol-chloroform extraction or column purification (e.g., chromatography-based purification).

V.D.4. Delivery Via Lipid Nanoparticle

[0326] An important aspect to consider in vaccine vector design is immunity against the vector itself (Riley 2017). This may be in the form of preexisting immunity to the vector itself, such as with certain human adenovirus systems, or in the form of developing immunity to the vector following administration of the vaccine. The latter is an important consideration if multiple administrations of the same vaccine are performed, such as separate priming and boosting doses, or if the same vaccine vector system is to be used to deliver different antigen cassettes.

[0327] In the case of alphavirus vectors, the standard delivery method is the previously discussed helper virus system that provides capsid, E1, and E2 proteins in trans to produce infectious viral particles. However, it is important to note that the E1 and E2 proteins are often major targets of neutralizing antibodies (Strauss 1994). Thus, the efficacy of using alphavirus vectors to deliver antigens of interest to target cells may be reduced if infectious particles are targeted by neutralizing antibodies.

[0328] An alternative to viral particle mediated gene delivery is the use of nanomaterials to deliver expression vectors (Riley 2017). Nanomaterial vehicles, importantly, can be made of non-immunogenic materials and generally avoid eliciting immunity to the delivery vector itself. These materials can include, but are not limited to, lipids, inorganic nanomaterials, and other polymeric materials. Lipids can be cationic, anionic, or neutral. The materials can be synthetic or naturally derived, and in some instances biodegradable. Lipids can include fats, cholesterol, phospholipids, lipid conjugates including, but not limited to, polyethyleneglycol (PEG) conjugates (PEGylated lipids), waxes, oils, glycerides, and fat soluble vitamins.

[0329] Lipid nanoparticles (LNPs) are an attractive delivery system due to the amphiphilic nature of lipids enabling formation of membranes and vesicle like structures (Riley 2017). In general, these vesicles deliver the expression vector by absorbing into the membrane of target cells and releasing nucleic acid into the cytosol. In addition, LNPs can be further modified or functionalized to facilitate targeting of specific cell types. Another consideration in LNP design is the balance between targeting efficiency and cytotoxicity. Lipid compositions generally include defined mixtures of cationic, neutral, anionic, and amphipathic lipids. In some instances, specific lipids are included to prevent LNP aggregation, prevent lipid oxidation, or provide functional chemical groups that facilitate attachment of additional moieties. Lipid composition can influence overall LNP size and stability. In an example, the lipid composition comprises dilinoleylmethyl-4-dimethylaminobutyrate (MC3) or MC3-like molecules. MC3 and MC3-like lipid compositions can be formulated to include one or more other lipids, such as a PEG or PEG-conjugated lipid, a sterol, or neutral lipids.

[0330] Nucleic-acid vectors, such as expression vectors, exposed directly to serum can have several undesirable consequences, including degradation of the nucleic acid by serum nucleases or off-target stimulation of the immune system by the free nucleic acids. Therefore, encapsulation of the alphavirus vector can be used to avoid degradation, while also avoiding potential off-target affects. In certain examples, an alphavirus vector is fully encapsulated within the delivery vehicle, such as within the aqueous interior of an LNP. Encapsulation of the alphavirus vector within an LNP can be carried out by techniques well-known to those skilled in the art, such as microfluidic mixing and droplet generation carried out on a microfluidic droplet generating device. Such devices include, but are not limited to, standard T-junction devices or flow-focusing devices. In an example, the desired lipid formulation, such as MC3 or MC3-like containing compositions, is provided to the droplet generating device in parallel with the alphavirus delivery vector and other desired agents, such that the delivery vector and desired agents are fully encapsulated within the interior of the MC3 or MC3-like based LNP. In an example, the droplet generating device can control the size range and size distribution of the LNPs produced. For example, the LNP can have a size ranging from 1 to 1000 nanometers in diameter, e.g., 1, 10, 50, 100, 500, or 1000 nanometers. Following droplet generation, the delivery vehicles encapsulating the expression vectors can be further treated or modified to prepare them for administration.

V.E. Chimpanzee Adenovirus (ChAd)

V.E.1. Viral Delivery with Chimpanzee Adenovirus

[0331] Vaccine compositions for delivery of one or more antigens (e.g., via an antigen cassette and including one or more neoantigens shown in Table A and/or AACR GENIE Results, and/or one or more antigens shown in Table 1.2) can

be created by providing adenovirus nucleotide sequences of chimpanzee origin, a variety of novel vectors, and cell lines expressing chimpanzee adenovirus genes. A nucleotide sequence of a chimpanzee C68 adenovirus (also referred to herein as ChAdV68) can be used in a vaccine composition for antigen delivery (See SEQ ID NO: 1). Use of C68 adenovirus derived vectors is described in further detail in U.S. Pat. No. 6,083,716, which is herein incorporated by reference in its entirety, for all purposes. ChAdV68-based vectors and delivery systems are described in detail in US App. Pub. No. US20200197500A1 and international patent application publication WO2020243719A1, each of which is herein incorporated by reference for all purposes.

[0332] In a further aspect, provided herein is a recombinant adenovirus comprising the DNA sequence of a chimpanzee adenovirus such as C68 and an antigen cassette operatively linked to regulatory sequences directing its expression. The recombinant virus is capable of infecting a mammalian, preferably a human, cell and capable of expressing the antigen cassette product in the cell. In this vector, the native chimpanzee E1 gene, and/or E3 gene, and/or E4 gene can be deleted. An antigen cassette can be inserted into any of these sites of gene deletion. The antigen cassette can include an antigen against which a primed immune response is desired.

[0333] In another aspect, provided herein is a mammalian cell infected with a chimpanzee adenovirus such as C68.

[0334] In still a further aspect, a novel mammalian cell line is provided which expresses a chimpanzee adenovirus gene (e.g., from C68) or functional fragment thereof.

[0335] In still a further aspect, provided herein is a method for delivering an antigen cassette into a mammalian cell comprising the step of introducing into the cell an effective amount of a chimpanzee adenovirus, such as C68, that has been engineered to express the antigen cassette.

[0336] Still another aspect provides a method for stimulating an immune response in a mammalian host to treat cancer. The method can comprise the step of administering to the host an effective amount of a recombinant chimpanzee adenovirus, such as C68, comprising an antigen cassette that encodes one or more antigens from the tumor against which the immune response is targeted.

[0337] Still another aspect provides a method for stimulating an immune response in a mammalian host to treat or prevent a disease in a subject. The method can comprise the step of administering to the host an effective amount of a recombinant chimpanzee adenovirus, such as C68, comprising an antigen cassette that encodes one or more antigens, such as from disease against which the immune response is targeted.

[0338] Also disclosed is a non-simian mammalian cell that expresses a chimpanzee adenovirus gene obtained from the sequence of SEQ ID NO: 1. The gene can be selected from the group consisting of the adenovirus E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 of SEQ ID NO: 1.

[0339] Also disclosed is a nucleic acid molecule comprising a chimpanzee adenovirus DNA sequence comprising a gene obtained from the sequence of SEQ ID NO: 1. The gene can be selected from the group consisting of said chimpanzee adenovirus E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 genes of SEQ ID NO: 1. In some aspects the nucleic acid molecule comprises SEQ ID NO: 1. In some aspects the nucleic acid molecule comprises the sequence of SEQ ID NO: 1, lacking at least one gene selected from the group consisting of E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 genes of SEQ ID NO: 1.

[0340] Also disclosed is a vector comprising a chimpanzee adenovirus DNA sequence obtained from SEQ ID NO: 1 and an antigen cassette operatively linked to one or more regulatory sequences which direct expression of the cassette in a heterologous host cell, optionally wherein the chimpanzee adenovirus DNA sequence comprises at least the cis-elements necessary for replication and virion encapsidation, the cis-elements flanking the antigen cassette and regulatory sequences. In some aspects, the chimpanzee adenovirus DNA sequence comprises a gene selected from the group consisting of E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 gene sequences of SEQ ID NO: 1. In some aspects the vector can lack the E1A and/or E1B gene.

[0341] Also disclosed herein is a adenovirus vector comprising: a partially deleted E4 gene comprising a deleted or partially-deleted E4orf2 region and a deleted or partially-deleted E4orf3 region, and optionally a deleted or partially-deleted E4orf4 region. The partially deleted E4 can comprise an E4 deletion of at least nucleotides 34,916 to 35,642 of the sequence shown in SEQ ID NO: 1, and wherein the vector comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1. The partially deleted E4 can comprise an E4 deletion of at least a partial deletion of nucleotides 34,916 to 34,942 of the sequence shown in SEQ ID NO:1, at least a partial deletion of nucleotides 34,952 to 35,305 of the sequence shown in SEQ ID NO:1, and at least a partial deletion of nucleotides 35,302 to 35,642 of the sequence shown in SEQ ID NO:1, and wherein the vector comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1 The partially deleted E4 can comprise an E4 deletion of at least nucleotides 34,980 to 36,516 of the sequence shown in SEQ ID NO:1, and wherein the vector comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1. The partially deleted E4 can comprise an E4 deletion of at least nucleotides 34,979 to 35,642 of the sequence shown in SEQ ID NO:1, and wherein the vector comprises at least nucleotides 2 to 36,518 of the sequence set forth in SEQ ID NO:1. The partially deleted E4 can comprise an E4 deletion of at least a partial deletion of E4Orf2, a fully deleted E4Orf3, and at least a partial deletion of E4Orf4. The partially deleted E4 can comprise an E4 deletion of at least a partial deletion of E4Orf2, at least a partial deletion of E4Orf3, and at least a partial deletion of E4Orf4. The partially deleted E4 can comprise an E4 deletion of at least a partial deletion of E4Orf1, a fully deleted E4Orf2, and at least a partial deletion of E4Orf3. The partially deleted E4 can comprise an E4 deletion

of at least a partial deletion of E4Orf2 and at least a partial deletion of E4Orf3. The partially deleted E4 can comprise an E4 deletion between the start site of E4Orf1 to the start site of E4Orf5. The partially deleted E4 can be an E4 deletion adjacent to the start site of E4Orf1. The partially deleted E4 can be an E4 deletion adjacent to the start site of E4Orf2. The partially deleted E4 can be an E4 deletion adjacent to the start site of E4Orf3. The partially deleted E4 can be an E4 deletion adjacent to the start site of E4Orf4. The E4 deletion can be at least 50, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, at least 1300, at least 1400, at least 1500, at least 1600, at least 1700, at least 1800, at least 1900, or at least 2000 nucleotides. The E4 deletion can be at least 700 nucleotides. The E4 deletion can be at least 1500 nucleotides. The E4 deletion can be 50 or less, 100 or less, 200 or less, 300 or less, 400 or less, 500 or less, 600 or less, 700 or less, 800 or less, 900 or less, 1000 or less, 1100 or less, 1200 or less, 1300 or less, 1400 or less, 1500 or less, 1600 or less, 1700 or less, 1800 or less, 1900 or less, or 2000 or less nucleotides. The E4 deletion can be 750 nucleotides or less. The E4 deletion can be at least 1550 nucleotides or less.

[0342] A partially deleted E4 gene can be the E4 gene sequence shown in SEQ ID NO: 1 that lacks at least nucleotides 34,916 to 35,642 of the sequence shown in SEQ ID NO: 1. A partially deleted E4 gene can be the E4 gene sequence shown in SEQ ID NO: 1 that lacks the E4 gene sequence shown in SEQ ID NO: 1 and that lacks at least nucleotides 34,916 to 34,942, nucleotides 34,952 to 35,305 of the sequence shown in SEQ ID NO:1, and nucleotides 35,302 to 35,642 of the sequence shown in SEQ ID NO: 1. A partially deleted E4 gene can be the E4 gene sequence shown in SEQ ID NO: 1 and that lacks at least nucleotides 34,980 to 36,516 of the sequence shown in SEQ ID NO:1. A partially deleted E4 gene can be the E4 gene sequence shown in SEQ ID NO:1 and that lacks at least nucleotides 34,979 to 35,642 of the sequence shown in SEQ ID NO:1. The adenovirus vector having A partially deleted E4 gene can have a cassette, wherein the cassette comprises at least one payload nucleic acid sequence, and wherein the cassette comprises at least one promoter sequence operably linked to the at least one payload nucleic acid sequence. The adenovirus vector having A partially deleted E4 gene can have one or more genes or regulatory sequences of the ChAdV68 sequence shown in SEQ ID NO: 1, optionally wherein the one or more genes or regulatory sequences comprise at least one of the chimpanzee adenovirus inverted terminal repeat (ITR), E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4, and L5 genes of the sequence shown in SEQ ID NO: 1. The adenovirus vector having A partially deleted E4 gene can have nucleotides 2 to 34,915 of the sequence shown in SEQ ID NO: 1, wherein A partially deleted E4 gene is 3′ of the nucleotides 2 to 34,915, and optionally the nucleotides 2 to 34,915 additionally lack nucleotides 577 to 3403 of the sequence shown in SEQ ID NO: 1 corresponding to an E1 deletion and/or lack nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO: 1 corresponding to an E3 deletion. The adenovirus vector having A partially deleted E4 gene can have nucleotides 35,643 to 36,518 of the sequence shown in SEQ ID NO: 1, and wherein A partially deleted E4 gene is 5′ of the nucleotides 35,643 to 36,518. The adenovirus vector having A partially deleted E4 gene can have nucleotides 2 to 34,915 of the sequence shown in SEQ ID NO: 1, wherein A partially deleted E4 gene is 3′ of the nucleotides 2 to 34,915, the nucleotides 2 to 34,915 additionally lack nucleotides 577 to 3403 of the sequence shown in SEQ ID NO: 1 corresponding to an E1 deletion and lack nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO: 1 corresponding to an E3 deletion. The adenovirus vector having A partially deleted E4 gene can have nucleotides 2 to 34,915 of the sequence shown in SEQ ID NO: 1, wherein A partially deleted E4 gene is 3′ of the nucleotides 2 to 34,915, the nucleotides 2 to 34,915 additionally lack nucleotides 577 to 3403 of the sequence shown in SEQ ID NO:1 corresponding to an E1 deletion and lack nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO: 1 corresponding to an E3 deletion, and have nucleotides 35,643 to 36,518 of the sequence shown in SEQ ID NO:1, and wherein A partially deleted E4 gene is 5′ of the nucleotides 35,643 to 36,518.

[0343] A partially deleted E4 gene can be the E4 gene sequence shown in SEQ ID NO: 1 that lacks at least nucleotides 34,916 to 35,642 of the sequence shown in SEQ ID NO:1, nucleotides 2 to 34,915 of the sequence shown in SEQ ID NO: 1, wherein A partially deleted E4 gene is 3′ of the nucleotides 2 to 34,915, the nucleotides 2 to 34,915 additionally lack nucleotides 577 to 3403 of the sequence shown in SEQ ID NO: 1 corresponding to an E1 deletion and lack nucleotides 27,125 to 31,825 of the sequence shown in SEQ ID NO: 1 corresponding to an E3 deletion, and have nucleotides 35,643 to 36,518 of the sequence shown in SEQ ID NO:1, and wherein A partially deleted E4 gene is 5′ of the nucleotides 35,643 to 36,518.

[0344] Also disclosed herein is a host cell transfected with a vector disclosed herein such as a C68 vector engineered to expression an antigen cassette. Also disclosed herein is a human cell that expresses a selected gene introduced therein through introduction of a vector disclosed herein into the cell.

[0345] Also disclosed herein is a method for delivering an antigen cassette to a mammalian cell comprising introducing into said cell an effective amount of a vector disclosed herein, such as a ChAd vector or self-amplifying RNA vector engineered to express an antigen cassette.

[0346] Also disclosed herein is a method for producing an antigen comprising introducing a vector disclosed herein into a mammalian cell, culturing the cell under suitable conditions and producing the antigen.

V.E.2. E1-Expressing Complementation Cell Lines

[0347] To generate recombinant chimpanzee adenoviruses (Ad) deleted in any of the genes described herein, the function of the deleted gene region, if essential to the replication and infectivity of the virus, can be supplied to the recombinant virus by a helper virus or cell line, i.e., a complementation or packaging cell line. For example, to generate

a replication-defective chimpanzee adenovirus vector, a cell line can be used which expresses the E1 gene products of the human or chimpanzee adenovirus; such a cell line can include HEK293 or variants thereof. The protocol for the generation of the cell lines expressing the chimpanzee E1 gene products (Examples 3 and 4 of U.S. Pat. No. 6,083,716) can be followed to generate a cell line which expresses any selected chimpanzee adenovirus gene.

[0348] An AAV augmentation assay can be used to identify a chimpanzee adenovirus E1-expressing cell line. This assay is useful to identify E1 function in cell lines made by using the E1 genes of other uncharacterized adenoviruses, e.g., from other species. That assay is described in Example 4B of U.S. Pat. No. 6,083,716.

[0349] A selected chimpanzee adenovirus gene, e.g., E1, can be under the transcriptional control of a promoter for expression in a selected parent cell line. Inducible or constitutive promoters can be employed for this purpose. Among inducible promoters are included the sheep metallothionine promoter, inducible by zinc, or the mouse mammary tumor virus (MMTV) promoter, inducible by a glucocorticoid, particularly, dexamethasone. Other inducible promoters, such as those identified in International patent application WO95/13392, incorporated by reference herein can also be used in the production of packaging cell lines. Constitutive promoters in control of the expression of the chimpanzee adenovirus gene can be employed also.

[0350] A parent cell can be selected for the generation of a novel cell line expressing any desired C68 gene. Without limitation, such a parent cell line can be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells. Other suitable parent cell lines can be obtained from other sources. Parent cell lines can include CHO, HEK293 or variants thereof, 911, HeLa, A549, LP-293, PER.C6, or AE1-2a.

[0351] An E1-expressing cell line can be useful in the generation of recombinant chimpanzee adenovirus E1 deleted vectors. Cell lines constructed using essentially the same procedures that express one or more other chimpanzee adenoviral gene products are useful in the generation of recombinant chimpanzee adenovirus vectors deleted in the genes that encode those products. Further, cell lines which express other human Ad E1 gene products are also useful in generating chimpanzee recombinant Ads.

V.E.3. Recombinant Viral Particles as Vectors

[0352] The compositions disclosed herein can comprise viral vectors, that deliver at least one antigen to cells. Such vectors comprise a chimpanzee adenovirus DNA sequence such as C68 and an antigen cassette operatively linked to regulatory sequences which direct expression of the cassette. The C68 vector is capable of expressing the cassette in an infected mammalian cell. The C68 vector can be functionally deleted in one or more viral genes. An antigen cassette comprises at least one antigen under the control of one or more regulatory sequences such as a promoter. Optional helper viruses and/or packaging cell lines can supply to the chimpanzee viral vector any necessary products of deleted adenoviral genes.

[0353] The term "functionally deleted" means that a sufficient amount of the gene region is removed or otherwise altered, e.g., by mutation or modification, so that the gene region is no longer capable of producing one or more functional products of gene expression. Mutations or modifications that can result in functional deletions include, but are not limited to, nonsense mutations such as introduction of premature stop codons and removal of canonical and non-canonical start codons, mutations that alter mRNA splicing or other transcriptional processing, or combinations thereof. If desired, the entire gene region can be removed.

[0354] Modifications of the nucleic acid sequences forming the vectors disclosed herein, including sequence deletions, insertions, and other mutations may be generated using standard molecular biological techniques and are within the scope of this invention.

V.E.4. Construction of the Viral Plasmid Vector

[0355] The chimpanzee adenovirus C68 vectors useful in this invention include recombinant, defective adenoviruses, that is, chimpanzee adenovirus sequences functionally deleted in the E1a or E1b genes, and optionally bearing other mutations, e.g., temperature-sensitive mutations or deletions in other genes. It is anticipated that these chimpanzee sequences are also useful in forming hybrid vectors from other adenovirus and/or adeno-associated virus sequences. Homologous adenovirus vectors prepared from human adenoviruses are described in the published literature [see, for example, Kozarsky I and II, cited above, and references cited therein, U.S. Pat. No. 5,240,846].

[0356] In the construction of useful chimpanzee adenovirus C68 vectors for delivery of an antigen cassette to a human (or other mammalian) cell, a range of adenovirus nucleic acid sequences can be employed in the vectors. A vector comprising minimal chimpanzee C68 adenovirus sequences can be used in conjunction with a helper virus to produce an infectious recombinant virus particle. The helper virus provides essential gene products required for viral infectivity and propagation of the minimal chimpanzee adenoviral vector. When only one or more selected deletions of chimpanzee adenovirus genes are made in an otherwise functional viral vector, the deleted gene products can be supplied in the viral vector production process by propagating the virus in a selected packaging cell line that provides the deleted gene functions in trans.

V.E.5. Recombinant Minimal Adenovirus

[0357] A minimal chimpanzee Ad C68 virus is a viral particle containing just the adenovirus cis-elements necessary for replication and virion encapsidation. That is, the vector contains the cis-acting 5′ and 3′ inverted terminal repeat (ITR) sequences of the adenoviruses (which function as origins of replication) and the native 5′ packaging/enhancer domains

(that contain sequences necessary for packaging linear Ad genomes and enhancer elements for the E1 promoter). See, for example, the techniques described for preparation of a "minimal" human Ad vector in International Patent Application WO96/13597 and incorporated herein by reference.

V.E.6. Other Defective Adenoviruses

[0358] Recombinant, replication-deficient adenoviruses can also contain more than the minimal chimpanzee adenovirus sequences. These other Ad vectors can be characterized by deletions of various portions of gene regions of the virus, and infectious virus particles formed by the optional use of helper viruses and/or packaging cell lines.

[0359] As one example, suitable vectors may be formed by deleting all or a sufficient portion of the C68 adenoviral immediate early gene Ela and delayed early gene E1b, so as to eliminate their normal biological functions. Replication-defective E1-deleted viruses are capable of replicating and producing infectious virus when grown on a chimpanzee adenovirus-transformed, complementation cell line containing functional adenovirus Ela and E1b genes which provide the corresponding gene products in trans. Based on the homologies to known adenovirus sequences, it is anticipated that, as is true for the human recombinant E1-deleted adenoviruses of the art, the resulting recombinant chimpanzee adenovirus is capable of infecting many cell types and can express antigen(s), but cannot replicate in most cells that do not carry the chimpanzee E1 region DNA unless the cell is infected at a very high multiplicity of infection.

[0360] As another example, all or a portion of the C68 adenovirus delayed early gene E3 can be eliminated from the chimpanzee adenovirus sequence which forms a part of the recombinant virus.

[0361] Chimpanzee adenovirus C68 vectors can also be constructed having a deletion of the E4 gene. Still another vector can contain a deletion in the delayed early gene E2a.

[0362] Deletions can also be made in any of the late genes L1 through L5 of the chimpanzee C68 adenovirus genome. Similarly, deletions in the intermediate genes IX and IVa2 can be useful for some purposes. Other deletions may be made in the other structural or non-structural adenovirus genes.

[0363] The above discussed deletions can be used individually, i.e., an adenovirus sequence can contain deletions of E1 only. Alternatively, deletions of entire genes or portions thereof effective to destroy or reduce their biological activity can be used in any combination. For example, in one exemplary vector, the adenovirus C68 sequence can have deletions of the E1 genes and the E4 gene, or of the E1, E2a and E3 genes, or of the E1 and E3 genes, or of E1, E2a and E4 genes, with or without deletion of E3, and so on. As discussed above, such deletions can be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

[0364] The cassette comprising antigen(s) be inserted optionally into any deleted region of the chimpanzee C68 Ad virus. Alternatively, the cassette can be inserted into an existing gene region to disrupt the function of that region, if desired.

V.E.7. Helper Viruses

[0365] Depending upon the chimpanzee adenovirus gene content of the viral vectors employed to carry the antigen cassette, a helper adenovirus or non-replicating virus fragment can be used to provide sufficient chimpanzee adenovirus gene sequences to produce an infective recombinant viral particle containing the cassette.

[0366] Useful helper viruses contain selected adenovirus gene sequences not present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. A helper virus can be replication-defective and contain a variety of adenovirus genes in addition to the sequences described above. The helper virus can be used in combination with the E1-expressing cell lines described herein.

[0367] For C68, the "helper" virus can be a fragment formed by clipping the C terminal end of the C68 genome with SspI, which removes about 1300 bp from the left end of the virus. This clipped virus is then co-transfected into an E1-expressing cell line with the plasmid DNA, thereby forming the recombinant virus by homologous recombination with the C68 sequences in the plasmid.

[0368] Helper viruses can also be formed into poly-cation conjugates as described in Wu et al, J. Biol. Chem., 264:16985-16987 (1989); K. J. Fisher and J. M. Wilson, Biochem. J., 299:49 (Apr. 1, 1994). Helper virus can optionally contain a reporter gene. A number of such reporter genes are known to the art. The presence of a reporter gene on the helper virus which is different from the antigen cassette on the adenovirus vector allows both the Ad vector and the helper virus to be independently monitored. This second reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

V.E.8. Assembly of Viral Particle and Infection of a Cell Line

[0369] Assembly of the selected DNA sequences of the adenovirus, the antigen cassette, and other vector elements into various intermediate plasmids and shuttle vectors, and the use of the plasmids and vectors to produce a recombinant viral particle can all be achieved using conventional techniques. Such techniques include conventional cloning techniques of cDNA, in vitro recombination techniques (e.g., Gibson assembly), use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence. Standard transfection and co-transfection techniques are employed, e.g., CaPO4 precipitation techniques or liposome-mediated transfection methods such as lipofectamine. Other conventional methods employed include homologous recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

[0370] For example, following the construction and assembly of the desired antigen cassette-containing viral vector, the

vector can be transfected in vitro in the presence of a helper virus into the packaging cell line. Homologous recombination occurs between the helper and the vector sequences, which permits the adenovirus-antigen sequences in the vector to be replicated and packaged into virion capsids, resulting in the recombinant viral vector particles.

[0371] The resulting recombinant chimpanzee C68 adenoviruses are useful in transferring an antigen cassette to a selected cell. In in vivo experiments with the recombinant virus grown in the packaging cell lines, the E1-deleted recombinant chimpanzee adenovirus demonstrates utility in transferring a cassette to a non-chimpanzee, preferably a human, cell.

V.E.9. Use of the Recombinant Virus Vectors

[0372] The resulting recombinant chimpanzee C68 adenovirus containing the antigen cassette (produced by cooperation of the adenovirus vector and helper virus or adenoviral vector and packaging cell line, as described above) thus provides an efficient gene transfer vehicle which can deliver antigen(s) to a subject in vivo or ex vivo.

[0373] The above-described recombinant vectors are administered to humans according to published methods for gene therapy. A chimpanzee viral vector bearing an antigen cassette can be administered to a patient, preferably suspended in a biologically compatible solution or pharmaceutically acceptable delivery vehicle. A suitable vehicle includes sterile saline. Other aqueous and non-aqueous isotonic sterile injection solutions and aqueous and non-aqueous sterile suspensions known to be pharmaceutically acceptable carriers and well known to those of skill in the art may be employed for this purpose.

[0374] The chimpanzee adenoviral vectors are administered in sufficient amounts to transduce the human cells and to provide sufficient levels of antigen transfer and expression to provide a therapeutic benefit without undue adverse or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the liver, intranasal, intravenous, intramuscular, subcutaneous, intradermal, oral and other parental routes of administration. Routes of administration may be combined, if desired.

[0375] Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. The dosage will be adjusted to balance the therapeutic benefit against any side effects and such dosages may vary depending upon the therapeutic application for which the recombinant vector is employed. The levels of expression of antigen(s) can be monitored to determine the frequency of dosage administration.

[0376] Recombinant, replication defective adenoviruses can be administered in a "pharmaceutically effective amount," that is, an amount of recombinant adenovirus that is effective in a route of administration to transfect the desired cells and provide sufficient levels of expression of the selected gene to provide a vaccinal benefit, i.e., some measurable level of protective immunity. C68 vectors comprising an antigen cassette can be co-administered with adjuvant. Adjuvant can be separate from the vector (e.g., alum) or encoded within the vector, in particular if the adjuvant is a protein. Adjuvants are well known in the art.

[0377] Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, intranasal, intramuscular, intratracheal, subcutaneous, intradermal, rectal, oral and other parental routes of administration. Routes of administration may be combined, if desired, or adjusted depending upon the immunogen or the disease. For example, in prophylaxis of rabies, the subcutaneous, intratracheal and intranasal routes are preferred. The route of administration primarily will depend on the nature of the disease being treated.

[0378] The levels of immunity to antigen(s) can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in the serum, for example, optional booster immunizations may be desired

VI. Therapeutic and Manufacturing Methods

[0379] Also provided is a method of stimulating a tumor specific immune response in a subject, vaccinating against a tumor, treating and or alleviating a symptom of cancer in a subject by administering to the subject one or more antigens such as a plurality of antigens identified using methods disclosed herein.

[0380] In some aspects, a subject has been diagnosed with cancer or is at risk of developing cancer. A subject can have been previously treated for cancer, such as previously undergone surgery to remove a tumor and/or cancerous tissue, chemotherapy, immunotherapy (e.g., immune checkpoint inhibitor therapy), radiation therapy, or combinations thereof. A subject can be a human, dog, cat, horse or any animal in which a tumor specific immune response is desired. A tumor can be any solid tumor such as breast, ovarian, prostate, lung, kidney, gastric, colon, testicular, head and neck, pancreas, brain, melanoma, and other tumors of tissue organs and hematological tumors, such as lymphomas and leukemias, including acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, T cell lymphocytic leukemia, and B cell lymphomas.

[0381] An antigen can be administered in an amount sufficient to stimulate a CTL response. An antigen can be administered in an amount sufficient to stimulate a T cell response. An antigen can be administered in an amount sufficient to stimulate a B cell response. An antigen can be administered in an amount sufficient to stimulate both a T cell response and a B cell response.

[0382] An antigen can be administered alone or in combination with other therapeutic agents. The therapeutic agent is for example, a chemotherapeutic agent, radiation, or immunotherapy. Any suitable therapeutic treatment for a particular cancer can be administered. A therapeutically effective amount of the therapeutic agent can be administered. An amount

of the therapeutic agent can be administered that alone is not generally considered a therapeutically effective amount but demonstrates a beneficial property when co-administered with any of the vaccine compositions described herein.

[0383] In addition, a subject can be further administered an anti-immunosuppressive/immunostimulatory agent such as a checkpoint inhibitor. For example, the subject can be further administered an anti-CTLA antibody or anti-PD-1 or anti-PD-L1. Blockade of CTLA-4 or PD-L1 by antibodies can enhance the immune response to cancerous cells in the patient. In particular, CTLA-4 blockade has been shown effective when following a vaccination protocol.

[0384] The optimum amount of each antigen to be included in a vaccine composition and the optimum dosing regimen can be determined. For example, an antigen or its variant can be prepared for intravenous (i.v.) injection, sub-cutaneous (s.c.) injection, intradermal (i.d.) injection, intraperitoneal (i.p.) injection, intramuscular (i.m.) injection. Methods of injection include s.c., i.d., i.p., i.m., and i.v. Methods of DNA or RNA injection include i.d., i.m., s.c., i.p. and i.v. Other methods of administration of the vaccine composition are known to those skilled in the art.

[0385] A vaccine can be compiled so that the selection, number and/or amount of antigens present in the composition is/are tissue, cancer, and/or patient-specific. For instance, the exact selection of peptides can be guided by expression patterns of the parent proteins in a given tissue or guided by mutation status of a patient. The selection can be dependent on the specific type of cancer, the status of the disease, earlier treatment regimens, the immune status of the patient, and, of course, the HLA-haplotype of the patient. Furthermore, a vaccine can contain individualized components, according to personal needs of the particular patient. Examples include varying the selection of antigens according to the expression of the antigen in the particular patient or adjustments for secondary treatments following a first round or scheme of treatment.

[0386] A patient can be identified for administration of an antigen vaccine through the use of various diagnostic methods, e.g., patient selection methods described further below. Patient selection can involve identifying mutations in, or expression patterns of, one or more genes. In some cases, patient selection involves identifying the haplotype of the patient. The various patient selection methods can be performed in parallel, e.g., a sequencing diagnostic can identify both the mutations and the haplotype of a patient. The various patient selection methods can be performed sequentially, e.g., one diagnostic test identifies the mutations and separate diagnostic test identifies the haplotype of a patient, and where each test can be the same (e.g., both high-throughput sequencing) or different (e.g., one high-throughput sequencing and the other Sanger sequencing) diagnostic methods.

[0387] For a composition to be used as a vaccine for cancer, antigens with similar normal self-peptides that are expressed in high amounts in normal tissues can be avoided or be present in low amounts in a composition described herein. On the other hand, if it is known that the tumor of a patient expresses high amounts of a certain antigen, the respective pharmaceutical composition for treatment of this cancer can be present in high amounts and/or more than one antigen specific for this particularly antigen or pathway of this antigen can be included.

[0388] Compositions comprising an antigen can be administered to an individual already suffering from cancer. In therapeutic applications, compositions are administered to a subject in an amount sufficient to stimulate an immune response, such as stimulating an effective CTL response to the tumor antigen and to cure or at least partially arrest symptoms and/or complications. An immune response can include a reduction in tumor size or volume. Reduction in tumor size or volume can include at least a 5%, at least a 10%, at least a 15%, at least a 20%, at least a 25%, at least a 30%, at least a 35%, at least a 40%, at least a 45%, at least a 50%, at least a 55%, at least a 60%, at least a 65%, at least a 70%, at least a 75%, at least a 80%, at least a 85%, at least a 90%, or at least a 95% reduction. Reduction in tumor size or volume can include at least a 15% reduction. Reduction in tumor size or volume can include at least a 20% reduction. An immune response can include stabilization of tumor size or volume. An immune response can result in amelioration of a subject's disease, such a complete response (CR), partial response (PR), or stable disease (SD) (e.g., as assessed by criteria set forth in a clinical study). An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend on, e.g., the composition, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician. It should be kept in mind that compositions can generally be employed in serious disease states, that is, life-threatening or potentially life threatening situations, especially when the cancer has metastasized. In such cases, in view of the minimization of extraneous substances and the relative nontoxic nature of an antigen, it is possible and can be felt desirable by the treating physician to administer substantial excesses of these compositions.

[0389] For therapeutic use, administration can begin at the detection or surgical removal of tumors. This can be followed by boosting doses until at least symptoms are substantially abated and for a period thereafter, or immunity is considered to be provided (e.g., a memory B cell or T cell population, or antigen specific B cells or antibodies are produced).

[0390] Compositions comprising an antigen (e.g., any of the compositions for delivery of a self-replicating alphavirus-based expression system or a chimpanzee adenovirus (ChAdV)-based expression system described herein) can be administered as an adjuvant therapy to a subject having already received a primary therapy. Compositions comprising an antigen can be administered 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 days following a primary therapy, or 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more weeks following a primary therapy. For example, compositions comprising an antigen can be administered as an adjuvant therapy following surgery to remove tumors and/or cancerous tissues, including 1, 2, 3, 4,

5, 6, 7, 8, 9, 10, 11, 12, 13, 14, days following surgery, or 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more weeks following surgery. Compositions comprising an antigen can be administered as an adjuvant therapy as a combination therapy with an additional therapy, such as administered in combination with chemotherapy, immune checkpoint inhibitior therapy, radiation therapy, or combinations thereof.

[0391] Compositions comprising an antigen (e.g., any of the compositions for delivery of a self-replicating alphavirus-based expression system or a chimpanzee adenovirus (ChAdV)-based expression system described herein, such as compositions including a therapeutically effective amount comprising 30 µg or less of each of self-amplifying alphavirus-based expression systems described herein) can be administered as maintenance therapy to a subject having already received a primary therapy, such as where the subject's cancer is in remission (e.g., complete remission) after the primary therapy.

[0392] Primary therapies can include surgery to remove a tumor and/or cancerous tissue, chemotherapy, immunotherapy (e.g., immune checkpoint inhibitor therapy), radiation therapy, or combinations thereof. Primary therapy can include surgery. Primary therapy can include chemotherapies, such as oxaliplatin, fluoropyrimidine, irinotecan, and/or bevacizumab. Primary therapy can include a combination therapy of oxaliplatin, fluoropyrimidine, and bevacizumab. Primary therapy can include irinotecan. Primary therapy can include a combination therapy of oxaliplatin, fluoropyrimidine, irinotecan, and bevacizumab. Fluoropyrimidines can include 5-FU or capecitabine. Primary therapy can include the chemotherapy regimen "FOLFOX" including folinic acid (e.g., leucovorin), fluorouracil (5-FU), and oxaliplatin. Primary therapy can include the chemotherapy regimen "FOLFOXIRI" including folinic acid, fluorouracil, oxaliplatin, and irinotecan. Primary therapy can include FOLFOX and bevacizumab. Primary therapy can include FOLFOXIRI and bevacizumab. Primary therapy can include a subject receiving oxaliplatin, fluoropyrimidine, and bevacizumab for up to 24 weeks. Primary therapy can include a subject receiving oxaliplatin, fluoropyrimidine, irinotecan, and bevacizumab for up to 24 weeks. Primary therapy can include a subject receiving FOLFOX, FOLFOXIRI, irinotecan, and/or bevacizumab for up to 24 weeks. Primary therapy can include a subject receiving irinotecan for up to 24 weeks in any of the primary therapies described here.

[0393] Compositions comprising an antigen (e.g., composition including a therapeutically effective amount comprising 30 µg or less of each of self-amplifying alphavirus-based expression systems described herein) can be administered as an adjuvant therapy or maintenance therapy as a combination therapy with an additional therapy, such as administered in combination with chemotherapy, immune checkpoint inhibitor therapy, radiation therapy, or combinations thereof. Combination therapies can include fluoropyrimidine (e.g., 5-FU or capecitabine), bevacizumab, and/or an immune checkpoint inhibitor therapy. Combination therapies can include fluoropyrimidine and bevacizumab. Combination therapies can include fluoropyrimidine, bevacizumab, and an immune checkpoint inhibitor therapy (e.g., an anti-PD-1 or anti-PD-L1 antibody).

[0394] Immune checkpoint inhibitors can include (1) an anti-PD-1 antibody or an antigen-binding fragment thereof, (2) an anti-PD-L1 antibody or an antigen-binding fragment thereof, and/or (3) an anti-CTLA-4 antibody or an antigen-binding fragment thereof. Immune checkpoint inhibitor therapy can include administration of an anti-CTLA-4 antibody or an antigen-binding fragment thereof only with the priming dose and the first boosting dose. Immune checkpoint inhibitor therapy can include the anti-CTLA-4 antibody ipilimumab. Immune checkpoint inhibitor therapy can include ipilimumab administered at a dose of 30 mg subcutaneously. Immune checkpoint inhibitor therapy can include administration of an anti-PD-L1 antibody or an antigen-binding fragment thereof every 4 weeks (Q4W). Immune checkpoint inhibitor therapy can include the anti-PD-L1 antibody atezolizumab or nivolumab. Atezolizumab can be administered at a dose of 1680 mg intravenously. Nivolumab is administered at a dose of 480 mg intravenously.

[0395] Immune checkpoint inhibitor therapy can include at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 separate administrations, e.g., at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 separate administrations on or about every 28 days (or every 4 weeks and/or every month). Immune checkpoint inhibitor therapy can include administration of the anti-PD-L1 antibody or an antigen-binding fragment thereof including at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 administrations, e.g., on or about every 28 days (or every 4 weeks and/or every month). Immune checkpoint inhibitor therapy can include at least 13 separate administrations, e.g., on or about every 28 days (or every 4 weeks and/or every month). In some aspects, the administration of the anti-PD-L1 antibody or an antigen-binding fragment thereof comprises at least 13 administrations, e.g., on or about every 28 days (or every 4 weeks and/or every month).

[0396] A ChAdV-based expression system can be administered as a boosting dose on or about day 140 after the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or about week 20 after the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or about month 5 after the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or after day 140 after the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or after week 20 after the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or after month 5 after the priming dose of the ChAdV-based expression system.

[0397] A self-replicating alphavirus-based expression system can be administered as at least two boosting doses. A self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 28

days apart. A self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 4 weeks (Q4W) apart. A self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least one month apart. A self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 56 days apart. A self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 8 weeks (Q8W) apart. A self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 2 months apart.

[0398] A self-replicating alphavirus-based expression system can be administered as at least two boosting doses on or about days 28 and 84 after the priming dose of the ChAdV-based expression system. A self-replicating alphavirus-based expression system can be administered as at least two boosting doses on or about weeks 4 and 12 after the priming dose of the ChAdV-based expression system. A self-replicating alphavirus-based expression system can be administered as at least two boosting doses on or about months 1 and 3 after the priming dose of the ChAdV-based expression system.

[0399] A self-replicating alphavirus-based expression system can be administered as at least four boosting doses. A self-replicating alphavirus-based expression system can be administered on or about days 28, 84, 224, and 308 relative to the priming dose of the ChAdV-based expression system. A self-replicating alphavirus-based expression system can be administered on or about weeks 4, 12, 32, and 44 relative to the priming dose of the ChAdV-based expression system. A self-replicating alphavirus-based expression system can be administered on or about months 1, 3, 8, and 11 relative to the priming dose of the ChAdV-based expression system.

[0400] A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two boosting doses. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 28 days apart. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 4 weeks (Q4W) apart. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least one month apart. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 56 days apart. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 8 weeks (Q8W) apart. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two or more boosting doses at least 2 months apart.

[0401] A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two boosting doses on or about days 28 and 84 after the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two boosting doses on or about weeks 4 and 12 after the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least two boosting doses on or about months 1 and 3 after the priming dose of the ChAdV-based expression system.

[0402] A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered as at least four boosting doses. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered on or about days 28, 84, 224, and 308 relative to the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or about, or after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered on or about weeks 4, 12, 32, and 44 relative to the priming dose of the ChAdV-based expression system. A ChAdV-based expression system can be administered as a boosting dose on or about, or

after day 140 (or week 20 and/or month 5) after the priming dose of the ChAdV-based expression system, and a self-replicating alphavirus-based expression system can be administered on or about months 1, 3, 8, and 11 relative to the priming dose of the ChAdV-based expression system.

[0403] The pharmaceutical compositions (e.g., vaccine compositions) for therapeutic treatment are intended for parenteral, topical, nasal, oral or local administration. A pharmaceutical compositions can be administered parenterally, e.g., intravenously, subcutaneously, intradermally, or intramuscularly. The compositions can be administered at a site of surgical excision to stimulate a local immune response to a tumor. The compositions can be administered to target specific tissues, organs, and/or cells of a subject. Disclosed herein are compositions for parenteral administration which comprise a solution of the antigen and vaccine compositions are dissolved or suspended in an acceptable carrier, e.g., an aqueous carrier. A variety of aqueous carriers can be used, e.g., water, buffered water, 0.9% saline, 0.3% glycine, hyaluronic acid and the like. These compositions can be sterilized by conventional, well known sterilization techniques, or can be sterile filtered. The resulting aqueous solutions can be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

[0404] Antigens can also be administered via liposomes, which target them to a particular cells tissue, such as lymphoid tissue. Liposomes are also useful in increasing half-life. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations the antigen to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to, e.g., a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes filled with a desired antigen can be directed to the site of lymphoid cells, where the liposomes then deliver the selected therapeutic/immunogenic compositions. Liposomes can be formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, e.g., liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al., Ann. Rev. Biophys. Bioeng. 9; 467 (1980), U.S. Pat. Nos. 4,235,871, 4,501,728, 4,501,728, 4,837,028, and 5,019,369.

[0405] For targeting to the immune cells, a ligand to be incorporated into the liposome can include, e.g., antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension can be administered intravenously, locally, topically, etc. in a dose which varies according to, inter alia, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

[0406] For therapeutic or immunization purposes, nucleic acids encoding a peptide and optionally one or more of the peptides described herein can also be administered to the patient. A number of methods are conveniently used to deliver the nucleic acids to the patient. For instance, the nucleic acid can be delivered directly, as "naked DNA". This approach is described, for instance, in Wolff et al., Science 247:1465-1468 (1990) as well as U.S. Pat. Nos. 5,580,859 and 5,589,466. The nucleic acids can also be administered using ballistic delivery as described, for instance, in U.S. Pat. No. 5,204,253. Particles comprised solely of DNA can be administered. Alternatively, DNA can be adhered to particles, such as gold particles. Approaches for delivering nucleic acid sequences can include viral vectors, mRNA vectors, and DNA vectors with or without electroporation.

[0407] The nucleic acids can also be delivered complexed to cationic compounds, such as cationic lipids. Lipid-mediated gene delivery methods are described, for instance, in 9618372WOAWO 96/18372; 9324640WOAWO 93/24640; Mannino & Gould-Fogerite, BioTechniques 6 (7): 682-691 (1988); U.S. Pat. No. 5,279,833 Rose U.S. Pat. Nos. 5,279,833; 9,106,309WOAWO 91/06309; and Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7414 (1987).

[0408] Antigens can also be included in viral vector-based vaccine platforms, such as vaccinia, fowlpox, self-replicating alphavirus, marabavirus, adenovirus (See, e.g., Tatsis et al., Adenoviruses, *Molecular Therapy* (2004) 10, 616-629), or lentivirus, including but not limited to second, third or hybrid second/third generation lentivirus and recombinant lentivirus of any generation designed to target specific cell types or receptors (See, e.g., Hu et al., Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases, *Immunol Rev*. (2011) 239 (1): 45-61, Sakuma et al., Lentiviral vectors: basic to translational, *Biochem J*. (2012) 443 (3): 603-18, Cooper et al., Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter, *Nucl. Acids Res.* (2015) 43 (1): 682-690, Zufferey et al., Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery, *J. Virol.* (1998) 72 (12): 9873-9880). Dependent on the packaging capacity of the above mentioned viral vector-based vaccine platforms, this approach can deliver one or more nucleotide sequences that encode one or more antigen peptides. The sequences may be flanked by non-mutated sequences, may be separated by linkers or may be preceded with one or more sequences targeting a subcellular compartment (See, e.g., Gros et al., Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients, *Nat Med*. (2016) 22 (4): 433-8, Stronen et al., Targeting of cancer neoantigens with donor-derived T cell receptor repertoires, *Science*. (2016) 352 (6291): 1337-41, Lu et al., Efficient identification of mutated cancer antigens recognized by T cells

associated with durable tumor regressions, *Clin Cancer Res.* (2014) 20 (13): 3401-10). Upon introduction into a host, infected cells express the antigens, and thereby elicit a host immune (e.g., CTL) response against the peptide(s). Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Pat. No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (Nature 351:456-460 (1991)). A wide variety of other vaccine vectors useful for therapeutic administration or immunization of antigens, e.g., *Salmonella typhi* vectors, and the like will be apparent to those skilled in the art from the description herein.

[0409] A means of administering nucleic acids uses minigene constructs encoding one or multiple epitopes. To create a DNA sequence encoding the selected CTL epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes are reverse translated. A human codon usage table is used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences are directly adjoined, creating a continuous polypeptide sequence. To optimize expression and/or immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequence that could be reverse translated and included in the minigene sequence include: helper T lymphocyte, epitopes, a leader (signal) sequence, and an endoplasmic reticulum retention signal. In addition, MHC presentation of CTL epitopes can be improved by including synthetic (e.g. poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL epitopes. The minigene sequence is converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) are synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides are joined using T4 DNA ligase. This synthetic minigene, encoding the CTL epitope polypeptide, can then cloned into a desired expression vector.

[0410] Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). A variety of methods have been described, and new techniques can become available. As noted above, nucleic acids are conveniently formulated with cationic lipids. In addition, glycolipids, fusogenic liposomes, peptides and compounds referred to collectively as protective, interactive, non-condensing (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

[0411] Also disclosed is a method of manufacturing a tumor vaccine, comprising performing the steps of a method disclosed herein; and producing a tumor vaccine comprising a plurality of antigens or a subset of the plurality of antigens.

[0412] Antigens disclosed herein can be manufactured using methods known in the art. For example, a method of producing an antigen or a vector (e.g., a vector including at least one sequence encoding one or more antigens) disclosed herein can include culturing a host cell under conditions suitable for expressing the antigen or vector wherein the host cell comprises at least one polynucleotide encoding the antigen or vector, and purifying the antigen or vector. Standard purification methods include chromatographic techniques, electrophoretic, immunological, precipitation, dialysis, filtration, concentration, and chromatofocusing techniques.

[0413] Host cells can include a Chinese Hamster Ovary (CHO) cell, NS0 cell, yeast, or a HEK293 cell. Host cells can be transformed with one or more polynucleotides comprising at least one nucleic acid sequence that encodes an antigen or vector disclosed herein, optionally wherein the isolated polynucleotide further comprises a promoter sequence operably linked to the at least one nucleic acid sequence that encodes the antigen or vector. In certain embodiments the isolated polynucleotide can be cDNA.

VII. Antigen Use and Administration

[0414] A vaccination protocol can be used to dose a subject with one or more antigens. A priming vaccine and a boosting vaccine can be used to dose the subject. The priming vaccine can be based on C68 (e.g., the sequences shown in SEQ ID NO: 1 or 2) or srRNA (e.g., the sequences shown in SEQ ID NO:3 or 4) and the boosting vaccine can be based on C68 (e.g., the sequences shown in SEQ ID NO: 1 or 2) or srRNA (e.g., the sequences shown in SEQ ID NO:3 or 4). Each vector typically includes a cassette that includes antigens. Cassettes can include about 20 antigens, separated by spacers such as the natural sequence that normally surrounds each antigen or other non-natural spacer sequences such as AAY. Cassettes can also include MHCII antigens such a tetanus toxoid antigen and PADRE antigen, which can be considered universal class II antigens. Cassettes can also include a targeting sequence such as a ubiquitin targeting sequence. In addition, each vaccine dose can be administered to the subject in conjunction with (e.g., concurrently, before, or after) a checkpoint inhibitor (CPI). CPI's can include those that inhibit CTLA4, PD1, and/or PDL1 such as antibodies or antigen-binding portions thereof. Such antibodies can include tremelimumab or durvalumab.

[0415] A priming vaccine can be injected (e.g., intramuscularly) in a subject. Bilateral injections per dose can be used. For example, one or more injections of ChAdV68 (C68) can be used (e.g., total dose 1×10^{12} viral particles); one or more injections of self-amplifying RNA (SAM or samRNA) at a dose of 30 μg or less RNA can be used. For ChAdV68 priming, 1×10^{12} or less of viral particles can be administered. For ChAdV68 priming, 3×10^{11} or less of the viral particles can be administered. For ChAdV68 priming, at least 1×10^{11} of the viral particles can be administered. For ChAdV68 priming, between 1×10^{11} and 1×10^{12}, between 3×10^{11} and 1×10^{12}, or between 1×10^{11} and 3×10^{11} of the viral particles can be administered. For ChAdV68 priming, 1×10^{11}, 3×10^{11}, or 1×10^{12} of the viral particles can be administered. For ChAdV68 priming, the viral particles can be

at a concentration of at 5×10.sup.11 vp/mL.

[0416] A vaccine boost (boosting vaccine) can be injected (e.g., intramuscularly) after prime vaccination. A boosting vaccine can be administered about every 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 weeks, e.g., every 4 weeks and/or 8 weeks after the prime. A boosting vaccine can be every 8 weeks after the prime. A boosting vaccine can be at least every 8 weeks after the prime. A boosting vaccine can be a period longer than every 8 weeks after the prime. Bilateral injections per dose can be used. One or more injections of samRNA at a dose of 30 µg or less can be used. A dose of 30 µg or less can represent the total content of RNA/samRNA administered. A dose of 30 µg or less can represent the total content of RNA/samRNA administered and include only a single distinct samRNA construct. A therapeutically effective amount of samRNA can be at least 1 µg, at least 3 µg, at least 10 µg, about 30 µg or less, or is 30 µg. A therapeutically effective amount of samRNA can be at least 1 µg. A therapeutically effective amount of samRNA can be at least 3 µg. A therapeutically effective amount of samRNA can be at least 10 µg. A therapeutically effective of samRNA amount can be about 30 µg or less. A therapeutically effective of samRNA amount can be 30 µg.

[0417] A vaccine can be administered using only a single distinct vector at at time (e.g., as an individual prime or boosting dose).

[0418] A therapeutically effective amount can include those doses capable of stimulating a T cell response following administration to a subject. T cell responses can be monitored as described herein. A therapeutically effective amount can include those doses that do not stimulate an inhibitory IFNα response following administration to a subject. A therapeutically effective amount can include those doses that reduce and/or minimize an inhibitory IFNα response following administration to a subject.

[0419] Anti-CTLA-4 (e.g., tremelimumab) can also be administered to the subject. For example, anti-CTLA4 can be administered subcutaneously near the site of the intramuscular vaccine injection (ChAdV68 prime or srRNA low doses) to ensure drainage into the same lymph node. Tremelimumab is a selective human IgG2 mAb inhibitor of CTLA-4. Target Anti-CTLA-4 (tremelimumab) subcutaneous dose is typically 70-75 mg (in particular 75 mg) with a dose range of, e.g., 1-100 mg or 5-420 mg.

[0420] In certain instances an anti-PD-L1 antibody can be used such as durvalumab (MEDI 4736). Durvalumab is a selective, high affinity human IgG1 mAb that blocks PD-L1 binding to PD-1 and CD80. Durvalumab is generally administered at 20 mg/kg i.v. every 4 weeks.

[0421] Immune monitoring can be performed before, during, and/or after vaccine administration. Such monitoring can inform safety and efficacy, among other parameters.

[0422] To perform immune monitoring, PBMCs are commonly used. PBMCs can be isolated before prime vaccination, and after prime vaccination (e.g. 4 weeks and 8 weeks). PBMCs can be harvested just prior to boost vaccinations and after each boost vaccination (e.g. 4 weeks and 8 weeks).

[0423] T cell responses can be assessed as part of an immune monitoring protocol. For example, the ability of a vaccine composition described herein to stimulate an immune response can be monitored and/or assessed. As used herein, "stimulate an immune response" refers to any increase in a immune response, such as initiating an immune response (e.g., a priming vaccine stimulating the initiation of an immune response in a naïve subject) or enhancement of an immune response (e.g., a boosting vaccine stimulating the enhancement of an immune response in a subject having a pre-existing immune response to an antigen, such as a pre-existing immune response initiated by a priming vaccine). T cell responses can be measured using one or more methods known in the art such as ELISpot, intracellular cytokine staining, cytokine secretion and cell surface capture, T cell proliferation, MHC multimer staining, or by cytotoxicity assay. T cell responses to epitopes encoded in vaccines can be monitored from PBMCs by measuring induction of cytokines, such as IFN-gamma, using an ELISpot assay. Specific CD4 or CD8 T cell responses to epitopes encoded in vaccines can be monitored from PBMCs by measuring induction of cytokines captured intracellularly or extracellularly, such as IFN-gamma, using flow cytometry. Specific CD4 or CD8 T cell responses to epitopes encoded in the vaccines can be monitored from PBMCs by measuring T cell populations expressing T cell receptors specific for epitope/MHC class I complexes using MHC multimer staining. Specific CD4 or CD8 T cell responses to epitopes encoded in the vaccines can be monitored from PBMCs by measuring the ex vivo expansion of T cell populations following 3H-thymidine, bromodeoxyuridine and carboxyfluoresceine-diacetate-succinimidylester (CFSE) incorporation. The antigen recognition capacity and lytic activity of PBMC-derived T cells that are specific for epitopes encoded in vaccines can be assessed functionally by chromium release assay or alternative colorimetric cytotoxicity assays.

[0424] B cell responses can be measured using one or more methods known in the art such as assays used to determine B cell differentiation (e.g., differentiation into plasma cells), B cell or plasma cell proliferation, B cell or plasma cell activation (e.g., upregulation of costimulatory markers such as CD80 or CD86), antibody class switching, and/or antibody production (e.g., an ELISA). Antibodies can also be assessed for function.

[0425] Disease status of a subject can be monitored following administration of any of the vaccine compositions described herein. For example, disease status may be monitored using isolated cell-free DNA (cfDNA) from a subject (also referred to as circulating tumor DNA "ctDNA"). In addition, the efficacy of a vaccine therapy may be monitored using isolated cfDNA from a subject. cfDNA minotoring can include the steps of: a. isolating or having isolated cfDNA from a subject; b. sequencing or having sequenced the isolated cfDNA; c. determining or having determined a frequency of one or more mutations in the cfDNA relative to a wild-type germline nucleic acid sequence of the subject,

and d. assessing or having assessed from step (c) the status of a disease in the subject. The method can also include, following step (c) above, d. performing more than one iteration of steps (a)-(c) for the given subject and comparing the frequency of the one or more mutations determined in the more than one iterations; and f. assessing or having assessed from step (d) the status of a disease in the subject. The more than one iterations can be performed at different time points, such as a first iteration of steps (a)-(c) performed prior to administration of the vaccine composition and a second iteration of steps (a)-(c) is performed subsequent to administration of the vaccine composition. Step (c) can include comparing: the frequency of the one or more mutations determined in the more than one iterations, or the frequency of the one or more mutations determined in the first iteration to the frequency of the one or more mutations determined in the second iteration. An increase in the frequency of the one or more mutations determined in subsequent iterations or the second iteration can be assessed as disease progression. A decrease in the frequency of the one or more mutations determined in subsequent iterations or the second iteration can be assessed as a response. In some aspects, the response is a Complete Response (CR) or a Partial Response (PR). A therapy can be administered to a subject following an assessment step, such as where assessment of the frequency of the one or more mutations in the cfDNA indicates the subject has the disease. The cfDNA isolation step can use centrifugation to separate cfDNA from cells or cellular debris. cfDNA can be isolated from whole blood, such as by separating the plasma layer, buffy coat, and red bloods. cfDNA sequencing can use next generation sequencing (NGS), Sanger sequencing, duplex sequencing, whole-exome sequencing, whole-genome sequencing, de novo sequencing, phased sequencing, targeted amplicon sequencing, shotgun sequencing, or combinations thereof, and may include enriching the cfDNA for one or more polynucleotide regions of interest prior to sequencing (e.g., polynucleotides known or suspected to encode the one or more mutations, coding regions, and/or tumor exome polynucleotides). Enriching the cfDNA may include hybridizing one or more polynucleotide probes, which may be modified (e.g., biotinylated), to the one or more polynucleotide regions of interest. In general, any number of mutations may be monitored simultaneously or in parallel.

[0426] Response to treatment (i.e., therapeutic response) may be determined by radiological assessment and/or a molecular response by monitoring neoantigen ctDNA (e.g., variant allele frequency "VAF") as described, for example, in Zhang et al. *Cancer Discov.* 2020; 10:1842-1853, Parikh et al. *Clin Cancer Res.* 2020; 26:1877-1885, and Vega et al. *JCO Precision Oncology* 2022; 6: e2100372 or neoantigen mutated haploid genome equivalents (mutated hGE) as described, for example, in Palmer et al. *Nat. Med.* 2022; 28:1619-1629 and Chabon et al. *Nature* 2020; 580 (7802): 245-251. Molecular response may be defined as a decrease in ctDNA relative to baseline ctDNA, for example a ≥20%, ≥30%, ≥40%, ≥50%, ≥60%, ≥70%, ≥80%, ≥90%, or ≥95% decrease in ctDNA relative to baseline ctDNA. In some embodiments, molecular response is defined as ≥30% decrease in ctDNA relative to baseline ctDNA. In some embodiments, molecular response is defined as ≥30% decrease in ctDNA relative to baseline ctDNA, wherein the decrease in ctDNA occurs by 24 months, by 18 months, by 12 months, or by 6 months of treatment. In some embodiments, molecular response is defined as ≥30% decrease in ctDNA relative to baseline ctDNA, wherein the decrease in ctDNA occurs by 12 weeks, by 8 weeks, by 4 weeks, or by 2 weeks of treatment. In other embodiments, molecular response is defined as >50% decrease in ctDNA relative to baseline ctDNA, wherein the decrease in ctDNA occurs by 24 months, by 18 months, by 12 months, or by 6 months of treatment. In other embodiments, molecular response is defined as ≥50% decrease in ctDNA relative to baseline ctDNA. In some embodiments, molecular response is defined as ≥50% decrease in ctDNA relative to baseline ctDNA, wherein the decrease in ctDNA occurs by 12 weeks, by 8 weeks, by 4 weeks, or by 2 weeks of treatment. Molecular response may be observed during early on-treatment radiological stable disease (SD) and/or in patients with radiological progressive disease (PD).

VIII. Antigen Identification

VIII.A. Antigen Candidate Identification

[0427] Research methods for NGS analysis of tumor and normal exome and transcriptomes have been described and applied in the antigen identification space. 6,14,15 Certain optimizations for greater sensitivity and specificity for antigen identification in the clinical setting can be considered. These optimizations can be grouped into two areas, those related to laboratory processes and those related to the NGS data analysis. Examples of optimizations are known to those skilled in the art, for example the methods described in more detail in international patent application publications WO/2017/106638, WO/2018/195357, and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes.

VIII.B. Isolation and Detection of HLA Peptides

[0428] Isolation of HLA-peptide molecules was performed using classic immunoprecipitation (IP) methods after lysis and solubilization of the tissue sample (55-58). A clarified lysate was used for HLA specific IP.

[0429] Immunoprecipitation was performed using antibodies coupled to beads where the antibody is specific for HLA molecules. For a pan-Class I HLA immunoprecipitation, a pan-Class I CR antibody is used, for Class II HLA-DR, an HLA-DR antibody is used. Antibody is covalently attached to NHS-sepharose beads during overnight incubation. After covalent attachment, the beads were washed and aliquoted for IP. (59, 60) Immunoprecipitations can also be performed with antibodies that are not covalently attached to beads. Typically this is done using sepharose or magnetic beads coated with Protein A and/or Protein G to hold the antibody to the column. Some antibodies that can be used to selectively enrich MHC/peptide complex are listed below.

TABLE-US-00001 Antibody Name Specificity W6/32 Class I HLA-A, B, C L243 Class II - HLA-DR Tu36 Class II -

HLA-DR LN3 Class II - HLA-DR Tu39 Class II - HLA-DR, DP, DQ SPVL3 Class II - HLA - DQ B7/21 Class II - HLA - DP

[0430] The clarified tissue lysate is added to the antibody beads for the immunoprecipitation. After immunoprecipitation, the beads are removed from the lysate and the lysate stored for additional experiments, including additional IPs. The IP beads are washed to remove non-specific binding and the HLA/peptide complex is eluted from the beads using standard techniques. The protein components are removed from the peptides using a molecular weight spin column or C18 fractionation. The resultant peptides are taken to dryness by SpeedVac evaporation and in some instances are stored at −20 C prior to MS analysis. HLA IPs can also be performed in 96 well plate format using plates that contain filter bottoms. Use of the plates allows for multiple IPs to be performed in tandem.

[0431] Dried peptides are reconstituted in an HPLC buffer suitable for reverse phase chromatography and loaded onto a C-18 microcapillary HPLC column for gradient elution in a Fusion Lumos mass spectrometer (Thermo). MS1 spectra of peptide mass/charge (m/z) were collected in the Orbitrap detector at high resolution followed by MS2 low resolution scans collected in the ion trap detector after HCD fragmentation of the selected ion. Additionally, MS2 spectra can be obtained using either CID or ETD fragmentation methods or any combination of the three techniques to attain greater amino acid coverage of the peptide. MS2 spectra can also be measured with high resolution mass accuracy in the Orbitrap detector.

[0432] MS2 spectra from each analysis are searched against a protein database using Comet (61, 62) and the peptide identification are scored using Percolator (63-65). Additional sequencing is performed using PEAKS studio (Bioinformatics Solutions Inc.) and other search engines or sequencing methods can be used including spectral matching and de novo sequencing (97).

VIII.B.1. MS Limit of Detection Studies in Support of Comprehensive HLA Peptide Sequencing

[0433] Using the peptide YVYVADVAAK (SEQ ID NO: 29364) it was determined what the limits of detection are using different amounts of peptide loaded onto the LC column. The amounts of peptide tested were 1 μmol, 100 fmol, 10 fmol, 1 fmol, and 100 amol. (Table 1) These results indicate that the lowest limit of detection (LoD) is in the attomol range (10-18), that the dynamic range spans five orders of magnitude, and that the signal to noise appears sufficient for sequencing at low femtomol ranges (10-15).

TABLE-US-00002 TABLE 1 Peptide m/z Loaded on Column Copies/Cell in 1e9cells 566.830 1 pmol 600 562.823 100 fmol 60 559.816 10 fmol 6 556.810 1 fmol 0.6 553.802 100 amol 0.06

IX. Presentation Model

[0434] Presentation models can be used to identify likelihoods of peptide presentation in patients. Various presentation models are known to those skilled in the art, for example the presentation models described in more detail in international patent application publications WO/2017/106638, WO/2018/195357, WO/2018/208856, WO2016187508, and US patent application US20110293637, each herein incorporated by reference, in their entirety, for all purposes.

X. Training Module

[0435] Training modules can be used to construct one or more presentation models based on training data sets that generate likelihoods of whether peptide sequences will be presented by MHC alleles associated with the peptide sequences. Various training modules are known to those skilled in the art, for example the presentation models described in more detail in international patent application publications WO/2017/106638, WO/2018/195357, and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes. A training module can construct a presentation model to predict presentation likelihoods of peptides on a per-allele basis. A training module can also construct a presentation model to predict presentation likelihoods of peptides in a multiple-allele setting where two or more MHC alleles are present.

XI. Prediction Module

[0436] A prediction module can be used to receive sequence data and select candidate antigens in the sequence data using a presentation model. Specifically, the sequence data may be DNA sequences, RNA sequences, and/or protein sequences extracted from tumor tissue cells of patients. A prediction module may identify candidate neoantigens that are mutated peptide sequences by comparing sequence data extracted from normal tissue cells of a patient with the sequence data extracted from tumor tissue cells of the patient to identify portions containing one or more mutations. A prediction module may identify candidate antigens that have altered expression in a tumor cell or cancerous tissue in comparison to a normal cell or tissue by comparing sequence data extracted from normal tissue cells of a patient with the sequence data extracted from tumor tissue cells of the patient to identify improperly expressed candidate antigens.

[0437] A presentation module can apply one or more presentation model to processed peptide sequences to estimate presentation likelihoods of the peptide sequences. Specifically, the prediction module may select one or more candidate antigen peptide sequences that are likely to be presented on tumor HLA molecules by applying presentation models to the candidate antigens. In one implementation, the presentation module selects candidate antigen sequences that have estimated presentation likelihoods above a predetermined threshold. In another implementation, the presentation model selects the N candidate antigen sequences that have the highest estimated presentation likelihoods (where N is generally the maximum number of epitopes that can be delivered in a vaccine). A vaccine including the selected candidate antigens for a given patient can be injected into the patient to induce immune responses.

XI.B.Cassette Design Module

XI.B.1 Overview

[0438] A cassette design module can be used to generate a vaccine cassette sequence based on selected candidate peptides for injection into a patient. Various cassette design modules are known to those skilled in the art, for example the cassette design modules described in more detail in international patent application publications WO/2017/106638, WO/2018/195357, and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes.

[0439] A set of therapeutic epitopes may be generated based on the selected peptides determined by a prediction module associated with presentation likelihoods above a predetermined threshold, where the presentation likelihoods are determined by the presentation models. However it is appreciated that in other embodiments, the set of therapeutic epitopes may be generated based on any one or more of a number of methods (alone or in combination), for example, based on binding affinity or predicted binding affinity to HLA class I or class II alleles of the patient, binding stability or predicted binding stability to HLA class I or class II alleles of the patient, random sampling, and the like.

[0440] Therapeutic epitopes may correspond to selected peptides themselves Therapeutic epitopes may also include C- and/or N-terminal flanking sequences in addition to the selected peptides. N- and C-terminal flanking sequences can be the native N- and C-terminal flanking sequences of the therapeutic vaccine epitope in the context of its source protein. Therapeutic epitopes can represent a fixed-length epitope Therapeutic epitopes can represent a variable-length epitope, in which the length of the epitope can be varied depending on, for example, the length of the C- or N-flanking sequence. For example, the C-terminal flanking sequence and the N-terminal flanking sequence can each have varying lengths of 2-5 residues, resulting in 16 possible choices for the epitope.

[0441] A cassette design module can also generate cassette sequences by taking into account presentation of junction epitopes that span the junction between a pair of therapeutic epitopes in the cassette. Junction epitopes are novel non-self but irrelevant epitope sequences that arise in the cassette due to the process of concatenating therapeutic epitopes and linker sequences in the cassette. The novel sequences of junction epitopes are different from the therapeutic epitopes of the cassette themselves.

[0442] A cassette design module can generate a cassette sequence that reduces the likelihood that junction epitopes are presented in the patient. Specifically, when the cassette is injected into the patient, junction epitopes have the potential to be presented by HLA class I or HLA class II alleles of the patient, and stimulate a CD8 or CD4 T-cell response, respectively. Such reactions are often times undesirable because T-cells reactive to the junction epitopes have no therapeutic benefit, and may diminish the immune response to the selected therapeutic epitopes in the cassette by antigenic competition..sup.76

[0443] A cassette design module can iterate through one or more candidate cassettes, and determine a cassette sequence for which a presentation score of junction epitopes associated with that cassette sequence is below a numerical threshold. The junction epitope presentation score is a quantity associated with presentation likelihoods of the junction epitopes in the cassette, and a higher value of the junction epitope presentation score indicates a higher likelihood that junction epitopes of the cassette will be presented by HLA class I or HLA class II or both.

[0444] In one embodiment, a cassette design module may determine a cassette sequence associated with the lowest junction epitope presentation score among the candidate cassette sequences.

[0445] A cassette design module may iterate through one or more candidate cassette sequences, determine the junction epitope presentation score for the candidate cassettes, and identify an optimal cassette sequence associated with a junction epitope presentation score below the threshold.

[0446] A cassette design module may further check the one or more candidate cassette sequences to identify if any of the junction epitopes in the candidate cassette sequences are self-epitopes for a given patient for whom the vaccine is being designed. To accomplish this, the cassette design module checks the junction epitopes against a known database such as BLAST. In one embodiment, the cassette design module may be configured to design cassettes that avoid junction self-epitopes.

[0447] A cassette design module can perform a brute force approach and iterate through all or most possible candidate cassette sequences to select the sequence with the smallest junction epitope presentation score. However, the number of such candidate cassettes can be prohibitively large as the capacity of the vaccine increases. For example, for a vaccine capacity of 20 epitopes, the cassette design module has to iterate through ˜1018 possible candidate cassettes to determine the cassette with the lowest junction epitope presentation score. This determination may be computationally burdensome (in terms of computational processing resources required), and sometimes intractable, for the cassette design module to complete within a reasonable amount of time to generate the vaccine for the patient. Moreover, accounting for the possible junction epitopes for each candidate cassette can be even more burdensome. Thus, acassette design module may select a cassette sequence based on ways of iterating through a number of candidate cassette sequences that are significantly smaller than the number of candidate cassette sequences for the brute force approach.

[0448] A cassette design module can generate a subset of randomly or at least pseudo-randomly generated candidate cassettes, and selects the candidate cassette associated with a junction epitope presentation score below a predetermined threshold as the cassette sequence. Additionally, the cassette design module may select the candidate cassette from the subset with the lowest junction epitope presentation score as the cassette sequence. For example, the cassette design module may generate a subset of ˜1 million candidate cassettes for a set of 20 selected epitopes, and select the candidate cassette with the smallest junction epitope presentation score. Although generating a subset of random

cassette sequences and selecting a cassette sequence with a low junction epitope presentation score out of the subset may be suboptimal relative to the brute force approach, it requires significantly less computational resources thereby making its implementation technically feasible. Further, performing the brute force method as opposed to this more efficient technique may only result in a minor or even negligible improvement in junction epitope presentation score, thus making it not worthwhile from a resource allocation perspective. A cassette design module can determine an improved cassette configuration by formulating the epitope sequence for the cassette as an asymmetric traveling salesman problem (TSP). Given a list of nodes and distances between each pair of nodes, the TSP determines a sequence of nodes associated with the shortest total distance to visit each node exactly once and return to the original node. For example, given cities A, B, and C with known distances between each other, the solution of the TSP generates a closed sequence of cities, for which the total distance traveled to visit each city exactly once is the smallest among possible routes. The asymmetric version of the TSP determines the optimal sequence of nodes when the distance between a pair of nodes are asymmetric. For example, the "distance" for traveling from node A to node B may be different from the "distance" for traveling from node B to node A. By solving for an improved optimal cassette using an asymmetric TSP, the cassette design module can find a cassette sequence that results in a reduced presentation score across the junctions between epitopes of the cassette. The solution of the asymmetric TSP indicates a sequence of therapeutic epitopes that correspond to the order in which the epitopes should be concatenated in a cassette to minimize the junction epitope presentation score across the junctions of the cassette. A cassette sequence determined through this approach can result in a sequence with significantly less presentation of junction epitopes while potentially requiring significantly less computational resources than the random sampling approach, especially when the number of generated candidate cassette sequences is large. Illustrative examples of different computational approaches and comparisons for optimizing cassette design are described in more detail in international patent application publications WO/2017/106638, WO/2018/195357, and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes.

XI.B.2 Shared Antigen Vaccine Sequence Selection

[0449] Shared antigen sequences for inclusion in a shared antigen vaccine and appropriate patients for treatment with such vaccine can be chosen by one of skill in the art using the detailed disclosure provided herein. For example, Tables: A, 1.2, Additional MS Validated Neoantigens, or AACR GENIE Results can be used for sequence selection. In certain instances a particular mutation and HLA allele combination can be preferred (e.g., based on sequencing data available from a given subject indicating that each are present in the subject) and subsequently used in combination together to identify a shared neoantigen sequence using Table A, Additional MS Validated Neoantigens, or AACR GENIE Results for inclusion in a vaccine. Exemplary mutations and their matched HLA alleles are shown in Tables 32A, 32B, and 34.

[0450] For example, for KRAS_G13D, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list KRAS_G13D and C0802 and A1101.

[0451] For example, for KRAS_Q61K or NRAS_Q61K, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list (1) KRAS_Q61K and A0101; or (2) NRAS Q61K, and A0101.

[0452] For example, for TP53_R249M, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list TP53_R249M and at least one of B3512, B3503, and B3501.

[0453] For example, for CTNNB1_S45P, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list CTNNB1_S45P and at least one of A0301, A6801, A0302, and A1101. For example, see relevant sequences shown in Table 32A and Table 32B.

[0454] For example, for CTNNB1_S45F, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list CTNNB1_S45F and at least one of A0301, A1101, and A6801.

[0455] For example, for ERBB2_Y772_A775dup, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list ERBB2_Y772_A775dup and B1801.

[0456] For example, for KRAS_G12D or NRAS_G12D, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list (1) KRAS_G12D and at least one of A1101, A0301, and C0802; or (2) NRAS_G12D and at least one of A1101, A0301, and C0802. For example, see relevant sequences shown in Table 32A or Table 32B.

[0457] For example, for KRAS_Q61R or NRAS_Q61R, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list (1)

KRAS_Q61R and A0101; or (2) NRAS_Q61R and A0101. For example, see relevant sequence shown in Table 32B.

[0458] For example, for CTNNB1_T41A, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list CTNNB1_T41A and at least one of A0301, A0302, A1101, B1510, C0303, and C0304. For example, see relevant sequence shown in Table 32B.

[0459] For example, for TP53_K132N, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list TP53_K132N and at least one of A2402 and A2301. For example, see relevant sequence shown in Table 32A.

[0460] For example, for KRAS_G12A, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list KRAS_G12A and at least one of A0301 and A1101.

[0461] For example, for KRAS_Q61L or NRAS_Q61L, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list (1) KRAS_Q61L and A0101; or (2) NRAS_Q61L and A0101.

[0462] For example, for TP53_R213L, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list TP53_R213L and at least one of A0207, C0802, and A0201. For example, see relevant sequence shown in Table 32B.

[0463] For example, for BRAF_G466V, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list BRAF_G466V and at least one of B1501 and B1503.

[0464] For example, for KRAS_G12V, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list KRAS_G12V and at least one of A0301, A1101, C0102, and A0302. For example, see relevant sequences shown in Table 32A and Table 32B.

[0465] For example, for KRAS_Q61H or NRAS_Q61H, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list (1) KRAS_Q61H and A0101; or (2) NRAS_Q61H and A0101.

[0466] For example, for CTNNB1_S37F, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list CTNNB1_S37F and at least one of A0101, A2301, A2402, B1510, B3906, C0501, C1402, and C1403.

[0467] For example, for TP53_S127Y, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list TP53_S127Y and at least one of A1101 and A0301.

[0468] For example, for TP53_K132E, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list TP53_K132E and at least one of A2402, C1403, and A2301.

[0469] For example, for KRAS_G12C or NRAS_G12C, a shared neoantigen or shared neoantigen-encoding sequence for inclusion in a vaccine can be selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion is selected by identifying all rows that list (1) KRAS_G12C and at least one of A0201, A0301, and A1101; or (2) NRAS_G12C and at least one of A0201, A0301, and A1101. For example, see relevant sequences shown in Table 32A.

XIII. Example Computer

[0470] A computer can be used for any of the computational methods described herein. One skilled in the art will recognize a computer can have different architectures. Examples of computers are known to those skilled in the art, for example the computers described in more detail in international patent application publications WO/2017/106638, WO/2018/195357, and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes.

XIV. Antigen Delivery Vector Example

[0471] Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

[0472] The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, biochemistry, recombinant DNA techniques and pharmacology, within the skill of the art. Such techniques

are explained fully in the literature. Sec, e.g., T. E. Creighton, *Proteins: Structures and Molecular Properties* (W.H. Freeman and Company, 1993); A. L. Lehninger, *Biochemistry* (Worth Publishers, Inc., current addition); Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *Methods In Enzymology* (S. Colowick and N. Kaplan eds., Academic Press, Inc.); *Remington's Pharmaceutical Sciences*, 18th Edition (Easton, Pennsylvania: Mack Publishing Company, 1990); Carey and Sundberg *Advanced Organic Chemistry* 3.sup.rd Ed. (Plenum Press) Vols A and B (1992).

XIV.A. Neoantigen Cassette Design

[0473] Through vaccination, multiple class I MHC restricted tumor-specific neoantigens (TSNAs) that stimulate the corresponding cellular immune response(s) can be delivered. In one example, a vaccine cassette was engineered to encode multiple epitopes as a single gene product where the epitopes were either embedded within their natural, surrounding peptide sequence or spaced by non-natural linker sequences. Several design parameters were identified that could potentially impact antigen processing and presentation and therefore the magnitude and breadth of the TSNA specific CD8 T cell responses. In the present example, several model cassettes were designed and constructed to evaluate: (1) whether robust T cell responses could be generated to multiple epitopes incorporated in a single expression cassette; (2) what makes an optimal linker placed between the TSNAs within the expression cassette—that leads to optimal processing and presentation of all epitopes; (3) if the relative position of the epitopes within the cassette impact T cell responses; (4) whether the number of epitopes within a cassette influences the magnitude or quality of the T cell responses to individual epitopes; (5) if the addition of cellular targeting sequences improves T cell responses.

[0474] Two readouts were developed to evaluate antigen presentation and T cell responses specific for marker epitopes within the model cassettes: (1) an in vitro cell-based screen which allowed assessment of antigen presentation as gauged by the activation of specially engineered reporter T cells (Aarnoudse et al., 2002; Nagai et al., 2012); and (2) an in vivo assay that used HLA-A2 transgenic mice (Vitiello et al., 1991) to assess post-vaccination immunogenicity of cassette-derived epitopes of human origin by their corresponding epitope-specific T cell responses (Cornet et al., 2006; Depla et al., 2008; Ishioka et al., 1999).

XV. ChAd Antigen Cassette Delivery Vector

XV.A. ChAd Antigen Cassette Delivery Vector Construction

[0475] In one example, Chimpanzee adenovirus (ChAd) was engineered to be a delivery vector for antigen cassettes. In a further example, a full-length ChAdV68 vector was synthesized based on AC_000011.1 (sequence 2 from U.S. Pat. No. 6,083,716) with E1 (nt 457 to 3014) and E3 (nt 27,816-31,332) sequences deleted. Reporter genes under the control of the CMV promoter/enhancer were inserted in place of the deleted E1 sequences. Transfection of this clone into HEK293 cells did not yield infectious virus. To confirm the sequence of the wild-type C68 virus, isolate VR-594 was obtained from the ATCC, passaged, and then independently sequenced (SEQ ID NO: 10). When comparing the AC_000011.1 sequence to the ATCC VR-594 sequence (SEQ ID NO:10) of wild-type ChAdV68 virus, 6 nucleotide differences were identified. In one example, a modified ChAdV68 vector was generated based on AC_000011.1, with the corresponding ATCC VR-594 nucleotides substituted at five positions (ChAdV68.5WTnt SEQ ID NO:1).

[0476] In another example, a modified ChAdV68 vector was generated based on AC_000011.1 with E1 (nt 577 to 3403) and E3 (nt 27,816-31,332) sequences deleted and the corresponding ATCC VR-594 nucleotides substituted at four positions. A GFP reporter (ChAdV68.4WTnt.GFP; SEQ ID NO:11) or model neoantigen cassette (ChAdV68.4WTnt.MAG25mer; SEQ ID NO: 12) under the control of the CMV promoter/enhancer was inserted in place of deleted E1 sequences.

[0477] In another example, a modified ChAdV68 vector was generated based on AC_000011.1 with E1 (nt 577 to 3403) and E3 (nt 27,125-31,825) sequences deleted and the corresponding ATCC VR-594 nucleotides substituted at five positions. A GFP reporter (ChAdV68.5WTnt.GFP; SEQ ID NO:13) or model neoantigen cassette (ChAdV68.5WTnt.MAG25mer; SEQ ID NO:2) under the control of the CMV promoter/enhancer was inserted in place of deleted E1 sequences.

[0478] Relevant vectors are described below: [0479] Full-Length ChAdVC68 sequence "ChAdV68.5WTnt" (SEQ ID NO:1); AC_000011.1 sequence with corresponding ATCC VR-594 nucleotides substituted at five positions. [0480] ATCC VR-594 C68 (SEQ ID NO:10); Independently sequenced; Full-Length C68 [0481] ChAdV68.4WTnt.GFP (SEQ ID NO:11); AC_000011.1 with E1 (nt 577 to 3403) and E3 (nt 27,816-31,332) sequences deleted; corresponding ATCC VR-594 nucleotides substituted at four positions; GFP reporter under the control of the CMV promoter/enhancer inserted in place of deleted E1 [0482] ChAdV68.4WTnt.MAG25mer (SEQ ID NO:12); AC_000011.1 with E1 (nt 577 to 3403) and E3 (nt 27,816-31,332) sequences deleted; corresponding ATCC VR-594 nucleotides substituted at four positions; model neoantigen cassette under the control of the CMV promoter/enhancer inserted in place of deleted E1 [0483] ChAdV68.5WTnt.GFP (SEQ ID NO:13); AC_000011.1 with E1 (nt 577 to 3403) and E3 (nt 27,125-31,825) sequences deleted; corresponding ATCC VR-594 nucleotides substituted at five positions; GFP reporter under the control of the CMV promoter/enhancer inserted in place of deleted E1

XVI. Alphavirus Antigen Cassette Delivery Vector

[0484] In one implementation of the present invention, a RNA alphavirus backbone for the antigen expression system was generated from a Venezuelan Equine Encephalitis (VEE) (Kinney, 1986, Virology 152:400-413) based self-

replicating RNA (srRNA) vector. In one example, the sequences encoding the structural proteins of VEE located 3′ of the 26S subgenomic promoter were deleted (VEE sequences 7544 to 11,175 deleted; numbering based on Kinney et al 1986; SEQ ID NO:6) and replaced by antigen sequences. (SEQ ID NO:14 and SEQ ID NO: 4) or a luciferase reporter (e.g., VEE-Luciferase, SEQ ID NO:15). Evaluation of the RNA alphavirus backbone, including those encoding antigen cassettes, is described in international patent application publication WO/2021/092095, herein incorporated by reference for all purposes.

XVIII. Non-Human Primate Studies

[0485] Various dosing protocols using ChAdV68 and self-replicating RNA (srRNA) were evaluated in non-human primates (NHP).

Assessment of IFNα Levels Following samRNA Immunization

[0486] A priming vaccine was injected intramuscularly (IM) in each NHP to initiate the study (vaccine prime). One or more boosting vaccines (vaccine boost) were also injected intramuscularly in each NHP. Bilateral injections per dose were administered according to groups outlined in tables and summarized below.

Immunizations

[0487] Mamu-A*01 Indian rhesus macaques were immunized bilaterally with 30 μg of VEE-MAG25MER samRNA, 300 μg of VEE-MAG25mer samRNA, or 1000 μg of VEE-MAG25mer samRNA formulated in with an LNP. Vaccine boosts of 30 μg, 300 μg or 1000 μg VEE-MAG25mer srRNA were administered intramuscularly at 4-weeks after prime vaccination.

Immune Monitoring

[0488] Scrum IFNa2a levels assessed by ELISA at 8 hours following the first samRNA dose. As shown in FIG. **1**, there was a dose-dependent induction of IFNα following samRNA immunization that increased as the dose of samRNA increased.

XX. Identification of Shared Neoantigens

[0489] We identified shared neoantigens using a series of steps. We obtained a list of common driver mutations classified as "confirmed somatic" from the COSMIC database. For each mutation, we generated candidate neoepitopes (8 to 11-mer peptides), used a TPM of 100, and ran our EDGE prediction model (a deep learning model trained on HLA presented peptides sequenced by MS/MS, as described in international patent application publications WO/2017/106638, WO/2018/195357, and WO/2018/208856, each herein incorporated by reference, in their entirety, for all purposes) across all modeled HLA alleles. Note that each peptide contained at least one mutant amino acid and was not a self-peptide. We then recorded any peptide with an HLA allele that has an EDGE score>0.001. The results are shown in Table A. A total of 10261 shared neoantigen sequences were thus identified and are described in SEQ ID NO: 10,755-21,015. The corresponding HLA allele(s) for each sequence are shown.

[0490] The initial list provided in Table A was further analyzed for level of neoantigen/HLA prevalence in the patient population. "Antigen/HLA prevalence" is calculated as the frequency of antigen (A) in a given population multiplied by the frequency of an HLA allele (B) in the given population. Antigen/HLA prevalence can also refer to mutation/HLA prevalence or neoantigen/HLA prevalence. As part of this analysis, for each mutation, its (A) frequency was obtained across common tumor types in TCGA and recorded at its highest frequency amongst tumor types. (B) For each HLA allele in EDGE, the HLA allele frequency TCGA (a predominantly Caucasian population) was recorded. HLA allele frequencies are described in more detail in Shukla, S. A. et al. (*Nat. Biotechnol.* 33, 1152-1158 2015), herein incorporated by reference for all purposes. The neoantigen/HLA prevalence was calculated as (A) multiplied by (B). Any neocpitopc/HLA pair in Table A that is >0.1% prevalence using this methodology is identified with a "Most Common 1" (2387/10261).

[0491] Additionally, we characterized the prevalence of cancer driver mutations across a large cohort of patient samples representative of the advanced cancer patient population relevant for potential clinical studies. EDGE prediction was performed using the publicly-released AACR Genie v4.1 dataset, which has over 40,000 patients sequenced on NGS cancer gene panels ranging from 50 to 500 genes, from major academic cancer centers including Dana-Farber, Johns Hopkins, MD Anderson, MSKCC, and Vanderbilt. We selected base substitution and indel mutations in lung, microsatellite stable colon, and pancreatic cancers, and required coverage across multiple gene panels. We analyzed each neoantigen peptide paired with each of greater than 90 Class I HLA alleles covered in our EDGE antigen presentation prediction model and recorded the epitopes and corresponding HLA alleles with an EDGE probability of HLA presentation score of >0.001. We then determined the neoantigen/HLA prevalence of those peptide with an EDGE score>0.001, calculated as A*B, where A is the highest frequency of the mutation amongst the three tumor types and B is the HLA allele frequency. We used HLA allele frequencies representative of the population in the USA by examining HLA alleles from the TCGA population and tabulating the frequency for each HLA allele (Shukla, S. A. et al.). Peptides and corresponding HLA alleles that demonstrated a neoantigen/HLA prevalence >0.01% from the analysis are described in SEQ ID NO:21,016-29,357 and referred to as AACR GENIE Results.

[0492] In addition, we analyzed the EDGE prediction run on the available AACR GENIE database of tumor mutations for specific cancer subsets across the following modeled HLA alleles: [0493] A0101; A0201; A0205; A0206; A0301; A1101; A2301; A2402; A2501; A2601; A2902; A3001; A3002; A3101; A3201; A3301; A3303; A6801; A6802; B0702; B0801; B1302; B1401; B1402; B1501; B1503; B1510; B1801; B2705; B3501; B3502; B3503; B3508; B3701; B3801;

B3901; B3906; B4001; B4002; B4402; B4403; B4901; B5001; B5101; B5501; B5701; B5801; C0102; C0202; C0302; C0303; C0304; C0401; C0501; C0602; C0701; C0702; C0704; C0801; C0802; C1203; C1402; C1502; C1601; C1701

[0494] We then recorded any peptide with an HLA allele that has an EDGE score>0.1. The results are described in international patent application publication WO/2021/092095, herein incorporated by reference for all purposes.

XXI. A. Validation of Shared Neoantigen Presentation in Human Tumors

[0495] Mass spectrometry (MS) validation of candidate shared neoantigens was performed using targeted mass spectrometry methods. Nearly 500 frozen resected lung, colorectal and pancreatic tumor samples were homogenized and used for both RNASeq transcriptome sequencing and immunoprecipitation of the HLA/peptide complexes. A peptide target list was generated for each sample by analysis of the transcriptome, whereby recurrent cancer driver mutations, as defined in the AACR Genie v4.1 dataset, were identified and RNA expression levels assessed. The EDGE model of antigen presentation was then applied to the mutation sequence and expression data to prioritize peptides for the targeting list. The peptides from the HLA molecules were eluted and collected using size exclusion to isolate the presented peptides prior to mass spectrometry. Synthetic heavy labeled peptide with the same amino acid sequence was co-loaded with each sample for targeted mass spectrometry. Both coelution of the heavy labeled peptide with the experimental peptide and analysis of the fragmentation pattern we were used to validate a candidate epitope. Mass spectrometry analysis methods are described in more detail in Gillete et al. (*Nat Methods.* 2013 January; 10 (1): 28-34), herein incorporated by reference in its entirety for all purposes. Shared neoantigen epitopes from driver mutations validated in this manner with sufficient prevalence for further consideration are summarized in Table 32A below, along with sample tumor type and associated HLA alleles.

TABLE-US-00003 TABLE 32A MS-validated neoantigen neoepitopes in human tumors Peptide Targeted Patient Tumor Presenting Gene_ Elution Time Peptide ID Type HLA* Mutation (min.) A0002082 CRC HLA-CTNNB1_S4 17.2 TTAPPLSGK A*03:01 5P (SEQ ID NO: 29393) A0001877 Lung HLA- KRAS_G12C 46.7 KLVVVGACG A*02:01 V (SEQ ID NO: 29395) A0002129 Lung HLA- KRAS_G12C 46.5 KLVVVGACG A*02:01 V (SEQ ID NO: 29395) A0001947 CRC HLA- KRAS_G12D 12 (pep. #1) VVGADGVGK A*11:01 (SEQ ID NO: 29396) A0001947 CRC HLA- KRAS_G12D 23.4 (pep. #2) VVVGADGVG A*11:01 K (SEQ ID NO: 29397) A0001474 Gastric HLA- KRAS_G12V 30.4 VVVGAVGVG A*11:01 K (SEQ ID NO: 29398) A0001730 Pancreatic HLA- KRAS_G12V 18.4 (pep. #1) VVGAVGVGK A*11:01 (SEQ ID NO: 29399) A0001730 Pancreatic HLA- KRAS_G12V 31.6 (pep. #2) VVVGAVGVG A*11:01 K (SEQ ID NO: 29398) A0001896 Lung HLA- KRAS_G12V 23.6 AVGVGKSAL C*01:02 (SEQ ID NO: 29400) A0001966 CRC HLA- KRAS_G12V 27.8 VVVGAVGVG A*03:01 K (SEQ ID NO: 29398) A0001973 Ovarian HLA-TP53_K132N 97.7 TYSPALNNM A*24:02 F (SEQ ID NO: 29401) A0001983 Ovarian HLA- CTNNB1_S37Y 50.9 YLDSGIHYGA A*02:01 A0000799 CRC HLA- CHD4_K73fs 43.3 TVRAATIL B*08:01 *When the same peptide was predicted to be presented by multiple HLA alleles for a patient and detected by MS/MS, it was inferred to be presented by the highest scoring HLA allele by EDGE or both alleles if the scores were sufficiently close

[0496] We further evaluated the MS data with respect to mutations for which peptides were not detected in order to assess the value of narrowly targeting patients with specific HLAs for treatment, e.g., requiring patients to have at least one validated or predicted HLA allele that presents a neoantigen contained in a vaccine cassette.

[0497] For example, in the case of KRAS, we counted the number of patient samples in which KRAS epitope peptides for particular HLA alleles were detected or not detected. (When the same peptide was predicted to be presented by multiple HLA alleles for a patient and detected by MS/MS, it was inferred to be presented by the highest scoring HLA allele by EDGE or both alleles if the scores were sufficiently close). Results are presented in Table 33. Based on these results, several common HLA alleles are not expected to present a given KRAS neoantigen and these KRAS neoantigen/HLA pair can be excluded for purposes of selection criteria for a vaccine cassette design and patient selection in this instance. For example, Table 34 directed to a specific vaccine cassette (see Section XXII below) does not include predicted neoantigen/HLA pair G12D/A*02:01 on the basis that the peptide was not detected in 17 samples tested, and likewise did not include G12V/A*02:01 on the basis that the peptide was not detected in 9 samples tested. In contrast, neoantigen/HLA pair G12D/A*11:01 was considered validated on the basis that the peptide was detected in 1/5 samples tested, and likewise G12V/A*11:01 was considered validated on the basis that the peptide was detected in 2/6 samples tested.

[0498] These results highlight the importance of identifying the relevant neoantigen/HLA pairs for proper HLA type selection in patient selection for treatment with a shared neoantigen vaccine, such as that described in Table 34. Specifically, several common KRAS neoantigen/HLA pairs were excluded for purposes of selection criteria in this case as the MS data suggested a shared neoantigen vaccine would unlikely provide a benefit to a patient with that predicted KRAS neoantigen/HLA pair (e.g., G12D/A*02:01 or G12V/A*02:01).

TABLE-US-00004 TABLE 33 Presentation of KRAS Neoantigens HLA Peptide Number of tested patient Gene.sub.—confirmation samples with mutation HLA by MS/MS MS/MS result KRAS_G12A HLA-A*02:01 N 2 KRAS_G12A HLA-A*02:06 N 1 KRAS_G12A HLA-A*03:01 N 3 KRAS_G12A HLA-B*27:05 N 1 KRAS_G12A HLA-B*35:01 N 1 KRAS_G12A HLA-B*41:02 N 1 KRAS_G12A HLA-B*48:01 N 1 KRAS_G12A HLA-C*08:03 N 1 KRAS_G12C HLA-A*02:01 Y 2 KRAS_G12C HLA-A*02:01 N 2 KRAS_G12C HLA-A*03:01 N 8 KRAS_G12C HLA-A*03:02 N 1 KRAS_G12C HLA-A*68:01 N 1 KRAS_G12C HLA-B*27:05 N 1 KRAS_G12D HLA-A*02:01 N 17 KRAS_G12D

HLA-A*02:05 N 2 KRAS_G12D HLA-A*03:01 N 4 KRAS_G12D HLA-A*11:01 Y 1 KRAS_G12D HLA-A*11:01 N 4 KRAS_G12D HLA-A*26:01 N 2 KRAS_G12D HLA-A*31:01 N 4 KRAS_G12D HLA-A*68:01 N 3 KRAS_G12D HLA-B*07:02 N 4 KRAS_G12D HLA-B*08:01 N 1 KRAS_G12D HLA-B*13:02 N 1 KRAS_G12D HLA-B*15:01 N 5 KRAS_G12D HLA-B*27:05 N 2 KRAS_G12D HLA-B*35:01 N 2 KRAS_G12D HLA-B*37:01 N 1 KRAS_G12D HLA-B*38:01 N 2 KRAS_G12D HLA-B*40:01 N 3 KRAS_G12D HLA-B*40:02 N 3 KRAS_G12D HLA-B*44:02 N 1 KRAS_G12D HLA-B*44:03 N 4 KRAS_G12D HLA-B*48:01 N 1 KRAS_G12D HLA-B*50:01 N 1 KRAS_G12D HLA-B*57:01 N 1 KRAS_G12D HLA-C*01:02 N 1 KRAS_G12D HLA-C*02:02 N 1 KRAS_G12D HLA-C*03:03 N 3 KRAS_G12D HLA-C*03:04 N 2 KRAS_G12D HLA-C*04:01 N 8 KRAS_G12D HLA-C*05:01 N 2 KRAS_G12D HLA-C*07:04 N 1 KRAS_G12D HLA-C*08:02 N 2 KRAS_G12D HLA-C*08:03 N 1 KRAS_G12D HLA-C*16:01 N 1 KRAS_G12D HLA-C*17:01 N 1 KRAS_G12R HLA-B*41:02 N 1 KRAS_G12R HLA-C*07:04 N 1 KRAS_G12V HLA-A*02:01 N 9 KRAS_G12V HLA-A*02:05 N 1 KRAS_G12V HLA-A*02:06 N 1 KRAS_G12V HLA-A*03:01 Y 1 KRAS_G12V HLA-A*03:01 N 4 KRAS_G12V HLA-A*11:01 Y 2 KRAS_G12V HLA-A*11:01 N 4 KRAS_G12V HLA-A*25:01 N 3 KRAS_G12V HLA-A*26:01 N 2 KRAS_G12V HLA-A*30:01 N 2 KRAS_G12V HLA-A*31:01 N 2 KRAS_G12V HLA-A*32:01 N 1 KRAS_G12V HLA-A*68:02 N 1 KRAS_G12V HLA-B*07:02 N 6 KRAS_G12V HLA-B*08:01 N 1 KRAS_G12V HLA-B*13:02 N 2 KRAS_G12V HLA-B*14:02 N 1 KRAS_G12V HLA-B*15:01 N 2 KRAS_G12V HLA-B*27:05 N 2 KRAS_G12V HLA-B*39:01 N 1 KRAS_G12V HLA-B*40:01 N 1 KRAS_G12V HLA-B*40:02 N 1 KRAS_G12V HLA-B*41:02 N 1 KRAS_G12V HLA-B*44:05 N 1 KRAS_G12V HLA-B*50:01 N 1 KRAS_G12V HLA-B*51:01 N 1 KRAS_G12V HLA-C*01:02 Y 1 KRAS_G12V HLA-C*01:02 N 1 KRAS_G12V HLA-C*03:03 N 1 KRAS_G12V HLA-C*03:04 N 2 KRAS_G12V HLA-C*08:02 N 2 KRAS_G12V HLA-C*14:02 N 1 KRAS_G12V HLA-C*17:01 N 2 KRAS_G13D HLA-A*02:01 N 2 KRAS_G13D HLA-B*07:02 N 1 KRAS_G13D HLA-B*08:01 N 1 KRAS_G13D HLA-B*35:01 N 1 KRAS_G13D HLA-B*35:03 N 1 KRAS_G13D HLA-B*35:08 N 1 KRAS_G13D HLA-B*38:01 N 1 KRAS_G13D HLA-C*04:01 N 3 KRAS_Q61H HLA-A*01:01 N 1 KRAS_Q61H HLA-A*02:01 N 2 KRAS_Q61H HLA-A*23:01 N 2 KRAS_Q61H HLA-A*29:01 N 1 KRAS_Q61H HLA-A*30:02 N 1 KRAS_Q61H HLA-A*33:01 N 1 KRAS_Q61H HLA-A*68:01 N 1 KRAS_Q61H HLA-B*07:02 N 1 KRAS_Q61H HLA-B*08:01 N 1 KRAS_Q61H HLA-B*18:01 N 1 KRAS_Q61H HLA-B*35:01 N 1 KRAS_Q61H HLA-B*38:01 N 1 KRAS_Q61H HLA-B*40:01 N 1 KRAS_Q61H HLA-B*44:02 N 1 KRAS_Q61H HLA-C*03:04 N 1 KRAS_Q61H HLA-C*05:01 N 2 KRAS_Q61H HLA-C*08:02 N 1

XXI. B. Validation of Shared Neoantigen Presentation Using and In Vitro System

[0499] Mass spectrometry (MS) validation for HLA-presentation for candidate shared neoantigens was also performed using targeted mass spectrometry methods and an in vitro single-HLA expression system. Briefly, cell lines were engineered to express a single specific HLA alleles and inducibly express a candidate shared neoantigen.

[0500] Single HLA Allele expressing K562 cell lines were created by traditional transfection methods using reagent kits and the instructions provided. Details are provided below for the methods used but other similar methods could be used.

[0501] To create virus particles for transduction of the HLA genes into K562 cells the plasmids were transfection into Phoenix-ampho cells. Phoenix-ampho cells were introduced into 6 well plates at a density of 5×10.sup.5 cells per well and incubated at 37 C overnight prior to transfection. 10 μg of purified DNA was mixed with 10 uL Plus Reagent and brought to 100 uL with pre-warmed Opti-MEM media. Lipofectamine reagent was prepared by mixing 8 μL of Lipofectamine with 92 μL of the pre-warmed Opti-MEM. Both mixtures were incubated at room temperature for 15 prior to mixing the 100 μL of Lipofectamine reagent with the 100 μL of DNA solution and allowing the combined solution to incubate at room temperature for another 15 min. The Phoenix-ampho cells were washed gently by aspirating the media and adding 6 mL of pre-warmed Opti-MEM media to wash the cells. The media was removed from the plated cells. 800 μL of the pre-warmed Opti-MEM was added to the DNA/Lipofectamine mixture to make 1 mL and that solution was added to the plated cells. After the plate was incubated for 3 hrs at 37 C, 3 mL of complete media was added and the cells were incubated overnight at 37 C. The complete media was exchanged after the incubation and the cells incubated for another 2 days. Virus particles were collected after the supernatant was passed through a 45 μm filter into a new 6 well plate. 20 μL of Plus reagent and 8 μL of Lipofectamine was added to each well with a 15 min room temperature incubation after each addition.

[0502] K562 cells were suspended complete media at a concentration of 5×106 per mL. 100 μL of K562 cells were added to each well of the 6 well plate containing the virus particles. The plate was centrifuged at 700×g for 20 min and the cells were incubated for 6 hrs at 37 C. Cells and virus were collected and transferred to T25 flasks with the addition of 7 mL of complete media. The cells were incubated for 3 days prior to a media change which included selection with Puromyocin antibiotic. Live cells were collected and passaged to create stocks of K562 cells expressing a single HLA allele. Overall 25 of these cell lines were created, each with a different HLA expressed, to provide a reagent pool for future experiments.

[0503] A shared neoantigen cassette was created to express 20 shared neoantigens with the mutation centered in a 25mer amino acid chain and was created with no linkers between the entries. This cassette was subcloned into a lentiviral Tet-One Inducible Expression vector system (Clonetech) and lenitvirus was produced in 293T cells by contransfecting the shared neoantigen expression vector with ViraPower (Thermo) packaging plasmids according to manufacturer's specifications. Single HLA Allele expressing K562 cell lines were then transduced with this virus as

described above and single cell clones were characterized for shared neoantigen expression. Briefly, expression of the shared neoantigen cassette was placed under control of a doxycycline (DOX)-controlled TRE3G promoter, where administration of DOX leads to expression of the neoantigens via stabilization of the Tet-On 3G transactivator protein that is constitutively expressed on the same plasmid. The TREG3 promoter-Tet-On 3G transactivator system allows titration of DOX to control the level of expression.

[0504] Cells containing both the single HLA allele and the shared neoantigen cassette were grown to ˜2.5×10.sup.8 cells and pelleted into 15 ml vials. Additionally, cells were plated in limited dilution to prepare single clones of the HLA/Cassette pairing. These single clones were tested to achieve a variety of expression levels of the cassette. Use of cell lines with differing expression levels of the cassette allows for analysis of the system at close to endogenous expression levels. Single clones were also grown to ˜2.5×10.sup.8 cells and pelleted into 15 ml vials. All pellets were washed 2× with cold PBS and frozen to allow for processing for mass spectrometry detection of HLA peptides. Expression levels of the HLA and the cassette was performed using SmartSeq or Taqman assays with appropriate probes.

[0505] For isolation of HLA peptides, each cell pellet was lysed with lysis buffer and centrifuged at 20,000×g for 1 hr to clarify the lysate and the HLA peptide complexes were enriched as previously described (see Section VIII.B. Isolation and Detection of HLA Peptides"). Heavy peptides—peptides synthesized with amino acids containing isotopically heavy amino acids—were added to the peptides prior to analysis by MS to aid in confirmation of the identity of the peptides detected. Results are described in international patent application publication WO/2021/092095, herein incorporated by reference for all purposes.

[0506] Results demonstrating validated HLA-specific presentation of predicted neoantigens are shown in Table 32B. TABLE-US-00005 TABLE 32B MS-validated neoantigen neoepitopes in a single-HLA K562 in vitro system Gene HLA Type Size Sequence CTNNB1_S45P A*11:01 9-mer TTAPPLSGK CTNNB1_T41A A*11:01 9-mer ATAPSLSGK KRAS_G12D A*11:01 10-mer VVVGADGVGK KRAS_G12V A*03:01 9-mer VVGAVGVGK KRAS_G12V A*03:01 10-mer VVVGAVGVGK KRAS_G12V A*11:01 9-mer VVGAVGVGK KRAS_G12V A*11:01 10-mer VVVGAVGVGK KRAS_Q61R A*01:01 10-mer ILDTAGREEY TP53_R213L A*02:01 9-mer YLDDRNTFL XXII. A. Selection of Shared Neoantigens for Vaccine Cassette

[0507] A vaccine cassette ("GO-005") containing 20 shared neoantigens was constructed. Table 34 describes features of the neoantigens selected for the cassette. Shared neoantigens directly detected on the surface of tumor cells by mass spectrometry, as described above in Table 32A, were included in the cassette and the HLA of the epitope was added to the eligible HLA list for the mutations. Neoantigens not independently verified as being presented in our assays were considered validated and added to the cassette if there was compelling literature evidence of tumor presentation (e.g., tumor-infiltrating lymphocytes (TIL) recognizing the neoantigen). KRAS G12D presented by HLA-C*08:02 was considered validated and added based on literature evidence of adoptive cell therapy targeting this neoantigen causing tumor regression in a patient with CRC (Tran et al. N Engl J Med. 2016 Dec. 8; 375 (23): 2255-2262.). Neoantigens with validated HLA alleles occupied 6 out of 20 slots.

[0508] Additional, rarer neoantigens predicted to be presented by tumor cells, but not yet validated by MS, were used to complement the initial set. Mutations with high EDGE scores were prioritized for inclusion as predicted neoantigens given the strong dependence we observed between EDGE score and probability of detection of candidate shared neoantigen peptides by targeted mass spectrometry (MS) validation experiments (see Section XXI above). Results showing the correlation between EDGE score and the probability of detection of candidate shared neoantigen peptides by targeted MS are described in international patent application publication WO/2021/092095, herein incorporated by reference for all purposes. Specifically, the remaining slots were filled with predicted neoantigens with an EDGE HLA presentation score of at least 0.3 and the highest cumulative neoantigen/HLA prevalence across NSCLC, CRC and Pancreatic cancer. For each slot in the cassette, combined HLA frequency was required to be at least 5-10% (e.g., there are 11% of the American population harboring HLA alleles B1501 or B1503). Of note, as KRAS and NRAS harbors the same cassette sequence around codons 12, 13, and 61, incorporation of prevalent NRAS mutations did not require additional slots. Validated HLAs, predicted HLAs with an EDGE score of at least 0.3, the mean EDGE score of the predicted HLAs, and neoantigen/HLA prevalence in the three cancer populations are also presented in Table 34. TABLE-US-00006 TABLE 34 Selected Shared Neoantigens in Vaccine Cassette GO-005 Mean Neoantigen/ Neoantigen/ Neoantigen/ Validated Predicted Predicted HLA Prevalence HLA Prevalence HLA Prevalence Slot Mutation HLA HLA EDGE Score in Lung in CRC in Pancreas 1 KRAS_G13D A1101 C0802 0.306 0.07% 0.39% 0.06% 2 KRAS_Q61K A0101 — N/A - validated 0.05% 0.37% 0.00% NRAS_Q61K only 3 TP53_R249M — B3512. 0.524 0.04% 0.00% 0.00% B3503, B3501 4 CTNNB1_S45P A0301, A6801, 0.894 0.13% 0.00% 0.00% A1101 A0302, 5 CTNNB1_S45F A1101 A0301, 0.478 0.08% 0.27% 0.00% A6801 6 ERBB2_Y772_A775dup.sup.a — B1801 0.758 0.11% 0.00% 0.00% 7 KRAS_G12D A1101, — N/A - validated 0.79% 2.28% 5.45% NRAS_G12D A0301, only C0802 8 KRAS_Q61R A0101 — N/A - validated 0.06% 0.33% 0.47% NRAS_Q61R only 9 CTNNB1_T41A A1101 A0301, 0.545 0.00% 0.27% 0.00% A0302, B1510, C0303, C0304 10 TP53_K132N A2402 A2301 0.659 0.04% 0.00% 0.00% 11 KRAS_G12A A0301, — N/A - validated 0.58% 0.49% 0.00% A1101 only 12 KRAS_Q61L — A0101 0.792 0.22% 0.08% 0.00% NRAS_Q61L 13 TP53_R213L A0201 A0207, 0.558 0.09% 0.18% 0.00% C0802 14 BRAF_G466V — B1501, 0.604 0.03% 0.05% 0.00% B1503 15 KRAS_G12V A0301, A0302 0.759 2.56% 3.43%

9.89% A1101, C0102 16 KRAS_Q61H A0101 — N/A - validated 0.42% 0.28% 0.91% NRAS_Q61H only 17 CTNNB1_S37F A0101 A2301, 0.554 0.29% 0.00% 0.00% A2402 B1510, B3906 C0501, C1402 C1403 18 TP53_S127Y — A1101, 0.522 0.04% 0.00% 0.00% A0301 19 TP53_K132E — A2402, 0.445 0.00% 0.05% 0.00% C1403, A2301 20 KRAS_G12C A0201, — N/A - validated 5.00% 1.46% 0.63% NRAS_G12C A0301, only A1101 .sup.aERBB2_Y772_A775dup is also commonly annotated as A775_G776insYVMA, M774_A775insAYVM, and E770delinsEAYVM. Also, ERBB2 may be referred to as HER2

[0509] Additionally, we determined the total population of patients with at least one HLA allele identified (i.e., either validated or predicted) to present at least one shared neoantigen contained in the shared neoantigen vaccine GO-005 and compared it to the population of patients with the mutations agnostic of whether the patient had the identified allele. Given the GO-005 vaccine cassette from Table 34, to estimate the GO-005 targeted patient population, we collected patient mutation data from AACR Genie. As such patients do not have matching HLA alleles, we sampled HLA alleles from the TCGA population and paired it to the AACR Genie dataset. Then given a tumor type, any patient from AACR Genie with matching both mutation and HLA is labeled positive, and any patient that doesn't meet the criteria is labeled negative. The percent positives give the overall addressable patient population, per tumor type, in Table 35.

[0510] It can be readily appreciated from Table 35 that only a subset of patients who carry a particular mutation also carry the HLA allele likely to present that mutation as a neoantigen. Patients with the mutation, but without the appropriate HLA allele are less likely to benefit from therapy. As an example, whereas an estimated ˜60% of pancreatic cancer patients carry appropriate mutations/neoantigens, more than 2 out of 3 of these patients do not carry the corresponding HLA allele(s). Therefore, a vaccination strategy that considers the relevant mutation and HLA allele pairs as proposed will target just those patients who may benefit. Thus, consideration of epitope presentation by validated or high-scoring predicted HLA is an important step in determining the potential efficacy of a shared neoantigen vaccine.

TABLE-US-00007 TABLE 35 Neoantigen/HLA Prevalence in Target Populations MSS CRC Pancreas Lung GO-005 Targeted Patient 9.0% 17.4% 10.6% Population (cumulative neoantigen/HLA prevalence) HLA agnostic patient 35.1% 60.8% 32.0% population (mutation frequency only)

XXII. B. Shared Neoantigens Vaccine Cassette Sequence Selection

[0511] Shared neoantigen sequences for inclusion in a shared neoantigen vaccine were chosen.

[0512] For KRAS_G13D, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list KRAS_G13D and C0802.

[0513] For KRAS_Q61K or NRAS_Q61K, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list (1) KRAS_Q61K and A0101; or (2) NRAS Q61K, and A0101.

[0514] For TP53_R249M, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list TP53_R249M and at least one of B3512, B3503, and B3501.

[0515] For CTNNB1_S45P, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list CTNNB1_S45P and at least one of A0301, A6801, A0302, and A1101. For example, see relevant sequences shown in Table 32A and Table 32B.

[0516] For CTNNB1_S45F, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list CTNNB1_S45F and at least one of A0301, A1101, and A6801.

[0517] For ERBB2_Y772_A775dup, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list ERBB2_Y772_A775dup and B1801.

[0518] For KRAS_G12D or NRAS_G12D, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list (1) KRAS_G12D and at least one of A1101 and C0802; or (2) NRAS_G12D and at least one of A1101 and C0802. For example, see relevant sequences shown in Table 32A or Table 32B.

[0519] For KRAS_Q61R or NRAS_Q61R, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list (1) KRAS_Q61R and A0101; or (2) NRAS_Q61R and A0101. For example, see relevant sequence shown in Table 32B.

[0520] For CTNNB1_T41A, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list CTNNB1_T41A and at least one of A0301, A0302, A1101, B1510, C0303, and C0304. For example, see relevant sequence shown in Table 32B.

[0521] For TP53_K132N, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list TP53_K132N and at least one of A2402 and A2301. For example, see relevant sequence shown in Table 32A.

[0522] For KRAS_G12A, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list KRAS_G12A.

[0523] For KRAS_Q61L or NRAS_Q61L, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list (1) KRAS_Q61L and A0101; or (2) NRAS_Q61L and A0101.

[0524] For TP53_R213L, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list TP53_R213L and at least one of A0207, C0802, and A0201. For example, see relevant sequence shown in Table 32B.

[0525] For BRAF_G466V, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list BRAF_G466V and at least one of B1501 and B1503.

[0526] For KRAS_G12V, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A, Additional MS Validated Neoantigens, or AACR GENIE Results, where each relevant sequence considered for inclusion was selected by identifying all rows that list KRAS_G12V and at least one of A0301, A1101, C0102, and A0302. For example, see relevant sequences shown in Table 32A and Table 32B.

[0527] For KRAS_Q61H or NRAS_Q61H, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered was selected by identifying all rows that list (1) KRAS_Q61H and A0101; or (2) NRAS_Q61H and A0101.

[0528] For CTNNB1_S37F, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered was selected by identifying all rows that list CTNNB1_S37F and at least one of A2301, A2402, B1510, B3906, C0501, C1402, and C1403.

[0529] For TP53_S127Y, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered was selected by identifying all rows that list TP53_S127Y and at least one of A1101 and A0301.

[0530] For TP53_K132E, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A or AACR GENIE Results, where each relevant sequence considered was selected by identifying all rows that list TP53_K132E and at least one of A2402, C1403, and A2301.

[0531] For KRAS_G12C or NRAS_G12C, the shared neoantigen-encoding sequence for inclusion in the vaccine was selected by reference to Table A 32, or AACR GENIE Results, where each relevant sequence considered was selected by identifying all rows that list (1) KRAS_G12C and A0201; or (2) NRAS_G12C and A0201. For example, see relevant sequences shown in Table 32A.

XXIII. Evaluation of T Cell Recognition of Shared Neoantigens

[0532] We evaluated whether neoantigens induce an immune response in patients. We obtained dissociated tumor cells from a patient with lung adenocarcinoma. Tumor cells were sequenced to determine the patient's HLA and identify mutations. The patient expressed HLA-A*1101 and we identified the KRAS G12V mutation in the tumor. Simultaneously, we sorted and expanded CD45+ cells from the tumor which represent tumor infiltrating lymphocytes (TIL). Expanded TILs were stained with mutated peptide HLA-A*11:01 tetramers to assess immunogenicity of this mutation in the patient. Results are described in international patent application publication WO/2021/092095, herein incorporated by reference for all purposes.

TABLE-US-00008 TABLE 36 Assessment of Neoantigen Reactive Naïve T cells Precursors Number of healthy donors with Number of precursors neoantigen- (% among specific Mutation/ HLA- donors clonotypes per neoantigen Peptide restriction tested) donor KRAS G12V 1 HLA-A*0301 1 (100%) 46 KRAS G12V 2 HLA-A*0301 2 (100%) Donor 1: 25, Donor 2: No sequence data KRAS G12V 1 HLA-A*1101 2 (100%) No sequence data KRAS G12V 2 HLA-A*1101 2 (100%)   Donor 1: 100, Donor 2: No sequence data KRAS G12C 3 HLA-A*0201 1 (100%) No sequence data CTNNB1.sub.— 4 HLA-A*0301 1 (100%) 987 S45P

XXIV. Selection of Shared Neoantigens and Patient Populations

[0533] One or more of the antigens provided in Table 34, Table A, Table 1.2, Additional MS Validated Neoantigens, or the AACR GENIE Results described herein (SEQ ID NOs: 57-29,364) are used to formulate a vaccine composition as described herein. The vaccine is administered to a patient, e.g., to treat cancer. In certain instances the patient is selected, e.g., using a companion diagnostic or a commonly use cancer gene panel NGS assay such as FoundationOne, FoundationOne CDx, Guardant 360, Guardant OMNI, or MSK IMPACT. Exemplary patient selection criteria are described below. An exemplary shared neoantigen vaccine composition GO-005 targets the mutations described in Table 34.

Patient Selection

[0534] Patient selection for shared neoantigen vaccination is performed by consideration of tumor gene expression, somatic mutation status, and patient HLA type. Specifically, a patient is considered eligible for the vaccine therapy if:

[0535] (a) the patient carries an HLA allele predicted or known to present an epitope included in a vaccine and the patient tumor expresses a gene with the epitope sequence, or [0536] (b) the patient carries an HLA allele predicted or known to present an epitope included in a vaccine, and the patient tumor carries the mutation giving rise to the epitope sequence, or [0537] (c) Same as (b), but also requiring that the patient tumor expresses the gene with the mutation above a certain threshold (e.g., 1 TPM or 10 TPM), or [0538] (d) Same as (b), but also requiring that the patient tumor expresses the mutation above a certain threshold (e.g., at least 1 mutated read observed at the level of RNA) [0539] (c) Same as (b), but also requiring both additional criteria in (c) and (d) [0540] (f) Any of the above, but also optionally requiring that loss of the presenting HLA allele is not detected in the tumor

[0541] Gene expression is measured at the RNA or protein level by any of the established methods including RNASeq, microarray, PCR, Nanostring, ISH, Mass spectrometry, or IHC. Thresholds for positivity of gene expression is established by several methods, including: (1) predicted probability of presentation of the epitope by the HLA allele at various gene expression levels, (2) correlation of gene expression and HLA epitope presentation as measured by mass spectrometry, and/or (3) clinical benefits of vaccination attained for patients expressing the genes at various levels. Patient selection is further extended to require positivity for greater than 1 epitope, for examples, at least 2, 3, 4 or 5 epitopes included in the vaccine.

[0542] Somatic mutational status is assessed by any of the established methods, including exome sequencing (NGS DNASeq), targeted exome sequencing (panel of genes), transcriptome sequencing (RNASeq), Sanger sequencing, PCR-based genotyping assays (e.g., Taqman or droplet digital PCR), Mass-spectrometry based methods (e.g., by Sequenom), or any other method known to those skilled in the art.

[0543] Additional new shared neoantigens are identified using any of the methods described, e.g., by mass spectrometry. These newly identified shared neoantigens are incorporated into the vaccine cassettes described herein.

[0544] Previously validated neoantigens are additionally validated as being presented by additional HLA alleles and informs neoantigen selection for the vaccine cassette and/or expands the potential treatable population.

[0545] Inclusions of a new neoantigen enables the broadening of addressable tumor type (eg, EGFR mutated NSCLC) or inclusion of patients with a new tumor type.

XXV. Identification of Shared Antigens

[0546] We identified shared antigen gene-based targets using three computational steps: First, we identified genes with low or no expression in most normal tissues using data available through the Genotype-Tissue Expression (GTEx) Project [1]. We obtained aggregated gene expression data from the Genotype-Tissue Expression (GTEx) Project (version V7p2). This dataset comprised greater than 11,000 post-mortem samples from greater than 700 individuals and greater than 50 different tissue types. Expression was measured using RNA-Seq and computationally processed according to the GTEx standard pipeline (https://www.gtexportal.org/home/documentationPage). Gene expression was calculated using the sum of isoform expressions that were calculated using RSEM v1.2.22 [2].

[0547] Next, we identified which of those genes are aberrantly expressed in cancer samples using data from The Cancer Genome Atlas (TCGA) Research Network: http://cancergenome.nih.gov/. We examined greater than 11,000 samples available from TCGA (Data Release 6.0).

[0548] Finally, in these genes, we identified which peptides are likely to be presented as cell surface antigens by MHC Class I proteins using a deep learning model trained on HLA presented peptides sequenced by MS/MS, as described in international patent application no. PCT/US2016/067159, herein incorporated by reference, in its entirety, for all purposes.

[0549] To identify the common tumor antigens (CTAs; shared antigens), we sought to define criteria to exclude genes that were expressed in normal tissue that were strict enough to ensure tumor specificity, but would account for potential artifacts such as read misalignment. Genes were eligible for inclusion as CTAs if they met the following criteria: The median GTEx expression in each organ that was a part of the brain, heart, or lung was less than 0.1 transcripts per million (TPM) with no one sample exceeding 5 TPM. The median GTEx expression in other essential organs was less than 2 TPM with no one sample exceeding 10 TPM. Expression was ignored for organs classified as non-essential (testis, thyroid, and minor salivary gland). Genes were considered expressed in tumor samples if they had expression in TCGA of greater than 10 TPM in at least 30 samples. Based on literature review, we also added the genes MAGEB4 and MAGEB6.

[0550] We also added the gene CTAGIA/CTAGIB (NY-ESO-1). As it's expression is not accurately quantified using the computational methodology in TCGA Data Release 6.0, we relied on the RSEM calculations available in the TCGA legacy archive (https://portal.gdc.cancer.gov/legacy-archive) which accounts for multiply mapped reads.

[0551] We then examined the distribution of the expression of the remaining genes across TCGA samples. When we examined the known CTAs, e.g. the MAGE family of genes, we observed that the expression of these genes in log space was generally characterized by a bimodal distribution. This distribution included a left mode around a lower expression value and a right mode (or thick tail) at a higher expression level. This expression pattern is consistent with a biological model in which some minimal expression can be detected at baseline in many samples and higher expression of the gene is observed in a subset of tumors experiencing epigenetic dysregulation. We reviewed the

distribution of expression of each gene across TCGA samples and discarded those where we observed only a unimodal distribution with no significant right-hand tail. A small number of genes were eliminated by hand curation, for example, if they were likely to be expressed in a tissue not available in GTEx. This resulted in a set of 59 genes. See Table 46. In Table 46 an X is used to indicate cancers in which the gene is expressed at greater than 10 TPM in at least 1% of cases. TABLE-US-00009 TABLE 46 Analysis of Expression Levels in Cancer Subtypes Cervical Brain Squamous Cell Acute Bladder Lower Breast Carcinoma and Myeloid Adrenocortical Urothelial Grade Invasive Endocervical Colon Esophageal Gene Leukemia Carcinoma Carcinoma Glioma Carcinoma Adenocarcinoma Cholangiocarcinoma Adenocarcinoma Carcinoma ACTL8 X X X X X X ADAM2 X ADAM7 AMELX X BPIFA2 X CRYGC CT83 X X X X X X X CTAG1A/ X X X X X X X CTAG1B CTCFL X X DCAF12L1 X DCAF4L2 X X X X DEFB126 X X X DMRT1 X X DPPA2 X X FMR1NB X FTHL17 X X X X GAGE1 X GAGE12J X X GAGE2A X X X GLYATL3 X X GPRC6A X GSG1L2 IFNK X X X IL22RA2 X X X X INSL4 X X INSL6 X X X X KCNU1 LIN28A X X LUZP4 MAGEA1 X X X X X X X MAGEA10 X X X X MAGEA11 X X X X MAGEA3 X X X X X X X MAGEA4 X X X X X MAGEA6 X X X X X X MAGEA9 X X MAGEB1 X MAGEB2 X X X X X MAGEB4 X MAGEB6 X MAGEC1 X X X X MAGEC2 X X X X X MORC1 X PAGE1 X X X PAGE5 X X X X X PASD1 X X PRDM7 PROKR1 R3HDML X SLC7A13 SMC1B X X X X X SMR3A SSX1 X X X X STRA8 X X X X TFDP3 X XAGE3 X XAGES ZFP42 X X X X X ZNF560 X X X X Lymphoid Head and Kidney Kidney Neoplasm Neck Renal Renal Lung Diffuse Squamous Clear Papillary Liver Squamous Large Glioblastoma Cell Kidney Cell Cell Hepatocellular Lung Cell B-cell Gene Multiforme Carcinoma Chromophobe Carcinoma Carcinoma Carcinoma Carcinoma Adenocarcinoma Carcinoma Lymphoma ACTL8 X X X X X X ADAM2 X ADAM7 X AMELX BPIFA2 X X X X X CRYGC CT83 X X X X X CTAG1A/ X X X X X X CTAG1B CTCFL X X X DCAF12L1 X X DCAF4L2 X X X X DEFB126 X DMRT1 X X DPPA2 X X X X FMR1NB X FTHL17 X X X X GAGE1 X X GAGE12J X GAGE2A X X X X X X GLYATL3 X GPRC6A X X GSG1L2 X IFNK X X IL22RA2 X X X X INSL4 X X INSL6 X X KCNU1 X X LIN28A LUZP4 MAGEA1 X X X X X X MAGEA10 X X X MAGEA11 X X X X MAGEA3 X X X X MAGEA4 X X X X X MAGEA6 X X X X X X MAGEA9 X X MAGEB1 X X X X MAGEB2 X X X X X X MAGEB4 MAGEB6 MAGEC1 X X X X X MAGEC2 X X X X X X X MORC1 PAGE1 X X X X PAGE5 X X X X X X X PASD1 X X PRDM7 PROKR1 R3HDML SLC7A13 X X X SMC1B X X X X X X SMR3A X X SSX1 X X X X STRA8 X X X X X TFDP3 X XAGE3 X XAGES X X ZFP42 X X X ZNF560 X X X Ovarian Serous Pancreatic Pheochromocytoma Prostate Rectum Gene Mesothelioma Cystadenocarcinoma Adenocarcinoma and Paraganglioma Adenocarcinoma Adenocarcinoma Sarcoma ACTL8 X X X X X X ADAM2 X ADAM7 X AMELX X X BPIFA2 X X CRYGC X CT83 X X X CTAG1A/ X X X X X X CTAG1B CTCFL X DCAF12L1 X X DCAF4L2 DEFB126 X DMRT1 X DPPA2 X FMR1NB FTHL17 GAGE1 X X GAGE12J X GAGE2A X X X GLYATL3 X X GPRC6A GSG1L2 IFNK IL22RA2 INSL4 X INSL6 X KCNU1 LIN28A LUZP4 MAGEA1 X X X X X X MAGEA10 X X X MAGEA11 X X X MAGEA3 X X X X X X X MAGEA4 X X X MAGEA6 X X X X X MAGEA9 MAGEB1 X MAGEB2 X X X X MAGEB4 MAGEB6 MAGEC1 X X X MAGEC2 X X X X X MORC1 PAGE1 X X X X PAGE5 X X X PASD1 X X PRDM7 PROKR1 X R3HDML X SLC7A13 SMC1B X X SMR3A SSX1 X X STRA8 X TFDP3 XAGE3 X X XAGES X ZFP42 X X ZNF560 X X X X X Uterine Skin Corpus Cutaneous Stomach Thyroid Uterine Endometrial Uveal Total Gene Melanoma Adenocarcinoma Thymoma Carcinoma Carcinosarcoma Carcinoma Melanoma Indications ACTL8 X X X X X 22 ADAM2 3 ADAM7 2 AMELX X 4 BPIFA2 X X 10 CRYGC X 2 CT83 X X X X 19 CTAG1A/ X X X X X 24 CTAG1B CTCFL X X X X 10 DCAF12L1 X X X 8 DCAF4L2 X 10 DEFB126 X 6 DMRT1 X X 7 DPPA2 X 8 FMR1NB X 3 FTHL17 8 GAGE1 X X X 8 GAGE12J X X 6 GAGE2A X X X X 16 GLYATL3 X X 7 GPRC6A 3 GSG1L2 1 IFNK 5 IL22RA2 8 INSL4 X X 7 INSL6 X X X 10 KCNU1 X 3 LIN28A X X X 5 LUZP4 X X 2 MAGEA1 X X X X X 24 MAGEA10 X X X X 14 MAGEA11 X X X X 15 MAGEA3 X X X X 24 MAGEA4 X X X X X 18 MAGEA6 X X X 23 MAGEA9 X 6 MAGEB1 X X X 10 MAGEB2 X X X X 21 MAGEB4 1 MAGEB6 1 MAGEC1 X X X X 17 MAGEC2 X X X X 22 MORC1 X 2 PAGE1 X X X X 16 PAGE5 X X X X 19 PASD1 X X X X 10 PRDM7 X 1 PROKR1 X 2 R3HDML 2 SLC7A13 3 SMC1B X X 15 SMR3A X X 4 SSX1 X X X X X 15 STRA8 X X 12 TFDP3 X 3 XAGE3 X X X 7 XAGES X X 5 ZFP42 X X X X 14 ZNF560 X X X X X 17

[0552] To identify peptides that are likely to be presented as cell surface antigens by MHC Class I proteins, we used a sliding window to parse each of these proteins into its constituent 8-11 amino acid sequences. We processed these peptides and their flanking sequences with the HLA peptide presentation deep learning model to calculate the probability of presentation of each peptide at the 99.9.sup.th percentile expression level observed for this gene in TCGA. We considered a peptide likely to be presented (i.e., a candidate target) if its quantile normalized probability of presentation calculated by our model was greater than 0.001.

[0553] To prioritize genes that are likely to be relevant to a given indication, we selected genes where the gene is expressed at a level of at least 10 TPM in at least 0.98% of cancer cases.

[0554] The results are shown in Table 1.2. A total of 10698 shared antigen sequences were identified. The corresponding HLA allele(s) for each sequence are shown.

XXVI. Heterologous Prime/Boost Stimulates CD8.SUP.+ T-Cell Responses in Human Subjects

[0555] An open-label, multi-center, multi-dose Phase 1/2 study was performed to assess the dose, safety and tolerability, immunogenicity, and early clinical activity of a heterologous prime/boost vaccination strategy. Two vaccine

programs, GRANITE and SLATE, were assessed.

[0556] A personalized neoantigen cancer vaccine ("GRANITE") was administered in combination with immune checkpoint blockade in patients with advanced cancer. The GRANITE heterologous prime/boost vaccine regimen included (1) a ChAdV that is used as a prime vaccination [GRT-C901] and (2) a SAM formulated in a LNP that is used for boost vaccinations [GRT-R902] following GRT-C901. The ChAdV vector is based on a modified ChAdV68 sequence having the sequence of SEQ ID NO: 1 with an E1 (nt 577 to 3403) deletion and an E3 (nt 27,125-31,825) deletion. The SAM vector is based on an RNA alphavirus backbone having the nucleic acid sequence set forth in SEQ ID NO:6. Both GRT-C901 and GRT-R902 expressed the same 20 personalized neoantigens as well as two universal CD4 T-cell epitopes (PADRE and Tetanus Toxoid). Tumors were used for whole-exome and transcriptome sequencing to detect somatic mutations, and blood was used for HLA typing and detection/subtraction of germline exome variants to generate the personalized neoantigen cassette using the EDGE algorithm for 10 subjects (Patients 1-11, referred to herein as patients G1-G11). The neoantigens included in subject G1-G10's cassette are shown in Table 47A and the full-length cassettes for subjects G1-G11 are shown in Table 47B. The determined HLAs of GRANITE subjects are shown in Table 48.

TABLE-US-00010 TABLE 47A GRANITE Subject-Specific Neoantigens

Ag Patient G1 Patient G2 Patient G3 Patient G4

1 QLQAEQKGLTEVIQ YPLAAHTYTPITGS ERYHFLEERLARLG KRYRVRSYEAVAV SLKMENEEFKK VSTIRQYPVSA LTEVYQWYLDL AAAASAYNYAEQ

2 GLFEGDEYATLMM RYYDRVTGVSSFH KQGKDLTVLDHTL EDPYLEDSEASLVP QCKEGAQKEGLM VTWVSQASADQR PTALTSPELSSW LAYQPIPESQS

3 APFPAWSAFGKEE VKELRRLEDQLASL MASMGTLAFHEYG ERAQGELFPDRASS HDEALKNTWNLH QQELAALALKQ RPFLIIKDQDRK SWGNAAAAAGD

4 LAEQVQKADEFIQ SSLDAVLSQDLILA NWVQQQAYGMLA DVKKLITDEFVKH ANATNKLTVIAE DGPAVEVGDSL GIPVTICTYPFVF KYLDYARVPNSN

5 EEQKLPEADVAVL RRMNRVGNESLNP CSPPSPSALIQEMEE HLLQSKSGTQPHTA RTFLRQQASILS AVAKAAETIIDT RLWEKEQEVA TRLSLQQPRAE

6 VSQKQIADRAAHN LYSLSVLYKGEAR VKGGYQESSLLRV KIPASQHTQNWSTT ESPGNNVVDISV VVLLKAAYDVSS LVTLNDLEPKAA WTKDSKRRDKR

7 LNTAQDELVTFSDE IMEIQLNGGTVAKK PNYADTLISQESFE PKPFRPILGLHLNL LAQLYHHVCLC VAWAQARLEKQ KKGFLSAPQSL GILYYVYMGLL

8 HNMKLSISYLREYA GFAAEFKLVDQSRP HSSFVAVTEEGTKA AILSFGELIFCGMPA KANDWLQFIIH PHEPKFVYQAK AAATGIGFTVT EMLDFFNDCG

9 GPKVDIDVPDVNFE DQAVGAYLPLTTA AVKLLIERATNQG ANMCRCLVTLKVL APEGKLKGPKF RYSTLLAGNLRI WMLRVKWLYHLA MADKPGNTVNFR

10 NEDVAGYIVGANT KRQELESENKKLN FFGMHVQEYGSDW SLQHQLPQPGAQH ETYLLEKSRAIR NDLNELRKAVAD PPPNTRRVYISY FATRGYPMEDMK

11 DTCVCSTVAEFSHQ DVTRQADCKMKL DTRIAKRMAMLIY TEPGDYNINILFTDT CSHAGGRPGNW ADFVKYYYSGKRE TDFLCMAPISFF HIPGSPFKAH

12 QGDGCSGHKNQNQ LRNDAKNLYAANS VRASSDGEGTMSL YKCLSCTKTFPNVP IDTLQEVRLRFL IGA VLAHKGYFR PASVDGSPVSPS RAARHAATHGP

13 PMNTIAEAVIEMVN DMFDCVFPTRTALF SKELVEKSEAVRK PLKIFAQDGEGQHI RGQIQITINGF GSALVPTGNLQ MSSSGKLWKPRQ DIQMKNRMDGT

14 GKISRHEMLQVLH MSNLRQVGLKKPM MDARRVPPKDLRV VYCDFSTGETCIQA LMVGVQVTEEQL ERSSVLDRYPPA KKNLKKFRYVKL QPENIPAKNWY

15 TDAYLLYTPSQIVL QLMPLKTFPAAIW VKKNRQNGMHTL NSLATVTMEDLIQP TAILSSASRAG GVIQSELNYSVI YLLDIKVKEQSLE WFPEFSEARAI

16 GIGWLITFCSKGED SEVSKAIINIFFFYG LFIYQSLDAIDGRQ KVHLAKSSGSALM CLTSQTRLRLS PPQKIIMDQR ARRTNSSSPLG VWGTGNPRRQFI

17 GQLSQDRDFIMTLN LQMTDIVQACHDSI YLSLLRMYLSPPCI SFVIPFIHLPEIINLY TSLHRSWWMEN KAALDISIKSD HCLGPIKLELL NLSEQNDVF

18 SEEAEVDVREREIQ LEAFLTQKAKVGK SLSLAEKSILHEECI SYQPAVTAPEPDSS RDREPKRARDL LKDDDFERISEL KSPVVETVPA PRLAVDFALPK

19 QGRVKDLESLFHQ KLPGLNTATILLMG PPDPLEPPSLPAQRP VFYIEACESGSMVN SEVELAAALSDK TEDALLQQLAD GPPTPAAAHS HLPDNINVYAT

20 CLAFKNDATEILHS SNLKVLNHSPMSG GDGGLVCPYMEFL RDSVKSCRVHVPG HVVKPVPAHPS ASVNFDYKSPSP KNENNELPKLQW KTLTCGKGRRDF

Ag Patient G5 Patient G6 Patient G7 Patient G8

1 RIEVMALPKPGGTH EQTEAAAREEILTN QIGALILATDISCON GSRKERQVYSKAL SLALVTVPSMG GGSLSHHHGVG EYLSLFRSHL NRLFGVEASGRR

2 VDPEDQSQVTLAS FQYNYEYLGNSGQ HMSSMEHTEEGLQ PYHRSQSSSSVLNN YGFAFRYCPSGK LVITPLTDRCYM ERLADAMAESPS KSMDSINYPSD

3 STLYGRVLAPLPER MDKDGYFWFVAR REKVDNVWTEESS NANFDLRTSQSVR AGGAASGGGGN SDDVILSSGYRIG NQFLQELETGQR AMLAYNMRDQHL

4 LRKRHRKHTQQRV AVRTPYTMGYGEP LRDVDAKALVRSH EVLLLSMRDMNIT AALSLSTLASPK MYYNDAYGALDY FLLVYGDVISNI KLTSVDAPLFNA

5 SARIPEQLDVTAHG MTEYKLVVVGAV IIAKLLADELISLKD AQPHVNGGPLYSK VYAPEDVYRFR GVGKSALTIQLIQ IDPVQRYVPL NSISIAPPPPPM

6 GTGGSGPGLSRTRA IEMWISILELNELEY VAASEKSTGKIFLY AKMKWTNLLYKA LSTGALPPLAP AAVELHQAKD DGRGDNQPLHI GGVKIRDDERLLQ

7 VEEDESAMPKKDS LLLAAHRFLATELS MFEVCFESKGTGPI EQSAERCVSTLLNL RTLLARFNEQIE QFSPSLISEKI PDQLVILDMKH IQTKVNYVVQE

8 GASAPGKSMEASM ELKHMVMSFRVSA QDVCENFWMVWS SYQHKFQDDDQTH DVSAPKVEADVS LQVLLGFAGRNK DAMAHGCFLNAKR VKGSLKAGFFGT

9 MNCREVTFVPGLC SGSKKKGPPPDADSP IIGVTLFSSTKEDFQ

EHKVLLTGTPLQKT KIFDEIL   VNAAD LYLPYKTLVST QYAAALQAGM VEELFSLLHFL 10
PTMASVPALQLTLP STLATIQALASGCS EALRSQAALGRVL PGHDKMLSPNFDV THHPRRCPIPK
LPITSLDATGN LDCGASCAADPG HHTAMLTRGEQH 11 KLGTRLVPAERRK NKQWDQHFRSMK
DSQFLAPDVTSTPV PPAVISQMLLLDSP KLKTSRDKLRKS HQYEQKITELRQK NTVVSGALDRL
HKEPIRLRYKL 12 SSDQTNDSEGASRP LLSSRSPLAEHRPD TSSTSTPSSTPGMT LSDQENAPPILPRRP
APFEMPSSVSE AYQDVSMVLLP WILTELTTAAT SEGLGPSPHL 13 IEDLLPASYFSTILL FIQRDYSSGTRCLF
ASLLTSGLELYCKK PKVRERHFSESTYI GVQTDQRVLR QTKFPAELENR GLSMTVEADPA DNALSRLTLGN
14 SKPDLTAALRDIW KPKEDKLRLAQEV CAMLRRLEKRILID PPACSSSSSSLFLAV AQYETIAAKNIS
SDMVIYCKSVHF LPSREARQAMI VARRLGRGSV 15 EIA   VLELAKSCPHV HYNYMCNSSCMGS
MTEYKLVVVGAGR GIALVMRVDGAVS INLHEVYENTS MNRRPILTIITL VGKSALTIQLIQ SCFRQLFLAQQR
16 MPAQRLLLLLTLLL EKKKPHIKKPLNVF ELYQRMLKGPPPP LHKSLKLYQVIFKG PGLGIFGSTST
MLYMKEMRAKV AARARSQRRPRH EIGKGNLGGIA 17 LSTMERSELAKKT KQQPQDNFKNNVQ
SDNSGCDAPGNSY APRLGSYSGTTILH GTTPDIILDDLL KSQLPVQLDLGG PSLSVPSSAESE LLNSTSNNLYL
18 MSGREGGKRKPLK KVTQTELMRESFT EHRKDLGLTTVWK APAPPARTSRGQVS QPKKQAKEMDEE
QKQEATESLKCQ DQLSYLLSPALA ERANEAGGQVG 19 LRQITVNDLPVGHS MAAADGALLEAA
NCPTKKYMPAVTL MTEYKLVVVGAD VDEALRLVQAF ALEQPAELPASVR TPTVNQHETSTS
GVGKSALTIQLIQ 20 SLHEICSKIMAEWK DEMVNTICDVLQK ELYQRMLKGPPPPP KDLAKRLLVGKSD
NAKALACSSLQ PEQFPLVQGVAI PEPAASAAQGT SVDAEKSMLSKL Ag Patient   G9 Patient   G10 1
KVAKPKKAAKSAT TSGSQGQSTQSSGYS KAVKPKAAKPKV SSYGYPPSSL 2 ECTDGSSFVDEVEK
ISHLIEPLANAAQAEA VMKCGCTRCVS SQLGHKVSQ 3 SVLSPPPEKAGATA PQTLLDQKVKVVNV
ATLLPHEVAPL ARNAKDVAVSY 4 RCRQGATRLPLPLL YCKFKNTEDITFSSV LRDYLLLRVEG
YVGLKDKLSG 5 EAHVAKGKSVFLH VDSYEDEWGRLHES QMKKFVEWLQNA LWDSAGVLREG 6
ALKEKYDSAAAMG AGAQRVGLPGPPAPP LNSRGKLVQVNS GPPGKPGQDG 7 AKLLGLLAELRSLN
ADTERFYTRESTQFL EAYGYQIQHIQ QQNPVTEYMK 8 FEHLKWNDMMEE KESRPPRKFPSDIIFE
VYQTLFLLQHLWS AISSMFPDK 9 LLLNLAENPAMTSE LQFSQIYPVVCVLSSF LLRAQVPSSLG FIGDSGIPL
10 RQLFLSENRRKEIL EALLHPFFAGLTLEE QRIIVELVEFI RSFHTSRNPS 11 ISLLKLTQGETLCK
VDQSPKPLIIGPVEDY LYEQHHVVQDM DPGYFNNES 12 EEHLSIEDFTQASG STPCWIELHLNGLLQ
MTPAAFSALPR WLDKVLTQMG 13 ELDSLITAITTNRAH QSLESSSLEGSHMGV PSKCVTIQRT
YFSAHWCPPC 14 EESREYTEDGQVT GEVLSRRCVNLLNTA KETRYSYNTEWR LRPDMWPKSE 15
TNITTETSKILLRKN HLCIVDDWICEEVIT MQIANHTLKY GTDALLRRML 16 PPARKKFVIPLDVD
GGFGLALGTHANCL EVPPGVAKPLF DSPPMFAGAGL 17 TLVQNLVNNGYV LRGNHESRQITQFYG
WDETVRAAPYDWR FYDECQTKYG 18 NLHGDGIALWYTL NEQACEDMDILKLES DRLVPGPVFGSK
YGTVVRISPQ 19 YMCNSSCMGGMN PRSAPSLRPKDYDVD QRPILTIITLEDS ATLKSLNNQI 20
LSTSRPPQHLGGLK QTNHCYIAILNISQGE KPTYDPVSEDQ VDPTQVHKS
TABLE-US-00011 TABLE   47B GRANITE   Subject-Specific   Neoantigen   Cassettes Patient Cassette   Sequence
G1 LAEQVQKADEFIQANATNKLTVIAEQGDGCSGHKNQNQIDTLQEVRLRFLGQLSQDR
DFIMTLNTSLHRSWWMENDTCVCSTVAEFSHQCSHAGGRPGNWPMNTIAEAVIEMV
NRGQIQITINGFVSQKQIADRAAHNESPGNNVVDISVSEEAEVDVREREIQRDREPKR
ARDLQLQAEQKGLTEVIQSLKMENEEFKKQGRVKDLESLFHQSEVELAAALSDKGKI
SRHEMLQVLHLMVGVQVTEEQLGLFEGDEYATLMMQCKEGAQKEGLMCLAFKND
ATEILHSHVVKPVPAHPSGIGWLITFCSKGEDCLTSQTRLRLSHNMKLSISYLREYAKA
NDWLQFIIHLNTAQDELVTFSDELAQLYHHVCLCAPFPAWSAFGKEEHDEALKNTW
NLHTDAYLLYTPSQIVLTAILSSASRAGGPKVDIDVPDVNFEAPEGKLKGPKFNFDVA
GYIVGANTETYLLEKSRAIREEQKLPEADVAVLRTFLRQQASILS G2
LEAFLTQKAKVGKLKDDDFERISELYPLAAHTYTPITGSVSTIRQYPVSAIMEIQLNGG
TVAKKVAWAQARLEKQKLPGLNTATILLMGTEDALLQQLADGFAAEFKLVDQSRPP
HEPKFVYQAKMSNLRQVGLKKPMERSSVLDRYPPAQLMPLKTFPAAIWGVIQSELN
YSVILYSLSVLYKGEARVVLLKAAYDVSSDMFDCVFPTRTALFGSALVPTGNLQSNL
KVLNHSPMSGASVNFDYKSPSPDQAVGAYLPLTTARYSTLLAGNLRIRYYDRVTGVS
SFHVTWVSQASADQRDVTRQADCKMKLADFVKYYYSGKRESEVSKAIINIFFFYGPP
QKIIMDQRLRNDAKNLYAANSIGAVLAHKGYFRKRQELESENKKLNNDLNELRKAV
ADLQMTDIVQACHDSIKAALDISIKSDSSLDAVLSQDLILADGPAVEVGDSLVKELRR
LEDQLASLQQELAALALKQRRMNRVGNESLNPAVAKAAETIIDT G3
SKELVEKSEAVRKMSSSGKLWKPRQMASMGTLAFHEYGRPFLIIKDQDRKDTRIAKR
MAMLIYTDFLCMAPISFFYLSLLRMYLSPPCIHCLGPIKLELLAVKLLIERATNQGWM
LRVKWLYHLAVRASSDGEGTMSLPASVDGSPVSPSCSPPSPSALIQEMEERLWEKEQ
EVAERYHFLEERLARLGLTEVYQWYLDLFFGMHVQEYGSDWPPPNTRRVYISYLFIY
QSLDAIDGRQARRTNSSSPLGNWVQQQAYGMLAGIPVTICTYPFVFMDARRVPPKDL
RVKKNLKKFRYVKLVKKNRQNGMHTLYLLDIKVKEQSLEPNYADTLISQESFEKKGF
LSAPQSLGDGGLVCPYMEFLKNENNELPKLQWSLSLAEKSILHEECIKSPVVETVPAP

PDPLEPPSLPAQRPGPPTPAAAHSVKGGYQESSLLRVLVTLNDLEPKAAKQGKDLTVL
DHTLPTALTSPELSSWHSSFVAVTEEGTKAAAATGIGFTVT G4
SYQPAVTAPEPDSSPRLAVDFALPKEDPYLEDSEASLVPLAYQPIPFSQSPLKIFAQDG
EGQHIDIQMKNRMDGTTEPGDYNINILFTDTHIPGSPFKAHNSLATVTMEDLIQPWFP
EFSEARAIKRYRVRSYEAVAVAAAASAYNYAEQANMCRCLVTLKVLMADKPGNTV
NFRHLLQSKSGTQPHTATRLSLQQPRAESLQHQLPQPGAQHFATRGYPMEDMKDVK
KLITDEFVKHKYLDYARVPNSNSFVIPFIHLPEIINLYNLSEQNDVFRDSVKSCRVHVP
GKTLTCGKGRRDFYKCLSCTKTFPNVPRAARHAATHGPERAQGELFPDRASSSWGN
AAAAAGDKIPASQHTQNWSTTWTKDSKRRDKRKVHLAKSSGSALMVWGTGNPRRQ
FIVYCDFSTGETCIQAQPENIPAKNWYAILSFGELIFCGMPAEMLDFFNDCGPKPFRPIL
GLHLNLGILYYVYMGLLVFYIEACESGSMVNHLPDNINVYAT G5
STLYGRVLAPLPERAGGAASGGGGNSKPDLTAALRDIWAQYETIAAKNISGTGGSGP
GLSRTRALSTGALPPLAPIEDLLPASYFSTILLGVQTDQRVLRRIEVMALPKPGGTHSL
ALVTVPSMGVEEDESAMPKKDSRTLLARFNEQIEPTMASVPALQLTLPTHHPRRCPIP
KLRKRHRKHTQQRVAALSLSTLASPKVDPEDQSQVTLASYGFAFRYCPSGKMNCRE
VTFVPGLCKIFDEILVNAADSLHEICSKIMAEWKNAKALACSSLQMSGREGGKRKPL
KQPKKQAKEMDEEEIAVLELAKSCPHVINLHEVYENTSGASAPGKSMEASMDVSAP
KVEADVSSSDQTNDSEGASRPAPFEMPSSVSESARIPEQLDVTAHGVYAPEDVYRFRK
LGTRLVPAERRKKLKTSRDKLRKSLRQITVNDLPVGHSVDEALRLVQAFLSTMERSE
LAKKTGTTPDIILDDLLMPAQRLLLLLTLLLPGLGIFGSTST G6
LLSSRSPLAEHRPDAYQDVSMVLLPMAAADGALLEAAALEQPAELPASVRHYNYMC
NSSCMGSMNRRPILTIITLFQYNYEYLGNSGQLVITPLTDRCYMFIQRDYSSGTRCLFQ
TKFPAELENRAVRTPYTMGYGEPMYYNDAYGALDYLLLAAHRFLATELSQFSPSLIS
EKIEQTEAAAREEILTNGGSLSHHHGVGSTLATIQALASGCSLPITSLDATGNKQQPQD
NFKNNVQKSQLPVQLDLGGMDKDGYFWFVARSDDVILSSGYRIGDEMVNTICDVLQ
KPEQFPLVQGVAIMTEYKLVVVGAVGVGKSALTIQLIQIEMWISILELNELEYAAVEL
HQAKDNKQWDQHFRSMKHQYEQKITELRQKKPKEDKLRLAQEVSDMVIYCKSVHF
ELKHMVMSFRVSALQVLLGFAGRNKKVTQTELMRESFTQKQEATESLKCQEKKKPH
IKKPLNVFMLYMKEMRAKVSGSKKKGPPDADSPLYLPYKTLVST G7
EALRSQAALGRVLLDCGASCAADPGTSSTSTPSSTPGMTWILTELTTAATQDVCENF
WMVWSDAMAHGCFLNAKRMTEYKLVVVGAGRVGKSALTIQLIQQIGALILATDISC
QNEYLSLFRSHLELYQRMLKGPPPPPPEPAASAAQGTIIAKLLADELISLKDIDPVQRY
VPLASLLTSGLELYCKKGLSMTVEADPANCPTKKYMPAVTLTPTVNQHETSTSREKV
DNVWTEESSNQFLQELETGQRDSQFLAPDVTSTPVNTVVSGALDRLHMSSMEHTEEG
LQERLADAMAESPSSDNSGCDAPGNSYPSLSVPSSAESEVAASEKSTGKIFLYDGRGD
NQPLHIELYQRMLKGPPPPAARARSQRRPRHIIGVTLFSSTKEDFQQYAAALQAGMC
AMLRRLEKRILIDLPSREARQAMIMFEVCFESKGTGPIPDQLVILDMKHLRDVDAKAL
VRSHFLLVYGDVISNIEHRKDLGLTTVWKDQLSYLLSPALA G8
APRLGSYSGTTILHLLNSTSNNLYLPPACSSSSSSLFLAVVARRLGRGSVSYQHKFQDD
DQTHVKGSLKAGFFGTPPAVISQMLLLDSPHKEPIRLRYKLPKVRERHFSESTYIDNAL
SRLTLGNEHKVLLTGTPLQKTVEELFSLLHFLPYHRSQSSSSVLNNKSMDSINYPSDLS
DQENAPPILPRRPSEGLGPSPHLEVLLLSMRDMNITKLTSVDAPLFNAPGHDKMLSPN
FDVHHTAMLTRGEQHLHKSLKLYQVIFKGEIGKGNLGGIAAQPHVNGGPLYSKNSISI
APPPPPMAPAPPARTSRGQVSERANEAGGQVGAKMKWTNLLYKAGGVKIRDDERLL
QNANFDLRTSQSVRAMLAYNMRDQHLMTEYKLVVVGADGVGKSALTIQLIQGIALV
MRVDGAVSSCFRQLFLAQQRGSRKERQVYSKALNRLFGVEASGRRKDLAKRLLVGK
SDSVDAEKSMLSKLEQSAERCVSTLLNLIQTKVNYVVQE G9
TNITTETSKILLRKNMQIANHTLKYSVLSPPPEKAGATAATLLPHEVAPLRCRQGATR
LPLPLLLRDYLLLRVEGELDSLITAITTNRAHPSKCVTIQRTYMCNSSCMGGMNQRPIL
TIITLEDSECTDGSSFVDEVEKVMKCGCTRCVSLLLNLAENPAMTSELLRAQVPSSLG
TLVQNLVNNGYVWDETVRAAPYDWRKVAKPKKAAKSATKAVKPKAAKPKVEAHV
AKGKSVFLHQMKKFVEWLQNAAKLLGLLAELRSLNEAYGYQIQHIQFEHLKWNDM
MEEVYQTLFLLQHLWSNLHGDGIALWYTLDRLVPGPVFGSKISLLKLTQGETLCKLY
EQHHVVQDMLSTSRPPQHLGGLKKPTYDPVSEDQEESREYTEDGQVTKETRYSYNTE
WRPPARKKFVIPLDVDEVPPGVAKPLFEEHLSIEDFTQASGMTPAAFSALPRRQLFLSE
NRRKEILQRIIVELVEFIALKEKYDSAAAMGLNSRGKLVQVNS G10
QSLESSSLEGSHMGVYFSAHWCPPCGGFGLALGTHANCLDSPPMFAGAGLKESRPPR
KFPSDIIFEAISSMFPDKVDQSPKPLIIGPVEDYDPGYFNNESTSGSQGQSTQSSGYSSSY
GYPPSSLSTPCWIELHLNGLLQWLDKVLTQMGLRGNHESRQITQFYGFYDECQTKYG
YCKFKNTEDITFSSVYVGLKDKLSGGEVLSRRCVNLLNTALRPDMWPKSEEALLHPF

FAGLTLEERSFHTSRNPSQTNHCYIAILNISQGEVDPTQVHKSAGAQRVGLPGPPAPPG
PPGKPGQDGVDSYEDEWGRLHESLWDSAGVLREGHLCIVDDWICEEVITGTDALLRR
MLLQFSQIYPVVCVLSSFFIGDSGIPLADTERFYTRESTQFLQQNPVTEYMKISHLIEPL
ANAAQAEASQLGHKVSQPQTLLDQKVKVVNVARNAKDVAVSYNEQACEDMDILKL
ESYGTVVRISPQPRSAPSLRPKDYDVDATLKSLNNQI G11
AAGPSCALKAGKTASGAGEGGALSVDSESHWVAPEHTKRSVPQDPDFAEVGTYWPV
NSDQAVLELDLGWPLTFGFVQVLPDADTLLHLPRKVLQMDFLVHNSPRSTPGFILDT
TSNKQTPQKKQLREHQLRLMSARACYERYSGNQVLFCMQSALAKERRELRQAQREA
EMDSIPIKSDLVNEEATGTVPCIPGAAQALHKNCPHIVVGTPGHILALARNKSLNLAA
QQGEPHPEFSVDSPSPGLNMELEVLLIVLVMMSLVLGKFVPVNWEPPQADSALRMPS
TGSIAPCVPPLSTSLLFRFNFLAPELPAFSEFSTSETMGHSTVTSAQKSSQIGNSQLLKR
HVQRTECQNSQSRKEQVNTELKAVPAQEVKMNEPKLSRPTAAIVLKGEKGDRGPKG
YEWQEVEKGSLPRVLLQTPAAPQRCTTIHYNYMCNSSYMGGMNRRPILTICKECGKA
FAVSSHLSRHERIHTGEKIASGSELSESAFSYFITACA   VIILT

TABLE-US-00012 TABLE 48A GRANITE Subject HLAs

Patient	HLA-A	HLA-A	HLA-B	HLA-B	HLA-C	HLA-C
G1	A*02:01	A*32:01	B*18:01	B*44:05	C*02:02	C*07:04
G2	A*30:04	A*02:01	B*08:01	B*57:01	C*07:01	C*06:02
G3	A*02:01	A*11:01	B*27:02	B*35:01	C*04:01	C*02:02
G4	A*25:01	A*03:01	B*15:17	B*18:01	C*07:01	C*07:01
G5	A*01:01	A*03:01	B*07:02	B*08:01	C*07:01	C*07:02
G6	A*01:01	A*31:01	B*37:01	B*40:01	C*03:04	C*06:02
G7	A*01:01	A*02:01	B*08:01	B*40:01	C*03:04	C*07:01
G8	A*03:01	A*33:01	B*42:01	B*53:01	C*04:01	C*17:01
G9	A*02:01	A*03:01	B*35:12	B*45:01	C*04:01	C*06:02
G10	A*01:01	A*02:01	B*08:01	B*15:01	C*04:01	C*07:01
G11	A*01:01	A*02:01	B*07:02	B*15:01	C*03:03	C*07:02

[0557] A shared neoantigen cancer vaccine ("SLATE") was administered in combination with immune checkpoint blockade in patients with advanced cancer. The SLATE heterologous prime/boost vaccine regimen included (1) a ChAdV that is used as a prime vaccination [GRT-C903] and (2) a SAM formulated in a LNP that is used for boost vaccinations [GRT-R904] following GRT-C903. Both GRT-C903 and GRT-R904 expressed the same 20 shared neoantigens derived from a specific list of oncogenic mutations (see Table 34) as well as two universal CD4 T-cell epitopes (PADRE and Tetanus Toxoid). For subject inclusion, tumors were used for whole-exome and transcriptome sequencing to detect somatic mutations, and blood was used for HLA typing. Enrolled SLATE subjects were determined to have HLA A02:01 and KRAS mutation G12C predicted to be presented by HLA A02:01 (Patients S1, S2, and S3), HLA A01:01 and KRAS mutation Q61H predicted to be presented by HLA A01:01 (Patients S4 and S7), or HLA A03:01 or A11:01 and KRAS mutation G12V predicted to be presented by HLA A03:01 or A11:01 (A03:01 for Patient S9; A11:01 for Patients S11 and S15).

[0558] Both treatment studies (ie, the GRANITE and SLATE vaccine regimens) administered the vaccine via IM injection bilaterally (eg, in each deltoid muscle) in combination with immune checkpoint blockade, specifically SC ipilimumab and IV nivolumab. The studies followed two sequential phases.

[0559] GRT-C901 and GRT-C903 are replication-defective, E1 and E3 deleted adenoviral vectors based on chimpanzee adenovirus 68. The vector contained an expression cassette encoding 20 neoantigens as well as two universal CD4 T-cell epitopes (PADRE and Tetanus Toxoid). GRT-C901 and GRT-C903 were formulated in solution at 5×10.sup.11 vp/mL and 1.0 mL was injected IM at each of 2 bilateral vaccine injection sites in opposing deltoid muscles. The GRT-C901 and GRT-C903 vectors differ only by the encoded neoantigens within the cassette.

[0560] GRT-R902 and GRT-R904 are SAM vectors derived from an alphavirus. The GRT-R902 and GRT-R904 vectors encoded the viral proteins and the 5′ and 3′ RNA sequences required for RNA amplification but encoded no structural proteins. The SAM vectors were formulated in LNPs that included 4 lipids: an ionizable amino lipid, a phosphatidylcholine, cholesterol, and a PEG-based coat lipid to encapsulate the SAM and form LNPs. The GRT-R902 vector contained the same neoantigen expression cassette as used in GRT-C901 for each patient, respectively. The GRT-R904 vector contained the same neoantigen expression cassette as used in GRT-C903. GRT-R902 and GRT-R904 were formulated in solution at 1 mg/mL and was injected IM at each of 2 bilateral vaccine injection sites in opposing deltoid muscles (deltoid muscle preferred, gluteus [dorso or ventro] or rectus femoris on each side may be used). The boost vaccination sites were as close to the prime vaccination site as possible. The injection volume was based on the dose to be administered. The dose level amount refers explicitly to the amount of the SAM vector, i.e., it does not refer to other components, such as the LNP. The ratio of LNP:SAM was approximately 24:1. Accordingly, the dose of LNP was 720 µg, 2400 µg, and 7200 µg for each respective GRT-R902/GRT-R904 dose level (see below).

[0561] Ipilimumab is a human monoclonal IgG1 antibody (see SEQ ID NO: 29520 and 29521) that binds to the cytotoxic T-lymphocyte associated antigen 4 (CTLA-4). Ipilimumab was formulated in solution at 5 mg/mL and was injected SC proximally (within ˜2 cm) to each of the bilateral vaccination sites. Ipilimumab was administered at a dose of 30 mg of antibody in four 1.5 mL (7.5 mg) injections proximal to the vaccine draining LN at each of the bilateral vaccination sites (ie, 1.5 mL below the vaccination site and 1.5 mL above the vaccination site on each bilateral side in each deltoid, ventrogluteal, dorsogluteal, or rectus femoris [deltoid preferred, but dependent on clinical site and patient preference])

[0562] Nivolumab is a human monoclonal IgG4 antibody (see SEQ ID NO: 29522 and 29523) that blocks the

interaction of PD-1 and its ligands, PD-L1 and PD-L2. Nivolumab was formulated in solution at 10 mg/mL and was administered as an IV infusion (480 mg) through a 0.2-micron to 1.2-micron pore size, low-protein binding in-line filter at the protocol-specified doses. It was not administered as an IV push or bolus injection. Nivolumab infusion was promptly followed by a flush of diluent to clear the line. Nivolumab was administered following each vaccination (i.e., each of GRT-C901, GRT-R902, GRT-C903, or GRT-R904) with or without ipilimumab on the same day. The dose and route of nivolumab was based on the Food and Drug Administration approved dose and route.

[0563] Initial evaluations of the subjects G1-G10 international are described in international patent application publication WO/2021/092095, herein incorporated by reference for all purposes. Further evaluations were performed, such as additional time points for subjects G1-G10 and additional, e.g., subject G11.

[0564] Serum IFNa2a levels assessed by ELISA at 24 hours following the first samRNA dose. As shown in FIG. **2**, there was a dose-dependent induction of IFNα following samRNA immunization in human that increased as the dose of samRNA increased.

[0565] The CD8 T cell pools responsive to vaccination for GRANITE subjects were further assessed and characterized. PBMCs were collected and expanded for 2 weeks in the presence of peptide pools (In Vitro Stimulation (IVS) culture), and T-cell responses were assessed by IFN-gamma ELISpot for each possible vaccine-encoded peptide (40 total to cover different distinct epitopes for each of the neoantigens, i.e., two or more distinct epitopes from the same neoantigen that may both elicit and immune response to that neoantigen) and reported as spot-forming cells (SFC) per 106 splenocytes. Assessment of GRANTE subjects G1, G2, G10, and G11 is shown in FIG. **3**. Comparing subjects G1 and G2, who both received 30 µg doses, to subjects G10 and G11, who both received 300 µg doses, there was a clear increase in T cell response with a samRNA boost at a 30 µg dose that was not observed at the 300 µg dose. Further, long intervals (>2mo) between samRNA doses revealed highly durable responses. FIG. **4** shows further comparisons for subjects who received 30 µg, 100 µg, and/or 300 µg doses and demonstrating T cell response increased with lower doses of samRNA. Overall, the data suggest dose dependent induction of IFNα has an inhibitory impact of early innate immune activation on samRNA amplification or expression, which may result in blunted T cell response with higher doses of samRNA.

XXVII. Clinical Assessment of a Neoantigen Vaccine

[0566] An open-label, multi-center, multi-dose Phase 1/2 study is performed to assess the dose, safety and tolerability, immunogenicity, and early clinical activity of a heterologous prime/boost vaccination strategy. Two vaccine programs, GRANITE and SLATE, are assessed.

[0567] A personalized neoantigen cancer vaccine ("GRANITE") is administered in combination with immune checkpoint blockade in patients with advanced cancer strategies. The GRANITE heterologous prime/boost vaccine regimen involves (1) a ChAdV that is used as a primo vaccination [GRT-C901] and (2) a SAM formulated in a LNP that is used for boost vaccinations [GRT-R902] following GRT-C901. Both GRT-C901 and GRT-R902 express the same 20 personalized neoantigens as well as two universal CD4 T-cell epitopes (PADRE and Tetanus Toxoid). Tumors are used for whole-exome and transcriptome sequencing to detect somatic mutations, and blood is used for HLA typing and detection/subtraction of germline exome variants to generate the personalized neoantigen cassette using the EDGE algorithm.

[0568] A shared neoantigen cancer vaccine ("SLATE") is administered in combination with immune checkpoint blockade in patients with advanced cancer. The SLATE heterologous prime/boost vaccine regimen involves (1) a ChAdV that is used as a prime vaccination [GRT-C903] and (2) a SAM formulated in a LNP that is used for boost vaccinations [GRT-R904] following GRT-C903. Both GRT-C903 and GRT-R904 express the same 20 shared neoantigens derived from a specific list of oncogenic mutations (see Table 34) as well as two universal CD4 T-cell epitopes (PADRE and Tetanus Toxoid). For subject inclusion, tumors are used for whole-exome and transcriptome sequencing to detect somatic mutations, and blood is used for HLA typing.

[0569] GRT-C901 and GRT-C903 are replication-defective, E1 and E3 deleted adenoviral vectors based on chimpanzee adenovirus 68. The vectors contain an expression cassette encoding 20 neoantigens as well as two universal CD4 T-cell epitopes (PADRE and Tetanus Toxoid). GRT-C901 and GRT-C903 are formulated in solution at 5×10.sup.11 vp/mL and 1.0 mL is injected IM at each of 2 bilateral vaccine injection sites in opposing deltoid muscles (deltoid muscle preferred, gluteus [dorso or ventro] or rectus femoris on each side may be used). The GRT-C901 and GRT-C903 vectors differ only by the encoded neoantigens.

[0570] GRT-R902 and GRT-R904 are SAM vectors derived from an alphavirus. The GRT-R902 and GRT-R904 vectors encod the viral proteins and the 5′ and 3′ RNA sequences required for RNA amplification but encoded no structural proteins. The SAM vectors are formulated in LNPs composed of 4 lipids: an ionizable amino lipid, a phosphatidylcholine, cholesterol, and a PEG-based coat lipid to encapsulate the SAM and form LNPs. The GRT-R902 vector contains the same neoantigen expression cassette as used in GRT-C901 for each patient, respectively. The GRT-R904 vector contains the same neoantigen expression cassette as used in GRT-C903. GRT-R902 and GRT-R904 are formulated in solution at 1 mg/mL and was injected IM at each of 2 bilateral vaccine injection sites in opposing deltoid muscles (deltoid muscle preferred, gluteus [dorso or ventro] or rectus femoris on each side may be used). The boost vaccination sites is as close to the prime vaccination site as possible. The injection volume is based on the dose to be administered. The dose level amount refers explicitly to the amount of the SAM vector, i.e., it does not refer to other

components, such as the LNP. The ratio of LNP:SAM is approximately 24:1. Accordingly, the dose of LNP is 720 µg, 2400 µg, and 7200 µg for each respective GRT-R902/GRT-R904 dose level (see below).

[0571] Ipilimumab is a human monoclonal IgG1 antibody that binds to the cytotoxic T-lymphocyte associated antigen 4 (CTLA-4). Ipilimumab is formulated in solution at 5 mg/ml and is injected SC proximally (within ˜2 cm) to each of the bilateral vaccination sites. The SC route of ipilimumab is distinct from the approved IV route of administration. Ipilimumab is administered at a does of 30 mg in one of two methods listed below:

[0572] 1. Four 1.5 mL (7.5 mg) injections proximal to the vaccine draining LN at each of the bilateral vaccination sites (ie, 1.5 mL below the vaccination site and 1.5 mL above the vaccination site on each bilateral side in each deltoid, ventrogluteal, dorsogluteal, or rectus femoris [deltoid preferred, but dependent on clinical site and patient preference]).

[0573] 2. Six 1 mL (5 mg) injections proximal to the vaccine draining LN at each of the bilateral vaccination sites (ie, 1 mL below the vaccination site, 1 mL to the side of the vaccination site, and I mL above the vaccination site on each bilateral side in each deltoid, ventrogluteal, dorsogluteal, or rectus femoris [deltoid preferred, but dependent on clinical site and patient preference]).

[0574] Nivolumab is a human monoclonal IgG4 antibody that blocks the interaction of PD-1 and its ligands, PD-L1 and PD-L2. Nivolumab is formulated in solution at 10 mg/mL and is administered as an IV infusion through a 0.2-micron to 1.2-micron pore size, low-protein binding in-line filter at the protocol-specified doses. It is not administered as an IV push or bolus injection. When the dose is fixed (eg, 240 mg flat dose), nivolumab injection may be infused undiluted or diluted so as not to exceed a total infusion volume of 160 mL. Nivolumab infusion is promptly followed by a flush of diluent to clear the line. Nivolumab is administered following each vaccination (i.e., each of GRT-C901, GRT-R902, GRT-C903, or GRT-R904) with or without ipilimumab on the same day. The dose and route of nivolumab is based on the Food and Drug Administration approved dose and route. Doses of nivolumab may be interrupted, delayed, or discontinued depending on how well the participant tolerates the treatment. Dosing visits are not skipped, only delayed. Vaccination does not occur in the absence of nivolumab unless the Investigator and Sponsor believe treatment with GRT-R902/GRT-R904 in the absence of nivolumab is in the best interest of the patient.

[0575] Phase 1 of GRANITE examines prime vaccination with GRT-C901 followed by multiple dose levels of GRT-R902 with all patients receiving IV nivolumab. Eligible patients include those with advanced or metastatic NSCLC, GEA, mUC or CRC-MSS. Following the initial demonstration of the safety and tolerability of Dose Level 1 (30 µg), the dose of GRT-R902 is escalated to Dose Level 2 (100 µg). If this combination is safe and well-tolerated, patients are treated at Dose Level 3 which incorporates SC ipilimumab (30 mg) at the same dose of GRT-R902 as in Dose Level 2 (100 µg). If this combination is safe and well-tolerated, then the dose of GRT-R902 is escalated to Dose Level 4 (300 µg) with all patients continuing to receive SC ipilimumab (30 mg). Phase 2 involves tumor-specific expansion cohorts.

[0576] Phase 1 of SLATE examines prime vaccination with GRT-C903 followed by multiple dose levels of GRT-R904 with all patients receiving IV nivolumab. Eligible patients include those with advanced or metastatic MSS-CRC, NSCLC, PDA, and other mutation-positive solid tumors. Following the initial demonstration of the safety and tolerability of Dose Level 1 (30 µg), patients are treated at Dose Level 2 which incorporates SC ipilimumab (30 mg) at the same dose of GRT-R904 as in Dose Level 1 (30 µg). If this combination is safe and well-tolerated, then the dose of GRT-R904 is escalated to Dose Level 3 (100 µg) followed by Dose Level 4 (300 µg) with all patients continuing to receive SC ipilimumab (30 mg). Phase 2 involves tumor-specific expansion cohorts and a cohort of mutation-positive tumors outside of tumor types already represented by other expansion cohorts. Cohorts 1 and 4 enroll patients who have not experienced disease progression to routine therapy and Cohorts 2, 3, 5, and 6 enroll patients who experienced disease progression on, or after, routine therapy. Patients in Phase 2 receive SC ipilimumab at the dose determined to be well tolerated in Phase 1. Patients receive IV nivolumab at 480 mg Q4W throughout Phase 1 and 2.

GRANITE Protocol Stages

[0577] For GRANITE, the personalized nature of the vaccine requires a specialized manufacturing process and thus, the patient's participation in the study is conducted in 2 stages:

[0578] 1) Vaccine Production Stage [0579] a. Neoantigen prediction: NGS of a patient's tumor and blood followed by prediction of a patient's neoantigens using the EDGE machine learning model. [0580] b. Vaccine manufacturing: generating a patient-specific heterologous prime/boost vaccine that incorporates the identified neoantigens.

[0581] 2) Study Treatment Stage: delivering the vaccine regimen to the patient consisting of the heterologous prime/boost vaccine in combination with immune checkpoint blockade.

[0582] Neoantigen prediction and vaccine manufacturing occur while the patient is receiving routine therapy for metastatic disease. The Study Treatment Stage will begin when the vaccine is ready and the patient meets eligibility criteria for the Study Treatment Stage.

[0583] The neoantigen vaccine strategy is outlined in the following steps: [0584] Obtain formalin-fixed paraffin-embedded (FFPE) tumor specimen and blood from patient [0585] Sequence the tumor and normal DNA and tumor RNA to identify non-synonymous exome mutations [0586] Perform transcriptome analysis to determine expression level of mutated proteins/peptides [0587] Determine the patient's HLA type from normal DNA [0588] Predict which of the mutant peptides are likely to be presented by the patient's class I HLA alleles using the EDGE model [0589] Assemble a prioritized list of candidate neoantigens based on predicted HLA presentation (including removal of epitopes identical to self-proteins to prevent potential autoimmune reactions) [0590] Insert the top 20 patient-specific

predicted neoantigens into an expression cassette [0591] Generate two plasmids encoding the identical neoantigen expression cassette; one plasmid for GRT-C901 and one plasmid for GRT-R902 [0592] Utilize plasmids to generate the ChAdV68 and SAM vectors [0593] Formulate the SAM in LNP [0594] Confirm the potency, identity, and sterility of ChAdV and SAM-LNP [0595] Immunize the patient with the patient-specific, selected neoantigen vaccine, in combination with checkpoint blockade

GRANITE Vaccine Production Stage

[0596] The first step in manufacturing GRT-C901 and GRT-R902 is to determine the presence of neoantigens in a patient's tumor. Patients have available and sufficient tumor tissue and blood for genomic sequencing and HLA typing. For inclusion in the study, each patient has an available FFPE tumor specimen (FFPE block/slides/scrolls sufficient to yield approximately 40 microns tissue thickness and preferably from the most recent available biopsy) and blood. When equally recent FFPE tumor specimens are available from both primary tumor and distant metastases, the primary tumor specimen is preferred. The tumor is used for whole-exome and transcriptome sequencing to detect somatic mutations, and the blood for HLA typing and detection/subtraction of germline exome variants.

GRANITE Study Treatment Stage

[0597] When a patient's vaccine is ready for administration, patients receive study treatment according to the study Phase/cohort in which they are enrolled, as described above. Patients who experience disease progression on or after, or intolerance to, routine therapy while vaccine is being manufactured receive study treatment as their next line of therapy. For patients who have SD or better response to their routine therapy while the vaccine is being manufactured, study treatment is either administered concurrent with their routine therapy or these patients interrupt or stop routine therapy and receive only study treatment if deemed clinically appropriate (eg, in the setting of cumulative toxicity, patient preference/refusal to continue routine therapy).

[0598] The study uses a sequential approach in Phase I to assess the safety of the intended GRANITE vaccine regimen consisting of a heterologous prime/boost vaccine in combination with nivolumab and ipilimumab. Patients receive a prime vaccination with GRT-C901 followed by boost vaccinations with escalating doses of GRT-R902. All patients receive IV nivolumab. As part of the dose escalation of GRT-R902, SC ipilimumab is combined with the GRT-C901 prime vaccination and GRT-R902 boost vaccinations for patients treated in Dose Levels 3 and 4. Given the immunogenicity data observed in NHPs (data not shown), the addition of SC ipilimumab is expected to boost T-cell responses providing the best opportunity for a patient to generate the most robust T-cell response.

[0599] Phase 2 consists of tumor-specific expansion cohorts that each comprise two treatment settings for each tumor type; an "A" arm that enrolls patients who have SD or better response to routine therapy and a "B" arm that enrolls patients who experience disease progression on, or after, routine therapy (ie, post-progression). Patients in Phase 2 receive SC ipilimumab at the dose determined to be well tolerated in Phase 1. Patients receive IV nivolumab at 480 mg Q4W throughout Phase 1 and 2. In one GRANITE phase 2 arm, MSS-CRC cohorts, including following FOLFOX/FOLFIRI treatment, are assessed using the RP2D determined. In another GRANITE phase 2 arm, Gastro-esophageal cancer (GEA) cohorts, including following 2nd line chemotherapy, are assessed using the RP2D determined.

[0600] If patients progress prior to vaccine availability, the patient can begin treatment with nivolumab until the vaccine is available (ic, nivolumab bridging). Once the vaccine is available, the vaccine is administered upon the next treatment with nivolumab to align vaccine administration with nivolumab administration.

SLATE Protocol Stages

[0601] For SLATE, as shared tumor neoantigens are presented via specific HLA proteins on the surface of the tumor cells for recognition by CD8 T cells, the patient's HLA Class I alleles are identified to confirm the patient's tumor-specific neoantigen can be presented. Therefore, patient's participation in the study is conducted in 2 stages: [0602] 1) HLA Screening Stage [0603] 2) Study Treatment Stage

[0604] Given that ~33% of patients with a relevant mutation are expected to have the matching HLA allele, patients are HLA typed prior to being eligible for study treatment. Patients eligible for the HLA Screening Stage are briefly described below; these patients include those with advanced or metastatic: [0605] MSS-CRC who are currently receiving systemic chemotherapy OR who have experienced disease progression following systemic chemotherapy, but have not initiated a new line of therapy [0606] NSCLC who are currently receiving an anti-PD-(L) 1 antibody in combination with platinum-based chemotherapy OR who have experienced disease progression following treatment with an anti-PD-(L) 1 antibody in combination with platinum-based chemotherapy, but have not initiated a new line of therapy [0607] PDA who are currently receiving first-line systemic chemotherapy for metastatic disease OR who have experienced disease progression on 1 L systemic cytotoxic chemotherapy, but have not initiated a new line of therapy [0608] Other solid tumor histology where the patient's tumor has the specified mutation and has experienced disease progression on available therapies known to confer clinical benefit

[0609] Patients with a matching specified mutation and HLA allele are eligible for study treatment.

SLATE HLA Screening Stage

[0610] The HLA Screening Stage of the study identifies patients who may be eligible for study treatment. A blood sample is submitted to the Sponsor to perform HLA typing to determine whether a patient possesses an HLA allele that matches the specified oncogenic mutation in that patient's tumor. Patients who do not possess the appropriate HLA

alleles for the tumor-specific mutation are not be eligible to participate in the Study Treatment Stage.

[0611] Patients are assessed in the HLA Screening Stage if their tumors are known to have at least one mutation included in the expression cassette of GRT-C903/GRT-R904 based on a suitable molecular profiling assay performed according to institutional standards in a Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory by a validated assay (eg, institutional or commercially available gene panel) (see Table 34). The mutation is clearly identified in the report and subclonal mutations are not considered eligible. HLA typing is performed by a centralized third-party CLIA-certified laboratory designated by the Sponsor. The Sponsor confirms the mutation result based on the mutation test report provided by the site and determines whether the patient's HLA Class I proteins are adequate for antigen presentation of the patient's tumor shared neoantigen(s).

SLATE Study Treatment Stage

[0612] Patients that meet eligibility criteria are administered the vaccine prime and boosts in combination with ipilimumab and nivolumab to augment T-cell responses. Patients receive study treatment according to the study Phase/cohort in which they are enrolled once their eligibility for the Study Treatment Stage has been confirmed.

[0613] Study treatment are administered either as maintenance therapy or as a new line of therapy as briefly described below: [0614] MSS-CRC 1 L/2 L maintenance: patients who have not experienced PD with at least 16 weeks of routine 1 L or 2 L chemotherapy [0615] MSS-CRC 3 L: patients who have experienced PD on 2 L chemotherapy [0616] NSCLC 2 L: patients who have experienced PD following 1 L treatment with anti-PD-(L)1 in combination with platinum-based chemotherapy (or PD following 1 L treatment with anti-PD-(L) 1 alone if patient refused platinum-based chemotherapy) [0617] PDA 1 L maintenance: patients who have not experienced PD with at least 16 weeks of routine 1 L chemotherapy [0618] PDA 2 L: patients who have experienced PD on 1 L chemotherapy [0619] Mutation positive solid tumor: patients who have experienced PD with all available therapies known to confer benefit

[0620] Some patients may be eligible to receive the shared neoantigen vaccine in the maintenance setting. For patients with MSS-CRC who have received >16 weeks of 1 L or 2 L chemotherapy or with PDA who have received >16 weeks of 1 L chemotherapy, but have not experienced disease progression, study treatment may be administered concurrent with their routine therapy or these patients may interrupt or stop routine therapy and receive only study treatment if deemed clinically appropriate by the Investigator (eg, in the setting of cumulative toxicity, patient preference/refusal to continue routine therapy).

[0621] Phase I utilizes adaptive statistical methods to guide dose selection with the goal of determining the highest vaccine dose with a dose-limiting toxicity rate below 30%. This target toxicity rate takes into account the known safety profile of nivolumab and ipilimumab while optimizing the number of patients needed to establish a preliminary toxicity profile of the vaccine in combination with nivolumab and ipilimumab compared to other statistical approaches (eg, a 3+3 design). The dosing decisions in Phase 1 is based on a safety review of the treatment by the Study Committee (eg, to determine whether a DLT has occurred). An established adaptive method termed the modified toxicity probability interval (mTPI) incorporates the currently available safety information and indicates what dose subsequent patients should receive (ic, dose escalation, de-escalation, continue at the same dose, or close that dose level). An improved version of the dose selection design, referred to as mTPI-2, is used. This design allows for real-time adjustment of the dose based on observed toxicities and has a lower risk of treating patients at a dose higher than the target toxicity rate compared to a traditional 3+3 design and is more likely to select the most tolerable dose. The mTPI-2 model incorporates available information for all patients treated at a particular dose level to guide dosing of subsequent patients. The mTPI-2 model may be regarded as providing guidance and information to be integrated with a clinical assessment of the toxicity profiles observed at the time of analysis in determining the next dose to be investigated.

[0622] Phase 1 enrolls a maximum sample size of 24 patients. The full sample size may not be reached if either of the following criteria are met: [0623] There are no DLTs observed in ≥5 patients at the highest available dose level [0624] At least 8 subjects have been treated at the highest available dose level and the mTPI-2 model continues to recommend dose escalation

[0625] If either of these criteria are met, then the study may proceed to the next Phase (ic, from Phase 1 to Phase 2) prior to reaching the maximum sample size provided that higher dose levels are not available (eg, after ≥5 patients have been treated at Dose Level 4 in Phase 1 with no DLTs, the study can proceed to Phase 2).

[0626] A dose is excluded from further consideration due to unacceptable toxicity (DU) if at least 3 patients have been evaluated at that dose and the posterior probability that the true DLT rate is >30% exceeds 95%. Patients who are receiving study treatment concurrent with chemotherapy may be considered separately if there is evidence that toxicity may be due to the concurrent treatment with chemotherapy.

[0627] For all new dose levels, there is >24 hours between the first and second patients receiving GRT-C903 or the first dose of GRT-R904 to allow for observation of any severe acute toxicities as an initial assessment of the safety of this therapeutic approach and to inform subsequent treatment of patients.

[0628] Phase 1 begins with patients receiving GRT-C903 at a fixed dose of 1×10^{12} viral particles (vp) with no planned dose escalation in combination with IV nivolumab. Based on the existing safety profile of adenovirus-based vaccines in multiple clinical studies in thousands of patients, GRT-C903 is anticipated to be well-tolerated at this dose. However, if DLTs are observed and the mTPI model calls for de-escalation, the dose of adenovirus can be de-escalated to 3×10^{11} vp with a further de-escalation to 1×10^{11} vp if indicated by the mTPI model. Once patients treated

with GRT-C903 have completed the first 28-day DLT observation period, patients receive GRT-R904 as vaccine boosts administered IM bilaterally every 4 weeks (Q4W). The first 28 days after the first treatment with GRT-R904 as a boost vaccination (ie, following priming with GRT-C903) are treated as a second DLT window for the assessment of GRT-R904-related toxicity. Patients continue to receive GRT-R904 as vaccine boosts.

[0629] The mTPI-2 model determines whether doses of GRT-R904 can be increased from 30 to 100 to 300 μg in subsequent patients. Patients beginning study treatment at Dose Level 1 receive their first dose (Dose 1) of 30 μg of GRT-R904 in combination with IV nivolumab followed by the DLT observation period of 28 days per patient. Dose escalation decisions are made based on the mTPI-2 model once at least two patients have completed the DLT observation period at any dose level. For example, in the absence of any observed DLTs in the first two patients treated at Dose Level 1, dose escalation is permitted as indicated by the mTPI-2 model.

[0630] If the mTPI-2 model dictates dose escalation, then subsequent patients are treated at Dose Level 2 (ie, 30 μg of GRT-R904 in combination with 30 mg of SC ipilimumab). Once at least two patients have completed the DLT observation period, the mTPI model determines the subsequent dose of GRT-R904 to treat additional patients in combination with 30 mg of SC ipilimumab. For example, if no DLTs are observed in at least two patients in Dose Level 2, the mTPI model dictates dose escalation resulting in subsequent patients being administered GRT-R904 at the next dose level of 100 μg in combination with 30 mg of SC ipilimumab (ie, Dose Level 3) followed by the DLT observation period of 28 days. Similarly, if no DLTs are observed in at least two patients receiving 100 μg, subsequent patients receive GRT-R904 at the next dose level of 300 μg in combination with 30 mg of SC ipilimumab (ie, Dose Level 4) followed by a DLT observation period of 28 days. If DLTs are observed at the initial 30 μg dose and the mTPI model indicates de-escalation, a lower dose (Dose Level-1) of 10 μg of GRT-R904 may be evaluated in these patients.

[0631] If a DLT is observed at a particular dose level and the mTPI-2 model dictates dose de-escalation, then subsequent patients are treated at the next lower dose level below the dose level at which the DLT was observed. For example, if a DLT was observed within the 28-day DLT observation period in Dose Level 4 (ie, 300 μg dose of GRT-R904 in combination with 30 mg of SC ipilimumab) and the mTPI-2 model dictates dose de-escalation, subsequent patients start Dose 1 at Dose Level 3 (ie, 100 μg of GRT-R904 in combination with 30 mg of SC ipilimumab).

[0632] Each patient could receive a total of 9 vaccine doses including one dose of GRT-C903 (Dose 1) and 8 doses of GRT-R904 (Doses 2 to 7, 10, and 13). For doses 8, 9, 11, 12, 14, and onward patients only receive nivolumab.

[0633] The RP2D for GRT-R904 is declared once the criteria for moving from Phase 1 to Phase 2 have been met as described above (no DLTs in at least 5 patients, observed DLT rate <30% in at least 8 patients treated at the highest dose and mTPI-2 continues to recommend escalation, or maximum number of patients have been treated).

[0634] Once the study has moved to Phase 2, patients treated at Dose Levels 1 and 2 may receive the RP2D for GRT-R904 in combination with SC ipilimumab to minimize patient exposure to potentially sub-therapeutic doses and increase the amount of safety data obtained with the RP2D dose.

[0635] Following evaluation of the safety and tolerability of the vaccine regimen in Phase 1, Phase 2 enrolls tumor-specific cohorts in Phase 2 to assess early signs of clinical activity. Patients are treated in Phase 2 as in Phase 1. The vaccine regimen includes patients receiving GRT-C903/GRT-R904 IM bilaterally in combination with IV nivolumab with SC ipilimumab based on the tolerability in Phase 1. The dose of GRT-C903/GRT-R904 and SC ipilimumab is the RP2D based on data from Phase 1. Each expansion cohort enrolls approximately 20 patients. In one SLATE Phase 2 arm, patients with TP53 mutations, including ovarian cancer cohorts and others, are assessed using dose level 4 (300 μg+SC IPI). In another SLATE Phase 2 arm, NSCLC cohorts, including patients following immunotherapy and/or chemotherapy, are assessed using dose level 4 (300 μg+SC IPI) with either the original SLATE cassette or a new SLATE cassette design (such as one that features repeated epitopes, including KRAS epitopes).

XXVIII. Clinical Assessment of a Neoantigen Vaccine in Adjuvant and Maintenance Therapies

[0636] Personalized cancer vaccines encoding neoepitope cassettes (as described throughout) is administered in combination with immune checkpoint blockade in patients with advanced cancer. The heterologous prime/boost vaccine regimen involves (1) a ChAdV that is used as a prime vaccination and (2) a SAM formulated in a LNP that is used for boost vaccinations following the ChAdV vector. Both the ChAdV and SAM vectors encode the same personalized neoepitope cassette specific for each subject, which also encodes two universal CD4 T-cell epitopes (PADRE and Tetanus Toxoid). For subject inclusion, tumors are used for whole-exome and transcriptome sequencing to detect somatic mutations, and blood is used for HLA typing.

[0637] The ChAdV vector is a replication-defective, E1, E3, E4 open reading frames 2-4 (ORF2-4) deleted adenoviral vector based on chimpanzee adenovirus 68 (C68, 68/SAdV-25, originally designated as Pan 9), which is a subgroup E adenovirus [ChAdV68-Empty-E4deleted; see SEQ ID NO:29365 which represents SEQ ID NO: 1 with an E1 deletion (577 to 3403), E3 deletion (27,125-31,825), and a partial E4 deletion spanning ORF2-4 (34,916 to 35,642)]. The ChAdV vector is formulated in solution at 5×10^{11} vp/mL and 1.0 mL is injected IM at each of 2 bilateral vaccine injection sites in opposing deltoid muscles (deltoid muscle preferred, gluteus [dorso or ventro] or rectus femoris on each side may be used).

[0638] The SAM vector (GRT-R902) is derived from an alphavirus. The SAM vector encodes the viral proteins and the 5′ and 3′ RNA sequences required for RNA amplification but encoded no structural proteins. The SAM vector is formulated in LNPs composed of 4 lipids: an ionizable amino lipid, a phosphatidylcholine, cholesterol, and a PEG-

based coat lipid to encapsulate the SAM and form LNPs. The SAM vector contains the same neoantigen expression cassette as used in the ChAdV vector. The SAM vector is formulated in solution at 1 mg/mL and was injected IM at each of 2 bilateral vaccine injection sites in opposing deltoid muscles (deltoid muscle preferred, gluteus [dorso or ventro] or rectus femoris on each side may be used). The boost vaccination sites is as close to the prime vaccination site as possible. The injection volume is based on the dose to be administered. A dose level amount refers explicitly to the amount of the SAM vector, i.e., it does not refer to other components, such as the LNP. The ratio of LNP:SAM is approximately 24:1.

[0639] Ipilimumab is a human monoclonal IgG1 antibody that binds to the cytotoxic T-lymphocyte associated antigen 4 (CTLA-4). Ipilimumab is formulated in solution at 5 mg/mL and is injected SC proximally (within ~2 cm) to each of the bilateral vaccination sites. The SC route of ipilimumab is distinct from the approved IV route of administration. Ipilimumab is administered at a dose of 30 mg in one of two methods listed below:

[0640] 1. Four 1.5 mL (7.5 mg) injections proximal to the vaccine draining LN at each of the bilateral vaccination sites (i.e., 1.5 mL below the vaccination site and 1.5 mL above the vaccination site on each bilateral side in each deltoid, ventrogluteal, dorsogluteal, or rectus femoris [deltoid preferred, but dependent on clinical site and patient preference]).

[0641] 2. Six 1 mL (5 mg) injections proximal to the vaccine draining LN at each of the bilateral vaccination sites (i.e., 1 mL below the vaccination site, 1 mL to the side of the vaccination site, and 1 mL above the vaccination site on each bilateral side in each deltoid, ventrogluteal, dorsogluteal, or rectus femoris [deltoid preferred, but dependent on clinical site and patient preference]).

[0642] Nivolumab is a human monoclonal IgG4 antibody that blocks the interaction of PD-1 and its ligands, PD-L1 and PD-L2. Nivolumab is formulated in solution at 10 mg/mL and is administered as an IV infusion through a 0.2-micron to 1.2-micron pore size, low-protein binding in-line filter at the protocol-specified doses. It is not administered as an IV push or bolus injection. When the dose is fixed (e.g., 240 mg flat dose), nivolumab injection may be infused undiluted or diluted so as not to exceed a total infusion volume of 160 mL. Nivolumab infusion is promptly followed by a flush of diluent to clear the line. Nivolumab is administered following each vaccination (i.e., each of the SAM or ChAdV vaccines) with or without ipilimumab on the same day. The dose and route of nivolumab is based on the Food and Drug Administration approved dose and route. Doses of nivolumab may be interrupted, delayed, or discontinued depending on how well the participant tolerates the treatment. Dosing visits are not skipped, only delayed. Vaccination does not occur in the absence of nivolumab unless the Investigator and Sponsor believe treatment with SAM vectors in the absence of nivolumab is in the best interest of the patient. Atezolizumab or cemiplimab may be administered in place of nivolumab, e.g., according to manufacturer's directions and/or appropriate dosing determined as recognized by one skilled in the art. Other PD1 and/or PD-L1 checkpoint inhibitors may alternatively be used. For example, atezolizumab can be administered via intravenous infusion at a dose of 1680 mg once every 4 weeks.

[0643] A Phase 2 study is preformed to characterize the clinical activity of maintenance therapy with GRT-C901/GRT-R902 in combination with checkpoint inhibitors in addition to fluoropyrimidine/bevacizumab versus fluoropyrimidine/bevacizumab alone as assessed by changes in circulating tumor (ct) DNA. A Phase 3 study is performed to demonstrate clinical efficacy of the regime as assessed by progression-free survival. Tumors harboring non-synonymous deoxyribonucleic acid (DNA) mutations can present peptides containing these mutations as non-self antigens in the context of human leukocyte antigens (HLAs) on the tumor cell surface. A fraction of mutated peptides result in neoantigens capable of generating T-cell responses that exclusively target tumor cells. Sensitive detection of these mutations allows for the identification of neoantigens unique to each patient's tumor to be included in a personalized cancer vaccine that targets these neoantigens. This vaccine regimen uses two vaccine vectors as a heterologous prime/boost approach (GRT-C901 first followed by GRT-R902) to stimulate an immune response. The studies explore the anti-tumor activity of this patient-specific immunotherapy in combination with checkpoint inhibitors in addition to fluoropyrimidine/bevacizumab. The study arms are shown in Table 37.

[0644] There are two stages to the study. In the vaccine production stage, while patients receive FOLFOX/bev induction therapy, neoantigen prediction is performed using a tumor biopsy and Gritstone's EDGE™ neoantigen prediction model. For patients randomized to the vaccine arm the top 20 predicted neoantigens are included in the vaccine vectors. After completing oxaliplatin, patients will enter the study treatment stage. Patients in the control arm will continue with maintenance therapy whereas patients in the vaccine arm will add the vaccine regimen to maintenance therapy. The vaccine regimen consists of GRT-C901/GRT-R902 as well as SC ipilimumab (30 mg) and IV atezolizumab (1680 mg). Over the first year of treatment, 6 vaccinations will occur. Ipilimumab will be administered SC with the first doses of GRT-C901 and GRT-R902. Atezolizumab will be administered every 4 weeks for up to 2 years.

Inclusion Criteria

[0645] Patients with histologically confirmed metastatic colorectal cancer (CRC) who are planned for, or have received <30 days of first-line treatment in the metastatic setting with FOLFOX/bev, CAPEOX/bev, FOLFOXIRI/bev, or CAPOXIRI/bev per SOC. [0646] Measurable and unresectable metastatic disease according to RECIST v1.1 [0647] Availability of formalin-fixed paraffin-embedded (FFPE) tumor specimens. [0648] Eastern Cooperative Oncology Group (ECOG) performance status of O or 1 [0649] Patient has adequate organ function per defined criteria [0650] If women of childbearing potential (WCBP), must be willing to undergo pregnancy testing and agrees to the use at least 1

highly effective contraceptive method during the study treatment period and for 150 days after last investigational study treatment

Exclusion Criteria

[0651] Patients with deficient mismatch repair (dMMR) or microsatellite instability (MSI-H) phenotype [0652] Patient has a known tumor mutation burden <1 non-synonymous mutations/megabase [0653] Known DNA Polymerase Epsilon mutations [0654] Patients with known BRAFV600E mutations [0655] Bleeding disorder or history of significant bruising or bleeding following IM injections or blood draws [0656] Immunosuppression anticipated at time of study treatment [0657] History of allogeneic tissue/solid organ transplant [0658] Active or history of autoimmune disease or immune deficiency [0659] Patient with symptomatic or actively progressing central nervous system (CNS) metastases, carcinomatous meningitis, or has been treated with whole brain radiation [0660] History of other cancer within 2 years with the exception of neoplasm that has undergone potentially curative therapy [0661] Any severe concurrent non-cancer disease that, in the judgment of the Investigator, would make the patient inappropriate for the current study [0662] Active tuberculosis or recent (<2 weeks) clinically significant infection, evidence of active hepatitis B or hepatitis C, or known history of positive test for HIV [0663] History of pneumonitis requiring systemic steroids for treatment (with the exception of prior resolved in-field radiation pneumonitis) [0664] Myocardial infarction within previous 3 months, unstable angina, serious uncontrolled cardiac arrhythmia, history of myocarditis, or congestive heart failure [0665] Pregnant, planning to become pregnant, or nursing

TABLE-US-00013 TABLE 37 Arms and Interventions Arms Assigned Interventions Experimental: Vaccine Arm Drug: GRT-C901 After receiving up to 24 weeks induction A patient-specific neoantigen cancer vaccine therapy with a administered via intramuscular (IM) injection as fluoropyrimidine/oxaliplatin/bevacizumab prime and single boost at a dose of 1 × 10.sup.12 viral (with or without irinotecan), per standard of particles 2 times over the course of the first year. care and after completing vaccine production Drug: GRT-R902 screening, patients will receive a total of 6 A patient-specific neoantigen cancer vaccine boost, administrations of GRT-C901/GRT-R902 plus administered via IM injection at a dose of 30 ug 4 ipilimumab co-administered only with the first times over the course of the first year. dose of GRT- C901 and GRT-R902. All Drug: Atezolizumab patients will receive atezolizumab in addition Atezolizumab will be administered via intravenous to maintenance therapy of a fluoropyrimidine (I.V.) infusion at a dose of 1680 mg once every 4 and bevacizumab according to standard of care. weeks. Other Names: Tecentriq Drug: Ipilimumab Ipilimumab will be administered via subcutaneous (SC) injection at a dose of 30 mg with the first dose of GRT-C901 and GRT-R902. Other Names: Yervoy Drug: Fluoropyrimidine plus leucovorin Fluoropyrimidine (infusional 5-FU or capecitabine) and leucovorin administered as maintenance therapy per standard of care. Other Names: Xeloda Drug: Bevacizumab Bevacizumab is administered as maintenance therapy per standard of care. Other Names: Avastin Drug: Oxaliplatin Oxaliplatin is administered as induction therapy per standard of care. Other Names: Eloxatin Active Comparator: Control Arm Drug: Fluoropyrimidine plus leucovorin After receiving up to 24 weeks induction Fluoropyrimidine (infusional 5-FU or capecitabine) therapy with a and leucovorin administered as maintenance therapy fluoropyrimidine/oxaliplatin/bevacizumab, per standard of care. (with or without irinotecan) per standard of Other Names: Xeloda care and after completing vaccine production Drug: Bevacizumab screening, patients will receive maintenance Bevacizumab administered as maintenance therapy therapy of a fluoropyrimidine and bevacizumab per standard of care. according to standard of care. Other Names: Avastin

Sequences

[0666] References to SEQ ID NOs are made with respect to the sequence listing in U.S. Application Pub. No. US20210196806A1, which is hereby incorporated by reference in its entirety.

Table A

[0667] Refer to Sequence Listing, SEQ ID NOS. 10,755-21,015. For clarity, each peptide that was predicted to associate with a given HLA allele peptide with an HLA allele with an EDGE score>0.001 is assigned a unique SEQ ID. NO. Each of the above sequence identifiers is associated with the amino acid sequence of the peptide, HLA subtype, the gene name corresponding to the peptide, the mutation associated with the peptide, and whether the prevelance of the peptide:HLA pair was greater than 0.1% (noted as "1") or less than 0.1% (noted as "0").

[0668] Table A is disclosed in U.S. Application Pub. No. US20210196806A1, which is hereby incorporated by reference. Sequences found in Table A are also contained in the ASCII copy, created on Nov. 4, 2020, named GSO-082WO_Sequence_Listing.txt filed together with U.S. Provisional Application No. 63/322,143 filed Mar. 21, 2022, which is hereby incorporated by reference for all purposes.

AACR GENIE Results

[0669] Refer to Sequence Listing, SEQ ID NOS. 21,016-29,357. For clarity, each peptide that was predicted to associate with a given HLA allele peptide with an HLA allele with an EDGE score>0.001 and prevelance >0.1% is assigned a unique SEQ ID. NO. Each of the above sequence identifiers is associated with the the gene name and mutations corresponding to the peptide, HLA subtype, and amino acid sequence of the peptide.

[0670] AACR GENIE results are disclosed in U.S. Application Pub. No. US20210196806A1, which is hereby incorporated by reference. Sequences found in AACR GENIE results are also contained in the ASCII copy, created on Nov. 4, 2020, named GSO-082WO_Sequence_Listing.txt filed together with U.S. Provisional Application No. 63/322,143 filed Mar. 21, 2022, which is hereby incorporated by reference for all purposes.

TABLE-US-00014 TABLE 1.2 Additional MS Validated Neoantigens (SEQ ID NOs 29513-29519) HLA Class I Restricted Peptide SEQ ID NO: Gene Mutation subtype amino acid sequence 29513 CTNNB1 S37Y point HLA-A*02:01 YLDSGIHYGA mutation 29514 CHD4 CHD4_K73fs HLA-B*08:01 TVRAATIL 29515 CTNNB1 CTNNB1_S45P A*11:01 TTAPPLSGK 29516 CTNNB1 CTNNB1_T41A A*11:01 ATAPSLSGK 29517 KRAS KRAS_G12V A*03:01 VVGAVGVGK 29518 KRAS KRAS_Q61R A*01:01 ILDTAGREEY 29519 TP53 TP53_R213L A*02:01 YLDDRNTFL

[0671] Refer to Sequence Listing, SEQ ID NOS. 57-10,754. Predicted shared antigens associated with gene expressed at a level of at least 10 TPM in at least 0.98% of cancer cases. Each of the above sequence identifiers is associated with the the gene name, amino acid sequence of the peptide, Ensembl ID, and corresponding HLA allele(s).

[0672] Table 1.2 is disclosed in U.S. Application Pub. No. US20210196806A1, which is hereby incorporated by reference. Sequences found in Table 1.2 are also contained in the ASCII copy, created on Nov. 4, 2020, named GSO-082WO_Sequence_Listing.txt filed together with U.S. Provisional Application No. 63/322,143 filed Mar. 21, 2022, which is hereby incorporated by reference for all purposes.

Certain Sequences

[0673] Vectors, cassettes, and antibodies referred to herein are described below and referred to by SEQ ID NO. Sequences throughout referred to by SEQ ID NO. are also contained in the ASCII copy, created on Nov. 4, 2020, named GSO-082WO_Sequence_Listing.txt filed together with U.S. Provisional Application No. 63/322,143 filed Mar. 21, 2022, which is hereby incorporated by reference for all purposes.

TABLE-US-00015 (SEQ ID NO: 16) Tremelimumab VL (SEQ ID NO: 17) Tremelimumab VH (SEQ ID NO: 18) Tremelimumab VH CDR1 (SEQ ID NO: 19) Tremelimumab VH CDR2 (SEQ ID NO: 20) Tremelimumab VH CDR3 (SEQ ID NO: 21) Tremelimumab VL CDR1 (SEQ ID NO: 22) Tremelimumab VL CDR2 (SEQ ID NO: 23) Tremelimumab VL CDR3 (SEQ ID NO: 24) Durvalumab (MEDI4736) VL (SEQ ID NO: 25) MEDI4736 VH (SEQ ID NO: 26) MEDI4736 VH CDR1 (SEQ ID NO: 27) MEDI4736 VH CDR2 (SEQ ID NO: 28) MEDI4736 VH CDR3 (SEQ ID NO: 29) MEDI4736 VL CDR1 (SEQ ID NO: 30) MEDI4736 VL CDR2 (SEQ ID NO: 31) MEDI4736 VL CDR3 (SEQ ID NO: 32) UbA76-25merPDTT nucleotide (SEQ ID NO: 33) UbA76-25merPDTT polypeptide (SEQ ID NO: 34) MAG-25merPDTT nucleotide (SEQ ID NO: 35) MAG-25merPDTT polypeptide (SEQ ID NO: 36) Ub7625merPDTT_NoSFL nucleotide (SEQ ID NO: 37) Ub7625merPDTT_NoSFL polypeptide ChAdV68.5WTnt.MAG25mer (SEQ ID NO: 2); AC_000011.1 with E1 (nt 577 to 3403) and E3 (nt 27,125-31,825) sequences deleted; corresponding ATCC VR-594 nucleotides substituted at five positions; model neoantigen cassette under the control of the CMV promoter/enhancer inserted in place of deleted E1; SV40 polyA 3′ of cassette Venezuelan equine encephalitis virus [VEE] (SEQ ID NO: 3) GenBank: L01442.2 VEE-MAG25mer (SEQ ID NO: 4); contains MAG-25merPDTT nucleotide (bases 30-1755) Venezuelan equine encephalitis virus strain TC-83 [TC-83](SEQ ID NO: 5) GenBank: L01443.1

ATAGGCGGCGCATGAGAGAAGCCCAGACCAATTACCTACCCAAAATGGAGAAAG
TTCACGTTGACATCGAGGAAGACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTT
CCCGCAGTTTGAGGTAGAAGCCAAGCAGGTCACTGATAATGACCATGCTAATGCC
AGAGCGTTTTCGCATCTGGCTTCAAAACTGATCGAAACGGAGGTGGACCCATCCG
ACACGATCCTTGACATTGGAAGTGCGCCCGCCCGCAGAATGTATTCTAAGCACAA
GTATCATTGTATCTGTCCGATGAGATGTGCGGAAGATCCGGACAGATTGTATAAG
TATGCAACTAAGCTGAAGAAAAACTGTAAGGAAATAACTGATAAGGAATTGGAC
AAGAAAATGAAGGAGCTCGCCGCCGTCATGAGCGACCCTGACCTGGAAACTGAG
ACTATGTGCCTCCACGACGACGAGTCGTGTCGCTACGAAGGGCAAGTCGCTGTTT
ACCAGGATGTATACGCGGTTGACGGACCGACAAGTCTCTATCACCAAGCCAATAA
GGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACCACCCCTTTTATGTTTAAG
AACTTGGCTGGAGCATATCCATCATACTCTACCAACTGGGCCGACGAAACCGTGT
TAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATGGAGCGGTCACGTAG
AGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAACAATGTTCTATTC
TCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAGGAGCTGGCACC
TGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCGGTGTGAGAC
TATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCCAGGCCTG
TATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTTGTGCT
GCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCACGTA
TGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAGT
GCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAAC
GGTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGG
CCCAGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAA
GGCCACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAG
AAGGCACAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAA

GTGAACAGCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGG
AGATCGGGCTGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGT
CACCTCTCATTACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGG
CTAAGGAGGTGCGTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGC
TGATGTTGAGGAGCCCACTCTGGAAGCCGATGTCGACTTGATGTTACAAGAGGCT
GGGGCCGGCTCAGTGGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATG
GCGAGGACAAGATCGGCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAG
TGAAAAATTATCTTGCATCCACCCTCTCGCTGAACAAGTCATAGTGATAACACAC
TCTGGCCGAAAAGGGCGTTATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGC
CAGAGGGACATGCAATACCCGTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCAC
CATTGTGTACAACGAACGTGAGTTCGTAAACAGGTACCTGCACCATATTGCCACA
CATGGAGGAGCGCTGAACACTGATGAAGAATATTACAAAACTGTCAAGCCCAGC
GAGCACGACGGCGAATACCTGTACGACATCGACAGGAAACAGTGCGTCAAGAAA
GAACTAGTCACTGGGCTAGGGCTCACAGGCGAGCTGGTGGATCCTCCCTTCCATG
AATTCGCCTACGAGAGTCTGAGAACACGACCAGCCGCTCCTTACCAAGTACCAAC
CATAGGGGTGTATGGCGTGCCAGGATCAGGCAAGTCTGGCATCATTAAAAGCGC
AGTCACCAAAAAGATCTAGTGGTGAGCGCCAAGAAAGAAACTGTGCAGAAAT
TATAAGGGACGTCAAGAAAATGAAAGGGCTGGACGTCAATGCCAGAACTGTGGA
CTCAGTGCTCTTGAATGGATGCAAACACCCGTAGAGACCCTGTATATTGACGAA
GCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCATAGCCATTATAAGACCTAA
AAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTTTTTTTAACATGATGTGC
CTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCTTCCACAAAAGCATCT
CTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTTGTTTTACGACAAA
AAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTGACACTACCGGC
AGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAGGGTGGGTGA
AGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCTGCCTCTC
AAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGAAAATCC
TCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGAGGAC
CGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTGCC
AAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT
GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATA
AGGCAAACGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCAT
AGACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCT
CACTCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATC
TGGACTCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCAC
TGGGATAACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTC
AGCTCTCTCGCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTA
TGACATGAACACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCT
GTAAACAGAAGACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGA
GTGACTTTTCTTCATTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGG
GGAAAAGTTGTCCGTCCCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAG
GCTACCTTCAGAGCTCGGCTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATG
ACATAATATTTGTTAATGTGAGGACCCCATATAAATACCATCACTATCAGCAGTG
TGAAGACCATGCCATTAAGCTTAGCATGTTGACCAAGAAAGCTTGTCTGCATCTG
AATCCCGGCGGAACCTGTGTCAGCATAGGTTATGGTTACGCTGACAGGGCCAGCG
AAAGCATCATTGGTGCTATAGCGCGGCAGTTCAAGTTTTCCCGGGTATGCAAACC
GAAATCCTCACTTGAAGAGACGGAAGTTCTGTTTGTATTCATTGGGTACGATCGC
AAGGCCCGTACGCACAATCCTTACAAGCTTTCATCAACCTTGACCAACATTTATA
CAGGTTCCAGACTCCACGAAGCCGGATGTGCACCCTCATATCATGTGGTGCGAGG
GGATATTGCCACGGCCACCGAAGGAGTGATTATAAATGCTGCTAACAGCAAAGG
ACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAAGAAATTCCCGGAAAGCTT
CGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGGTCAAAGGTGCAGCTAA
ACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTTCGGAGGTTGAAGGT
GACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATTGTCAACGATAACA
ATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTTTCCGGGAACAA
AGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGACACCACTGAT
GCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCTCAAGGAA
GCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGACTCTTCA
GTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTGGCTG
GAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGGGA

CCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCGT
TGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAG
CAGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCCACCTAGC
ACGCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAA
AAGCCTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTAT
AGAATCACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGA
AAGTGCCTGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGA
CGAGACTCCGGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACA
ACCACCACTTATAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCAT
CATCGAAGAGGAAGAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCA
CCAGGTGCTGCAAGTCGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCA
TCCTGGTCCATTCCTCATGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGA
CACCCTGGAGGGAGCTAGCGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTC
TTACTTCGCAAAGAGTATGGAGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACA
GTATTCAGGAACCCTCCACATCCCGCTCCGCGCACAAGAACACCGTCACTTGCAC
CCAGCAGGGCCTGCTCGAGAACCAGCCTAGTTTCCACCCCGCCAGGCGTGAATAG
GGTGATCACTAGAGAGGAGCTCGAGGCGCTTACCCCGTCACGCACTCCTAGCAGG
TCGGTCTCGAGAACCAGCCTGGTCTCCAACCCGCCAGGCGTAAATAGGGTGATTA
CAAGAGAGGAGTTTGAGGCGTTCGTAGCACAACAACAATGACGGTTTGATGCGG
GTGCATACATCTTTTCCTCCGACACCGGTCAAGGGCATTTACAACAAAAATCAGT
AAGGCAAACGGTGCTATCCGAAGTGGTGTTGGAGAGGACCGAATTGGAGATTTC
GTATGCCCCGCGCCTCGACCAAGAAAAGAAGAATTACTACGCAAGAAATTACA
GTTAAATCCCACACCTGCTAACAGAAGCAGATACCAGTCCAGGAAGGTGGAGAA
CATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGCCTAGGGCATTATTTGAAG
GCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCCTGTTCCTTTGTATTCAT
CTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTGGAAGCCTGTAACGC
CATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATTATTCCAGAGTACG
ATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGACACTGCCAGTTTT
TGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTGGAACCCACAA
TACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGTCCTGGCAGC
TGCCACAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCGTATTGGAT
TCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATGAATATT
GGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAACGTGGTAAATT
ACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACATAA
TTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAGA
GACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG
GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAG
AGCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGAT
ATGTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATT
GTGTTCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGC
TCTGACCGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACG
CTGATTGAGGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTA
AATTTAAATTCGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAA
CACAGTCATTAACATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGG
ATCACCATGTGCAGCATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCG
GACAAATTAATGGCAGACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATT
ATAGATGCTGTGGTGGGCGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGT
GTGACTCCGTGACCGGCACAGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTT
TAAGCTTGGCAAACCTCTGGCAGCAGACGATGAACATGATGATGACAGGAGAAG
GGCATTGCATGAAGAGTCAACACGCTGGAACCGAGTGGGTATTCTTTCAGAGCTG
TGCAAGGCAGTAGAATCAAGGTATGAAACCGTAGGAACTTCCATCATAGTTATGG
CCATGACTACTCTAGCTAGCAGTGTTAAATCATTCAGCTACCTGAGAGGGGCCCC
TATAACTCTCTACGGCTAACCTGAATGGACTACGACATAGTCTAGTCCGCCAAGA
TGTTCCCGTTCCAGCCAATGTATCCGATGCAGCCAATGCCCTATCGCAACCCGTTC
GCGGCCCCGCGCAGGCCCTGGTTCCCCAGAACCGACCCTTTTCTGGCGATGCAGG
TGCAGGAATTAACCCGCTCGATGGCTAACCTGACGTTCAAGCAACGCCGGGACGC
GCCACCTGAGGGGCCATCCGCTAAGAAACCGAAGAAGGAGGCCTCGCAAAAACA
GAAAGGGGGAGGCCAAGGGAAGAAGAAGAAGAACCAAGGGAAGAAGAAGGCT
AAGACAGGGCCGCCTAATCCGAAGGCACAGAATGGAAACAAGAAGAAGACCAA
CAAGAAACCAGGCAAGAGACAGCGCATGGTCATGAAATTGGAATCTGACAAGAC

```
GTTCCCAATCATGTTGGAAGGGAAGATAAACGGCTACGCTTGTGTGGTCGGAGGG
AAGTTATTCAGGCCGATGCATGTGGAAGGCAAGATCGACAACGACGTTCTGGCCG
CGCTTAAGACGAAGAAAGCATCCAAATACGATCTTGAGTATGCAGATGTGCCACA
GAACATGCGGGCCGATACATTCAAATACACCCATGAGAAACCCCAAGGCTATTAC
AGCTGGCATCATGGAGCAGTCCAATATGAAAATGGGCGTTTCACGGTGCCGAAA
GGAGTTGGGGCCAAGGGAGACAGCGGACGACCCATTCTGGATAACCAGGGACGG
GTGGTCGCTATTGTGCTGGGAGGTGTGAATGAAGGATCTAGGACAGCCCTTTCAG
TCGTCATGTGGAACGAGAAGGGAGTTACCGTGAAGTATACTCCGGAGAACTGCG
AGCAATGGTCACTAGTGACCACCATGTGTCTGCTCGCCAATGTGACGTTCCCATG
TGCTCAACCACCAATTTGCTACGACAGAAAACCAGCAGAGACTTTGGCCATGCTC
AGCGTTAACGTTGACAACCCGGGCTACGATGAGCTGCTGGAAGCAGCTGTTAAGT
GCCCCGGAAGGAAAAGGAGATCCACCGAGGAGCTGTTTAATGAGTATAAGCTAA
CGCGCCCTTACATGGCCAGATGCATCAGATGTGCAGTTGGGAGCTGCCATAGTCC
AATAGCAATCGAGGCAGTAAAGAGCGACGGGCACGACGGTTATGTTAGACTTCA
GACTTCCTCGCAGTATGGCCTGGATTCCTCCGGCAACTTAAAGGGCAGGACCATG
CGGTATGACATGCACGGGACCATTAAAGAGATACCACTACATCAAGTGTCACTCT
ATACATCTCGCCCGTGTCACATTGTGGATGGGCACGGTTATTTCCTGCTTGCCAGG
TGCCCGGCAGGGGACTCCATCACCATGGAATTTAAGAAAGATTCCGTCAGACACT
CCTGCTCGGTGCCGTATGAAGTGAAATTTAATCCTGTAGGCAGAGAACTCTATAC
TCATCCCCCAGAACACGGAGTAGAGCAAGCGTGCCAAGTCTACGCACATGATGC
ACAGAACAGAGGAGCTTATGTCGAGATGCACCTCCCGGGCTCAGAAGTGGACAG
CAGTTTGGTTTCCTTGAGCGGCAGTTCAGTCACCGTGACACCTCCTGATGGGACTA
GCGCCCTGGTGGAATGCGAGTGTGGCGGCACAAAGATCTCCGAGACCATCAACA
AGACAAAACAGTTCAGCCAGTGCACAAAGAAGGAGCAGTGCAGAGCATATCGGC
TGCAGAACGATAAGTGGGTGTATAATTCTGACAAACTGCCCAAAGCAGCGGGAG
CCACCTTAAAAGGAAAACTGCATGTCCCATTCTTGCTGGCAGACGGCAAATGCAC
CGTGCCTCTAGCACCAGAACCTATGATAACCTTCGGTTTCAGATCAGTGTCACTG
AAACTGCACCCTAAGAATCCCACATATCTAATCACCCGCCAACTTGCTGATGAGC
CTCACTACACGCACGAGCTCATATCTGAACCAGCTGTTAGGAATTTTACCGTCAC
CGAAAAAGGGTGGGAGTTTGTATGGGGAAACCACCCGCCGAAAAGGTTTTGGGC
ACAGGAAACAGCACCCGGAAATCCACATGGGCTACCGCACGAGGTGATAACTCA
TTATTACCACAGATACCCTATGTCCACCATCCTGGGTTTGTCAATTTGTGCCGCCA
TTGCAACCGTTTCCGTTGCAGCGTCTACCTGGCTGTTTTGCAGATCTAGAGTTGCG
TGCCTAACTCCTTACCGGCTAACACCTAACGCTAGGATACCATTTTGTCTGGCTGT
GCTTTGCTGCGCCCGCACTGCCCGGGCCGAGACCACCTGGGAGTCCTTGGATCAC
CTATGGAACAATAACCAACAGATGTTCTGGATTCAATTGCTGATCCCTCTGGCCG
CCTTGATCGTAGTGACTCGCCTGCTCAGGTGCGTGTGCTGTGTCGTGCCTTTTTTA
GTCATGGCCGGCGCCGCAGGCGCCGGCGCCTACGAGCACGCGACCACGATGCCG
AGCCAAGCGGGAATCTCGTATAACACTATAGTCAACAGAGCAGGCTACGCACCA
CTCCCTATCAGCATAACACCAACAAAGATCAAGCTGATACCTACAGTGAACTTGG
AGTACGTCACCTGCCACTACAAAACAGGAATGGATTCACCAGCCATCAAATGCTG
CGGATCTCAGGAATGCACTCCAACTTACAGGCCTGATGAACAGTGCAAAGTCTTC
ACAGGGGTTTACCCGTTCATGTGGGGTGGTGCATATTGCTTTTGCGACACTGAGA
ACACCCAAGTCAGCAAGGCCTACGTAATGAAATCTGACGACTGCCTTGCGGATCA
TGCTGAAGCATATAAAGCGCACACAGCCTCAGTGCAGGCGTTCCTCAACATCACA
GTGGGAGAACACTCTATTGTGACTACCGTGTATGTGAATGGAGAAACTCCTGTGA
ATTTCAATGGGGTCAAAATAACTGCAGGTCCGCTTTCCACAGCTTGGACACCCTTT
GATCGCAAAATCGTGCAGTATGCCGGGGAGATCTATAATTATGATTTTCCTGAGT
ATGGGGCAGGACAACCAGGAGCATTTGGAGATATACAATCCAGAACAGTCTCAA
GCTCTGATCTGTATGCCAATACCAACCTAGTGCTGCAGAGACCCAAAGCAGGAGC
GATCCACGTGCCATACACTCAGGCACCTTCGGGTTTTGAGCAATGGAAGAAAGAT
AAAGCTCCATCATTGAAATTTACCGCCCCTTTCGGATGCGAAATATATACAAACC
CCATTCGCGCCGAAAACTGTGCTGTAGGGTCAATTCCATTAGCCTTTGACATTCCC
GACGCCTTGTTCACCAGGGTGTCAGAAACACCGACACTTTCAGCGGCCGAATGCA
CTCTTAACGAGTGCGTGTATTCTTCCGACTTTGGTGGGATCGCCACGGTCAAGTAC
TCGGCCAGCAAGTCAGGCAAGTGCGCAGTCCATGTGCCATCAGGGACTGCTACCC
TAAAAGAAGCAGCAGTCGAGCTAACCGAGCAAGGGTCGGCGACTATCCATTTCTC
GACCGCAAATATCCACCCGGAGTTCAGGCTCCAAATATGCACATCATATGTTACG
TGCAAAGGTGATTGTCACCCCCCGAAAGACCATATTGTGACACACCCTCAGTATC
ACGCCCAAACATTTACAGCCGCGGTGTCAAAAACCGCGTGGACGTGGTTAACATC
```

CCTGCTGGGAGGATCAGCCGTAATTATTATAATTGGCTTGGTGCTGGCTACTATTG
TGGCCATGTACGTGCTGACCAACCAGAAACATAATTGAATACAGCAGCAATTGGC
AAGCTGCTTACATAGAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATTTT
ATTTTTCTTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTTC VEE Delivery Vector (SEQ ID NO: 6); VEE genome with nucleotides 7544-11175 deleted [alphavirus structural proteins removed ATGggcggcgcatgagagaagcccagaccaattacctacccaaaATGGagaaagttcacgttgacatcgaggaagacagccc
attcctcagagctttgcagcggagcttcccgcagtttgaggtagaagccaagcaggtcactgataatgaccatgctaatgccagagcgtt
ttcgcatctggcttcaaaactgatcgaaacggaggtggacccatccgacacgatccttgacattggaagtgcgcccgcccgcagaatgt
attctaagcacaagtatcattgtatctgtccgatgagatgtgcggaagatccggacagattgtataagtatgcaactaagctgaagaaaaa
ctgtaaggaaataactgataaggaattggacaagaaaatgaaggagctcgccgccgtcatgagcgaccctgacctggaaactgagac
tatgtgcctccacgacgacgagtcgtgtcgctacgaagggcaagtcgctgtttaccaggatgtatacgcggttgacggaccgacaagtc
tctatcaccaagccaataagggagttagagtcgcctactggataggctttgacacccacccctttttatgtttaagaacttggctggagcatat
ccatcatactctaccaactgggccgacgaaaccgtgttaacggctcgtaacataggcctatgcagctctgacgttatggagcggtcacgt
agagggatgtccattcttagaaagaagtatttgaaaccatccaacaatgttctattctctgttggctcgaccatctaccacgagaagaggg
acttactgaggagctggcacctgccgtctgtatttcacttacgtggcaagcaaaattacacatgtcggtgtgagactatagttagttgcgac
gggtacgtcgttaaaagaatagctatcagtccaggcctgtatgggaagccttcaggctatgctgctacgatgcaccgcgagggattcttg
tgctgcaaagtgacagacacattgaacggggagaggggtctctttttcccgtgtgcacgtatgtgccagctacattgtgtgaccaaatgact
ggcatactggcaacagatgtcagtgcggacgacgcgcaaaaactgctggttgggctcaaccagcgtatagtcgtcaacggtcgcacc
cagagaaacaccaataccatgaaaaattacctttttgcccgtagtggcccaggcatttgctaggtgggcaaaggaatataaggaagatca
agaagatgaaaggccactaggactacgagatagacagttagtcatgggggtgttgttgggcttttagaaggcacaagataacatctatttat
aagcgcccggatacccaaaccatcatcaaagtgaacagcgatttccactcattcgtgctgcccaggataggcagtaacacattggagat
cgggctgagaacaagaatcaggaaaatgttagaggagcacaaggagccgtcacctctcattaccgccgaggacgtacaagaagcta
agtgcgcagccgatgaggctaaggaggtgcgtgaagccgaggagttgcgcgcagctctaccacctttggcagctgatgttgaggagc
ccactctggaagccgatgtcgacttgatgttacaagaggctggggccggctcagtggagacacctcgtggcttgataaaggttaccag
ctacgctggcgaggacaagatcggctcttacgctgtgctttctccgcaggctgtactcaagagtgaaaaattatcttgcatccaccctctc
gctgaacaagtcatagtgataacacactctggccgaaaagggcgttatgccgtggaaccataccatggtaaagtagtggtgccagagg
gacatgcaatacccgtccaggactttcaagctctgagtgaaagtgccaccattgtgtacaacgaacgtgagttcgtaaacaggtacctgc
accatattgccacacatggaggagcgctgaacactgatgaagaatattacaaaactgtcaagcccagcgagcacgacggcgaatacct
gtacgacatcgacaggaaacagtgcgtcaagaaagaactagtcactgggctagggctcacaggcgagctggtggatcctcccttccat
gaattcgcctacgagagtctgagaacacgaccagccgctccttaccaagtaccaaccatagggggtgtatggcgtgccaggatcaggc
aagtctggcatcattaaaagcgcagtcaccaaaaaagatctagtggtgagcgccaagaaagaaactgtgcagaaattataagggacg
tcaagaaaatgaaagggctggacgtcaatgccagaactgtggactcagtgctcttgaatggatgcaaacaccccgtagagaccctgtat
attgacgaagcttttgcttgtcatgcaggtactctcagagcgctcatagccattataagacctaaaaaggcagtgctctgcgggatccca
aacagtgcggtttttttaacatgatgtgcctgaaagtgcattttaaccacgagatttgcacacaagtcttccacaaaagcatctctcgccgtt
gcactaaatctgtgacttcggtcgtctcaaccttgtttacgacaaaaaaatgagaacgacgaatccgaaagagactaagattgtgattga
cactaccggcagtaccaaacctaagcaggacgatctcattctcacttgtttcagagggtggggtgaagcagttgcaaatagattacaaagg
caacgaaataatgacggcagctgcctctcaagggctgacccgtaaaggtgtgtatgccgttcggtacaaggtgaatgaaaatcctctgt
acgcacccacctcagaacatgtgaacgtcctactgacccgcacggaggaccgcatcgtgtggaaaacactagccggcgacccatgg
ataaaaacactgactgccaagtaccctgggaatttcactgccacgatagaggagtggcaagcagagcatgatgccatcatgaggcaca
tcttggagagaccggaccctaccgacgtcttccagaataaggcaaacgtgtgttgggccaaggctttagtgccggtgctgaagaccgct
ggcatagacatgaccactgaacaatggaacactgtggattattttgaaacggacaaagctcactcagcagagatagtattgaaccaacta
tgcgtgaggttctttggactcgatctggactccggtctattttctgcacccactgttccgttatccattaggaataatcactgggataactccc
cgtcgcctaacatgtacgggctgaataaagaagtggtccgtcagctctctcgcaggtacccacaactgcctcgggcagttgccactgga
agagtctatgacatgaacactggtacactgcgcaattatgatccgcgcataaacctagtacctgtaaacagaagactgcctcatgctttag
tcctccaccataatgaacacccacagagtgacttttcttcattcgtcagcaaattgaagggcagaactgtcctggtggtcggggaaaagtt
gtccgtcccaggcaaaatggttgactggttgtcagaccggcctgaggctaccttcagagctcggctggatttaggcatcccaggtgatgt
gcccaaatatgacataatatttgttaatgtgaggacccccatataaataccatcactatcagcagtgtgaagaccatgccattaagcttagca
tgttgaccaagaaagcttgtctgcatctgaatcccggcggaacctgtgtcagcataggttatggttacgctgacagggccagcgaaagc
atcattggtgctatagcgcggcagttcaagttttcccgggtatgcaaaccgaaatcctcacttgaagagacggaagttctgtttgtattcatt
gggtacgatcgcaaggcccgtacgcacaatccttacaagctttcatcaaccttgaccaacatttatacaggttccagactccacgaagcc
ggatgtgcaccctcatatcatgtggtgcgaggggatattgccacggccaccgaaggagtgattataaatgctgctaacagcaaaggac
aacctggcggaggggtgtgcggagcgctgtataagaaattcccggaaagcttcgatttacagccgatcgaagtaggaaaagcgcgac
tggtcaaaggtgcagctaaacatatcattcatgccgtaggaccaaacttcaacaaagtttcggaggttgaaggtgacaaacagttggca
gaggcttatgagtccatcgctaagattgtcaacgataacaattacaagtcagtagcgattccactgttgtccaccggcatctttccgggaa
caaagatcgactaacccaatcattgaaccatttgctgacagctttagacaccactgatgcagatgtagccatatactgcagggacaagaa
atgggaaatgactctcaaggaagcagtggctaggagagaagcagtggaggagatatgcatatccgacgactcttcagtgacagaacc
tgatgcagagctggtgagggtgcatccgaagagttctttggctggaaggaagggctacagcacaagcgatggcaaaactttctcatattt
ggaagggaccaagtttcaccaggcggccaaggatatagcagaaattaatgccatgtggcccgttgcaacggaggccaatgagcaggt
atgcatgtatatcctcggagaaagcatgagcagtattaggtcgaaatgccccgtcgaagagtcggaagcctccacaccacctagcacg
ctgccttgcttgtgcatccatgccatgactccagaaagagtacagcgccctaaaagcctcacgtccagaacaaattactgtgtgctcatcct

ttccattgccgaagtatagaatcactggtgtgcagaagatccaatgctcccagcctatattgttctcaccgaaagtgcctgcgtatattcatc
caaggaagtatctcgtggaaacaccaccggtagacgagactccggagccatcggcagagaaccaatccacagaggggacacctga
acaaccaccacttataaccgaggatgagaccaggactagaacgcctgagccgatcatcatcgaagaggaagaagaggatagcataa
gtttgctgtcagatggcccgacccaccaggtgctgcaagtcgaggcagacattcacgggccgccctctgtatctagctcatcctggtcc
attcctcatgcatccgactttgatgtggacagtttatccatacttgacaccctggagggagctagcgtgaccagcggggcaacgtcagcc
gagactaactcttacttcgcaaagagtatggagtttctggcgcgaccggtgcctgcgcctcgaacagtattcaggaaccctccacatccc
gctccgcgcacaagaacaccgtcacttgcacccagcagggcctgctcgagaaccagcctagtttccaccccgccaggcgtgaatagg
gtgatcactagagaggagctcgaggcgcttaccccgtcacgcactcctagcaggtcggtctcgagaaccagcctggtctccaacccg
ccaggcgtaaataggtgattacaagagaggagtttgaggcgttcgtagcacaacaacaatgacggtttgatgcgggtgcatacatcttt
tcctccgacaccggtcaagggcatttacaacaaaaatcagtaaggcaaacggtgctatccgaagtggtgttggagaggaccgaattgg
agatttcgtatgccccgcgcctcgaccaagaaaaagaagaattactacgcaagaaattacagttaaatcccacacctgctaacagaagc
agataccagtccaggaaggtggagaacatgaaagccataacagctagacgtattctgcaaggcctagggcattatttgaaggcagaag
gaaaagtggagtgctaccgaacccttgcatcctgttcctttgtattcatctagtgtgaaccgtgcctttcaagccccaaggtcgcagtgga
agcctgtaacgccatgttgaaagagaactttccgactgtggcttcttactgtattattccagagtacgatgcctatttggacatggttgacgg
agcttcatgctgcttagacactgccagttttgccctgcaaagctgcgcagctttccaaagaaacactcctatttggaacccacaatacgat
cggcagtgccttcagcgatccagaacacgctccagaacgtcctggcagctgccacaaaaagaaattgcaatgtcacgcaaatgagag
aattgcccgtattggattcggcggcctttaatgtggaatgcttcaagaaatatgcgtgtaataatgaatattgggaaacgtttaaagaaaac
cccatcaggcttactgaagaaaacgtggtaaattacattaccaaattaaaaggaccaaaagctgctgctctttttgcgaagacacataattt
gaatatgttgcaggacataccaatggacaggtttgtaatggacttaaagagagacgtgaaagtgactccaggaacaaaacatactgaag
aacggcccaaggtacaggtgatccaggctgccgatccgctagcaacagcgtatctgtgcggaatccaccgagagctggttaggagat
taaatgcggtcctgcttccgaacattcatacactgtttgatatgtcggctgaagactttgacgctattatagccgagcacttccagcctggg
gattgtgttctggaaactgacatcgcgtcgtttgataaaagtgaggacgacgccatggctctgaccgcgttaatgattctggaagacttag
gtgtggacgcagagctgttgacgctgattgaggcggctttcggcgaaatttcatcaatacatttgcccactaaaactaaatttaaattcgga
gccatgatgaaatctggaatgttcctcacactgtttgtgaacacagtcattaacattgtaatcgcaagcagagtgttgagagaacggctaa
ccggatcaccatgtgcagcattcattggagatgacaatatcgtgaaaggagtcaaatcggacaaattaatggcagacaggtgcgccac
ctggttgaatatggaagtcaagattatagatgctgtggtgggcgagaaagcgccttatttctgtggagggtttattttgtgtgactccgtgac
cggcacagcgtgccgtgtggcagacccctaaaaaggctgtttaagcttggcaaacctctggcagcagacgatgaacatgatgatgac
aggagaagggcattgcatgaagagtcaacacgctggaaccgagtgggtattctttcagagctgtgcaaggcagtagaatcaaggtatg
aaaccgtaggaacttccatcatagttatggccatgactactctagctagcagtgttaaatcattcagctacctgagaggggcccctataact
ctctacggc**TAA**cctgaatggactacgacgtatcacgcccaaacatttacagccgcggtgtcaaaaaccgcgtggacgtggttaaca
tccctgctgggaggatcagccgtaattattataattggcttggtgctggctactattgtggccatgtacgtgctgaccaaccagaaacataa
ttgaatacagcagcaattggcaagctgcttacatagaactcgcggcgattggcatgccgcccttaaaattttatttttatttttctttctttccg
aatcggattttgtttttaatatttc**AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA** TC-83 Delivery Vector (SEQ ID
NO: 7); TC-83 genome with nucleotides 7544-11175 deleted [alphavirus structural proteins removed
ATAGGCGGCGCATGAGAGAAGCCCAGACCAATTACCTACCCAAAATGGAGAAAG
TTCACGTTGACATCGAGGAAGACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTT
CCCGCAGTTTGAGGTAGAAGCCAAGCAGGTCACTGATAATGACCATGCTAATGCC
AGAGCGTTTTCGCATCTGGCTTCAAAACTGATCGAAACGGAGGTGGACCCATCCG
ACACGATCCTTGACATTGGAAGTGCGCCCGCCCGCAGAATGTATTCTAAGCACAA
GTATCATTGTATCTGTCCGATGAGATGTGCGGAAGATCCGGACAGATTGTATAAG
TATGCAACTAAGCTGAAGAAAAACTGTAAGGAAATAACTGATAAGGAATTGGAC
AAGAAAATGAAGGAGCTCGCCGCCGTCATGAGCGACCCTGACCTGGAAACTGAG
ACTATGTGCCTCCACGACGACGAGTCGTGTCGCTACGAAGGGCAAGTCGCTGTTT
ACCAGGATGTATACGCGGTTGACGGACCGACAAGTCTCTATCACCAAGCCAATAA
GGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACCACCCCTTTTATGTTTAAG
AACTTGGCTGGAGCATATCCATCATACTCTACCAACTGGGCCGACGAAACCGTGT
TAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATGGAGCGGTCACGTAG
AGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAACAATGTTCTATTC
TCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAGGAGCTGGCACC
TGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCGGTGTGAGAC
TATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCCAGGCCTG
TATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTTGTGCT
GCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCACGTA
TGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAGT
GCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAAC
GGTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGG
CCCAGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAA
GGCCACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAG
AAGGCACAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAA

```
GTGAACAGCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGG
AGATCGGGCTGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGT
CACCTCTCATTACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGG
CTAAGGAGGTGCGTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGC
TGATGTTGAGGAGCCCACTCTGGAAGCCGATGTCGACTTGATGTTACAAGAGGCT
GGGGCCGGCTCAGTGGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATG
GCGAGGACAAGATCGGCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAG
TGAAAAATTATCTTGCATCCACCCTCTCGCTGAACAAGTCATAGTGATAACACAC
TCTGGCCGAAAAGGGCGTTATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGC
CAGAGGGACATGCAATACCCGTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCAC
CATTGTGTACAACGAACGTGAGTTCGTAAACAGGTACCTGCACCATATTGCCACA
CATGGAGGAGCGCTGAACACTGATGAAGAATATTACAAAACTGTCAAGCCCAGC
GAGCACGACGGCGAATACCTGTACGACATCGACAGGAAACAGTGCGTCAAGAAA
GAACTAGTCACTGGGCTAGGGCTCACAGGCGAGCTGGTGGATCCTCCCTTCCATG
AATTCGCCTACGAGAGTCTGAGAACACGACCAGCCGCTCCTTACCAAGTACCAAC
CATAGGGGTGTATGGCGTGCCAGGATCAGGCAAGTCTGGCATCATTAAAAGCGC
AGTCACCAAAAAGATCTAGTGGTGAGCGCCAAGAAAGAAACTGTGCAGAAAT
TATAAGGGACGTCAAGAAAATGAAAGGGCTGGACGTCAATGCCAGAACTGTGGA
CTCAGTGCTCTTGAATGGATGCAAACACCCGTAGAGACCCTGTATATTGACGAA
GCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCATAGCCATTATAAGACCTAA
AAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTTTTTTAACATGATGTGC
CTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCTTCCACAAAAGCATCT
CTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTTGTTTTACGACAAA
AAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTGACACTACCGGC
AGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAGGGTGGGTGA
AGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCTGCCTCTC
AAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGAAAATCC
TCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGAGGAC
CGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTGCC
AAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT
GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATA
AGGCAAACGTGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCAT
AGACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCT
CACTCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATC
TGGACTCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCAC
TGGGATAACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTC
AGCTCTCTCGCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTA
TGACATGAACACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCT
GTAAACAGAAGACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGA
GTGACTTTTCTTCATTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGG
GGAAAAGTTGTCCGTCCCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAG
GCTACCTTCAGAGCTCGGCTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATG
ACATAATATTTGTTAATGTGAGGACCCCATATAAATACCATCACTATCAGCAGTG
TGAAGACCATGCCATTAAGCTTAGCATGTTGACCAAGAAAGCTTGTCTGCATCTG
AATCCCGGCGGAACCTGTGTCAGCATAGGTTATGGTTACGCTGACAGGGCCAGCG
AAAGCATCATTGGTGCTATAGCGCGGCAGTTCAAGTTTTCCCGGGTATGCAAACC
GAAATCCTCACTTGAAGAGACGGAAGTTCTGTTTGTATTCATTGGGTACGATCGC
AAGGCCCGTACGCACAATCCTTACAAGCTTTCATCAACCTTGACCAACATTTATA
CAGGTTCCAGACTCCACGAAGCCGGATGTGCACCCTCATATCATGTGGTGCGAGG
GGATATTGCCACGGCCACCGAAGGAGTGATTATAAATGCTGCTAACAGCAAAGG
ACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAAGAAATTCCCGGAAAGCTT
CGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGGTCAAAGGTGCAGCTAA
ACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTTCGGAGGTTGAAGGT
GACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATTGTCAACGATAACA
ATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTTTCCGGGAACAA
AGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGACACCACTGAT
GCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCTCAAGGAA
GCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGACTCTTCA
GTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTGGCTG
GAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGGGA
```

CCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCGT
TGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAG
CAGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCCACCTAGC
ACGCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAA
AAGCCTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTAT
AGAATCACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGA
AAGTGCCTGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGA
CGAGACTCCGGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACA
ACCACCACTTATAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCAT
CATCGAAGAGGAAGAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCA
CCAGGTGCTGCAAGTCGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCA
TCCTGGTCCATTCCTCATGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGA
CACCCTGGAGGGAGCTAGCGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTC
TTACTTCGCAAAGAGTATGGAGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACA
GTATTCAGGAACCCTCCACATCCCGCTCCGCGCACAAGAACACCGTCACTTGCAC
CCAGCAGGGCCTGCTCGAGAACCAGCCTAGTTTCCACCCCGCCAGGCGTGAATAG
GGTGATCACTAGAGAGGAGCTCGAGGCGCTTACCCCGTCACGCACTCCTAGCAGG
TCGGTCTCGAGAACCAGCCTGGTCTCCAACCCGCCAGGCGTAAATAGGGTGATTA
CAAGAGAGGAGTTTGAGGCGTTCGTAGCACAACAACAATGACGGTTTGATGCGG
GTGCATACATCTTTTCCTCCGACACCGGTCAAGGGCATTTACAACAAAAATCAGT
AAGGCAAACGGTGCTATCCGAAGTGGTGTTGGAGAGGACCGAATTGGAGATTTC
GTATGCCCCGCGCCTCGACCAAGAAAAGAAGAATTACTACGCAAGAAATTACA
GTTAAATCCCACACCTGCTAACAGAAGCAGATACCAGTCCAGGAAGGTGGAGAA
CATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGCCTAGGGCATTATTTGAAG
GCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCCTGTTCCTTTGTATTCAT
CTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTGGAAGCCTGTAACGC
CATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATTATTCCAGAGTACG
ATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGACACTGCCAGTTTT
TGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTGGAACCCACAA
TACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGTCCTGGCAGC
TGCCACAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCGTATTGGAT
TCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATGAATATT
GGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAACGTGGTAAATT
ACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACATAA
TTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAGA
GACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG
GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAG
AGCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGAT
ATGTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATT
GTGTTCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGC
TCTGACCGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACG
CTGATTGAGGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTA
AATTTAAATTCGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAA
CACAGTCATTAACATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGG
ATCACCATGTGCAGCATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCG
GACAAATTAATGGCAGACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATT
ATAGATGCTGTGGTGGGCGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGT
GTGACTCCGTGACCGGCACAGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTT
TAAGCTTGGCAAACCTCTGGCAGCAGACGATGAACATGATGATGACAGGAGAAG
GGCATTGCATGAAGAGTCAACACGCTGGAACCGAGTGGGTATTCTTTCAGAGCTG
TGCAAGGCAGTAGAATCAAGGTATGAAACCGTAGGAACTTCCATCATAGTTATGG
CCATGACTACTCTAGCTAGCAGTGTTAAATCATTCAGCTACCTGAGAGGGGCCCC
TATAACTCTCTACGGCTAACCTGAATGGACTACGACATAGTCTAGTCCGCCAAGA
TGTTCCCGTTCCAGCCAATGTATCCGATGCAGCCAATGCCCTATCGCAACCCGTTC
GCGGCCCCGCGCAGGCCCTGGTTCCCCAGAACCGACCCTTTTCTGGCGATGCAGG
TGCAGGAATTAACCCGCTCGATGGCTAACCTGACGTTCAAGCAACGCCGGGACGC
GCCACCTGAGGGGCCATCCGCTAAGAAACCGAAGAAGGAGGCCTCGCAAAAACA
GAAAGGGGGAGGCCAAGGGAAGAAGAAGAAGAACCAAGGGAAGAAGAAGGCT
AAGACAGGGCCGCCTAATCCGAAGGCACAGAATGGAAACAAGAAGAAGACCAA
CAAGAAACCAGGCAAGAGACAGCGCATGGTCATGAAATTGGAATCTGACAAGAC

GTTCCCAATCATGTTGGAAGGGAAGATAAACGGCTACGCTTGTGTGGTCGGAGGG
AAGTTATTCAGGCCGATGCATGTGGAAGGCAAGATCGACAACGACGTTCTGGCCG
CGCTTAAGACGAAGAAAGCATCCAAATACGATCTTGAGTATGCAGATGTGCCACA
GAACATGCGGGCCGATACATTCAAATACACCCATGAGAAACCCCAAGGCTATTAC
AGCTGGCATCATGGAGCAGTCCAATATGAAAATGGGCGTTTCACGGTGCCGAAA
GGAGTTGGGGCCAAGGGAGACAGCGGACGACCCATTCTGGATAACCAGGGACGG
GTGGTCGCTATTGTGCTGGGAGGTGTGAATGAAGGATCTAGGACAGCCCTTTCAG
TCGTCATGTGGAACGAGAAGGGAGTTACCGTGAAGTATACTCCGGAGAACTGCG
AGCAATGGTCACTAGTGACCACCATGTGTCTGCTCGCCAATGTGACGTTCCCATG
TGCTCAACCACCAATTTGCTACGACAGAAAACCAGCAGAGACTTTGGCCATGCTC
AGCGTTAACGTTGACAACCCGGGCTACGATGAGCTGCTGGAAGCAGCTGTTAAGT
GCCCCGGAAGGAAAAGGAGATCCACCGAGGAGCTGTTTAATGAGTATAAGCTAA
CGCGCCCTTACATGGCCAGATGCATCAGATGTGCAGTTGGGAGCTGCCATAGTCC
AATAGCAATCGAGGCAGTAAAGAGCGACGGGCACGACGGTTATGTTAGACTTCA
GACTTCCTCGCAGTATGGCCTGGATTCCTCCGGCAACTTAAAGGGCAGGACCATG
CGGTATGACATGCACGGGACCATTAAAGAGATACCACTACATCAAGTGTCACTCT
ATACATCTCGCCCGTGTCACATTGTGGATGGGCACGGTTATTTCCTGCTTGCCAGG
TGCCCGGCAGGGGACTCCATCACCATGGAATTTAAGAAAGATTCCGTCAGACACT
CCTGCTCGGTGCCGTATGAAGTGAAATTTAATCCTGTAGGCAGAGAACTCTATAC
TCATCCCCCAGAACACGGAGTAGAGCAAGCGTGCCAAGTCTACGCACATGATGC
ACAGAACAGAGGAGCTTATGTCGAGATGCACCTCCCGGGCTCAGAAGTGGACAG
CAGTTTGGTTTCCTTGAGCGGCAGTTCAGTCACCGTGACACCTCCTGATGGGACTA
GCGCCCTGGTGGAATGCGAGTGTGGCGGCACAAAGATCTCCGAGACCATCAACA
AGACAAACAGTTCAGCCAGTGCACAAAGAAGGAGCAGTGCAGAGCATATCGGC
TGCAGAACGATAAGTGGGTGTATAATTCTGACAAACTGCCCAAAGCAGCGGGAG
CCACCTTAAAAGGAAAACTGCATGTCCCATTCTTGCTGGCAGACGGCAAATGCAC
CGTGCCTCTAGCACCAGAACCTATGATAACCTTCGGTTTCAGATCAGTGTCACTG
AAACTGCACCCTAAGAATCCCACATATCTAATCACCCGCCAACTTGCTGATGAGC
CTCACTACACGCACGAGCTCATATCTGAACCAGCTGTTAGGAATTTTACCGTCAC
CGAAAAAGGGTGGGAGTTTGTATGGGGAAACCACCCGCCGAAAAGGTTTTGGGC
ACAGGAAACAGCACCCGGAAATCCACATGGGCTACCGCACGAGGTGATAACTCA
TTATTACCACAGATACCCTATGTCCACCATCCTGGGTTTGTCAATTTGTGCCGCCA
TTGCAACCGTTTCCGTTGCAGCGTCTACCTGGCTGTTTTGCAGATCTAGAGTTGCG
TGCCTAACTCCTTACCGGCTAACACCTAACGCTAGGATACCATTTTGTCTGGCTGT
GCTTTGCTGCGCCCGCACTGCCCGGGCCGAGACCACCTGGGAGTCCTTGGATCAC
CTATGGAACAATAACCAACAGATGTTCTGGATTCAATTGCTGATCCCTCTGGCCG
CCTTGATCGTAGTGACTCGCCTGCTCAGGTGCGTGTGCTGTGTCGTGCCTTTTTA
GTCATGGCCGGCGCCGCAGGCGCCGGCGCCTACGAGCACGCGACCACGATGCCG
AGCCAAGCGGGAATCTCGTATAACACTATAGTCAACAGAGCAGGCTACGCACCA
CTCCCTATCAGCATAACACCAACAAAGATCAAGCTGATACCTACAGTGAACTTGG
AGTACGTCACCTGCCACTACAAAACAGGAATGGATTCACCAGCCATCAAATGCTG
CGGATCTCAGGAATGCACTCCAACTTACAGGCCTGATGAACAGTGCAAAGTCTTC
ACAGGGGTTTACCCGTTCATGTGGGGTGGTGCATATTGCTTTTGCGACACTGAGA
ACACCCAAGTCAGCAAGGCCTACGTAATGAAATCTGACGACTGCCTTGCGGATCA
TGCTGAAGCATATAAAGCGCACACAGCCTCAGTGCAGGCGTTCCTCAACATCACA
GTGGGAGAACACTCTATTGTGACTACCGTGTATGTGAATGGAGAAACTCCTGTGA
ATTTCAATGGGGTCAAAATAACTGCAGGTCCGCTTTCCACAGCTTGGACACCCTTT
GATCGCAAAATCGTGCAGTATGCCGGGGAGATCTATAATTATGATTTTCCTGAGT
ATGGGGCAGGACAACCAGGAGCATTTGGAGATATACAATCCAGAACAGTCTCAA
GCTCTGATCTGTATGCCAATACCAACCTAGTGCTGCAGAGACCCAAAGCAGGAGC
GATCCACGTGCCATACACTCAGGCACCTTCGGGTTTTGAGCAATGGAAGAAAGAT
AAAGCTCCATCATTGAAATTTACCGCCCCTTTCGGATGCGAAATATATACAAACC
CCATTCGCGCCGAAAACTGTGCTGTAGGGTCAATTCCATTAGCCTTTGACATTCCC
GACGCCTTGTTCACCAGGGTGTCAGAAACACCGACACTTTCAGCGGCCGAATGCA
CTCTTAACGAGTGCGTGTATTCTTCCGACTTTGGTGGGATCGCCACGGTCAAGTAC
TCGGCCAGCAAGTCAGGCAAGTGCGCAGTCCATGTGCCATCAGGGACTGCTACCC
TAAAAGAAGCAGCAGTCGAGCTAACCGAGCAAGGGTCGGCGACTATCCATTTCTC
GACCGCAAATATCCACCCGGAGTTCAGGCTCCAAATATGCACATCATATGTTACG
TGCAAAGGTGATTGTCACCCCCGAAAGACCATATTGTGACACACCCTCAGTATC
ACGCCCAAACATTTACAGCCGCGGTGTCAAAAACCGCGTGGACGTGGTTAACATC

CCTGCTGGGAGGATCAGCCGTAATTATTATAATTGGCTTGGTGCTGGCTACTATTG
TGGCCATGTACGTGCTGACCAACCAGAAACATAATTGAATACAGCAGCAATTGGC
AAGCTGCTTACATAGAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATTTT
ATTTTTCTTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTTC VEE Production Vector (SEQ ID NO: 8); VEE genome with nucleotides 7544-11176 deleted, plus 5′ T7-promoter, plus 3′ restriction sites TAATACGACTCACTATAGGATGggcggcgcatgagagaagcccagaccaattacctacccaaaATGGagaa
agttcacgttgacatcgaggaagacagcccattcctcagagctttgcagcggagcttcccgcagtttgaggtagaagccaagcaggtc
actgataatgaccatgctaatgccagagcgttttcgcatctggcttcaaaactgatcgaaacggaggtggacccatccgacacgatcctt
gacattggaagtgcgcccgcccgcagaatgtattctaagcacaagtatcattgtatctgtccgatgagatgtgcggaagatccggacag
attgtataagtatgcaactaagctgaagaaaaactgtaaggaaataactgataaggaattggacaagaaaatgaaggagctcgccgcc
gtcatgagcgaccctgacctggaaactgagactatgtgcctccacgacgacgagtcgtgtcgctacgaagggcaagtcgctgtttacc
aggatgtatacgcggttgacggaccgacaagtctctatcaccaagccaataaggggagttagagtcgcctactggataggctttgacacc
acccctttatgtttaagaacttggctggagcatatccatcatactctaccaactgggccgacgaaaccgtgttaacggctcgtaacatag
gcctatgcagctctgacgttatggagcggtcacgtagagggatgtccattcttagaaagaagtatttgaaaccatccaacaatgttctattc
tctgttggctcgaccatctaccacgagaagagggacttactgaggagctggcacctgccgtctgtatttcacttacgtggcaagcaaaatt
acacatgtcggtgtgagactatagttagttgcgacgggtacgtcgttaaaagaatagctatcagtccaggcctgtatgggaagccttcag
gctatgctgctacgatgcaccgcgagggattcttgtgctgcaaagtgacagacacattgaacggggagagggtctcttttcccgtgtgca
cgtatgtgccagctacattgtgtgaccaaatgactggcatactggcaacagatgtcagtgcggacgacgcgcaaaaactgctggttggg
ctcaaccagcgtatagtcgtcaacggtcgcacccagagaaacaccaataccatgaaaaattaccttttgcccgtagtggcccaggcattt
gctaggtgggcaaaggaatataaggaagatcaagaagatgaaaggccactaggactacgagatagacagttagtcatggggtgttgtt
gggcttttagaaggcacaagataacatctatttataagcgcccggatacccaaaccatcatcaaagtgaacagcgatttccactcattcgt
gctgcccaggataggcagtaacacattggagatcgggctgagaacaagaatcaggaaaatgttagaggagcacaaggagccgtcac
ctctcattaccgccgaggacgtacaagaagctaagtgcgcagccgatgaggctaaggaggtgcgtgaagccgaggagttgcgcgca
gctctaccacctttggcagctgatgttgaggagcccactctggaagccgatgtcgacttgatgttacaagaggctggggccggctcagt
ggagacacctcgtggcttgataaaggttaccagctacgctggcgaggacaagatcggctcttacgctgtgctttctccgcaggctgtact
caagagtgaaaaattatcttgcatccaccctctcgctgaacaagtcatagtgataacacactctggccgaaaagggcgttatgccgtgga
accataccatggtaaagtagtggtgccagagggacatgcaatacccgtccaggactttcaagctctgagtgaaagtgccaccattgtgt
acaacgaacgtgagttcgtaaacaggtacctgcaccatattgccacacatggaggagcgctgaacactgatgaagaatattacaaaact
gtcaagcccagcgagcacgacggcgaatacctgtacgacatcgacaggaaacagtgcgtcaagaaagaactagtcactgggctagg
gctcacaggcgagctggtggatcctcccttccatgaattcgcctacgagagtctgagaacacgaccagccgctccttaccaagtaccaa
ccatagggggtgtatggcgtgccaggatcaggcaagtctggcatcattaaaagcgcagtcaccaaaaaagatctagtggtgagcgccaa
gaaagaaaactgtgcagaaattataagggacgtcaagaaaatgaaagggctggacgtcaatgccagaactgtggactcagtgctcttg
aatggatgcaaacaccccgtagagaccctgtatattgacgaagctttttgcttgtcatgcaggtactctcagagcgctcatagccattataa
gacctaaaaaggcagtgctctgcggggatcccaaacagtgcggttttttttaacatgatgtgcctgaaagtgcattttaaccacgagatttg
cacacaagtcttccacaaaagcatctctcgccgttgcactaaatctgtgacttcggtcgtctcaaccttgttttacgacaaaaaaatgagaa
cgacgaatccgaaagagactaagattgtgattgacactaccggcagtaccaaacctaagcaggacgatctcattctcacttgtttcagag
ggtgggtgaagcagttgcaaatagattacaaaggcaacgaaataatgacggcagctgcctctcaagggctgacccgtaaaggtgtgta
tgccgttcggtacaaggtgaatgaaaatcctctgtacgcacccacctcagaacatgtgaacgtcctactacccgcacggaggaccgc
atcgtgtggaaaacactagccggcgacccatggataaaaacactgactgccaagtaccctgggaatttcactgccacgatagaggagt
ggcaagcagagcatgatgccatcatgaggcacatcttggagagaccggaccctaccgacgtcttccagaataaggcaaacgtgtgttg
ggccaaggctttagtgccggtgctgaagaccgctggcatagacatgaccactgaacaatggaacactgtggattattttgaaacggaca
aagctcactcagcagagatagtattgaaccaactatgcgtgaggttctttggactcgatctggactccggtctattttctgcacccactgttc
cgttatccattaggaataatcactgggataactccccgtcgcctaacatgtacgggctgaataaagaagtggtccgtcagctctctcgca
ggtacccacaactgcctcgggcagttgccactggaagagtctatgacatgaacactggtacactgcgcaattatgatccgcgcataaac
ctagtacctgtaaacagaagactgcctcatgctttagtcctccaccataatgaacacccacagagtgactttcttcattcgtcagcaaattg
aagggcagaactgtcctggtggtcggggaaaagttgtccgtcccaggcaaaatggttgactggttgtcagaccggcctgaggctacct
tcagagctcggctggatttaggcatcccaggtgatgtgcccaaatatgacataatatttgttaatgtgaggaccccatataaataccatcac
tatcagcagtgtgaagaccatgccattaagcttagcatgttgaccaagaaagcttgtctgcatctgaatcccggcggaacctgtgtcagc
ataggttatggttacgctgacagggccagcgaaagcatcattggtgctatagcgcggcagttcaagttttcccgggtatgcaaaccgaa
atcctcacttgaagagacggaagttctgtttgtattcattgggtacgatcgcaaggcccgtacgcacaatccttacaagctttcatcaacctt
gaccaacatttatacaggttccagactccacgaagccggatgtgcaccctcatatcatgtggtgcgaggggatattgccacggccaccg
aaggagtgattataaatgctgctaacagcaaaggacaacctggcggaggggtgtgcggagcgctgtataagaaattcccggaaagctt
cgatttacagccgatcgaagtaggaaaagcgcgactggtcaaaggtgcagctaaacatatcattcatgccgtaggaccaaacttcaaca
aagtttcggaggttgaaggtgacaaacagttggcagaggcttatgagtccatcgctaagattgtcaacgataacaattacaagtcagtag
cgattccactgttgtccaccggcatctttccgggaacaaagatcgactaacccaatcattgaaccatttgctgacagctttagacaccact
gatgcagatgtagccatatactgcaggggacaagaaatgggaaatgactctcaaggaagcagtggctaggagagaagcagtggagga
gatatgcatatccgacgactcttcagtgacagaacctgatgcagagctggtgagggtgcatccgaagagttcttggctggaaggaagg
gctacagcacaagcgatggcaaaactttctcatatttggaagggaccaagtttcaccaggggccaaggatatagcagaaattaatgcc
atgtggcccgttgcaacggaggccaatgagcaggtatgcatgtatatcctcggagaaagcatgagcagtattaggtcgaaatgccccgt
cgaagagtcggaagcctccacaccacctagcacgctgcccttgcttgtgcatccatgccatgactccagaaagagtacagcgcctaaaa

gcctcacgtccagaacaaattactgtgtgctcatcctttccattgccgaagtatagaatcactggtgtgcagaagatccaatgctcccagc
ctatattgttctcaccgaaagtgcctgcgtatattcatccaaggaagtatctcgtggaaacaccaccggtagacgagactccggagccat
cggcagagaaccaatccacagaggggacacctgaacaaccaccacttataaccgaggatgagaccaggactagaacgcctgagcc
gatcatcatcgaagaggaagaagaggatagcataagtttgctgtcagatggcccgacccaccaggtgctgcaagtcgaggcagacatt
cacgggccgccctctgtatctagctcatcctggtccattcctcatgcatccgactttgatgtggacagtttatccatacttgacaccctggag
ggagctagcgtgaccagcggggcaacgtcagccgagactaactcttacttcgcaaagagtatggagtttctggcgcgaccggtgcctg
cgcctcgaacagtattcaggaaccctccacatcccgctccgcgcacaagaacaccgtcacttgcacccagcagggcctgctcgagaa
ccagcctagtttccaccccgccaggcgtgaataggggtgatcactagagaggagctcgaggcgcttaccccgtcacgcactcctagca
ggtcggtctcgagaaccagcctggtctccaacccgccaggcgtaaatagggtgattacaagagaggagtttgaggcgttcgtagcaca
acaacaatgacggtttgatgcgggtgcatacatcttttcctccgacaccggtcaagggcatttacaacaaaaatcagtaaggcaaacggt
gctatccgaagtggtgttggagaggaccgaattggagatttcgtatgccccgcgcctcgaccaagaaaaagaagaattactacgcaag
aaattacagttaaatcccacacctgctaacagaagcagataccagtccaggaaggtggagaacatgaaagccataacagctagacgta
ttctgcaaggcctagggcattatttgaaggcagaaggaaaagtggagtgctaccgaaccctgcatcctgttcctttgtattcatctagtgtg
aaccgtgcctttcaagccccaaggtcgcagtggaagcctgtaacgccatgttgaaagagaactttccgactgtggcttcttactgtattat
tccagagtacgatgcctatttggacatggttgacggagcttcatgctgcttagacactgccagttttgccctgcaaagctgcgcagctttc
caaagaaacactcctatttggaacccacaatacgatcggcagtgccttcagcgatccagaacacgctccagaacgtcctggcagctgc
cacaaaaagaaattgcaatgtcacgcaaatgagagaattgcccgtattggattcggcggcctttaatgtggaatgcttcaagaaatatgc
gtgtaataatgaatattgggaaacgtttaaagaaaaccccatcaggcttactgaagaaaacgtggtaaattacattaccaaattaaaagga
ccaaaagctgctgctctcttttgcgaagacacataatttgaatatgttgcaggacataccaatggacaggtttgtaatggacttaaagagaga
cgtgaaagtgactccaggaacaaaacatactgaagaacggcccaaggtacaggtgatccaggctgccgatccgctagcaacagcgt
atctgtgcggaatccaccgagagctggttaggagattaaatgcggtcctgcttccgaacattcatacactgtttgatatgtcggctgaaga
ctttgacgctattatagccgagcacttccagcctggggattgtgttctggaaactgacatcgcgtcgtttgataaaagtgaggacgacgcc
atggctctgaccgcgttaatgattctggaagacttaggtgtggacgcagagctgttgacgctgattgaggcggctttcggcgaaatttcat
caatacatttgcccactaaaactaaatttaaattcggagccatgatgaaatctggaatgttcctcacactgtttgtgaacacagtcattaacat
tgtaatcgcaagcagagtgttgagagaacggctaaccggatcaccatgtgcagcattcattggagatgacaatatcgtgaaaggagtca
aatcggacaaattaatggcagacaggtgcgccacctggttgaatatggaagtcaagattatagatgctgtggtgggc   gagaaagcgcc
ttatttctgtggagggtttattttgtgtgactccgtgaccggcacagcgtgccgtgtggcagaccccctaaaaaggctgtttaagcttggca
aacctctggcagcagacgatgaacatgatgatgacaggagaagggcattgcatgaagagtcaacacgctggaaccgagtgggtattc
tttcagagctgtgcaaggcagtagaatcaaggtatgaaaccgtaggaacttccatcatagttatggccatgactactctagctagcagtgt
taaatcattcagctacctgagaggggcccctataactctctacggcTAAcctgaatggactacgacgtatcacgcccaaacatttaca
gccgcggtgtcaaaaaccgcgtggacgtggttaacatccctgctgggaggatcagccgtaattattataattggcttggtgctggctact
attgtggccatgtacgtgctgaccaaccagaaacataattgaatacagcagcaattggcaagctgcttacatagaactcgcggcgattgg
catgccgcccttaaaattttttattttatttttttcttttcttttccgaatcggatttttgtttttaatatttcAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAA TC-
83 Production Vector(SEQ ID NO: 9); TC-83 genome with nucleotides 7544-11176 deleted,
plus 5′ T7-promoter, plus 3′ restriction sites
ATAGGCGGCGCATGAGAGAAGCCCAGACCAATTACCTACCCAAAATGGAGAAAG
TTCACGTTGACATCGAGGAAGACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTT
CCCGCAGTTTGAGGTAGAAGCCAAGCAGGTCACTGATAATGACCATGCTAATGCC
AGAGCGTTTTCGCATCTGGCTTCAAAACTGATCGAAACGGAGGTGGACCCATCCG
ACACGATCCTTGACATTGGAAGTGCGCCCGCCCGCAGAATGTATTCTAAGCACAA
GTATCATTGTATCTGTCCGATGAGATGTGCGGAAGATCCGGACAGATTGTATAAG
TATGCAACTAAGCTGAAGAAAAACTGTAAGGAAATAACTGATAAGGAATTGGAC
AAGAAAATGAAGGAGCTCGCCGCCGTCATGAGCGACCCTGACCTGGAAACTGAG
ACTATGTGCCTCCACGACGACGAGTCGTGTCGCTACGAAGGGCAAGTCGCTGTTT
ACCAGGATGTATACGCGGTTGACGGACCGACAAGTCTCTATCACCAAGCCAATAA
GGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACCACCCTTTTATGTTTAAG
AACTTGGCTGGAGCATATCCATCATACTCTACCAACTGGGCCGACGAAACCGTGT
TAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATGGAGCGGTCACGTAG
AGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAACAATGTTCTATTC
TCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAGGAGCTGGCACC
TGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCGGTGTGAGAC
TATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCCAGGCCTG
TATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTTGTGCT
GCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCACGTA
TGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAGT
GCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAAC
GGTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGG
CCCAGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAA
GGCCACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAG

```
AAGGCACAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAA
GTGAACAGCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGG
AGATCGGGCTGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGT
CACCTCTCATTACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGG
CTAAGGAGGTGCGTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGC
TGATGTTGAGGAGCCCACTCTGGAAGCCGATGTCGACTTGATGTTACAAGAGGCT
GGGGCCGGCTCAGTGGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATG
GCGAGGACAAGATCGGCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAG
TGAAAAATTATCTTGCATCCACCCTCTCGCTGAACAAGTCATAGTGATAACACAC
TCTGGCCGAAAAGGGCGTTATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGC
CAGAGGGACATGCAATACCCGTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCAC
CATTGTGTACAACGAACGTGAGTTCGTAAACAGGTACCTGCACCATATTGCCACA
CATGGAGGAGCGCTGAACACTGATGAAGAATATTACAAAACTGTCAAGCCCAGC
GAGCACGACGGCGAATACCTGTACGACATCGACAGGAAACAGTGCGTCAAGAAA
GAACTAGTCACTGGGCTAGGGCTCACAGGCGAGCTGGTGGATCCTCCCTTCCATG
AATTCGCCTACGAGAGTCTGAGAACACGACCAGCCGCTCCTTACCAAGTACCAAC
CATAGGGGTGTATGGCGTGCCAGGATCAGGCAAGTCTGGCATCATTAAAAGCGC
AGTCACCAAAAAGATCTAGTGGTGAGCGCCAAGAAAGAAACTGTGCAGAAAT
TATAAGGGACGTCAAGAAATGAAAGGGCTGGACGTCAATGCCAGAACTGTGGA
CTCAGTGCTCTTGAATGGATGCAAACACCCCGTAGAGACCCTGTATATTGACGAA
GCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCATAGCCATTATAAGACCTAA
AAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTTTTTTTAACATGATGTGC
CTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCTTCCACAAAAGCATCT
CTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTTGTTTTACGACAAA
AAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTGACACTACCGGC
AGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAGGGTGGGTGA
AGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCTGCCTCTC
AAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGAAAATCC
TCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGAGGAC
CGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTGCC
AAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT
GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATA
AGGCAAACGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCAT
AGACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCT
CACTCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATC
TGGACTCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCAC
TGGGATAACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTC
AGCTCTCTCGCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTA
TGACATGAACACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCT
GTAAACAGAAGACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGA
GTGACTTTTCTTCATTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGG
GGAAAAGTTGTCCGTCCCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAG
GCTACCTTCAGAGCTCGGCTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATG
ACATAATATTTGTTAATGTGAGGACCCCATATAAATACCATCACTATCAGCAGTG
TGAAGACCATGCCATTAAGCTTAGCATGTTGACCAAGAAAGCTTGTCTGCATCTG
AATCCCGGCGGAACCTGTGTCAGCATAGGTTATGGTTACGCTGACAGGGCCAGCG
AAAGCATCATTGGTGCTATAGCGCGGCAGTTCAAGTTTTCCCGGGTATGCAAACC
GAAATCCTCACTTGAAGAGACGGAAGTTCTGTTTGTATTCATTGGGTACGATCGC
AAGGCCCGTACGCACAATCCTTACAAGCTTTCATCAACCTTGACCAACATTTATA
CAGGTTCCAGACTCCACGAAGCCGGATGTGCACCCTCATATCATGTGGTGCGAGG
GGATATTGCCACGGCCACCGAAGGAGTGATTATAAATGCTGCTAACAGCAAAGG
ACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAAGAAATTCCCGGAAAGCTT
CGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGGTCAAAGGTGCAGCTAA
ACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTTCGGAGGTTGAAGGT
GACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATTGTCAACGATAACA
ATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTTTCCGGGAACAA
AGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGACACCACTGAT
GCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCTCAAGGAA
GCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGACTCTTCA
GTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTGGCTG
```

GAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGGGA
CCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCGT
TGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAG
CAGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCACCTAGC
ACGCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAA
AAGCCTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTAT
AGAATCACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGA
AAGTGCCTGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGA
CGAGACTCCGGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACA
ACCACCACTTATAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCAT
CATCGAAGAGGAAGAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCA
CCAGGTGCTGCAAGTCGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCA
TCCTGGTCCATTCCTCATGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGA
CACCCTGGAGGGAGCTAGCGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTC
TTACTTCGCAAAGAGTATGGAGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACA
GTATTCAGGAACCCTCCACATCCCGCTCCGCGCACAAGAACACCGTCACTTGCAC
CCAGCAGGGCCTGCTCGAGAACCAGCCTAGTTTCCACCCCGCCAGGCGTGAATAG
GGTGATCACTAGAGAGGAGCTCGAGGCGCTTACCCCGTCACGCACTCCTAGCAGG
TCGGTCTCGAGAACCAGCCTGGTCTCCAACCCGCCAGGCGTAAATAGGGTGATTA
CAAGAGAGGAGTTTGAGGCGTTCGTAGCACAACAACAATGACGGTTTGATGCGG
GTGCATACATCTTTTCCTCCGACACCGGTCAAGGGCATTTACAACAAAAATCAGT
AAGGCAAACGGTGCTATCCGAAGTGGTGTTGGAGAGGACCGAATTGGAGATTTC
GTATGCCCCGCGCCTCGACCAAGAAAAGAAGAATTACTACGCAAGAAATTACA
GTTAAATCCCACACCTGCTAACAGAAGCAGATACCAGTCCAGGAAGGTGGAGAA
CATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGCCTAGGGCATTATTTGAAG
GCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCCTGTTCCTTTGTATTCAT
CTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTGGAAGCCTGTAACGC
CATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATTATTCCAGAGTACG
ATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGACACTGCCAGTTTT
TGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTGGAACCCACAA
TACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGTCCTGGCAGC
TGCCACAAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCGTATTGGAT
TCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATGAATATT
GGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAACGTGGTAAATT
ACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACATAA
TTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAGA
GACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG
GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAG
AGCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGAT
ATGTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATT
GTGTTCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGC
TCTGACCGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACG
CTGATTGAGGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTA
AATTTAAATTCGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAA
CACAGTCATTAACATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGG
ATCACCATGTGCAGCATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCG
GACAAATTAATGGCAGACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATT
ATAGATGCTGTGGTGGGCGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGT
GTGACTCCGTGACCGGCACAGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTT
TAAGCTTGGCAAACCTCTGGCAGCAGACGATGAACATGATGATGACAGGAGAAG
GGCATTGCATGAAGAGTCAACACGCTGGAACCGAGTGGGTATTCTTTCAGAGCTG
TGCAAGGCAGTAGAATCAAGGTATGAAACCGTAGGAACTTCCATCATAGTTATGG
CCATGACTACTCTAGCTAGCAGTGTTAAATCATTCAGCTACCTGAGAGGGGCCCC
TATAACTCTCTACGGCTAACCTGAATGGACTACGACGTATCACGCCCAAACATTT
ACAGCCGCGGTGTCAAAAACCGCGTGGACGTGGTTAACATCCCTGCTGGGAGGAT
CAGCCGTAATTATTATAATTGGCTTGGTGCTGGCTACTATTGTGGCCATGTACGTG
CTGACCAACCAGAAACATAATTGAATACAGCAGCAATTGGCAAGCTGCTTACATA
GAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATTTTATTTTTCTTTTCTTT
TCCGAATCGGATTTTGTTTTTAATATTTCAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA VEE-UbAAY

(SEQ ID NO: 14); VEE delivery vector with MHC class I mouse tumor epitopes SIINFEKL (SEQ ID NO: 29362) and AH1-A5 inserted VEE-Luciferase (SEQ ID NO: 15); VEE delivery vector with luciferase gene inserted at 7545 ubiquitin (SEQ ID NO: 38) > UbG76 0-228 Ubiquitin A76 (SEQ ID NO: 39) > UbA76 0-228 HLA-A2 (MHC class I) signal peptide (SEQ ID NO: 40) > MHC SignalPep 0-78 HLA-A2 (MHC class I) Trans Membrane domain (SEQ ID NO: 41) > HLA A2 TM Domain 0-201 IgK Leader Seq (SEQ ID NO: 42) > IgK Leader Seq 0-60 Human DC-Lamp (SEQ ID NO: 43) > HumanDCLAMP 0-3178 Mouse LAMP1 (SEQ ID NO: 44) > MouseLamp1 0-1858 Human Lamp1 cDNA (SEQ ID NO: 45) > Human Lamp1 0-2339 Tetanus toxoid nulceic acid sequence (SEQ ID NO: 46) Tetanus toxoid amino acid sequence (SEQ ID NO: 47) PADRE nulceotide sequence (SEQ ID NO: 48) PADRE amino acid sequence (SEQ ID NO: 49) WPRE (SEQ ID NO: 50) > WPRE 0-593 IRES (SEQ ID NO: 51) > eGFP_IRES_SEAP_Insert 1746-2335 GFP (SEQ ID NO: 52) SEAP (SEQ ID NO: 53) Firefly Luciferase (SEQ ID NO: 54) FMDV 2A (SEQ ID NO: 55) Ipilimumab Heavy Chain (SEQ ID NO: 29520)
QVQLVESGGGVVQPGRSLRLSCAASGFTFSSYTMHWVRQAPGKGLEWVTFISYDGN
NKYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAIYYCARTGWLGPFDYWGQGT
LVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHT
FPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCP
PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV
HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVL
DSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK Ipilimumab Light Chain (SEQ ID NO: 29521) EIVLTQSPGTLSLSPGERATLSCRASQSVGSSYLAWYQQKPGQAPRLLIYGAFSRATGI
PDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYGSSPWTFGQGTKVEIKRTVAAPSVFI
FPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYS
LSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC Nivolumab Heavy Chain (SEQ ID NO: 29522) QVQLVESGGGVVQPGRSLRLDCKASGITFSNSGMHWVRQAPGKGLEWVAVIWYDG
SKRYYADSVKGRFTISRDNSKNTLFLQMNSLRAEDTAVYYCATNDDYWGQGTLVTV
SSASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAV
LQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRVESKYGPPCPPCPAPEF
LGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKP
REEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPREPQV
YTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL
YSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLGK Nivolumab Light Chain (SEQ ID NO: 29523) EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAPRLLIYDASNRATGIP
ARFSGSGSGTDFTLTISSLEPEDFAVYYCQQSSNWPRTFGQGTKVEIKRTVAAPSVFIF
PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSL
SSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC SAM in vitro transcription template DNA; VEE genome with nucleotides 7544-11175 deleted, plus minimal 5′ T7-promoter
*TAATACGACTCACTATA*ATGggcggcgcatgagagaagcccagaccaattacctacccaaaATGGagaaagttc
acgttgacatcgaggaagacagcccattcctcagagctttgcagcggagcttcccgcagtttgaggtagaagccaagcaggtcactga
taatgaccatgctaatgccagagcgttttcgcatctggcttcaaaactgatcgaaacggaggtggacccatccgacacgatccttgacatt
ggaagtgcgcccgcccgcagaatgtattctaagcacaagtatcattgtatctgtccgatgagatgtgcggaagatccggacagattgtat
aagtatgcaactaagctgaagaaaaactgtaaggaaataactgataaggaattggacaagaaaatgaaggagctcgccgccgtcatga
gcgaccctgacctggaaactgagactatgtgcctccacgacgacgagtcgtgtcgctacgaagggcaagtcgctgtttaccaggatgt
atacgcggttgacggaccgacaagtctctatcaccaagccaataagggagttagagtcgcctactggataggctttgacaccaccccttt
tatgtttaagaacttggctggagcatccatcatactctaccaactgggccgacgaaaccgtgttaacggctcgtaacataggcctatgc
agctctgacgttatggagcggtcacgtagagggatgtccattcttagaaagaagtatttgaaaccatccaacaatgttctattctctgttggc
tcgaccatctaccacgagaagagggacttactgaggagctggcacctgccgtctgtatttcacttacgtggcaagcaaaattacacatgt
cggtgtgagactatagttagttgcgacgggtacgtcgttaaaagaatagctatcagtccaggcctgtatgggaagccttcaggctatgct
gctacgatgcaccgcgagggattcttgtgctgcaaagtgacagacacattgaacggggagagggtctcttttcccgtgtgcacgtatgt
gccagctacattgtgtgaccaaatgactggcatactggcaacagatgtcagtgcggacgacgcgcaaaaactgctggttgggctcaac
cagcgtatagtcgtcaacggtcgcacccagagaaacaccaataccatgaaaaattaccttttgcccgtagtggcccaggcatttgctag
gtgggcaaaggaatataaggaagatcaagaagatgaaaggccactaggactacgagatagacagttagtcatggggtgttgttgggct
tttagaaggcacaagataacatctatttataagcgcccggataccaaaccatcatcaaagtgaacagcgatttccactcattcgtgctgc
ccaggataggcagtaacacattggagatcgggctgagaacaagaatcaggaaaatgttagaggagcacaaggagccgtcacctctca
ttaccgccgaggacgtacaagaagctaagtgcgcagccgatgaggctaaggaggtgcgtgaagccgaggagttgcgcgcagctcta
ccacctttggcagctgatgttgaggagcccactctggaagccgatgtcgacttgatgttacaagaggctggggccggctcagtggaga
cacctcgtggcttgataaaaggttaccagctacgctggcgaggacaagatcggctcttacgctgtgctttctccgcaggctgtactcaaga
gtgaaaaattatcttgcatccaccctctcgctgaacaagtcatagtgataacacactctggccgaaaagggcgttatgccgtggaaccat

accatggtaaagtagtggtgccagagggacatgcaatacccgtccaggactttcaagctctgagtgaaagtgccaccattgtgtacaac
gaacgtgagttcgtaaacaggtacctgcaccatattgccacacatggaggagcgctgaacactgatgaagaatattacaaaactgtcaa
gcccagcgagcacgacggcgaatacctgtacgacatcgacaggaaacagtgcgtcaagaaagaactagtcactgggctagggctca
caggcgagctggtggatcctcccttccatgaattcgcctacgagagtctgagaacacgaccagccgctccttaccaagtaccaaccata
ggggtgtatggcgtgccaggatcaggcaagtctggcatcattaaaagcgcagtcaccaaaaaagatctagtggtgagcgccaagaaa
gaaaactgtgcagaaattataagggacgtcaagaaaatgaaagggctggacgtcaatgccagaactgtggactcagtgctcttgaatg
gatgcaaacaccccgtagagaccctgtatattgacgaagcttttgcttgtcatgcaggtactctcagagcgctcatagccattataagacc
taaaaaggcagtgctctgcggggatcccaaacagtgcggttttttttaacatgatgtgcctgaaagtgcattttaaccacgagatttgcaca
caagtcttccacaaaagcatctctcgccgttgcactaaatctgtgacttcggtcgtctcaaccttgttttacgacaaaaaaatgagaacgac
gaatccgaaagagactaagattgtgattgacactaccggcagtaccaaacctaagcaggacgatctcattctcacttgtttcagagggtg
ggtgaagcagttgcaaatagattacaaaggcaacgaaataatgacggcagctgcctctcaagggctgacccgtaaaggtgtgtatgcc
gttcggtacaaggtgaatgaaaatcctctgtacgcacccacctcagaacatgtgaacgtcctactgacccgcacggaggaccgcatcgt
gtggaaaacactagccggcgacccatggataaaaacactgactgccaagtaccctgggaatttcactgccacgatagaggagtggca
agcagagcatgatgccatcatgaggcacatcttggagagaccggaccctaccgacgtcttccagaataaggcaaacgtgtgttgggcc
aaggctttagtgccggtgctgaagaccgctggcatagacatgaccactgaacaatggaacactgtggattattttgaaacggacaaagc
tcactcagcagagatagtattgaaccaactatgcgtgaggttctttggactcgatctggactccggtctattttctgcacccactgttccgtta
tccattaggaataatcactgggataactcccgtcgcctaacatgtacgggctgaataaagaagtggtccgtcagctctctcgcaggtac
ccacaactgcctcgggcagttgccactggaagagtctatgacatgaacactggtacactgcgcaattatgatccgcgcataaacctagt
acctgtaaacagaagactgcctcatgctttagtcctccaccataatgaacacccacagagtgactttttcttcattcgtcagcaaattgaagg
gcagaactgtcctggtggtcggggaaaagttgtccgtcccaggcaaaatggttgactggttgtcagaccggcctgaggctaccttcaga
gctcggctggatttaggcatcccaggtgatgtgcccaaatatgacataatatttgttaatgtgaggaccccatataaataccatcactatca
gcagtgtgaagaccatgccattaagcttagcatgttgaccaagaaagcttgtctgcatctgaatcccggcggaacctgtgtcagcatagg
ttatggttacgctgacagggccagcgaaagcatcattggtgctatagcgcggcagttcaagttttcccgggtatgcaaaccgaaatcctc
acttgaagagacggaagttctgtttgtattcattgggtacgatcgcaaggcccgtacgcacaatccttacaagctttcatcaaccttgacca
acatttatacaggttccagactccacgaagccggatgtgcaccctcatatcatgtggtgcgaggggatattgccacggccaccgaagga
gtgattataaatgctgctaacagcaaaggacaacctggcggaggggtgtgcggagcgctgtataagaaattcccggaaagcttcgattt
acagccgatcgaagtaggaaaagcgcgactggtcaaaggtgcagctaaacatatcattcatgccgtaggaccaaacttcaacaaagttt
cggaggttgaaggtgacaaacagttggcagaggcttatgagtccatcgctaagattgtcaacgataacaattacaagtcagtagcgattc
cactgttgtccaccggcatcttttccgggaacaaagatcgactaacccaatcattgaaccatttgctgacagctttagacaccactgatgc
agatgtagccatatactgcagggacaagaaatgggaaatgactctcaaggaagcagtggctaggagagaagcagtggaggagatat
gcatatccgacgactcttcagtgacagaacctgatgcagagctggtgagggtgcatccgaagagttctttggctggaaggaagggcta
cagcacaagcgatggcaaaactttctcatatttggaagggaccaagtttcaccaggcggccaaggatatagcagaaattaatgccatgt
ggcccgttgcaacgcgaggccaatgagcaggtatgcatgtatatcctcggagaaagcatgagcagtattaggtcgaaatgccccgtcga
agagtcggaagcctccacaccacctagcacgctgccttgcttgtgcatccatgccatgactccagaaagagtacagcgcctaaaagcc
tcacgtccagaacaaattactgtgtgctcatcctttccattgccgaagtatagaatcactggtgtgcagaagatccaatgctcccagcctat
attgttctcaccgaaagtgcctgcgtatattcatccaaggaagtatctcgtggaaacaccaccggtagacgagactcc   ggagccatcgg
cagagaaccaatccacagaggggacacctgaacaaccaccacttataaccgaggatgagaccaggactagaacgcctgagccgatc
atcatcgaagaggaagaagaggatagcataagtttgctgtcagatggcccgacccaccaggtgctgcaagtcgaggcagacattcac
gggccgccctctgtatctagctcatcctggtccattcctcatgcatccgactttgatgtggacagtttatccatacttgacaccctggaggg
agctagcgtgaccagcggggcaacgtcagccgagactaactcttacttcgcaaagagtatggagtttctggcgcgaccggtgcctgcg
cctcgaacagtattcaggaaccctccacatcccgctccgcgcacaagaacaccgtcacttgcacccagcagggcctgctcgagaacc
agcctagtttccaccccgccaggcgtgaatagggtgatcactagagaggagctcgaggcgcttaccccgtcacgcactcctagcaggt
cggtctcgagaaccagcctggtctccaacccgccaggcgtaaatagggtgattacaagagaggagtttgaggcgttcgtagcacaaca
acaatgacggtttgatgcggggtgcatacatctttcctccgacaccggtcaagggcatttacaacaaaaatcagtaaggcaaacggtgct
atccgaagtggtgttggagaggaccgaattggagatttcgtatgccccgcgcctcgaccaagaaaaagaagaattactacgcaagaaa
ttacagttaaatcccacacctgctaacagaagcagataccagtccaggaaggtggagaacatgaaagccataacagctagacgtattct
gcaaggcctagggcattatttgaaggcagaaggaaaagtggagtgctaccgaaccctgcatcctgttcctttgtattcatctagtgtgaac
cgtgcctttcaagccccaaggtcgcagtggaagcctgtaacgccatgttgaaagagaactttccgactgtggcttcttactgtattattcc
agagtacgatgcctatttggacatggttgacggagcttcatgctgcttagacactgccagttttttgccctgcaaagctgcgcagctttccaa
agaaacactcctatttggaacccacaatacgatcggcagtgccttcagcgatccagaacacgctccagaacgtcctggcagctgccac
aaaaagaaattgcaatgtcacgcaaatgagagaattgcccgtattggattcggcggcctttaatgtggaatgcttcaagaaatatgcgtgt
aataatgaatattgggaaacgtttaaagaaaaccccatcaggcttactgaagaaaacgtggtaaattacattaccaaattaaaaggacca
aaagctgctgctctttttgcgaagacacataatttgaatatgttgcaggacataccaatggacaggtttgtaatggacttaaagagagacgt
gaaagtgactccaggaacaaaacatactgaagaacggcccaaggtacaggtgatccaggctgccgatccgctagcaacagcgtatct
gtgcggaatccaccgagagctggttaggagattaaatgcggtcctgcttccgaacattcatacactgtttgatatgtcggctgaagacttt
gacgctattatagccgagcacttccagcctggggattgtgttctggaaactgacatcgcgtcgtttgataaaagtgaggacgacgccatg
gctctgaccgcgttaatgattctggaagacttaggtgtggacgcagagctgttgacgctgattgaggcggctttcggcgaaatttcatcaa
tacatttgccccactaaaactaaatttaaattcggagccatgatgaaatctggaatgttcctcacactgtttgtgaacacagtcattaacattgt
aatcgcaagcagagtgttgagagaacggctaaccggatcaccatgtgcagcattcattggagatgacaatatcgtgaaaggagtcaaat
cggacaaattaatggcagacaggtgcgccacctggttgaatatggaagtcaagattatagatgctgtggtgggcgagaaagcgccttat

ttctgtggagggtttattttgtgtgactccgtgaccggcacagcgtgccgtgtggcagaccccctaaaaaggctgtttaagcttggcaaac
ctctggcagcagacgatgaacatgatgatgacaggagaagggcattgcatgaagagtcaacacgctggaaccgagtgggtattctttc
agagctgtgcaaggcagtagaatcaaggtatgaaaccgtaggaacttccatcatagttatggccatgactactctagctagcagtgttaa
atcattcagctacctgagaggggcccctataactctctacggcTAAcctgaatggactacgacgTatcacgcccaaacatttacagc
cgcggtgtcaaaaaccgcgtggacgtggttaacatccctgctgggaggatcagccgtaattattataattggcttggtgctggctactatt
gtggccatgtacgtgctgaccaaccagaaacataattgaatacagcagcaattggcaagctgcttacatagaactcgcggcgattggca
tgccgccttaaaattttattttatttttctttctttccgaatcggatttttgtttttaatatttcAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

REFERENCES

[0674] 1. Desrichard, A., Snyder, A. & Chan, T. A. Cancer Neoantigens and Applications for Immunotherapy. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* (2015). doi: 10.1158/1078-0432.CCR-14-3175 [0675] 2. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* 348, 69-74 (2015). [0676] 3. Gubin, M. M., Artyomov, M. N., Mardis, E. R. & Schreiber, R. D. Tumor neoantigens: building a framework for personalized cancer immunotherapy. *J. Clin. Invest.* 125, 3413-3421 (2015). [0677] 4. Rizvi, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124-128 (2015). [0678] 5. Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* 371, 2189-2199 (2014). [0679] 6. Carreno, B. M. et al. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science* 348, 803-808 (2015). [0680] 7. Tran, E. et al. Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. *Science* 344, 641-645 (2014). [0681] 8. Hacohen, N. & Wu, C. J.-Y. United States Patent Application: 20110293637-COMPOSITIONS AND METHODS OF IDENTIFYING TUMOR SPECIFIC NEOANTIGENS. (A1). at <http://appft1.uspto.gov/netacgi/nph-Parser?Scct1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=/netahtml/PTO/srchnum.html&r=1 &f=G&1=50&s1=20110293637.PGNR.> [0682] 9. Lundegaard, C., Hoof, I., Lund, O. & Nielsen, M. State of the art and challenges in sequence based T-cell epitope prediction. *Immunome Res.* 6 Suppl 2, S3 (2010). [0683] 10. Yadav, M. et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 515, 572-576 (2014). [0684] 11. Bassani-Sternberg, M., Pletscher-Frankild, S., Jensen, L. J. & Mann, M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol. Cell. Proteomics MCP* 14, 658-673 (2015). [0685] 12. Van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350, 207-211 (2015). [0686] 13. Yoshida, K. & Ogawa, S. Splicing factor mutations and cancer. *Wiley Interdiscip. Rev. RNA* 5, 445-459 (2014). [0687] 14. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511, 543-550 (2014). [0688] 15. Rajasagi, M. et al. Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood* 124, 453-462 (2014). [0689] 16. Downing, S. R. et al. U.S. patent application Ser. No. 01/202,08706-OPTIMIZATION OF MULTIGENE ANALYSIS OF TUMOR SAMPLES. (A1). at <http://appft1.uspto.gov/nctacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=/netahtml/PTO/srchnum.html&r=1&f=G&1=50&s1=20120208706.PGNR.> [0690] 17. Target Capture for NextGen Sequencing-IDT. at <http://www.idtdna.com/pages/products/nextgen/target-capture> [0691] 18. Shukla, S. A. et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat. Biotechnol.* 33, 1152-1158 (2015). [0692] 19. Cieslik, M. et al. The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing. *Genome Res.* 25, 1372-1381 (2015). [0693] 20. Bodini, M. et al. The hidden genomic landscape of acute myeloid leukemia: subclonal structure revealed by undetected mutations. *Blood* 125, 600-605 (2015). [0694] 21. Saunders, C. T. et al. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinforma. Oxf. Engl.* 28, 1811-1817 (2012). [0695] 22. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 31, 213-219 (2013). [0696] 23. Wilkerson, M. D. et al. Integrated RNA and DNA sequencing improves mutation detection in low purity tumors. Nucleic Acids Res. 42, c107 (2014). [0697] 24. Mosc, L. E., Wilkerson, M. D., Hayes, D. N., Perou, C. M. & Parker, J. S. ABRA: improved coding indel detection via assembly-based realignment. *Bioinforma. Oxf. Engl.* 30, 2813-2815 (2014). [0698] 25. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinforma. Oxf. Engl.* 25, 2865-2871 (2009). [0699] 26. Lam, H. Y. K. et al. Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpoint library. *Nat. Biotechnol.* 28, 47-55 (2010). [0700] 27. Frampton, G. M. et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat. Biotechnol.* 31, 1023-1031 (2013). [0701] 28. Boegel, S. et al. HLA typing from RNA-Seq sequence reads. *Genome Med.* 4, 102 (2012). [0702] 29. Liu, C. et al. ATHLATES: accurate typing of human leukocyte antigen through exome sequencing. *Nucleic Acids Res.* 41, e142 (2013). [0703] 30. Mayor, N. P. et al. HLA Typing for the Next Generation. *PloS One* 10, c0127153 (2015). [0704] 31. Roy, C. K., Olson, S., Graveley, B. R., Zamore, P. D. & Moore, M. J. Assessing long-distance RNA sequence connectivity via RNA-templated DNA-DNA ligation. *eLife* 4, (2015). [0705] 32. Song, L. & Florea, L. CLASS: constrained transcript assembly of RNA-seq reads. *BMC Bioinformatics* 14 Suppl 5, S14 (2013). [0706] 33. Maretty, L., Sibbesen, J. A. & Krogh, A. Bayesian transcriptome assembly. *Genome Biol.* 15, 501 (2014). [0707] 34. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from

RNA-seq reads. *Nat. Biotechnol.* 33, 290-295 (2015). [0708] 35. Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinforma. Oxf. Engl.* (2011). doi: 10.1093/bioinformatics/btr355 [0709] 36. Vitting-Seerup, K., Porse, B. T., Sandelin, A. & Waage, J. spliceR: an R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinformatics* 15, 81 (2014). [0710] 37. Rivas, M. A. et al. Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science* 348, 666-669 (2015). [0711] 38. Skelly, D. A., Johansson, M., Madeoy, J., Wakefield, J. & Akcy, J. M. A powerful and flexible statistical framework for testing hypotheses of allele-specific gene expression from RNA-seq data. *Genome Res.* 21, 1728-1737 (2011). [0712] 39. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinforma. Oxf. Engl.* 31, 166-169 (2015). [0713] 40. Furney, S. J. et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov.* (2013). doi: 10.1158/2159-8290.CD-13-0330 [0714] 41. Zhou, Q. et al. A chemical genetics approach for the functional assessment of novel cancer genes. *Cancer Res.* (2015). doi: 10.1158/0008-5472.CAN-14-2930 [0715] 42. Maguire, S. L. et al. SF3B1 mutations constitute a novel therapeutic target in breast cancer. *J. Pathol.* 235, 571-580 (2015). [0716] 43. Carithers, L. J. et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation Biobanking* 13, 311-319 (2015). [0717] 44. Xu, G. et al. RNA COMPASS: a dual approach for pathogen and host transcriptome analysis of RNA-seq datasets. *PloS One* 9, e89445 (2014). [0718] 45. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinforma. Oxf. Engl.* (2015). doi: 10.1093/bioinformatics/btv639 [0719] 46. Jørgensen, K. W., Rasmussen, M., Buus, S. & Nielsen, M. NetMHCstab-predicting stability of peptide-MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery. *Immunology* 141, 18-26 (2014). [0720] 47. Larsen, M. V. et al. An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. *Eur. J. Immunol.* 35, 2295-2303 (2005). [0721] 48. Nielsen, M., Lundegaard, C., Lund, O. & Keşmir, C. The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of protcasomal cleavage. *Immunogenetics* 57, 33-41 (2005). [0722] 49. Boisvert, F.-M. et al. A Quantitative Spatial Proteomics Analysis of Proteome Turnover in Human Cells. *Mol. Cell. Proteomics* 11, M111.011429-M111.011429 (2012). [0723] 50. Duan, F. et al. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *J. Exp. Med.* 211, 2231-2248 (2014). [0724] 51. Janeway's Immunobiology: 9780815345312: Medicine & Health Science Books @ Amazon.com. at <http://www.amazon.com/Janeways-Immunobiology-Kenneth-Murphy/dp/0815345313> [0725] 52. Calis, J. J. A. et al. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLOS Comput. Biol.* 9, c1003266 (2013). [0726] 53. Zhang, J. et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* 346, 256-259 (2014) [0727] 54. Walter, M. J. et al. Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* 366, 1090-1098 (2012). [0728] 55. Hunt D F, Henderson R A, Shabanowitz J, Sakaguchi K, Michel H, Sevilir N, Cox A L, Appella E, Engelhard V H. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science 1992. 255:1261-1263. [0729] 56. Zarling A L, Polefrone J M, Evans A M, Mikesh L M, Shabanowitz J, Lewis S T, Engelhard V H, Hunt D F. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. Proc Natl Acad Sci USA. 2006 Oct. 3; 103 (40): 14889-94. [0730] 57. Bassani-Sternberg M, Pletscher-Frankild S, Jensen L J, Mann M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. Mol Cell Proteomics. 2015 March; 14 (3): 658-73. doi: 10.1074/mcp.M114.042812. [0731] 58. Abelin J G, Trantham P D, Penny S A, Patterson A M, Ward S T, Hildebrand W H, Cobbold M, Bai D L, Shabanowitz J, Hunt D F. Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry. Nat Protoc. 2015 Scp; 10 (9): 1308-18. doi: 10.1038/nprot.2015.086. Epub 2015 Aug. 6 [0732] 59. Barnstable C J, Bodmer W F, Brown G, Galfre G, Milstein C, Williams A F, Ziegler A. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. Cell. 1978 May; 14 (1): 9-20. [0733] 60. Goldman J M, Hibbin J, Kearney L, Orchard K, Th′ng KH. HLA-DR monoclonal antibodies inhibit the proliferation of normal and chronic granulocytic leukaemia myeloid progenitor cells. Br J Haematol. 1982 November; 52 (3): 411-20. [0734] 61. Eng J K, Jahan T A, Hoopmann M R. Comet: an open-source MS/MS sequence database search tool. Proteomics. 2013 January; 13 (1): 22-4. doi: 10.1002/pmic.201200439. Epub 2012 Dec. 4. [0735] 62. Eng J K, Hoopmann M R, Jahan T A, Egertson J D, Noble W S, MacCoss MJ. A deeper look into Comet—implementation and features. J Am Soc Mass Spectrom. 2015 November; 26 (11): 1865-74. doi: 10.1007/s13361-015-1179-x. Epub 2015 Jun. 27. [0736] 63. Lukas Kall, Jesse Canterbury, Jason Weston, William Stafford Noble and Michael J. MacCoss. Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nature Methods 4:923-925, November 2007 [0737] 64. Lukas Käll, John D. Storey, Michael J. MacCoss and William Stafford Noble. Assigning confidence measures to peptides identified by tandem mass spectrometry. Journal of Proteome Research, 7 (1): 29-34, January 2008 [0738] 65. Lukas Kall, John D. Storey and William Stafford Noble. Nonparametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry. Bioinformatics, 24 (16): 142-148, August 2008 [0739] 66. Kinney R M, BJ Johnson, VL Brown, DW Trent. Nucleotide Sequence of the 26 S mRNA of the Virulent Trinidad Donkey Strain of Venezuelan Equine Encephalitis Virus and Deduced Sequence of the Encoded Structural Proteins. Virology 152 (2), 400-413. 1986 Jul. 30. [0740] 67. Jill E Slansky, Frederique M Rattis,

Lisa F Boyd, Tarek Fahmy, Elizabeth M Jaffee, Jonathan P Schneck, David H Margulies, Drew M Pardoll. Enhanced Antigen-Specific Antitumor Immunity with Altered Peptide Ligands that Stabilize the MHC-Peptide-TCR Complex. Immunity, Volume 13, Issue 4, 1 Oct. 2000, Pages 529-538. [0741] 68. A Y Huang, P H Gulden, A S Woods, M C Thomas, C D Tong, W Wang, V H Engelhard, G Pasternack, R Cotter, D Hunt, D M Pardoll, and E M Jaffec. The immunodominant major histocompatibility complex class I-restricted antigen of a murine colon tumor derives from an endogenous retroviral gene product. Proc Natl Acad Sci USA.; 93 (18): 9730-9735, 1996 Sep. 3. [0742] 69. JOHNSON, BARBARA J. B., RICHARD M. KINNEY, CRYSTLE L. KOST AND DENNIS W. TRENT. Molecular Determinants of Alphavirus Neurovirulence: Nucleotide and Deduced Protein Sequence Changes during Attenuation of Venezuelan Equine Encephalitis Virus. J Gen Virol 67:1951-1960, 1986. [0743] 70. Aarnoudse, C. A., Kruse, M., Konopitzky, R., Brouwenstijn, N., and Schrier, P. I. (2002). TCR reconstitution in Jurkat reporter cells facilitates the identification of novel tumor antigens by cDNA expression cloning. Int J Cancer 99, 7-13. [0744] 71. Alexander, J., Sidney, J., Southwood, S., Ruppert, J., Oscroff, C., Macwal, A., Snoke, K., Serra, H. M., Kubo, R. T., and Sette, A. (1994). Development of high potency universal DR-restricted helper epitopes by modification of high affinity DR-blocking peptides. Immunity 1, 751-761. [0745] 72. Banu, N., Chia, A., Ho, Z. Z., Garcia, A. T., Paravasivam, K., Grotenbreg, G. M., Bertoletti, A., and Gehring, A. J. (2014). Building and optimizing a virus-specific T cell receptor library for targeted immunotherapy in viral infections. Scientific Reports 4, 4166. [0746] 73. Cornet, S., Miconnet, I., Mencz, J., Lemonnier, F., and Kosmatopoulos, K. (2006). Optimal organization of a polypeptide-based candidate cancer vaccine composed of cryptic tumor peptides with enhanced immunogenicity. Vaccine 24, 2102-2109. [0747] 74. Depla, E., van der Aa, A., Livingston, B. D., Crimi, C., Allosery, K., de Brabandere, V., Krakover, J., Murthy, S., Huang, M., Power, S., et al. (2008). Rational design of a multiepitope vaccine encoding T-lymphocyte epitopes for treatment of chronic hepatitis B virus infections. Journal of Virology 82, 435-450. [0748] 75. Ishioka, G. Y., Fikes, J., Hermanson, G., Livingston, B., Crimi, C., Qin, M., del Guercio, M. F., Oseroff, C., Dahlberg, C., Alexander, J., et al. (1999). Utilization of MHC class I transgenic mice for development of minigene DNA vaccines encoding multiple HLA-restricted CTL epitopes. J Immunol 162, 3915-3925. [0749] 76. Janetzki, S., Price, L., Schroeder, H., Britten, C. M., Welters, M. J. P., and Hoos, A. (2015). Guidelines for the automated evaluation of Elispot assays. Nat Protoc 10, 1098-1115. [0750] 77. Lyons, G. E., Moore, T., Brasic, N., Li, M., Roszkowski, J. J., and Nishimura, M. I. (2006). Influence of human CD8 on antigen recognition by T-cell receptor-transduced cells. Cancer Res 66, 11455-11461. [0751] 78. Nagai, K., Ochi, T., Fujiwara, H., An, J., Shirakata, T., Mineno, J., Kuzushima, K., Shiku, H., Melenhorst, J. J., Gostick, E., et al. (2012). Aurora kinase A-specific T-cell receptor gene transfer redirects T lymphocytes to display effective antileukemia reactivity. Blood 119, 368-376. [0752] 79. Panina-Bordignon, P., Tan, A., Termijtelen, A., Demotz, S., Corradin, G., and Lanzavecchia, A. (1989). Universally immunogenic T cell epitopes: promiscuous binding to human MHC class II and promiscuous recognition by T cells. Eur J Immunol 19, 2237-2242. [0753] 80. Vitiello, A., Marchesini, D., Furze, J., Sherman, L. A., and Chesnut, R. W. (1991). Analysis of the HLA-restricted influenza-specific cytotoxic T lymphocyte response in transgenic mice carrying a chimeric human-mouse class I major histocompatibility complex. J Exp Med 173, 1007-1015. [0754] 81. Yachi, P. P., Ampudia, J., Zal, T., and Gascoigne, N. R. J. (2006). Altered peptide ligands induce delayed CD8-T cell receptor interaction—a role for CD8 in distinguishing antigen quality. Immunity 25, 203-211. [0755] 82. Pushko P, Parker M, Ludwig G V, Davis N L, Johnston R E, Smith J F. Replicon-helper systems from attenuated Venezuelan equine encephalitis virus: expression of heterologous genes in vitro and immunization against heterologous pathogens in vivo. Virology. 1997 Dec. 22; 239 (2): 389-401. [0756] 83. Strauss, JH and E G Strauss. The alphaviruses: gene expression, replication, and evolution. Microbiol Rev. 1994 September; 58 (3): 491-562. [0757] 84. Rheme C, Ehrengruber M U, Grandgirard D. Alphaviral cytotoxicity and its implication in vector development. Exp Physiol. 2005 January; 90 (1): 45-52. Epub 2004 Nov. 12. [0758] 85. Riley, Michael K. II, and Wilfred Vermerris. Recent Advances in Nanomaterials for Gene Delivery-A Review. Nanomaterials 2017, 7 (5), 94. [0759] 86. Frolov I, Hardy R, Rice C M. Cis-acting RNA elements at the 5′ end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. RNA. 2001 November; 7 (11): 1638-51. [0760] 87. Jose J, Snyder J E, Kuhn R J. A structural and functional perspective of alphavirus replication and assembly. Future Microbiol. 2009 September; 4 (7): 837-56. [0761] 88. Bo Li and C. olin N. Dewey. RSEM: accurate transcript quantification from RNA-Seq data with or without a referenfe genome. BMC Bioinformatics, 12:323, August 2011 [0762] 89. Hillary Pearson, Tariq Daouda, Diana Paola Granados, Chantal Durette, Eric Bonneil, Mathieu Courcelles, Anja Rodenbrock, Jean-Philippe Laverdure, Caroline Côté, Sylvie Mader, Sébastien Lemieux, Pierre Thibault, and Claude Perreault. MHC class I-associated peptides derive from selective regions of the human genome. The Journal of Clinical Investigation, 2016, [0763] 90. Juliane Liepe, Fabio Marino, John Sidney, Anita Jeko, Daniel E. Bunting, Alessandro Sette, Peter M. Kloetzel, Michael P. H. Stumpf, Albert J. R. Heck, Michele Mishto. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. Science, 21, October 2016. [0764] 91. Mommen G P., Marino, F., Meiring H D., Poelen, MC., van Gaans-van den Brink, JA., Mohammed S., Heck A J., and van Els C A. Sampling From the Proteome to the Human Leukocyte Antigen-DR (HLA-DR) Ligandome Proceeds Via High Specificity. Mol Cell Proteomics 15 (4): 1412-1423 April 2016. [0765] 92. Sebastian Kreiter, Mathias Vormehr, Niels van de Roemer, Mustafa Diken, Martin Lower, Jan Dickmann, Sebastian Boegel, Barbara Schrörs, Fulvia Vascotto, John C. Castle, Arbel D. Tadmor, Stephen P. Schoenberger, Christoph Huber, Özlem Türeci, and Ugur Sahin. Mutant MHC class II epitopes drive therapeutic immune responses to caner. Nature 520, 692-696, April 2015. [0766] 93. Tran

E., Turcotte S., Gros A., Robbins P. F., Lu Y. C., Dudley M. E., Wunderlich J. R., Somerville R. P., Hogan K., Hinrichs C. S., Parkhurst M. R., Yang J. C., Rosenberg S. A. Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. Science 344 (6184) 641-645, May 2014. [0767] 94. Andreatta M., Karosiene E., Rasmussen M., Stryhn A., Buus S., Nielsen M. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. Immunogenetics 67 (11-12) 641-650, November 2015. [0768] 95. Nielsen, M., Lund, O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics 10:296, September 2009. [0769] 96. Nielsen, M., Lundegaard, C., Lund, O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. BMC Bioinformatics 8:238, July 2007. [0770] 97. Zhang, J., et al. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. Molecular & Cellular Proteomics. 11 (4): 1-8. Jan. 2, 2012. [0771] 98. Jensen, Kamilla Kjaergaard, et al. "Improved Methods for Prediting Peptide Binding Affinity to MHC Class II Molecules." Immunology, 2018, doi: 10.1111/imm.12889. [0772] 99. Carter, S. L., Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P. W., Onofrio, R. C., Winckler, W., Weir, B. A., et al. (2012). Absolute quantification of somatic DNA alterations in human cancer. Nat. Biotechnol. 30, 413-421 100. McGranahan, N., Rosenthal, R., Hiley, C. T., Rowan, A. J., Watkins, T. B. K., Wilson, G. A., Birkbak, N. J., Vecriah, S., Van Loo, P., Herrero, J., et al. (2017). Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell 171, 1259-1271.el1. [0773] 101. Shukla, S. A., Rooney, M. S., Rajasagi, M., Tiao, G., Dixon, P. M., Lawrence, M. S., Stevens, J., Lane, W. J., Dellagatta, J. L., Steelman, S., et al. (2015). Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nat. Biotechnol. 33, 1152-1158. [0774] 102. Van Loo, P., Nordgard, S. H., Lingjærde, O. C., Russnes, H. G., Rye, I. H., Sun, W., Weigman, V. J., Marynen, P., Zetterberg, A., Naume, B., et al. (2010). Allele-specific copy number analysis of tumors. Proc. Natl. Acad. Sci. U.S.A 107, 16910-16915. [0775] 103. Van Loo, P., Nordgard, S. H., Lingjærde, O. C., Russnes, H. G., Rye, I. H., Sun, W., Weigman, V. J., Marynen, P., Zetterberg, A., Naume, B., et al. (2010). Allele-specific copy number analysis of tumors. Proc. Natl. Acad. Sci. U.S.A 107, 16910-16915.

Claims

1. A composition for delivery of a self-amplifying alphavirus-based expression system, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises: (A) the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, wherein the one or more vectors comprises: (a) an RNA alphavirus backbone, wherein the RNA alphavirus backbone comprises: (i) at least one promoter nucleotide sequence, and (ii) at least one polyadenylation (poly(A)) sequence; and (b) a cassette, wherein the cassette comprises: (i) at least one antigen-encoding nucleic acid sequence comprising: a. an epitope-encoding nucleic acid sequence, optionally comprising at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence, b. optionally a 5′ linker sequence, and c. optionally a 3′ linker sequence; (ii) optionally, a second promoter nucleotide sequence operably linked to the at least one antigen-encoding nucleic acid sequence; and (iii) optionally, at least one second poly(A) sequence, wherein the second poly(A) sequence is a native poly(A) sequence or an exogenous poly(A) sequence to the alphavirus, and (B) a lipid-nanoparticle (LNP), wherein the LNP encapsulates the self-amplifying alphavirus-based expression system, and wherein the composition comprises a therapeutically effective amount comprising 30 µg or less of each of the one or more vectors.
2. (canceled)
3. The composition of claim 1, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises: (A) a therapeutically effective amount comprising 30 µg or less in total for all of the one or more vectors combined; (B) 30 µg of each of the one or more vectors; (C) 30 µg in total for all of the one or more vectors combined of each of the one or more vectors; or (D) 30 µg in total for all of the one or more vectors combined, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises only a single distinct vector.
4. (canceled)
5. (canceled)
6. (canceled)
7. The composition of claim 1, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises only a single distinct vector.
8. The composition of claim 1, wherein: (A) the therapeutically effective amount is capable of stimulating a T cell response following administration to a subject, optionally wherein the T cell response is assessed by monitoring ex vivo stimulation with the encoded epitope in PBMCs from the subject, optionally monitoring ex vivo stimulation comprises an ELISpot assay, further optionally wherein the ELISpot assay quantifies IFN-gamma production; and/or (B) the therapeutically effective amount does not stimulate an inhibitory IFNα response and/or reduces the inhibitory IFNα response following administration to a subject.
9. (canceled)

10. (canceled)

11. (canceled)

12. The composition of claim, wherein: (A) the weight to weight ratio of the LNP to total weight of the one or more vectors is between 10-40 to 1; 16-32 to 1; about 24 to 1; or 24 to 1; and/or (B) the one or more vectors is at a concentration of 1 mg/mL.

13. The composition of claim 1, wherein the therapeutically effective amount is at least 1 μg, at least 3 μg, at least 10 μg, about 30 μg or less, or is 30 μg.

14. (canceled)

15. (canceled)

16. (canceled)

17. (canceled)

18. (canceled)

19. (canceled)

20. The composition of claim 1, wherein the backbone comprises at least one nucleotide sequence of an Aura virus, a Fort Morgan virus, a Venezuelan equine encephalitis virus, a Ross River virus, a Semliki Forest virus, a Sindbis virus, or a Mayaro virus, optionally wherein the backbone comprises at least sequences for nonstructural protein-mediated amplification, a 26S promoter sequence, and a poly(A) sequence encoded by the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus; optionally wherein sequences for nonstructural protein-mediated amplification are selected from the group consisting of: an alphavirus 5′ UTR, a 51-nt CSE, a 24-nt CSE, a 26S subgenomic promoter sequence, a 19-nt CSE, an alphavirus 3′ UTR, or combinations thereof; and/or the backbone does not encode structural virion proteins capsid, E2 and E1, optionally wherein the antigen cassette is inserted in place of structural virion proteins within the nucleotide sequence of the Aura virus, the Fort Morgan virus, the Venezuelan equine encephalitis virus, the Ross River virus, the Semliki Forest virus, the Sindbis virus, or the Mayaro virus; and/or optionally wherein the Venezuelan equine encephalitis virus comprises: the sequence of SEQ ID NO:3 or SEQ ID NO:5, optionally further comprising a deletion between base pair 7544 and 11175, or the sequence set forth in SEQ ID NO:6 or SEQ ID NO:7, optionally wherein the antigen cassette is inserted at position 7544 to replace the deletion between base pairs 7544 and 11175 as set forth in the sequence of SEQ ID NO:3 or SEQ ID NO:5.

21. The composition of claim 1, wherein: each antigen-encoding nucleic acid sequence is linked directly to one another; and/or at least one of the at least one antigen-encoding nucleic acid sequences is linked to a distinct antigen-encoding nucleic acid sequence with a nucleic acid sequence encoding a linker, optionally wherein the linker links two MHC class I sequences or an MHC class I sequence to an MHC class II sequence, optionally wherein the linker is one or more native sequences flanking the antigen derived from the cognate protein of origin and that is at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 2-20 amino acid residues in length, or the linker links two MHC class II sequences or an MHC class II sequence to an MHC class I sequence, optionally wherein the linker comprises the sequence GPGPG.

22. The composition of claim 1, wherein: the at least one antigen-encoding nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleic acid sequences; or the at least one antigen-encoding nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 nucleic acid sequences; or the at least one antigen-encoding nucleic acid sequence comprises at least 2-400 nucleic acid sequences and wherein at least two of the antigen-encoding nucleic acid sequences encode polypeptide sequences or portions thereof that are presented by MHC class I on the tumor cell surface; and/or each MHC class I antigen-encoding nucleic acid sequence encodes a polypeptide sequence between 8 and 35 amino acids in length, optionally 9-17, 9-25, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35 amino acids in length; and/or the at least one promoter nucleotide sequence is a native 26S promoter nucleotide sequence encoded by the backbone or wherein the at least one promoter nucleotide sequence is an exogenous RNA promoter; and/or the second promoter nucleotide sequence is a 26S promoter nucleotide sequence, or comprises multiple 26S promoter nucleotide sequences, wherein each 26S promoter nucleotide sequence provides for transcription of one or more of the separate open reading frames.

23. The composition of claim 1, wherein: the antigen cassette comprises junctional epitope sequences formed by adjacent sequences in the antigen cassette, optionally wherein at least one or each junctional epitope sequence has an affinity of greater than 500 nM for MHC and/or each junctional epitope sequence is non-self; and/or the antigen cassette does not encode a non-therapeutic MHC class I or class II epitope nucleic acid sequence comprising a translated, wild-type nucleic acid sequence, wherein the non-therapeutic epitope is predicted to be displayed on an MHC allele of the subject, optionally wherein the non-therapeutic predicted MHC class I or class II epitope sequence is a junctional epitope sequence formed by adjacent sequences in the antigen cassette.

24. (canceled)

25. (canceled)

26. The composition of claim 1, wherein: the at least one nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleic acid sequences, optionally wherein each nucleic acid sequence encodes a distinct non-coding nucleic

acid sequence, a distinct polypeptide-encoding nucleic acid sequence, or a combination thereof; or the at least one nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 nucleic acid sequences, optionally wherein each nucleic acid sequence encodes a distinct non-coding nucleic acid sequence, a distinct polypeptide-encoding nucleic acid sequence, or a combination thereof; or the at least one nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 polypeptide-encoding nucleic acid sequences; or the at least one nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 polypeptide-encoding nucleic acid sequences; or the at least one nucleic acid sequence comprises at least 2-10, 2, 3, 4, 5, 6, 7, 8, 9, or 10 antigen-encoding nucleic acid sequences; or the at least one nucleic acid sequence comprises at least 11-20, 15-20, 11-100, 11-200, 11-300, 11-400, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or up to 400 antigen-encoding nucleic acid sequences; or the at least one nucleic acid sequence comprises at least 2-400 antigen-encoding nucleic acid sequences and wherein at least two of the antigen-encoding nucleic acid sequences encode polypeptide sequences or portions thereof that are presented by MHC class I on a cell surface.

27. (canceled)

28. The composition of claim 1, wherein: the epitope-encoding nucleic acid sequences comprises at least one MHC class I epitope-encoding nucleic acid sequence, and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence between 8 and 35 amino acids in length, optionally 9-17, 9-25, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35 amino acids in length; and/or the at least one MHC class II epitope-encoding nucleic acid sequence is present, optionally wherein the at least one MHC class II epitope-encoding nucleic acid sequence is present and comprises at least one MHC class II epitope-encoding nucleic acid sequence that comprises at least one alteration that makes the encoded epitope sequence distinct from the corresponding peptide sequence encoded by a wild-type nucleic acid sequence; and/or the epitope-encoding nucleic acid sequence comprises an MHC class II epitope-encoding nucleic acid sequence and wherein each antigen-encoding nucleic acid sequence encodes a polypeptide sequence that is 12-20, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 20-40 amino acids in length; and/or the epitope-encoding nucleic acid sequences comprises an MHC class II epitope-encoding nucleic acid sequence, wherein the at least one MHC class II epitope-encoding nucleic acid sequence is present, and wherein the at least one MHC class II epitope-encoding nucleic acid sequence comprises at least one universal MHC class II epitope-encoding nucleic acid sequence, optionally wherein the at least one universal sequence comprises at least one of Tetanus toxoid and PADRE; and/or the at least one promoter nucleotide sequence or the second promoter nucleotide sequence is inducible or non-inducible; and/or the at least one poly(A) sequence comprises a poly(A) sequence native to the alphavirus or exogenous to the alphavirus; and/or the at least one poly(A) sequence is operably linked to at least one of the at least one nucleic acid sequences; and/or wherein the at least one poly(A) sequence is at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, or at least 90 consecutive A nucleotides, or is at least 100 consecutive A nucleotides.

29. The composition of claim 1, wherein the epitope-encoding nucleic acid sequence comprises a MHC class I epitope-encoding nucleic acid sequence, and wherein the MHC class I epitope-encoding nucleic acid sequence is selected by performing the steps of: (a) obtaining at least one of exome, transcriptome, or whole genome tumor nucleotide sequencing data from the tumor, wherein the tumor nucleotide sequencing data is used to obtain data representing peptide sequences of each of a set of epitopes; (b) inputting the peptide sequence of each epitope into a presentation model to generate a set of numerical likelihoods that each of the epitopes is presented by one or more of the MHC alleles on the tumor cell surface of the tumor, the set of numerical likelihoods having been identified at least based on received mass spectrometry data; and (c) selecting a subset of the set of epitopes based on the set of numerical likelihoods to generate a set of selected epitopes which are used to generate the MHC class I epitope-encoding nucleic acid sequence. optionally wherein a number of the set of selected epitopes is 2-20; and/or the presentation model represents dependence between: (a) presence of a pair of a particular one of the MHC alleles and a particular amino acid at a particular position of a peptide sequence, and (b) likelihood of presentation on the tumor cell surface, by the particular one of the MHC alleles of the pair, of such a peptide sequence comprising the particular amino acid at the particular position; and/or selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being presented on the tumor cell surface relative to unselected epitopes based on the presentation model; and/or selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of inducing a tumor-specific immune response in the subject relative to unselected epitopes based on the presentation model; and/or selecting the set of selected epitopes comprises selecting epitopes that have an increased likelihood of being capable of being presented to naïve T cells by professional antigen presenting cells (APCs) relative to unselected epitopes based on the presentation model, optionally wherein the APC is a dendritic cell (DC); and/or selecting the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being subject to inhibition via central or peripheral tolerance relative to unselected epitopes based on the presentation model; and/or the set of selected epitopes comprises selecting epitopes that have a decreased likelihood of being capable of inducing an autoimmune response to normal tissue in the subject relative to unselected epitopes based on the presentation model; and/or exome or transcriptome nucleotide sequencing data is obtained by performing sequencing on the tumor tissue, optionally wherein the sequencing is next generation sequencing (NGS) or any massively parallel sequencing approach.

30. The composition of claim 1, wherein: the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have increased binding affinity to its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence; and/or the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have increased binding stability to its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence; and/or the epitope-encoding nucleic acid sequence comprises at least one alteration that makes the encoded epitope have an increased likelihood of presentation on its corresponding MHC allele relative to the translated, corresponding wild-type nucleic acid sequence; and/or the at least one alteration comprises a point mutation, a frameshift mutation, a non-frameshift mutation, a deletion mutation, an insertion mutation, a splice variant, a genomic rearrangement, or a proteasome-generated spliced antigen; and/or the subject is known or suspected to have cancer, optionally wherein stimulating the immune response treats the cancer and/or wherein the cancer is selected from the group consisting of: lung cancer, melanoma, breast cancer, ovarian cancer, prostate cancer, kidney cancer, gastric cancer, colon cancer, testicular cancer, head and neck cancer, pancreatic cancer, bladder cancer, brain cancer, B-cell lymphoma, acute myelogenous leukemia, adult acute lymphoblastic leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, T cell lymphocytic leukemia, non-small cell lung cancer, and small cell lung cancer, optionally wherein the cancer is a solid tumor, optionally wherein the cancer is selected from the group consisting of: MSS-CRC, NSCLC, and PDA; and/or wherein the subject has one or more tumors, optionally wherein stimulating the immune response reduces tumor volume of the one or more tumors.

31. (canceled)

32. The composition of claim 1, wherein the epitope-encoding nucleic acid sequence comprises an epitope selected from the group consisting of SEQ ID NO: 57-29,357 and SEQ ID NO: 29,512-29,519; and/or the at least one antigen-encoding nucleic acid sequence comprises at least each of: (A) a KRAS_G12C MHC class I epitope encoding nucleic acid sequence, (B) a KRAS_G12D MHC class I epitope encoding nucleic acid sequence, and (C) a KRAS_G12V MHC class I epitope encoding nucleic acid sequence.

33. A method for stimulating an immune response in a subject, the method comprising administering to the subject a composition for delivery of a self-amplifying alphavirus-based expression system and administering to the subject a composition for delivery of a chimpanzee adenovirus (ChAdV)-based expression system, and wherein either: a. the composition for delivery of the ChAdV-based expression system comprises the ChAdV-based expression system, wherein the ChAdV-based expression system comprises a viral particle comprising a ChAdV vector, and wherein the composition comprises 1×10.sup.12 or less of the viral particles, b, wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, and wherein the composition comprises a therapeutically effective amount comprising 30 μg or less of each of the one or more vectors, or c. the composition for delivery of the ChAdV-based expression system comprises the ChAdV-based expression system, wherein the ChAdV-based expression system comprises a viral particle comprising a ChAdV vector, and wherein the composition comprises 1×10.sup.12 or less of the viral particles and wherein the composition for delivery of the self-amplifying alphavirus-based expression system comprises the self-amplifying alphavirus-based expression system, wherein the self-amplifying alphavirus-based expression system comprises one or more vectors, and wherein the composition comprises a therapeutically effective amount comprising 30 μg or less of each of the one or more vectors.

34. The method of claim 33, wherein the composition for delivery of the ChAdV-based expression system is administered as a priming dose and the composition for delivery of the self-amplifying alphavirus-based expression system is administered as one or more boosting doses, optionally wherein the priming dose is administered on day 1 and the one or more boosting doses are administered at least every 8 weeks (Q8W) following the priming dose, further optionally wherein the one or more boosting doses are administered at least every 8 weeks for a time period, further optionally wherein the time period is the first 6 months following the priming dose, further optionally wherein one or more additional boosting doses are administered at a second interval following the time period, further optionally wherein the second interval is every 3 months.

35. The method of claim 34, wherein the priming dose is administered on day 1 and the one or more boosting doses are administered at least every 8 weeks (Q8W) following the priming dose.

36-127. (canceled)

128. A method for stimulating an immune response in a subject, the method comprising administering to the subject the composition of claim 1, wherein the self-replicating alphavirus-based expression system is administered as at least two boosting doses.

129. The method of claim 128, wherein the at least two or more boosting doses are administered (a) at least 28 days apart; (b) at least 4 weeks (Q4W) apart; (c) at least one month apart; (d) at least 56 days apart; (e) at least 8 weeks (Q8W) apart; or (f) at least 2 months apart.

130-141. (canceled)