TARGETED NEOEPITOPE VECTORS AND METHODS THEREFOR

[0001] This application claims priority to our copending US provisional patent application with the serial number 62/489,102, filed April 24, 2017.

Field of the Invention

[0002] The field of the invention is compositions and methods of improved neoepitope-based immune therapeutics, especially as it relates to recombinant nucleic acid therapeutics in the treatment of cancer.

Background of the Invention

[0003] The background description includes information that may be useful in understanding the present invention. It is not an admission that any of the information provided herein is prior art or relevant to the presently claimed invention, or that any publication specifically or implicitly referenced is prior art.

[0004] All publications and patent applications herein are incorporated by reference to the same extent as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. Where a definition or use of a term in an incorporated reference is inconsistent or contrary to the definition of that term provided herein, the definition of that term provided herein applies and the definition of that term in the reference does not apply.

[0005] Cancer immunotherapies targeting certain antigens common to a specific cancer have led to remarkable responses in some patients. Unfortunately, many patients failed to respond to such immunotherapy despite apparent expression of the same antigen. One possible reason for such failure could be that various effector cells of the immune system may not have been present in sufficient quantities, or may have been exhausted. Moreover, intracellular antigen processing and HLA variability among patients may have led to insufficient processing of the antigen and/or antigen display, leading to a therapeutically ineffective or lacking response.

[0006] To increase the selection of targets for immune therapy, random mutations have more recently been considered since some random mutations in tumor cells may give rise to unique tumor specific antigens (neoepitopes). As such, and at least conceptually, neoepitopes may provide a unique precision target for immunotherapy. Additionally, it has been shown that

cytolytic T-cell responses can be triggered by very small quantities of peptides (*e.g.*, Sykulev et al., *Immunity*, Volume 4, Issue 6, p565–571, 1 June 1996). Moreover, due to the relatively large number of mutations in many cancers, the number of possible targets is relatively high. In view of these findings, the identification of cancer neoepitopes as therapeutic targets has attracted much attention. Unfortunately, current data appear to suggest that all or almost all cancer neoepitopes are unique to a patient and specific tumor and fail to provide any specific indication as to which neoepitope may be useful for an immunotherapeutic agent that is therapeutically effective.

[0007] To overcome at least some of the problems associated with large numbers of possible targets for immune therapy, the neoepitopes can be filtered for the type of mutation (*e.g.*, to ascertain missense or nonsense mutation), the level of transcription to confirm transcription of the mutated gene, and to confirm protein expression. Moreover, the so filtered neoepitope may be further analyzed for specific binding to the patient's HLA system as described in WO 2016/172722. While such system advantageously reduces the relatively large number of potential neoepitopes, the significance of these neoepitopes with respect to treatment outcome remains uncertain. Still further, and especially where multiple peptides are to be expressed in an antigen presenting cell (*e.g.*, dendritic cell), processing of precursor proteins to generate the neoepitopes is not fully understood and contributes as such to the lack of predictability of therapeutic success.

[0008] Immune therapy can be performed using at least two conceptually distinct approaches, with the first approach based on DNA vaccination and the second approach based on use of a recombinant virus that encodes one or more antigens that are expressed in a cell infected with the virus. For example, clinical trials have suggested that plasmid DNA vaccines are safe and immunologically effective in humans at doses of 300 mcg of plasmid DNA encoding HIV rev and env proteins when administered intramuscularly. Such DNA vaccination elicited antigen-specific, IFN gamma-secreting T cell responses in HIV-seronegative patients (*J. Infect. Dis.* (2000) 181:476-83). In addition, results of a clinical trial targeting PSMA (prostate-specific membrane antigen) in patients with prostate cancer using intradermal injections of plasmid DNA and adenovirus have been reported (see *Eur. Urol.* (2000), 38:208 217). Here, 26 patients were immunized either in a prime/boost strategy with an adenoviral vector expressing PSMA followed by immunization with plasmid DNA encoding PSMA, or with plasmid DNA alone, and no significant toxicity were observed. However, therapeutic

efficacy of such vaccinations, particularly in treatment of cancer has not been demonstrated using such approaches. In still other examples, adenoviral expression of cancer neoepitopes has been reported as described in US 2017/0312351. While such approaches are highly specific towards the patient and tumor of the patient, generation of sufficient quantities of viral particles that encode one or more neoepitopes is time consuming. For example, virus production to generate a single dose of 10^{11} virus particles will often require 6-8 weeks and in some cases even longer, and multiple administrations are often required to elicit a therapeutic effect. Depending on the type of cancer and growth speed, such production time frame can be prohibitive. In addition, immune stimulation with virally expressed proteins only is often less effective, and additional treatment modalities such as cytokines are frequently required to elicit a desirable therapeutic effect.

[0009] Thus, even though multiple methods of identification and delivery of neoepitopes to various cells are known in the art, all or almost all of them suffer from various disadvantages, particularly in terms of efficacy and time requirements. Consequently, it would be desirable to have improved systems and methods for neoepitope selection and production that increases the likelihood of a therapeutic response in immune therapy in an expedient fashion.

Summary of The Invention

[0010] The inventive subject matter is directed to various immune therapeutic compositions and methods, and especially recombinant expression systems in which multiple neoepitopes are combined to form a rational-designed polypeptide with a trafficking signal to increase antigen processing and presentation and to so enhance therapeutic efficacy. Additionally, the systems and methods contemplated herein take advantage of multiple and distinct vaccination modalities that will provide both, significantly shortened time-to-first-vaccination periods and different and complementary modes of immune stimulation.

[0011] For example, where the first vaccination modality comprises a DNA vaccination that encodes a polytope (typically comprising multiple neoepitopes and/or TAAs), the vaccine can be prepared within a few days and will provide a TLR stimulus (*e.g.*, TLR9 stimulus), while the second vaccination modality may comprise a recombinant bacterial or yeast vaccine that encodes the polytope (typically the same polytope as the first) and will so provide a different TRL stimulus (*e.g.*, TRL1, TLR2, TLR5, etc.). In yet another example, the first vaccination modality may comprise a bacterial or yeast vaccine that encodes the polytope and the second

vaccination modality may comprise a recombinant virus that encodes the polytope, which once again will provide distinct (and typically complementary or even synergistic) innate immune stimuli.

[0012] As will be readily apparent, such multi-modality strategy will substantially reduce the time to generate the first vaccination as DNA, bacterial, and yeast vaccines can be prepared within a few days rather than several weeks as is the case with most viral vaccines. Moreover, due to the distinct forms of delivery, contemplated vaccine compositions and methods will also take advantage of the distinct immune stimulatory effects provided by the different modalities and will as such be particularly useful in a prime/boost regimen. Additionally, it should be recognized that contemplated systems and methods take advantage of substantially the same polytope across the different vaccine modalities. In other words, once antigens with potential therapeutic effect are determined for a patient, a recombinant nucleic acid encoding these antigens can be assembled into a polytope cassette that can then be used across multiple vaccine platforms.

[0013] In one aspect of the inventive subject matter, the inventors contemplate a method of generating recombinant expression constructs for use in immune therapy in a mammal. Such methods will typically include a step of generating a first recombinant nucleic acid having a sequence that encodes a polytope, wherein the polytope comprises a plurality of filtered neoepitope sequences, wherein the polytope further comprises a trafficking element that directs the polytope to a sub-cellular location selected from the group consisting of a recycling endosome, a sorting endosome, and a lysosome, and wherein the first recombinant nucleic acid comprises a first promoter operably linked to the sequence that encodes the polytope to drive expression of the polytope in the mammal. In another step, a second recombinant nucleic acid is generated having the same sequence that encodes the polytope, wherein the second recombinant nucleic acid comprises a second promoter operably linked to the sequence that encodes the polytope to drive expression of the polytope in a non-mammalian cell;

[0014] For example, in exemplary embodiments the first promoter may be a constitutive promoter or a promoter that is inducible by hypoxia, IFN-gamma, or IL-8. Additionally, the trafficking element may be a CD1b leader sequence, a CD1a tail, a CD1c tail, or a LAMP1-transmembrane sequence. Most typically, the filtered neoepitope sequences are filtered by comparing tumor versus matched normal of the same patient, and/or filtered to have binding

affinity to an MHC complex of equal or less than 200 nM. While in some aspects, the filtered neoepitope sequences are calculated to bind to MHC-I and the trafficking element directs the polytope to the recycling endosome, sorting endosome, or lysosome, in other aspects the filtered neoepitope sequences are calculated to bind to MHC-II and the trafficking element directs the polytope to the recycling endosome, sorting endosome, or lysosome.

[0015] Where desired, the first recombinant nucleic acid further may further comprise an additional sequence that encodes a second polytope, wherein the second polytope comprises a second trafficking element that directs the second polytope to a different sub-cellular location and wherein the second polytope comprises a second plurality of filtered neoepitope sequences. In some embodiments, at least one of the filtered neoepitope sequences and at least one of the second filtered neoepitope sequences may be the same.

[0016] It is still further contemplated that the first recombinant nucleic acid further comprises a sequence that encodes at least one of a co-stimulatory molecule, an immune stimulatory cytokine, and a protein that interferes with or down-regulates checkpoint inhibition. For example, suitable co-stimulatory molecules include CD80, CD86, CD30, CD40, CD30L, CD40L, ICOS-L, B7-H3, B7-H4, CD70, OX40L, 4-1BBL, GITR-L, TIM-3, TIM-4, CD48, CD58, TL1A, ICAM-1, and/or LFA3, while suitable immune stimulatory cytokine include IL-2, IL-12, IL-15, IL-15 super agonist (ALT803), IL-21, IPS1, and/or LMP1, and/or suitable proteins that interfere include antibodies against or antagonists of CTLA-4, PD-1, TIM1 receptor, 2B4, and/or CD160.

[0017] While not limiting to the inventive subject matter, the first recombinant nucleic acid may be replicated in a bacterial cell or a yeast cell, and/or the first recombinant nucleic acid may be an expression vector or a shuttle vector for generation of a recombinant virus (e.g., adenovirus, optionally with at least one of an E1 and an E2b gene deleted). It is also contemplated that such methods may also include a step of formulating the first recombinant nucleic acid into a pharmaceutical formulation for injection.

[0018] Most typically, the second promoter is a constitutive bacterial or a yeast promoter. Therefore, suitable non-mammalian cells include *E. coli* cell and *Saccharomyces cerevisiae*. In such cases, methods may include the additional steps of transfecting the second recombinant nucleic acid into a bacterial cell or a yeast cell; expressing the polytope in the

bacterial cell or the yeast cell; and formulating the bacterial cell or the yeast cell into a pharmaceutical formulation for injection.

[0019] Consequently, the inventors also contemplate a recombinant bacterial or yeast expression vector for immune therapy of a mammal. Most preferably, such vector will include a recombinant nucleic acid having a sequence that encodes a polytope operably linked to a bacterial or yeast promoter to drive expression of the polytope, wherein the polytope comprises a trafficking element that directs the polytope to a sub-cellular location of a mammalian immune competent cell selected from the group consisting of recycling endosome, sorting endosome, and lysosome; and wherein the polytope comprises a plurality of filtered neoepitope sequences.

[0020] Preferably, but not necessarily, the promoter is a constitutive promoter, while the trafficking element is selected from the group consisting of a CD1b leader sequence, a CD1a tail, a CD1c tail, and a LAMP1-transmembrane sequence. As noted earlier, the filtered neoepitope sequences may be filtered by comparing tumor versus matched normal of the same patient, and the filtered neoepitope sequences bind to MHC-I and/or MHC-II, and the trafficking element directs the polytope to the recycling endosome, sorting endosome, or lysosome. It is still further contemplated that the recombinant nucleic acid may also comprise an additional sequence that encodes a second polytope, wherein the second polytope comprises a second trafficking element that directs the second polytope to a different sub-cellular location and wherein the second polytope comprises a second plurality of filtered neoepitope sequences. As before, at least one of the filtered neoepitope sequences and at least one of the second filtered neoepitope sequences may be identical.

[0021] In still further contemplated aspects, the expression vector is a bacterial expression vector or a yeast expression vector. Therefore, recombinant yeast cells and bacterial cells transfected with the vector contemplated above are particularly contemplated. These cells may then be formulated into a pharmaceutical composition comprising the recombinant yeast cells or bacterial cells.

[0022] In a further aspect of the on inventive subject matter, the inventors also contemplate a method of preparing first and second treatment compositions for an individual having a tumor. Such methods will typically include a step of identifying a plurality of expressed neoepitope sequences from omics data of the tumor, wherein each of the expressed

neoepitope sequences have a calculated binding affinity of equal or less than 500 nM to at least one of MHC-I and MHC-II of the individual, and a further step of generating a first recombinant nucleic acid having a sequence that encodes a polytope, wherein the polytope comprises the plurality of expressed neoepitope sequences. Preferably, the polytope further comprises a trafficking element that directs the polytope to a sub-cellular location selected from the group consisting of a recycling endosome, a sorting endosome, and a lysosome, and the first recombinant nucleic acid comprises a first promoter operably linked to the sequence that encodes the polytope to drive expression of the polytope in a cell of the individual. In another step, the first recombinant nucleic is formulated into a DNA vaccine formulation to so obtain the first treatment composition. In yet another step, a second recombinant nucleic acid is generated that includes the sequence that encodes the polytope, wherein the second recombinant nucleic acid comprises a second promoter operably linked to the sequence that encodes the polytope to drive expression of the polytope in a bacterial cell or a yeast cell, and in a further step, the bacterial cell or the yeast cell is transfected with the second recombinant nucleic acid and expressing the polytope in the bacterial cell or the yeast cell. In a still further step, the transfected bacterial cell or the yeast cell is formulated into a cell-based vaccine formulation to so obtain the second treatment composition.

[0023] Most typically, the expressed neoepitope sequences are identified using incremental synchronous alignment of omics data from the tumor and omics data from a non-tumor sample of the same individual. It is further generally preferred that the first recombinant nucleic acid is an expression vector, and/or that the trafficking element is a CD1b leader sequence, a CD1a tail, a CD1c tail, or a LAMP1-transmembrane sequence. The second promoter is preferably a constitutive bacterial or a yeast promoter, and the bacterial cell or the yeast cell is preferably *E. coli* cell or *Saccharomyces cerevisiae.* Most typically, the cell-based vaccine formulation is formulated for injection. Furthermore, where desired, such method may further comprise a step of generating a third recombinant nucleic acid that is a viral expression vector that includes the sequence that encodes the polytope, wherein the third recombinant nucleic acid comprises a third promoter operably linked to the sequence that encodes the polytope to drive expression of the polytope in a cell of the individual.

[0024] In still another aspect of the on inventive subject matter, the inventors also contemplate a method of preparing first and second treatment compositions for an individual having a tumor that includes a step of identifying a plurality of expressed neoepitope

sequences from omics data of the tumor, wherein the expressed neoepitope sequences have a calculated binding affinity of equal or less than 500 nM to at least one of MHC-I and MHC-II of the individual; and a further step of generating a first recombinant nucleic acid having a sequence that encodes a polytope, wherein the polytope comprises the plurality of expressed neoepitope sequences, wherein the first recombinant nucleic acid is a viral expression vector. Preferably, the polytope further comprises a trafficking element that directs the polytope to a sub-cellular location selected from the group consisting of a recycling endosome, a sorting endosome, and a lysosome, and the first recombinant nucleic acid comprises a first promoter operably linked to the sequence that encodes the polytope to drive expression of the polytope in a cell of the individual. In a further step, viral particles are generated from the viral expression vector and the viral particles are formulated into a viral vaccine formulation to so obtain the first treatment composition. Such methods will also include a step of generating a second recombinant nucleic acid having the sequence that encodes the polytope, wherein the second recombinant nucleic acid comprises a second promoter operably linked to the sequence that encodes the polytope to drive expression of the polytope in a non-mammalian cell, and a further step of transfecting a bacterial cell or a yeast cell with the second recombinant nucleic acid and expressing the polytope in the bacterial cell or the yeast cell. The so transfected bacterial cell or the yeast cell is then formulated into a cell-based vaccine formulation to so obtain the second treatment composition.

[0025] Most typically, the plurality of expressed neoepitope sequences are identified using incremental synchronous alignment of omics data from the tumor and omics data from a non-tumor sample of the same individual, and/or the trafficking element is a CD1b leader sequence, a CD1a tail, a CD1c tail, or a LAMP1-transmembrane sequence. Likewise, it is preferred that the first promoter is a constitutive promoter or that the first promoter is inducible by hypoxia, IFN-gamma, or IL-8, and/or the second promoter is a constitutive bacterial or a yeast promoter. In further contemplated embodiments, the viral expression vector is an adenoviral expression vector, optionally having E1 and E2b genes deleted. While not limiting to the inventive subject matter, it is generally preferred that the non-mammalian cell or the yeast cell is an *E. coli* cell or a *Saccharomyces cerevisiae* cell, and/or that the viral vaccine formulation and the cell-based vaccine formulation are both formulated for injection.

[0026] Various objects, features, aspects and advantages of the inventive subject matter will become more apparent from the following detailed description of preferred embodiments, along with the accompanying drawing in which like numerals represent like components.

Brief Description of The Drawing

[0027] Figure 1 is a schematic representation of various neoepitope arrangements.

[0028] Figure 2 is an exemplary and partial schematic for selecting preferred arrangements of neoepitopes.

[0029] Prior Art Figure 3 is a schematic illustration of antigen processing in the cytoplasm and MHC-I presentation.

[0030] Prior Art Figure 4 is a schematic illustration of antigen processing in the lysosomal and endosomal compartment and MHC-II presentation.

[0031] Figures 5A-5C are exemplary sequence arrangements for class I antigen processing in the cytoplasm and MHC-I presentation.

[0032] Figures 6A-6C are exemplary sequence arrangements for class I antigen processing in the cytoplasm and MHC-II presentation.

[0033] Figures 7A-7C are exemplary sequence arrangements for class II antigen processing in the cytoplasm and MHC-II presentation.

[0034] Figure 8 is an exemplary prophylactic vaccination schedule for B16-F10 melanoma.

[0035] Figures 9A-9C are graphs depicting exemplary results for the anti-tumor vaccination using subcutaneous injection of the vaccine.

[0036] Figures 10A-10C are graphs depicting exemplary results for the anti-tumor vaccination using intravenous injection of the vaccine.

[0037] Figures 11A-11E are graphs depicting exemplary results for selected DNA and viral vaccine compositions and selected routes of administration.

[0038] Figure 12 is a graph depicting exemplary results for the anti-tumor effect using DNA and viral vaccines with various compositions and routes.

Detailed Description

[0039] The inventors have now discovered that various aspects in tumor antigen-based and/or neoepitope-based immune therapy can further be improved by not only targeting the antigens or neoepitopes towards specific processing and cell surface presentation pathways, but also by using different vaccine modalities that preferably trigger different immune stimulatory pathways.

[0040] Therefore, in addition to a viral cancer vaccine that is based on a recombinant virus that triggers in a host cell expression of tumor associated or tumor specific antigens, the same (and/or additional) antigens may be expressed from a DNA vaccine and/or provided in yeast and/or bacterial cells that are genetically engineered to express these antigens. For example, where a plasmid is used in a DNA vaccine, innate immune response mechanisms against free DNA (*e.g.*, TLR9- or STING-based) may be triggered along with the adaptive immune response based on *in vivo* expression of the free DNA. In another example, where a viral expression vector is employed as part of a viral vaccine in which a virus infects patient cells, such infection will typically trigger different innate immune response mechanisms (typically TLR2, TLR4, TLR 7, TLR 8, TLR 9). In still another example, where a bacterial or yeast vaccine is used in which the bacterium or yeast has expressed the antigen(s), such vaccine vaccination will once more trigger distinct innate immune response mechanisms (typically TLR1-3 for bacterial and TLR1-4 for yeast). As will be readily appreciated, triggering of various and distinct innate immune response mechanisms may provide complementary or even synergistic enhancement of the vaccine compositions.

[0041] Thus, it should be appreciated that the compositions and methods presented herein will preferably include use of at least two different vaccine modalities. For example, the first modality may be a modality selected from the group consisting of a DNA vaccine, protein vaccine, a bacterial vaccine, a yeast vaccine, and a viral vaccine, while the second/subsequent modality may be another, different, modality selected from the same group. Most preferably, the antigens present in any of the modalities will overlap or be the same to take advantage of a prime/boost effect upon repeated antigenic challenge.

[0042] In addition, it should be recognized that beyond the benefit of triggering multiple distinct innate immune pathways, contemplated compositions and methods also allow for a rapid start of treatment if a patient with respect to the point in time at which the tumor

relevant antigens in the patient were identified. Indeed, it should be recognized that a DNA vaccine can be prepared based on the relevant antigens within a few days, typically within less than a week and even less than 4 days, or even less 48 hours. Moreover, bacterial or yeast vaccines can be also prepared using (the same) antigens within a few days, typically within less than 2 weeks, and more typically within less than 1 week. At the same time, a viral vaccine can be prepared when the patient already received the first vaccines (*e.g.*, DNA vaccine, bacterial vaccine, and/or yeast vaccine).

[0043] Viewed from a different perspective, it should be appreciated that the compositions and methods presented herein will include one or more tumor associated antigens, tumor specific antigens, and/or neoepitopes that are specific to the patient and the tumor in the patient to allow for targeted treatment. Moreover, such treatment may advantageously be tailored to achieve one or more specific immune reactions, including an innate immune response, a $CD4^+$ biased immune response, a $CD8^+$ biased immune response, antibody biased immune response, and/or a stimulated immune response (*e.g.*, reducing checkpoint inhibition and/or by activation of immune competent cells using cytokines). Most typically, such effects are in achieved in the context of the neoepitopes originating from the recombinant nucleic acid that can be administered via one or more routes in one or more distinct formats (*e.g.*, as recombinant plasmid and as recombinant virus).

[0044] Antigens

[0045] With respect to suitable therapeutic antigens it is contemplated that various antigens are deemed suitable for use herein. However, particularly preferred antigens include tumor associated antigens (*e.g.*, CEA, MUC1, brachyury), tumor specific antigens (*e.g.*, HER2, PSA, PSMA, etc.), and especially tumor and patient specific antigens (*i.e.*, neoepitopes). Neoepitopes can be characterized as expressed random mutations in tumor cells that created unique and tumor specific antigens. Therefore, viewed from a different perspective, neoepitopes may be identified by considering the type (*e.g.*, deletion, insertion, transversion, transition, translocation) and impact of the mutation (*e.g.*, non-sense, missense, frame shift, etc.), which may as such serve as a content filter through which silent and other non-relevant (*e.g.*, non-expressed) mutations are eliminated. It should also be appreciated that neoepitope sequences can be defined as sequence stretches with relatively short length (*e.g.*, 8-12 mers or 14-20mers) wherein such stretches will include the change(s) in the amino acid sequences. Most typically, but not necessarily, the changed amino acid will be at or near the central

amino acid position. For example, a typical neoepitope may have the structure of A_4-N-A_4, or A_3-N-A_5, or A_2-N-A_7, or A_5-N-A_3, or A_7-N-A_2, where A is a proteinogenic wild type or normal (*i.e.*, from corresponding healthy tissue of the same patient) amino acid and N is a changed amino acid (relative to wild type or relative to matched normal). Therefore, the neoepitope sequences contemplated herein include sequence stretches with relatively short length (*e.g.*, 5-30 mers, more typically 8-12 mers, or 14-20 mers) wherein such stretches include the change(s) in the amino acid sequences. Where desired, additional amino acids may be placed upstream or downstream of the changed amino acid, for example, to allow for additional antigen processing in the various compartments (*e.g.*, for proteasome processing in the cytosol, or specific protease processing in the endosomal and/or lysosomal compartments) of a cell.

[0046] Thus, it should be appreciated that a single amino acid change may be presented in numerous neoepitope sequences that include the changed amino acid, depending on the position of the changed amino acid. Advantageously, such sequence variability allows for multiple choices of neoepitopes and as such increases the number of potentially useful targets that can then be selected on the basis of one or more desirable traits (*e.g.*, highest affinity to a patient HLA-type, highest structural stability, etc.). Most typically, neoepitopes will be calculated to have a length of between 2-50 amino acids, more typically between 5-30 amino acids, and most typically between 8-12 amino acids, or 14-20 amino acids, with the changed amino acid preferably centrally located or otherwise situated in a manner that improves its binding to MHC. For example, where the epitope is to be presented by the MHC-I complex, a typical neoepitope length will be about 8-12 amino acids, while the typical neoepitope length for presentation via MHC-II complex will have a length of about 14-20 amino acids. As will be readily appreciated, since the position of the changed amino acid in the neoepitope may be other than central, the actual peptide sequence and with that actual topology of the neoepitope may vary considerably, and the neoepitope sequence with a desired binding affinity to the MHC-I or MHC-II presentation and/or desired protease processing will typically dictate the particular sequence.

[0047] Of course, it should be appreciated that the identification or discovery of neoepitopes may start with a variety of biological materials, including fresh biopsies, frozen, or otherwise preserved tissue or cell samples, circulating tumor cells, exosomes, various body fluids (and especially blood), etc. Therefore, suitable methods of omics analysis include nucleic acid

sequencing, and particularly NGS methods operating on DNA (*e.g.*, Illumina sequencing, ion torrent sequencing, 454 pyrosequencing, nanopore sequencing, etc.), RNA sequencing (*e.g.*, RNAseq, reverse transcription based sequencing, etc.), and in some cases protein sequencing or mass spectroscopy based sequencing (*e.g.*, SRM, MRM, CRM, etc.).

[0048] As such, and particularly for nucleic acid based sequencing, it should be particularly recognized that high-throughput genome sequencing of a tumor tissue will allow for rapid identification of neoepitopes. However, it must be appreciated that where the so obtained sequence information is compared against a standard reference, the normally occurring inter-patient variation (*e.g.*, due to SNPs, short indels, different number of repeats, etc.) as well as heterozygosity will result in a relatively large number of potential false positive neoepitopes. Notably, such inaccuracies can be eliminated where a tumor sample of a patient is compared against a matched normal (*i.e.*, non-tumor) sample of the same patient.

[0049] In one especially preferred aspect of the inventive subject matter, DNA analysis is performed by whole genome sequencing and/or exome sequencing (typically at a coverage depth of at least 10x, more typically at least 20x) of both tumor and matched normal sample. Alternatively, DNA data may also be provided from an already established sequence record (*e.g.*, SAM, BAM, FASTA, FASTQ, or VCF file) from a prior sequence determination of the same patient. Therefore, data sets suitable for use herein include unprocessed or processed data sets, and exemplary preferred data sets include those having BAM format, SAM format, GAR format, FASTQ format, or FASTA format, as well as BAMBAM, SAMBAM, and VCF data sets. However, it is especially preferred that the data sets are provided in BAM format or as BAMBAM diff objects as is described in US2012/0059670A1 and US2012/0066001A1. Moreover, it should be noted that the data sets are reflective of a tumor and a matched normal sample of the same patient. Thus, genetic germ line alterations not giving rise to the tumor (*e.g.*, silent mutation, SNP, etc.) can be excluded. Of course, it should be recognized that the tumor sample may be from an initial tumor, from the tumor upon start of treatment, from a recurrent tumor and/or metastatic site, etc. In most cases, the matched normal sample of the patient is blood, or a non-diseased tissue from the same tissue type as the tumor.

[0050] Likewise, the computational analysis of the sequence data may be performed in numerous manners. In most preferred methods, however, analysis is performed *in silico* by incremental location-guided synchronous alignment of tumor and normal samples as, for example, disclosed in US 2012/0059670 and US 2012/0066001 using BAM files and BAM

servers. Such analysis advantageously reduces false positive neoepitopes and significantly reduces demands on memory and computational resources.

[0051] It should be noted that any language directed to a computer should be read to include any suitable combination of computing devices, including servers, interfaces, systems, databases, agents, peers, engines, controllers, or other types of computing devices operating individually or collectively. One should appreciate the computing devices comprise a processor configured to execute software instructions stored on a tangible, non-transitory computer readable storage medium (*e.g.*, hard drive, solid state drive, RAM, flash, ROM, etc.). The software instructions preferably configure the computing device to provide the roles, responsibilities, or other functionality as discussed below with respect to the disclosed apparatus. Further, the disclosed technologies can be embodied as a computer program product that includes a non-transitory computer readable medium storing the software instructions that causes a processor to execute the disclosed steps associated with implementations of computer-based algorithms, processes, methods, or other instructions. In especially preferred embodiments, the various servers, systems, databases, or interfaces exchange data using standardized protocols or algorithms, possibly based on HTTP, HTTPS, AES, public-private key exchanges, web service APIs, known financial transaction protocols, or other electronic information exchanging methods. Data exchanges among devices can be conducted over a packet-switched network, the Internet, LAN, WAN, VPN, or other type of packet switched network; a circuit switched network; cell switched network; or other type of network.

[0052] Viewed from a different perspective, a patient- and cancer-specific *in silico* collection of sequences can be established that encode neoepitopes having a predetermined length of, for example, between 5 and 25 amino acids and include at least one changed amino acid. Such collection will typically include for each changed amino acid at least two, at least three, at least four, at least five, or at least six members in which the position of the changed amino acid is not identical. Such collection advantageously increases potential candidate molecules suitable for immune therapy and can then be used for further filtering (*e.g.*, by sub-cellular location, transcription/expression level, MHC-I and/or II affinity, etc.) as is described in more detail below. Of course, it should be appreciated that these neoepitope sequences can be readily back-translated into the corresponding nucleic acid sequences to so generate a nucleic acid sequence that encodes the neoepitope. Most typically, but not necessarily, such back-

translation will take into account the proper codon usage of the organism in which the nucleic acid is being expressed.

[0053] For example, and using synchronous location guided analysis to tumor and matched normal sequence data, the inventors previously identified various cancer neoepitopes from a variety of cancers and patients, including the following cancer types: BLCA, BRCA, CESC, COAD, DLBC, GBM, HNSC, KICH, KIRC, KIRP, LAML, LGG, LIHC, LUAD, LUSC, OV, PRAD, READ, SARC, SKCM, STAD, THCA, and UCEC. Exemplary neoepitope data for these cancers can be found in International application WO 2016/172722, incorporated by reference herein.

[0054] Depending on the type and stage of the cancer, as well as the patient's immune status it should be recognized that not all of the identified neoepitopes will necessarily lead to a therapeutically equally effective reaction in a patient. Indeed, it is well known in the art that only a fraction of neoepitopes will generate an immune response. To increase likelihood of a therapeutically desirable response, the initially identified neoepitopes can be further filtered. Of course, it should be appreciated that downstream analysis need not take into account silent mutations for the purpose of the methods presented herein. However, preferred mutation analyses will provide in addition to the particular type of mutation (*e.g.*, deletion, insertion, transversion, transition, translocation) also information of the impact of the mutation (*e.g.*, non-sense, missense, etc.) and may as such serve as a first content filter through which silent mutations are eliminated. For example, neoepitopes can be selected for further consideration where the mutation is a frame-shift, non-sense, and/or missense mutation.

[0055] In a further filtering approach, neoepitopes may also be subject to detailed analysis for sub-cellular location parameters. For example, neoepitope sequences may be selected for further consideration if the neoepitopes are identified as having a membrane associated location (*e.g.*, are located at the outside of a cell membrane of a cell) and/or if an *in silico* structural calculation confirms that the neoepitope is likely to be solvent exposed, or presents a structurally stable epitope (*e.g.*, *J Exp Med* 2014), etc.

[0056] With respect to filtering neoepitopes, it is generally contemplated that neoepitopes are especially suitable for use herein where omics (or other) analysis reveals that the neoepitope is actually expressed. Identification of expression and expression level of a neoepitope can be performed in all manners known in the art and preferred methods include quantitative

RNA (hnRNA or mRNA) analysis and/or quantitative proteomics analysis. Most typically, the threshold level for inclusion of neoepitopes will be an expression level of at least 20%, at least 30%, at least 40%, or at least 50% of expression level of the corresponding matched normal sequence, thus ensuring that the (neo)epitope is at least potentially 'visible' to the immune system. Consequently, it is generally preferred that the omics analysis also includes an analysis of gene expression (transcriptomic analysis) to so help identify the level of expression for the gene with a mutation.

[0057] There are numerous methods of transcriptomic analysis known in the art, and all of the known methods are deemed suitable for use herein. For example, preferred materials include mRNA and primary transcripts (hnRNA), and RNA sequence information may be obtained from reverse transcribed polyA$^+$-RNA, which is in turn obtained from a tumor sample and a matched normal (healthy) sample of the same patient. Likewise, it should be noted that while polyA$^+$-RNA is typically preferred as a representation of the transcriptome, other forms of RNA (hn-RNA, non-polyadenylated RNA, siRNA, miRNA, etc.) are also deemed suitable for use herein. Preferred methods include quantitative RNA (hnRNA or mRNA) analysis and/or quantitative proteomics analysis, especially including RNAseq. In other aspects, RNA quantification and sequencing is performed using RNAseq, qPCR and/or rtPCR based methods, although various alternative methods (*e.g.*, solid phase hybridization-based methods) are also deemed suitable. Viewed from another perspective, transcriptomic analysis may be suitable (alone or in combination with genomic analysis) to identify and quantify genes having a cancer- and patient-specific mutation.

[0058] Similarly, proteomics analysis can be performed in numerous manners to ascertain actual translation of the RNA of the neoepitope, and all known manners of proteomics analysis are contemplated herein. However, particularly preferred proteomics methods include antibody-based methods and mass spectroscopic methods. Moreover, it should be noted that the proteomics analysis may not only provide qualitative or quantitative information about the protein *per se*, but may also include protein activity data where the protein has catalytic or other functional activity. One exemplary technique for conducting proteomic assays is described in US 7473532, incorporated by reference herein. Further suitable methods of identification and even quantification of protein expression include various mass spectroscopic analyses (*e.g.*, selective reaction monitoring (SRM), multiple reaction monitoring (MRM), and consecutive reaction monitoring (CRM)). Consequently, it

should be appreciated that the above methods will provide patient and tumor specific neoepitopes, which may be further filtered by sub-cellular location of the protein containing the neoepitope (*e.g.*, membrane location), the expression strength (*e.g.*, overexpressed as compared to matched normal of the same patient), etc.

[0059] In yet another aspect of filtering, the neoepitopes may be compared against a database that contains known human sequences (*e.g.*, of the patient or a collection of patients) to so avoid use of a human-identical sequence. Moreover, filtering may also include removal of neoepitope sequences that are due to SNPs in the patient where the SNPs are present in both the tumor and the matched normal sequence. For example, dbSNP (The Single Nucleotide Polymorphism Database) is a free public archive for genetic variation within and across different species developed and hosted by the National Center for Biotechnology Information (NCBI) in collaboration with the National Human Genome Research Institute (NHGRI). Although the name of the database implies a collection of one class of polymorphisms only (single nucleotide polymorphisms (SNPs)), it in fact contains a relatively wide range of molecular variation: (1) SNPs, (2) short deletion and insertion polymorphisms (indels/DIPs), (3) microsatellite markers or short tandem repeats (STRs), (4) multinucleotide polymorphisms (MNPs), (5) heterozygous sequences, and (6) named variants. The dbSNP accepts apparently neutral polymorphisms, polymorphisms corresponding to known phenotypes, and regions of no variation. Using such database and other filtering options as described above, the patient and tumor specific neoepitopes may be filtered to remove those known sequences, yielding a sequence set with a plurality of neoepitope sequences having substantially reduced false positives.

[0060] Once the desired level of filtering for the neoepitope is accomplished (*e.g.*, neoepitope filtered by tumor versus normal, and/or expression level, and/or sub-cellular location, and/or patient specific HLA-match, and/or known variants), a further filtering step is contemplated that takes into account the gene type that is affected by the neoepitope. For example, suitable gene types include cancer driver genes, genes associated with regulation of cell division, genes associated with apoptosis, and genes associated with signal transduction. However, in especially preferred aspects, cancer driver genes are particularly preferred (which may span by function a variety of gene types, including receptor genes, signal transduction genes, transcription regulator genes, etc.). In further contemplated aspects, suitable gene types may also be known passenger genes and genes involved in metabolism.

[0061] With respect to the identification or other determination (*e.g.*, prediction) of a gene as being a cancer driver gene, various methods and prediction algorithms are known in the art, and are deemed suitable for use herein. For example, suitable algorithms include MutsigCV (*Nature* 2014, 505(7484):495–501), ActiveDriver (*Mol Syst Biol* 2013, 9:637), MuSiC (*Genome Res* 2012, 22(8):1589–1598), OncodriveClust (*Bioinformatics* 2013, 29(18):2238–2244), OncodriveFM (*Nucleic Acids Res* 2012,40(21):e169), OncodriveFML (*Genome Biol* 2016, 17(1):128), Tumor Suppressor and Oncogenes (TUSON) (*Cell* 2013, 155(4):948–962), 20/20+ (https://github.com/KarchinLab/2020plus), and oncodriveROLE (*Bioinformatics* (2014) 30 (17): i549-i555). Alternatively, or additionally, identification of cancer driver genes may also employ various sources for known cancer driver genes and their association with specific cancers. For example, the Intogen Catalog of driver mutations (2016.5; URL: www.intogen.org) contains the results of the driver analysis performed by the Cancer Genome Interpreter across 6,792 exomes of a pan-cancer cohort of 28 tumor types.

[0062] Nevertheless, despite filtering, it should be recognized that not all neoepitopes will be visible to the immune system as the neoepitopes also need to be processed where present in a larger context (*e.g.*, within a polytope) and presented on the MHC complex of the patient. In that context, it must be appreciated that only a fraction of all neoepitopes will have sufficient affinity for presentation. Consequently, and especially in the context of immune therapy it should be apparent that neoepitopes will be more likely effective where the neoepitopes are properly processed, bound to, and presented by the MHC complexes. Viewed from another perspective, treatment success will be increased with an increasing number of neoepitopes that can be presented via the MHC complex, wherein such neoepitopes have a minimum affinity to the patient's HLA-type. Consequently, it should be appreciated that effective binding and presentation is a combined function of the sequence of the neoepitope and the particular HLA-type of a patient. Therefore, HLA-type determination of the patient tissue is typically required. Most typically, the HLA-type determination includes at least three MHC-I sub-types (*e.g.*, HLA-A, HLA-B, HLA-C) and at least three MHC-II sub-types (*e.g.*, HLA-DP, HLA-DQ, HLA-DR), preferably with each subtype being determined to at least 2-digit or at least 4-digit depth. However, greater depth (*e.g.*, 6 digit, 8 digit) is also contemplated.

[0063] Once the HLA-type of the patient is ascertained (using known chemistry or *in silico* determination), a structural solution for the HLA-type is calculated and/or obtained from a database, which is then used in a docking model *in silico* to determine binding affinity of the

(typically filtered) neoepitope to the HLA structural solution. As will be further discussed below, suitable systems for determination of binding affinities include the NetMHC platform (see *e.g.*, Nucleic Acids Res. 2008 Jul 1; 36(Web Server issue): W509–W512.). Neoepitopes with high affinity (*e.g.*, less than 100 nM, less than 75 nM, less than 50 nM) for a previously determined HLA-type are then selected for therapy creation, along with the knowledge of the patient's MHC-I/II subtype.

[0064] HLA determination can be performed using various methods in wet-chemistry that are well known in the art, and all of these methods are deemed suitable for use herein. However, in especially preferred methods, the HLA-type can also be predicted from omics data *in silico* using a reference sequence containing most or all of the known and/or common HLA-types. For example, in one preferred method according to the inventive subject matter, a relatively large number of patient sequence reads mapping to chromosome 6p21.3 (or any other location near/at which HLA alleles are found) is provided by a database or sequencing machine. Most typically the sequence reads will have a length of about 100-300 bases and comprise metadata, including read quality, alignment information, orientation, location, etc. For example, suitable formats include SAM, BAM, FASTA, GAR, etc. While not limiting to the inventive subject matter, it is generally preferred that the patient sequence reads provide a depth of coverage of at least 5x, more typically at least 10x, even more typically at least 20x, and most typically at least 30x.

[0065] In addition to the patient sequence reads, contemplated methods further employ one or more reference sequences that include a plurality of sequences of known and distinct HLA alleles. For example, a typical reference sequence may be a synthetic (without corresponding human or other mammalian counterpart) sequence that includes sequence segments of at least one HLA-type with multiple HLA-alleles of that HLA-type. For example, suitable reference sequences include a collection of known genomic sequences for at least 50 different alleles of HLA-A. Alternatively, or additionally, the reference sequence may also include a collection of known RNA sequences for at least 50 different alleles of HLA-A. Of course, and as further discussed in more detail below, the reference sequence is not limited to 50 alleles of HLA-A, but may have alternative composition with respect to HLA-type and number/composition of alleles. Most typically, the reference sequence will be in a computer readable format and will be provided from a database or other data storage device. For example, suitable reference sequence formats include FASTA, FASTQ, EMBL, GCG, or GenBank format, and may be

directly obtained or built from data of a public data repository (*e.g.*, IMGT, the International ImMunoGeneTics information system, or The Allele Frequency Net Database, EUROSTAM, URL: www.allelefrequencies.net). Alternatively, the reference sequence may also be built from individual known HLA-alleles based on one or more predetermined criteria such as allele frequency, ethnic allele distribution, common or rare allele types, etc.

[0066] Using the reference sequence, the patient sequence reads can now be threaded through a de Bruijn graph to identify the alleles with the best fit. In this context, it should be noted that each individual carries two alleles for each HLA-type, and that these alleles may be very similar, or in some cases even identical. Such high degree of similarity poses a significant problem for traditional alignment schemes. The inventor has now discovered that the HLA alleles, and even very closely related alleles can be resolved using an approach in which the de Bruijn graph is constructed by decomposing a sequence read into relatively small k-mers (typically having a length of between 10-20 bases), and by implementing a weighted vote process in which each patient sequence read provides a vote ("quantitative read support") for each of the alleles on the basis of k-mers of that sequence read that match the sequence of the allele. The cumulatively highest vote for an allele then indicates the most likely predicted HLA allele. In addition, it is generally preferred that each fragment that is a match to the allele is also used to calculate the overall coverage and depth of coverage for that allele.

[0067] Scoring may further be improved or refined as needed, especially where many of the top hits are similar (*e.g.*, where a significant portion of their score comes from a highly shared set of k-mers). For example, score refinement may include a weighting scheme in which alleles that are substantially similar (e.g., > 99%, or other predetermined value) to the current top hit are removed from future consideration. Counts for k-mers used by the current top hit are then re-weighted by a factor (*e.g.*, 0.5), and the scores for each HLA allele are recalculated by summing these weighted counts. This selection process is repeated to find a new top hit. The accuracy of the method can be even further improved using RNA sequence data that allows identification of the alleles expressed by a tumor, which may sometimes be just 1 of the 2 alleles present in the DNA. In further advantageous aspects of contemplated systems and methods, DNA or RNA, or a combination of both DNA and RNA can be processed to make HLA predictions that are highly accurate and can be derived from tumor or blood DNA or RNA. Further aspects, suitable methods and considerations for high-

accuracy *in silico* HLA typing are described in WO 2017/035392, incorporated by reference herein.

[0068] Once patient and tumor specific neoepitopes and HLA-type are identified, further computational analysis can be performed by *in silico* docking neoepitopes to the HLA and determining best binders (*e.g.*, lowest K_D, for example, less than 500nM, or less than 250nM, or less than 150nM, or less than 50nM), for example, using NetMHC. It should be appreciated that such approach will not only identify specific neoepitopes that are genuine to the patient and tumor, but also those neoepitopes that are most likely to be presented on a cell and as such most likely to elicit an immune response with therapeutic effect. Of course, it should also be appreciated that thusly identified HLA-matched neoepitopes can be biochemically validated *in vitro* prior to inclusion of the nucleic acid encoding the epitope as payload into the virus as is further discussed below.

[0069] Of course, it should be appreciated that matching of the patient's HLA-type to the patient- and cancer-specific neoepitope can be done using systems other than NetMHC, and suitable systems include NetMHC II, NetMHCpan, IEDB Analysis Resource (URL immuneepitope.org), RankPep, PREDEP, SVMHC, Epipredict, HLABinding, and others (see e.g., *J Immunol Methods* 2011;374:1–4). In calculating the highest affinity, it should be noted that the collection of neoepitope sequences in which the position of the altered amino acid is moved (supra) can be used. Alternatively, or additionally, modifications to the neoepitopes may be implemented by adding N- and/or C-terminal modifications to further increase binding of the expressed neoepitope to the patient's HLA-type. Thus, neoepitopes may be native as identified or further modified to better match a particular HLA-type. Moreover, where desired, binding of corresponding wild type sequences (*i.e.*, neoepitope sequence without amino acid change) can be calculated to ensure high differential affinities. For example, especially preferred high differential affinities in MHC binding between the neoepitope and its corresponding wild type sequence are at least 2-fold, at least 5-fold, at least 10-fold, at least 100-fold, at least 500-fold, at least 1000-fold, etc.).

[0070] Binding affinity and particularly differential binding affinity may also be determined *in vitro* using various systems and methods. For example, antigen presenting cells of a patient or cells with matched HLA-type can be transfected with a nucleic acid (*e.g.*, viral, plasmid, linear DNA, RNA, etc.) to express one or more neoepitopes using constructs as described in more detail below. Upon expression and antigen processing, the neoepitopes can then be

identified in the MHC complex on the outside of the cell, either using specific binders to the neoepitope or using a cell based system (*e.g.*, PBMC of the patient) in which T cell activation or cytotoxic NK cell activity can be observed *in vitro*. Neoepitopes with differential activity (elicit a stronger signal or immune response as compared to the corresponding wild type epitope) will then be selected for therapy creation.

[0071] Recombinant Nucleic Acids/Polytopes

[0072] Upon proper selection of filtered neoepitopes (or tumor associated antigens or tumor specific antigens), a recombinant nucleic acid can be constructed that forms the basis of all downstream vaccine compositions. Most typically, the desired nucleic acid sequences (for expression from virus infected cells) are under the control of appropriate regulatory elements well known in the art. As will also be readily appreciated, the choice of regulatory elements will be dictated by the system in which the recombinant nucleic acid is to be expressed. Therefore, suitable regulatory elements include constitutively active or inducible bacterial and yeast promoters (and associated inducer and/or repressor sequences where desired), as well as eukaryotic (and preferably mammal/human) promoter sequences. For example, where the recombinant nucleic acid is used in a DNA vaccine, suitable promoter elements include constitutive strong promoters (*e.g.*, SV40, CMV, UBC, EF1A, PGK, CAGG promoter). On the other hand, where the recombinant nucleic acid is part of a viral expression vector, contemplated promoters also include inducible promoters, particularly where induction conditions are typical for a tumor microenvironment. For example, inducible promoters include those sensitive to hypoxia and promoters that are sensitive to TGF-β or IL-8 (e.g., via TRAF, JNK, Erk, or other responsive elements promoter). In other examples, suitable inducible promoters include the tetracycline-inducible promoter, the myxovirus resistance 1 (Mx1) promoter, etc.

[0073] Similarly, where the recombinant nucleic acid is used to generate a bacterial and/or yeast vaccine in which the bacterium or yeast expresses the neoepitope or other therapeutic antigen, suitable promoters include strong constitutive or inducible bacterial and yeast promoters. For example, suitable bacterial promoters for expression of antigens/a polytope include the T7 promoter, the Tac promoter, the BAD promoter, the Trc promoter, etc. Likewise, yeast contemplated yeast promoters include the AOX1 promoter, the GAL promoter, the GDS promoter, the ADH promoter, etc.

[0074] In this context, it should be appreciated that the inventors have discovered that the manner of neoepitope arrangement and rational-designed trafficking of the neoepitopes can have a substantial impact on the efficacy of various immune therapeutic compositions as is further described in more detail below. For example, single neoepitopes can be expressed individually from the respective recombinant constructs that are delivered as a single plasmid, viral expression construct, etc. Alternatively, multiple neoepitopes can be separately expressed from individual promoters to form individual mRNA that are then individually translated into the respective neoepitopes, or from a single mRNA comprising individual translation starting points for each neoepitope sequence (*e.g.*, using 2A or IRES signals). Notably, while such arrangements are generally thought to allow for controlled delivery of proper neoepitope peptide, efficacy of such expression systems has been less than desirable (data not shown).

[0075] In contrast, where multiple neoepitopes were expressed from a single transcript to so form a single transcript that is then translated into a single polytope (*i.e.*, polypeptide with a series of concatemerically linked neoepitopes, optionally with intervening linker sequences) expression, processing, and antigen presentation was found to be effective. Notably, the expression of polytopes requires processing by the appropriate proteases (*e.g.*, proteasome, endosomal proteases, lysosomal proteases) within a cell to yield the neoepitope sequences, and polytopes led to improved antigen processing and presentation for most neoepitopes as compared to expression of individual neoepitopes, particularly where the individual neoepitopes had a relatively short length (*e.g.*, less than 25 amino acids; results not shown). Moreover, such approach also allows rational design of protease sensitive sequence motifs between the neoepitope peptide sequences to so assure or avoid processing by specific proteases as the proteasome, endosomal proteases, and lysosomal proteases have distinct cleavage preferences. Therefore, polytopes may be designed that include not only linker sequences to spatially separate neoepitopes, but also sequence portions (*e.g.*, between 3-15 amino acids) that will be preferentially cleaved by a specific protease.

[0076] Therefore, the inventors contemplate recombinant nucleic acids and expression vectors (*e.g.*, viral expression vectors) that comprise a nucleic acid segment that encodes a polytope wherein the polytope is operably coupled to a desired promoter element, and wherein individual neoepitopes are optionally separated by a linker and/or protease cleavage or recognition sequence. For example, **Figure 1** exemplarily illustrates various contemplated

arrangements for neoepitopes for expression from an adenoviral expression system (here: AdV5, with deletion of E1 and E2b genes). Here, Construct 1 exemplarily illustrates a neoepitope arrangement that comprises eight neoepitopes ('minigene') with a total length of 15 amino acids in concatemeric series without intervening linker sequences, while Construct 2 shows the arrangement of Construct 1 but with inclusion of nine amino acid linkers between each neoepitope sequence. Of course, and as already noted above, it should be recognized that the exact length of the neoepitope sequence is not limited to 15 amino acids, and that the exact length may vary considerably. However, in most cases, where neoepitope sequences of between 8-12 amino acids are flanked by additional amino acids, the total length will typically not exceed 25 amino acids, or 30 amino acids, or 50 amino acids. Likewise, it should be noted that while Figure 1 denotes G-S linkers, various other linker sequences are also suitable for use herein. Such relatively short neoepitopes are especially beneficial where presentation of the neoepitope is intended to be via the MHC-I complex.

[0077] In this context, it should be appreciated that suitable linker sequences will provide steric flexibility and separation of two adjacent neoepitopes. However, care must be taken to as to not choose amino acids for the linker that could be immunogenic/form an epitope that is already present in a patient. Consequently, it is generally preferred that the polytope construct is filtered once more for the presence of epitopes that could be found in a patient (*e.g.*, as part of normal sequence or due to SNP or other sequence variation). Such filtering will apply the same technology and criteria as already discussed above.

[0078] Similarly, Construct 3 exemplarily illustrates a neoepitope arrangement that includes eight neoepitopes in concatemeric series without intervening linker sequences, and Construct 4 shows the arrangement of Construct 3 with inclusion of nine amino acid linkers between each neoepitope sequence. As noted above, it should be recognized that the exact length of such neoepitope sequences is not limited to 25 amino acids, and that the exact length may vary considerably. However, in most cases, where neoepitope sequences of between 14-20 amino acids are flanked by additional amino acids, the total length will typically not exceed 30 amino acids, or 45 amino acids, or 60 amino acids. Likewise, it should be noted that while Figure 1 denotes G-S linkers for these constructs, various other linker sequences are also suitable for use herein. Such relatively long neoepitopes are especially beneficial where presentation of the neoepitope is intended to be via the MHC-II complex.

[0079] In this example, it should be appreciated that the 15-aa minigenes are MHC Class I targeted tumor mutations selected with 7 amino acids of native sequence on either side, and that the 25-aa minigenes are MHC Class II targeted tumor mutations selected with 12 amino acids of native sequence on either side. The exemplary 9 amino acid linkers are deemed to have sufficient length such that "unnatural" MHC Class I epitopes will not form between adjacent minigenes. Polytope sequences tended to be processed and presented more efficiently than single neoepitopes (data not shown), and addition of amino acids beyond 12 amino acids for MHC-I presentation and addition of amino acids beyond 20 amino acids for MHC-I presentation appeared to allow for somewhat improved protease processing.

[0080] To maximize the likelihood that customized protein sequences remain intracellular for processing and presentation by the HLA complex, neoepitope sequences may be arranged in a manner to minimize hydrophobic sequences that may direct trafficking to the cell membrane or into the extracellular space. Most preferably, hydrophobic sequence or signal peptide detection is done either by comparison of sequences to a weight matrix (see *e.g.*, *Nucleic Acids Res.* 1986 Jun 11; 14(11): 4683–4690) or by using neural networks trained on peptides that contain signal sequences (see *e.g.*, *Journal of Molecular Biology* 2004, Volume 338, Issue 5, 1027-1036). **Figure 2** depicts an exemplary scheme of arrangement selection in which a plurality of polytope sequences are analyzed. Here, all positional permutations of all neoepitopes are calculated to produce a collection of arrangements. This collection is then processed through a weight matrix and/or neural network prediction to generate a score representing the likelihood of presence and/or strength of hydrophobic sequences or signal peptides. All positional permutations are then ranked by score, and the permutation(s) with a score below a predetermined threshold or lowest score for likelihood of presence and/or strength of hydrophobic sequences or signal peptides is/are used to construct a customized neoepitope expression cassette.

[0081] With respect to the total number of neoepitope sequences in a polytope it is generally preferred that the polytope comprise at least two, or at least three, or at least five, or at least eight, or at least ten neoepitope sequences. Indeed, the payload capacity of the host organism of the recombinant DNA is generally contemplated the limiting factor, along with the availability of filtered and appropriate neoepitopes. For example, adenoviral expression vectors, and particularly Adv5 are especially preferred as such vectors can accommodate up to 14kb in recombinant payload. Likewise, bacterial and yeast systems can accommodate

even larger payloads, typically in excess of 50kb. On the other hand, where the recombinant DNA is used in a DNA vaccine, suitable sizes will typically range between 5kb and 20kb.

[0082] In still further contemplated aspects of the inventive subject matter, it should be noted that the neoepitopes/polytopes can be directed towards a specific sub-cellular compartment (*e.g.*, cytosol, endosome, lysosome), and with that, towards a particular MHC presentation type. Such directed expression, processing, and presentation is particularly advantageous as contemplated compositions may be prepared that direct an immune response towards a $CD8^+$ type response (where the polytope is directed to the cytoplasmic space) or towards a $CD4^+$ type response (where the polytope is directed to the endosomal/lysosomal compartment). Moreover, it should be recognized that polytopes that would ordinarily be presented via the MHC-I pathway can be presented via the MHC-II pathway (and thereby mimic cross-presentation of neoepitopes). Therefore, it should be appreciated that neoepitope and polytope sequences may be designed and directed to one or both MHC presentation pathways using suitable sequence elements. With respect to routing the so expressed neoepitopes to the desired MHC-system, it is noted that the MHC-I presented peptides will typically arise from the cytoplasm via proteasome processing and delivery through the endoplasmic reticulum. Thus, expression of the epitopes intended for MHC-I presentation will generally be directed to the cytoplasm as is further discussed in more detail below. On the other hand, MHC-II presented peptides will typically arise from the endosomal and lysosomal compartment via degradation and processing by acidic proteases (*e.g.*, legumain, cathepsin L and cathepsin S) prior to delivery to the cell membrane.

[0083] Moreover, it is contemplated that proteolytic degradation of the polytope can also be enhanced using various methods, and especially contemplated methods include addition of a cleavable or non-cleavable ubiquitin moiety to the N-terminus, and/or placement of one or more destabilizing amino acids (*e.g.*, N, K, C, F, E, R, Q) to the N-terminus of the polytope where the presentation is directed towards MHC-I. On the other hand, where presentation is directed towards MHC-II, cleavage sites for particular endosomal or lysosomal proteases can be engineered into the polytope to so facilitate of promote antigen processing.

[0084] Therefore, in contemplated aspects of the inventive subject matter, signal and/or leader peptides may be used for trafficking neoepitopes and/or polytopes to the endosomal and lysosomal compartment, or for retention in the cytoplasmic space. For example, where the polytope is to be exported to the endosomal and lysosomal compartment, a leader peptide

such as the CD1b leader peptide may be employed to sequester the (nascent) protein from the cytoplasm. Additionally, or alternatively, targeting presequences and/or targeting peptides can be employed. The presequences of the targeting peptide may be added to the N-terminus and/or C-terminus and typically comprise between 6-136 basic and hydrophobic amino acids. In case of peroxisomal targeting, the targeting sequence may be at the C-terminus. Other signals (*e.g.*, signal patches) may be used and include sequence elements that are separate in the peptide sequence and become functional upon proper peptide folding. In addition, protein modifications like glycosylations can induce targeting. Among other suitable targeting signals, the inventors contemplate peroxisome targeting signal 1 (PTS1), a C-terminal tripeptide, and peroxisome targeting signal 2 (PTS2), which is a nonapeptide located near the N-terminus.

[0085] In addition, sorting of proteins to endosomes and lysosomes may also be mediated by signals within the cytosolic domains of the proteins, typically comprising short, linear sequences. Some signals are referred to as tyrosine-based sorting signals and conform to the NPXY or YXXØ consensus motifs. Other signals known as dileucine-based signals fit [DE]XXXL[LI] or DXXLL consensus motifs. All of these signals are recognized by components of protein coats peripherally associated with the cytosolic face of membranes. YXXØ and [DE]XXXL[LI] signals are recognized with characteristic fine specificity by the adaptor protein (AP) complexes AP-1, AP-2, AP-3, and AP-4, whereas DXXLL signals are recognized by another family of adaptors known as GGAs. Also FYVE domain can be added, which has been associated with vacuolar protein sorting and endosome function. In still further aspects, endosomal compartments can also be targeted using human CD1 tail sequences (see *e.g.*, *Immunology*, 122, 522–531). For example, lysosomal targeting can be achieved using a LAMP1-TM (transmembrane) sequence, while recycling endosomes can be targeted via the CD1a tail targeting sequence, and sorting endosomes can be targeted via the CD1c tail targeting sequence as is shown in more detail further below.

[0086] Trafficking to or retention in the cytosolic compartment may not necessarily require one or more specific sequence elements. However, in at least some aspects, N- or C-terminal cytoplasmic retention signals may be added, including a membrane-anchored protein or a membrane anchor domain of a membrane-anchored protein such that the protein is retained in the cell facing the cytosol. For example, membrane-anchored proteins include SNAP-25, syntaxin, synaptoprevin, synaptotagmin, vesicle associated membrane proteins (VAMPs),

synaptic vesicle glycoproteins (SV2), high affinity choline transporters, Neurexins, voltage-gated calcium channels, acetylcholinesterase, and NOTCH.

[0087] In still further contemplated aspects of the inventive subject matter, the polytope may also comprise one or more transmembrane segments that will direct the neoepitope after processing to the outside of the cell membrane to so be visible to immune competent cells. There are numerous transmembrane domains known in the art, and all of those are deemed suitable for use herein, including those having a single alpha helix, multiple alpha helices, alpha/beta barrels, etc. For example, contemplated transmembrane domains can comprise comprises the transmembrane region(s) of the alpha, beta, or zeta chain of the T-cell receptor, CD28, CD3 epsilon, CD45, CD4, CD5, CD8 (*e.g.*, CD8 alpha, CD8 beta), CD9, CD16, CD22, CD33, CD37, CD64, CD80, CD86, CD134, CD137, CD154, KIRDS2, OX40, CD2, CD27, LFA-1 (CD11a, CD18), ICOS (CD278), 4-1BB (CD137), GITR, CD40, BAFFR, HVEM (LIGHTR), SLAMF7, NKp80 (KLRF1), CD160, CD19, IL2R beta, IL2R gamma, IL7R α, ITGA1, VLA1, CD49a, ITGA4, IA4, CD49D, ITGA6, VLA-6, CD49f, ITGAD, CD11d, ITGAE, CD103, ITGAL, CD11a, LFA-1, ITGAM, CD11b, ITGAX, CD11c, ITGB1, CD29, ITGB2, CD18, LFA-1, ITGB7, TNFR2, DNAM1 (CD226), SLAMF4 (CD244, 2B4), CD84, CEACAM1, CRTAM, Ly9 (CD229), CD160 (BY55), PSGL1, CD100 (SEMA4D), SLAMF6 (NTB-A, Ly108), SLAM (SLAMF1, CD150, IPO-3), BLAME (SLAMF8), SELPLG (CD162), LTBR, or PAG/Cbp. Where a fusion protein is desired, it is contemplated that the recombinant chimeric gene has a first portion that encodes the transmembrane region(s), wherein the first portion is cloned in frame with a second portion that encodes the inhibitory protein. It should be noted that such presentation will not result in MHC-complex presentation and as such provides a neoepitope presentation independent of MHC/T-cell receptor interaction, which may further open additional avenues for immune recognition and trigger antibody production against the neoepitopes.

[0088] Alternatively, or additionally, the polytope may also be designed to include signal sequences for protein export of one or more neoepitope to thereby force a transfected cell to produce and secrete one or more neoepitopes. For example, the SPARC leader sequence may be added to a neoepitope or polytope sequence, leading to *in vivo* secretion of the neoepitope or polytope sequence into the extracellular space. Advantageously, such secreted neoepitopes or polytopes are then taken up by immune competent cells, and especially antigen presenting

cells and dendritic cells that in turn process and display the neoepitopes, typically via MHC-II pathways.

[0089] In still further contemplated aspects, the polytope may also be designed as a chimeric polytope that includes at least a portion of, and more typically an entire tumor associated antigen (*e.g.*, CEA, PSMA, PSA, MUC1, AFP, MAGE, HER2, HCC1, p62, p90, etc.). Most notably, tumor associated antigens are generally processed and presented via the MHC-II pathway. Therefore, instead of using compartment specific signal sequences and/or leader sequences, the processing mechanism for tumor associated antigens can be employed for MHC-II targeting.

[0090] Therefore, it should be appreciated that immune therapeutic compositions may be prepared that can deliver one or more neoepitopes to various sub-cellular locations, and with that generate distinct immune responses. For example, **Prior Art Figure 3** schematically illustrates a scenario where the polytope is predominantly processed in the proteasome of the cytoplasm and presented via the MHC-I complex, which is recognized by the T-cell receptor of a $CD8^+$ T-cell. Consequently, targeting polytope processing to the cytosolic compartment will skew the immune response towards a $CD8^+$ type response. On the other hand, **Prior Art Figure 4** schematically illustrates a scenario where the polytope is predominantly processed in the endosomal compartment and presented via the MHC-II complex, which is recognized by the T-cell receptor of a $CD4^+$ T-cell. Consequently, targeting polytope processing to the endosomal or lysosomal compartment will skew the immune response towards a $CD4^+$ type response. In addition, it should be appreciated that such targeting methods allow for specific delivery of a polytope or neoepitope peptide to an MHC subtype having the highest affinity with the peptide, even if that peptide would otherwise not be presented by that MHC subtype. Therefore, and as noted earlier, peptides for MHC-I presentation will generally be designed to have 8-12 amino acids (plus additional amino acids for flexibility in protease processing), while peptides for MHC-II presentation will be designed to have 14-20 amino acids (plus additional amino acids for flexibility in protease processing). In the examples below, further amino acids were added to allow for processing flexibility in the cytoplasmic, proteasome, or endosomal compartments.

[0091] In still further contemplated aspects of the inventive subject matter, it should be noted that trafficking modes of the neoepitope or polytope may be combined to accommodate one or more specific purposes. For example, sequential administration of the same neoepitopes

or polytope with different targeting may be particularly beneficial in a prime-boost regimen where in a first administration the patient in inoculated with a recombinant virus to infect the patients cells, leading to antigen expression, processing, and presentation (*e.g.*, predominantly MHC-I presentation) that will result in a first immune response originating from within a cell. The second administration of the same neoepitopes bound to albumin may then be employed as a boost as the so delivered protein is taken up by antigen presenting cells, leading in most cases to a distinct antigen presentation (*e.g.*, predominantly MHC-II presentation). Where the same neoepitopes or polytope is trafficked to the cell surface for cell surface bound MHC-independent presentation, ADCC responses or NK mediated cell killing may be promoted. In still further contemplated aspects, and as illustrated in the examples below, immunogenicity of neoepitopes may be enhanced by cross presentation or MHC-II directed presentation. Notably, as cancer cell neoepitopes are typically internally generated and recycled, and with that preferentially presented via the MHC-I system, contemplated systems and methods now allow for presentation of such neoepitopes via MHC-II, which may be more immunogenic as is shown in more detail below. In addition, multiple and distinct trafficking of the same neoepitopes or polytopes may advantageously increase or supplement an immune response due to the stimulation of various and distinct components of the cellular and humoral immune system.

[0092] Of course, it should be appreciated that multiple and distinct trafficking of the same neoepitopes or polytopes may be achieved in numerous manners. For example, differently trafficked neoepitopes or polytopes may be administered separately using the same (*e.g.*, viral expression vector) or different (*e.g.*, viral expression vector and albumin bound) modality. Similarly, and especially where the therapeutic agent is an expression system (*e.g.*, viral or bacterial), the recombinant nucleic acid may include two distinct portions that encode the same, albeit differently trafficked neoepitope or polytope (*e.g.*, first portion trafficked to first location (*e.g.*, cytosol or endosomal or lysosomal), second portion trafficked to a second, distinct location (*e.g.*, cytosol or endosomal or lysosomal, secreted, membrane bound)). Likewise, a first administration may employ viral delivery of cytoplasm targeted neoepitopes or polytope, while a second administration is typically at least a day, two days, four days, a week, or two weeks after the first administration and may employ viral delivery of endosomal or lysosomal targeted or secreted neoepitopes or polytope.

[0093] In addition, it should be appreciated that where the recombinant nucleic acid is used in a DNA, bacterial, or yeast vaccine, the manner of uptake of these modalities will at least partially dictate intracellular trafficking. Most typically, DNA, bacterial, or yeast vaccines are taken up by endocytotic or related processes and as such will be preferentially routed to the endosomal or lysosomal compartments. Such routing can be further enhanced (at least in the case of DNA vaccines) using appropriate trafficking signals as already described above or counteracted by use of a cytoplasmic retention sequence. However, other embodiments, it should be appreciated that polytopes delivered via DNA, bacterial, or yeast vaccines need not have a trafficking signal at all. Such polytopes will then preferentially processed/presented via the MHC-II system.

[0094] Additionally, it is contemplated that the expression construct, and especially the recombinant viral expression vector or DNA plasmid for a DNA vaccine, may further encode at least one, more typically at least two, even more typically at least three, and most typically at least four co-stimulatory molecules to enhance the interaction between the infected cells (*e.g.*, antigen presenting cells) and T-cells. For example, suitable co-stimulatory molecules include CD80, CD86, CD30, CD40, CD30L, CD40L, ICOS-L, B7-H3, B7-H4, CD70, OX40L, 4-1BBL, while other stimulatory molecules with less defined (or understood) mechanism of action include GITR-L, TIM-3, TIM-4, CD48, CD58, TL1A, ICAM-1, LFA3, and members of the SLAM family. However, especially preferred molecules for coordinated expression with the cancer-associated sequences include CD80 (B7-1), CD86 (B7-2), CD54 (ICAM-1) and CD11 (LFA-1). In addition to co-stimulatory molecules, the inventors also contemplate that one or more cytokines or cytokine analogs may be expressed from the recombinant nucleic acid, and especially preferred cytokines and cytokine analogs include IL-2, IL-15, and IL-a5 superagonist (ALT-803). Moreover, it should be appreciated that expression of the co-stimulatory molecules and/or cytokines will preferably be coordinated such that the neoepitopes or polytope are expressed contemporaneously with one or more co-stimulatory molecules and/or cytokines. Thus, it is typically contemplated that the co-stimulatory molecules and/or cytokines are produced from a single transcript (which may or may not include the sequence portion encoding the polytope), for example, using an internal ribosome entry site or 2A sequence, or from multiple transcripts.

[0095] Likewise, it is contemplated that the viral vector may also include a sequence portion that encodes one or more peptide ligands that bind to a checkpoint receptor. Most typically,

binding will inhibit or at least reduce signaling via the receptor, and particularly contemplated receptors include CTLA-4 (especially for $CD8^+$ cells), PD-1 (especially for $CD4^+$ cells), TIM1 receptor, 2B4, and CD160. For example, suitable peptide binders can include antibody fragments and especially scFv, but also small molecule peptide ligands (*e.g.*, isolated via RNA display or phage panning) that specifically bind to the receptors. Once more, it should be appreciated that expression of the peptide molecules will preferably be coordinated such that the neoepitopes or polytope are expressed contemporaneously with one or more of the peptide ligands. Thus, it is typically contemplated that the peptide ligands are produced from a single transcript (which may or may not include the sequence portion encoding the polytope), for example, using an internal ribosome entry site or 2A sequence, or from multiple transcripts.

[0096] It should be appreciated that all of the above noted co-stimulatory genes and genes coding for inhibitory proteins that interfere with/down-regulate checkpoint inhibition are well known in the art, and sequence information of these genes, isoforms, and variants can be retrieved from various public resources, including sequence data bases accessible at the NCBI, EMBL, GenBank, RefSeq, etc. Moreover, while the above exemplary stimulating molecules are preferably expressed in full length form as expressed in human, modified and non-human forms are also deemed suitable so long as such forms assist in stimulating or activating T-cells. Therefore, muteins, truncated forms and chimeric forms are expressly contemplated herein.

[0097] Consequently, contemplated expression constructs will preferably include a sequence portion that encodes one or more polytopes, wherein at least one, and more typically at least two, or all of the polytopes will include a trafficking signal that will result in preferential trafficking of the polytope to at least one, and more typically at least two different sub-cellular locations. For example, the first polytope may be directed towards the cytoplasm (and may include an additional cleavable or non-cleavable ubiquitin) while the second polytope may be directed towards the endosomal or lysosomal compartment. Or the first polytope may be directed towards the endosomal or lysosomal compartment while the second polytope may be directed towards the cell membrane or be secreted. As noted before, the encoded polytope will comprise at least two neoepitopes, optionally separated by a linker. Moreover, such contemplated expression constructs will also include a sequence portion that encodes one or more co-stimulatory molecules and/or cytokines, and may also include one or

more inhibitory proteins that interfere with/down-regulate checkpoint inhibition. Most typically, the expression construct will also include regulatory sequences operably coupled to the above sequence portions to drive contemporaneous expression of the polytope and the co-stimulatory molecules, cytokines, and/or inhibitory proteins. Suitable promoter elements are known in the art, and especially preferred promoters include the constitutive and inducible promoters discussed above.

[0098] Vaccine Compositions

[0099] Upon identification of desired neoepitopes, one or more immune therapeutic agents may be prepared using the sequence information of the neoepitopes, preferably configured as a polytope as described above. Preferably, the immune therapeutic agents include at least two of a DNA vaccine that includes a recombinant nucleic acid that encodes at least one antigen (and more typically at least two, three, four, or more antigens) that is present in the tumor, a bacterial vaccine in which a bacterium expresses at least one antigen (and more typically at least two, three, four, or more antigens) that is present in the tumor, a yeast vaccine in which a bacterium expresses at least one antigen (and more typically at least two, three, four, or more antigens) that is present in the tumor, and a viral vaccine that comprises a viral expression vector that includes a recombinant nucleic acid that encodes at least one antigen (and more typically at least two, three, four, or more antigens) that is present in the tumor.

[00100] With respect to recombinant nucleic acids for expression and DNA vaccination systems it is contemplated that the recombinant nucleic acid may be an RNA or a DNA. With respect to the use of RNA, DNA, or other recombinant vectors that lead to the expression of the tumor antigens and/or neoepitopes, especially contemplated nucleic acids include plasmid vectors that may be supercoiled, coiled, relaxed, or even linearized. For example, and among other suitable choices, contemplated vectors include vectors used in cloning one or more sequence portions used in the preparation of the viral expression vector. Thus, especially contemplated vectors include transfer or shuttle vectors, and various general cloning vectors (*e.g.*, having a bacterial origin of replication, a selection marker (*e.g.*, antibiotic resistance or fluorescent protein), and a multiple cloning site). Suitable vectors are well known in the art and are typically based on a plasmid with replicative capability in bacteria for cloning and production of substantial quantities. Proper vector selection may further be determined by its particular use (*e.g.*, shuttle vector for adenovirus, lentivirus, or baculovirus, etc.), choice of inducible or constitutive promoter (*e.g.*, CMV, UbC), choice of permanent or transient

expression, manner of transfection (*e.g.*, lipofection, electroporation, etc.), capacity for the recombinant payload, etc.

[00101] It should still further be noted that plasmids may be methylated or unmethylated, which may be controlled by directed enzymatic *in vitro* reactions or more simply by *in vivo* replication of the plasmid in a methylation competent or a methylation deficient host. Still further, it should be appreciated that the plasmid may further have one or more nucleic acid portions that are known to trigger an innate immune response (*e.g.*, CpG islands or other sequence motifs that interact with Toll-like receptors such as TLR3, TLR 7, TLR 8, TLR 9, RIG-I-like receptors, STING, and/or intracellular DNA sensors such as NLRP3/CIAS1).

[00102] Most typically, and as already noted above, contemplated plasmids and other nucleic acids will include one or more sequence elements that encode the preferably patient- and tumor specific neoepitopes or the polytope, most preferably operably coupled to regulatory elements that permit or drive the expression of the neoepitope or polytope in eukaryotic cells, and especially mammalian cells (*e.g.*, human). Moreover, it should be noted that the neoepitope or polytope may include the trafficking signals for routing the peptide to the desired sub-cellular location as also discussed herein. Therefore, especially preferred plasmids include plasmids used in the production of a viral expression vector, and as such will already include all regulatory elements needed for expression and/or trafficking of the polytope in a mammalian cell. Thus cloning vectors and shuttle vectors are especially preferred.

[00103] As will be appreciated, plasmids contemplated herein may be administered in numerous manners known in the art, and suitable delivery modes include injection (intramuscular, intravenous, intradermal), delivery via gene gun or other ballistic transfer, or by liposome mediated transfer. Therefore, contemplated compositions will particularly include injectable formulations comprising nucleic acid lipoplexes and other DNA-lipid or DNA lipoprotein complexes. Advantageously, the choice of delivery may be used to polarize the immune response towards either Th1 (via injection using saline) or Th2 (via gene gun delivery) response as described elsewhere (see *e.g.*, *J Immunol* March 1, 1997, 158 (5) 2278-2284). In especially preferred aspects, vaccination with DNA will preferably be performed by intravenous injection as is discussed in more detail below, however, other routes (including intramuscular, intradermal, subcutaneous, intra-arterial) are also deemed suitable for use

herein, while administration of the viral expression vector (typically using a recombinant virus) will preferably be performed by subcutaneous injection.

[00104] Without wishing to be bound by any specific theory or hypothesis, it is believed that the use of plasmids and other 'naked' nucleic acids (*e.g.*, linear DNA RNA) will provide a first opportunity to generate an immune response that is both non-specific and specific to the patient's neoepitope. Non-specific reactions are deemed to arise from a component of the innate immune response against foreign nucleic acids, for example, where the nucleic acid is unmethylated (*e.g.*, via TLR-9). In addition, expression of the neoepitope or polytope in the cells transfected with the plasmid will also promote adaptive immune responses.

[00105] In addition to the use of DNA vaccination, contemplated plasmids may also be used in the production of a viral or yeast expression vector that can be employed to produce a recombinant virus (*e.g.*, lentivirus, adenovirus) or yeast for subsequent administration to the patient. Where the plasmid is used for production of a yeast or viral expression system, all known expression systems are deemed suitable for use herein. For example, suitable materials and protocols can be found in the non-profit plasmid repository Addgene or in the AdEasy adenoviral vector system (commercially available from Agilent). Similarly, there are numerous yeast expression systems known in the art and all of those are deemed suitable for use herein. Such second administration can be viewed as a boost regimen to the DNA vaccination as the mammal was already primed by the plasmid vector. Of course, it should be recognized that where the prime vaccination was a DNA vaccination with a plasmid as described above, the boost may use various alternate formats, including a vaccination with the neoepitope peptides or polytope using more conventional vaccine formulations. Further suitable DNA vaccines are described, for example, in US 2014/0178438.

[00106] With respect to bacterial expression and vaccination systems it is contemplated that all bacterial strains are deemed suitable, and especially include species from *Salmonella*, *Clostridium*, *Bacillus*, *Lactobacillus*, *Bifidobacterium*, etc., particularly where such strains are non-pathogenic, genetically engineered to have reduced toxicity, and/or were irradiated prior to administration. Historically, most bacteria strains have been deemed unsuitable for introducing into the blood stream or transplanting into an organ or tissue, as most bacteria express lipopolysaccharides that trigger immune responses and cause endotoxic responses, which can lead potentially fatal sepsis (*e.g.*, CD-14 mediated sepsis) in patients. Thus, one especially preferred bacterial strain is based on a genetically modified bacterium which

expresses endotoxins at a level low enough not to cause an endotoxic response in human cells and/or insufficient to induce a CD-14 mediated sepsis when introduced to the human body.

[00107] One preferred bacterial species is a genetically modified *Escherichia coli* (*E. coli*) strain due to its fast growth (*e.g.*, one complete cell cycle in 20 min) and availability of many strains optimized for protein overexpression upon induction (*e.g.*, *lac* promoter induction with IPTG, etc.). Most typically, the genetic modification will reduce or remove production of most lipopolysaccharide components leading to endotoxic response. For example, one exemplary bacteria strain with modified lipopolysaccharide synthesis includes ClearColi® BL21(DE3) electrocompetent cells. This bacterial strain is BL21 with a genotype F– ompT hsdSB (rB- mB-) gal dcm lon λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) msbA148 ΔgutQΔkdsD ΔlpxLΔlpxMΔpagPΔlpxPΔeptA. In this context, it should be appreciated that several specific deletion mutations (ΔgutQ ΔkdsD ΔlpxL ΔlpxMΔpagPΔlpxPΔeptA) encode the modification of LPS to Lipid IV_A, while one additional compensating mutation (msbA148) enables the cells to maintain viability in the presence of the LPS precursor lipid IVA. These mutations result in the deletion of the oligosaccharide chain from the LPS. More specifically, two of the six acyl chains are deleted. The six acyl chains of the LPS are the trigger which is recognized by the Toll-like receptor 4 (TLR4) in complex with myeloid differentiation factor 2 (MD-2), causing activation of NF-ƙB and production of proinflammatory cytokines. Lipid IV_A, which contains only four acyl chains, is not recognized by TLR4 and thus does not trigger the endotoxic response. While electrocompetent BL21 bacteria is provided as an example, the inventors contemplates that the genetically modified bacteria can be also chemically competent bacteria.

[00108] Alternatively, the inventors also contemplate that the patient's own endosymbiotic bacteria can be used as a vehicle to express human disease-related antigens *in vivo* to elicit immune response at least locally. As used herein, the patient's endosymbiotic bacteria refers bacteria residing in the patient's body regardless of the patient's health condition without invoking any substantial immune response. Thus, it is contemplated that the patient's endosymbiotic bacteria is a normal flora of the patient. For example, the patient's endosymbiotic bacteria may include *E. coli* or *Streptococcus* that can be commonly found in human intestine or stomach. In these embodiments, patient's own endosymbiotic bacteria can be obtained from the patient's biopsy samples from a portion of intestine, stomach, oral

mucosa, or conjunctiva, or in fecal samples. The patient's endosymbiotic bacteria can then be cultured *in vitro* and transfected with nucleotides encoding human disease-related antigen(s).

[00109] Therefore, it should be appreciated that the bacteria used in the methods presented herein may be from a strain that produces LPS, or that are genetically engineered to have reduced or abrogated expression of one or more enzymes leading to the formation of LPS that is recognized by a TLR, and particularly TLR4. Most typically, such bacteria will be genetically modified to express in an inducible manner at least one human disease-related antigen for immunotherapy. Among other options, induction of expression may be done with synthetic compounds that are not ordinarily found in a mammal (*e.g.*, IPTG, substituted benzenes, cyclohexanone-related compounds) or with compounds that naturally occur in a mammal (e.g., sugars (including l-arabinose, l-rhamnose, xylose, and sucrose), ε-caprolactam, propionate, or peptides), or induction may be under the control of one or more environmental factors (*e.g.*, temperature or oxygen sensitive promoter).

[00110] Contemplated recombinant nucleic acids that encode the tumor antigens or the polytope can be inserted into an expression vector that has a specific promoter (*e.g.*, inducible promoter, etc.) to drive expression of the antigens or polytope in the bacterium. The vector is then transfected into the bacterium (*e.g.*, ClearColi® BL21(DE3) electrocompetent cells) following conventional methods, or any other type of competent bacterium expressing low endotoxin level that is insufficient to induce a CD-14 mediated sepsis when introduced to the human body), or to patient's own endosymbiotic bacterium that is optionally cultured in vitro before transformation as described above.

[00111] With respect to yeast expression and vaccination systems it is contemplated that all known yeast strains are deemed suitable for use herein. However, it is preferred that the yeast is a recombinant *Saccharomyces* train that is genetically modified with a nucleic acid construct as discussed above that leads to expression of at least one of the tumor antigens to thereby initiate an immune response against the tumor. Nevertheless, it is noted that any yeast strain can be used to produce a yeast vehicle of the present invention. Yeasts are unicellular microorganisms that belong to one of three classes: *Ascomycetes*, *Basidiomycetes* and *Fungi Imperfecti*. One consideration for the selection of a type of yeast for use as an immune modulator is the pathogenicity of the yeast. In preferred embodiments, the yeast is a non-pathogenic strain such as *Saccharomyces cerevisiae* as non-pathogenic yeast strains minimize any adverse effects to the individual to whom the yeast vehicle is administered. However,

pathogenic yeast may also be used if the pathogenicity of the yeast can be negated using pharmaceutical intervention.

[00112] For example, suitable genera of yeast strains include *Saccharomyces*, *Candida*, *Cryptococcus, Hansenula, Kluyveromyces, Pichia, Rhodotorula, Schizosaccharomyces* and *Yarrowia*. In one aspect, yeast genera are selected from *Saccharomyces, Candida, Hansenula, Pichia or Schizosaccharomyces*, and in a preferred aspect, *Saccharomyces* is used. Species of yeast strains that may be used in the invention include *Saccharomyces cerevisiae, Saccharomyces carlsbergensis, Candida albicans, Candida kefyr, Candida tropicalis, Cryptococcus laurentii, Cryptococcus neoformans, Hansenula anomala, Hansenula polymorpha, Kluyveromyces fragilis, Kluyveromyces lactis, Kluyveromyces marxianus var. lactis, Pichia pastoris, Rhodotorula rubra, Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

[00113] It should further be appreciated that a number of these species include a variety of subspecies, types, subtypes, etc. that are intended to be included within the aforementioned species. In one aspect, yeast species used in the invention include *S. cerevisiae, C. albicans, H. polymorpha, P. pastoris and S. pombe. S. cerevisiae* is useful due to it being relatively easy to manipulate and being "Generally Recognized As Safe" or "GRAS" for use as food additives (GRAS, FDA proposed Rule 62FR18938, Apr. 17, 1997). Therefore, the inventors particularly contemplate a yeast strain that is capable of replicating plasmids to a particularly high copy number, such as a *S. cerevisiae* cir strain. The *S. cerevisiae* strain is one such strain that is capable of supporting expression vectors that allow one or more target antigen(s) and/or antigen fusion protein(s) and/or other proteins to be expressed at high levels. In addition, any mutant yeast strains can be used in the present invention, including those that exhibit reduced post-translational modifications of expressed target antigens or other proteins, such as mutations in the enzymes that extend N-linked glycosylation.

[00114] Expression of contemplated antigens/neoepitopes in yeast can be accomplished using techniques known to those skilled in the art. Most typically, a nucleic acid molecule encoding at least neoepitope or other protein is inserted into an expression vector such manner that the nucleic acid molecule is operatively linked to a transcription control sequence to be capable of effecting either constitutive or regulated expression of the nucleic acid molecule when transformed into a host yeast cell. As will be readily appreciated, nucleic acid molecules encoding one or more antigens and/or other proteins can be on one or more

expression vectors operatively linked to one or more expression control sequences. Particularly important expression control sequences are those which control transcription initiation, such as promoter and upstream activation sequences.

[00115] Any suitable yeast promoter can be used in the present invention and a variety of such promoters are known to those skilled in the art and have generally be discussed above. Promoters for expression in *Saccharomyces cerevisiae* include promoters of genes encoding the following yeast proteins: alcohol dehydrogenase I (ADH1) or II (ADH2), CUP1, phosphoglycerate kinase (PGK), triose phosphate isomerase (TPI), translational elongation factor EF-1 alpha (TEF2), glyceraldehyde-3-phosphate dehydrogenase (GAPDH; also referred to as TDH3, for triose phosphate dehydrogenase), galactokinase (GAL1), galactose-1-phosphate uridyl-transferase (GAL7), UDP-galactose epimerase (GAL10), cytochrome c1 (CYC1), Sec7 protein (SEC7) and acid phosphatase (PHO5), including hybrid promoters such as ADH2/GAPDH and CYC1/GAL10 promoters, and including the ADH2/GAPDH promoter, which is induced when glucose concentrations in the cell are low (e.g., about 0.1 to about 0.2 percent), as well as the CUP1 promoter and the TEF2 promoter. Likewise, a number of upstream activation sequences (UASs), also referred to as enhancers, are known. Upstream activation sequences for expression in *Saccharomyces cerevisiae* include the UASs of genes encoding the following proteins: PCK1, TPI, TDH3, CYC1, ADH1, ADH2, SUC2, GAL1, GAL7 and GAL10, as well as other UASs activated by the GAL4 gene product, with the ADH2 UAS being used in one aspect. Since the ADH2 UAS is activated by the ADR1 gene product, it may be preferable to overexpress the ADR1 gene when a heterologous gene is operatively linked to the ADH2 UAS. Transcription termination sequences for expression in Saccharomyces cerevisiae include the termination sequences of the alpha-factor, GAPDH, and CYC1 genes. Transcription control sequences to express genes in methyltrophic yeast include the transcription control regions of the genes encoding alcohol oxidase and formate dehydrogenase.

[00116] Likewise, transfection of a nucleic acid molecule into a yeast cell according to the present invention can be accomplished by any method by which a nucleic acid molecule administered into the cell and includes diffusion, active transport, bath sonication, electroporation, microinjection, lipofection, adsorption, and protoplast fusion. Transfected nucleic acid molecules can be integrated into a yeast chromosome or maintained on extrachromosomal vectors using techniques known to those skilled in the art. As discussed

above, yeast cytoplast, yeast ghost, and yeast membrane particles or cell wall preparations can also be produced recombinantly by transfecting intact yeast microorganisms or yeast spheroplasts with desired nucleic acid molecules, producing the antigen therein, and then further manipulating the microorganisms or spheroplasts using techniques known to those skilled in the art to produce cytoplast, ghost or subcellular yeast membrane extract or fractions thereof containing desired antigens or other proteins. Further exemplary yeast expression systems, methods, and conditions suitable for use herein are described in US20100196411A1, US2017/0246276, or US 2017/0224794, and US 2012/0107347.

[00117] With respect to viral expression and vaccination systems it is contemplated that all therapeutic recombinant viral expression systems are deemed suitable for use herein so long as such viruses are capable to lead to expression of the recombinant payload in an infected cell. For example, suitable viruses include genetically modified alphaviruses, adenoviruses, adeno-associated viruses, herpes viruses, lentiviruses, etc. However, adenoviruses are particularly preferred. For example, genetically modified adenoviruses are preferred that are suitable not only for multiple vaccinations but also vaccinations in individuals with preexisting immunity to the adenovirus (see *e.g.*, WO 2009/006479 and WO 2014/031178), typically achieved by deletion of the E2b gene and other late proteins to reduce immunogenicity. Moreover, due to these specific deletions, such genetically modified viruses were replication deficient and allowed for relatively large recombinant cargo. For example, WO 2014/031178 describes the use of such genetically modified viruses to express CEA (colorectal embryonic antigen) to provide an immune reaction against colon cancer. Moreover, relatively high titers of recombinant viruses can be achieved using genetically modified human 293 cells as has been reported (*e.g.*, *J Virol.* 1998 Feb; 72(2): 926–933).

[00118] Regardless of the type of recombinant virus it is contemplated that the virus may be used to infect patient (or non-patient) cells *ex vivo* or *in vivo*. For example, the virus may be injected subcutaneously or intravenously, or may be administered intranasaly or via inhalation to so infect the patients cells, and especially antigen presenting cells. Alternatively, immune competent cells (*e.g.*, NK cells, T cells, macrophages, dendritic cells, etc.) of the patient (or from an allogeneic source) may be infected *in vitro* and then transfused to the patient. Alternatively, immune therapy need not rely on a virus but may be effected with nucleic acid transfection or vaccination using RNA or DNA, or other recombinant vector that

leads to the expression of the neoepitopes (*e.g.*, as single peptides, tandem mini-gene, etc.) in desired cells, and especially immune competent cells.

[00119] As noted above, the desired nucleic acid sequences (for expression from virus infected cells) are under the control of appropriate regulatory elements well known in the art. For example, suitable promoter elements include constitutive strong promoters (*e.g.*, SV40, CMV, UBC, EF1A, PGK, CAGG promoter), but inducible promoters are also deemed suitable for use herein, particularly where induction conditions are typical for a tumor microenvironment. For example, inducible promoters include those sensitive to hypoxia and promoters that are sensitive to TGF-β or IL-8 (e.g., via TRAF, JNK, Erk, or other responsive elements promoter). In other examples, suitable inducible promoters include the tetracycline-inducible promoter, the myxovirus resistance 1 (Mx1) promoter, etc.

[00120] Alternatively, or additionally, it should be recognized that the antigen/neoepitope or polytope may also be administered as peptide, optionally bound to a carrier protein to so act as a peptide vaccine. Among other suitable carrier proteins, human albumin or lactoferrin are particularly preferred. Such carrier proteins may be in native conformation, or pretreated to form nanoparticles with exposed hydrophobic domains (see *e.g.*, *Adv Protein Chem Struct Biol*. 2015;98:121-43) to which the neoepitope or polytope can be coupled. Most typically, coupling of the neoepitope or polytope to the carrier protein will be non-covalent. Similar to the secreted neoepitopes or polytopes, carrier protein-bound neoepitopes or polytopes will be taken up by the immune competent cells, and especially antigen presenting cells and dendritic cells that in turn process and display the neoepitopes, typically via MHC-II pathways.

[00121] Formulations

[00122] Where the vaccine is a viral vaccine (*e.g.*, an adenovirus, and especially AdV with E1 and E2b deleted), it is contemplated that the recombinant viruses may then be individually or in combination used as a therapeutic vaccine in a pharmaceutical composition, typically formulated as a sterile injectable composition with a virus titer of between 10^6-10^{13} virus particles, and more typically between 10^9-10^{12} virus particles per dosage unit. Alternatively, virus may be employed to infect patient (or other HLA matched) cells *ex vivo* and the so infected cells are then transfused to the patient. In further examples, treatment of patients with the virus may be accompanied by allografted or autologous natural killer cells or T cells in a bare form or bearing chimeric antigen receptors expressing antibodies targeting

neoepitope, neoepitopes, tumor associated antigens or the same payload as the virus. The natural killer cells, which include the patient-derived NK-92 cell line, may also express CD16 and can be coupled with an antibody.

[00123] Similarly, where the vaccine is a bacterial or yeast vaccine, the bacteria or yeast cells are preferably irradiated prior to administration to prevent further propagation. Most typically, administration is as a therapeutic vaccine in a pharmaceutical composition, typically formulated as a sterile injectable composition with a cell titer of between 10^6-10^9 cells, and more typically between 10^8-10^{11} cells per dosage unit. Most preferably, the vaccine formulation is administered intramuscularly, subcutaneously, or intratumorally. DNA vaccines will typically be administered as is well known in the art, preferably by intravenous injection of the DNA in a buffered solution as a pharmaceutical composition. While not limiting to the inventive subject matter, the total dose of DNA per administration will typically be in the range of 0.1 mcg to several 10 mg, or between 10 mcg to several 1,000 mcg.

[00124] Where desired, additional therapeutic modalities may be employed which may be neoepitope based (*e.g.*, synthetic antibodies against neoepitopes as described in WO 2016/172722), alone or in combination with autologous or allogenic NK cells, and especially haNK cells or taNK cells (*e.g.*, both commercially available from NantKwest, 9920 Jefferson Blvd. Culver City, CA 90232). Where haNK or taNK cells are employed, it is particularly preferred that the haNK cell carries a recombinant antibody on the CD16 variant that binds to a neoepitope of the treated patient, and where taNK cells are employed it is preferred that the chimeric antigen receptor of the taNK cell binds to a neoepitope of the treated patient. The additional treatment modality may also be independent of neoepitopes, and especially preferred modalities include cell-based therapeutics such as activated NK cells (e.g., aNK cells, commercially available from NantKwest, 9920 Jefferson Blvd. Culver City, CA 90232), and non cell-based therapeutics such as chemotherapy and/or radiation. In still further contemplated aspects, immune stimulatory cytokines, and especially IL-2, IL15, and IL-21 may be administered, alone or in combination with one or more checkpoint inhibitors (*e.g.*, ipilimumab, nivolumab, etc.). Similarly, it is still further contemplated that additional pharmaceutical intervention may include administration of one or more drugs that inhibit immune suppressive cells, and especially MDSCs Tregs, and M2 macrophages. Thus, suitable drugs include IL-8 or interferon-γ inhibitors or antibodies binding IL-8 or interferon-

γ, as well as drugs that deactivate MDSCs (*e.g.*, NO inhibitors, arginase inhibitors, ROS inhibitors), that block development of or differentiation of cells to MDSCs (*e.g.*, IL-12, VEGF-inhibitors, bisphosphonates), or agents that are toxic to MDSCs (*e.g.*, gemcitabine, cisplatin, 5-FU). Likewise, drugs like cyclophosphamide, daclizumab, and anti-GITR or anti-OX40 antibodies may be used to inhibit Tregs.

[00125] Protocols

[00126] Therefore, the inventors contemplate various exemplary strategies for treatment with the compositions contemplated herein. Most typically, treatments will include at least two, or at least three distinct vaccine modalities that are administered in a sequential fashion. For example in some embodiments, the initial administration will be a DNA vaccine that is followed by administration of a bacterial vaccine. The bacterial vaccine may optionally be followed by a yeast vaccine. Following the bacterial or yeast vaccine administration may then be a viral vaccine administration. In another example, the initial administration will be a bacterial vaccine that is optionally followed by administration of a yeast vaccine. Following the bacterial or yeast vaccine administration may then be a viral vaccine administration. In still another example, the initial administration will be a DNA vaccine, which is followed by a viral vaccine administration. As will be readily appreciated, each of the modalities can be administered once, or repeatedly as desired. For example, the DNA vaccine may be given once, while a subsequent bacterial and/or yeast vaccine may be administered twice, three times or more, before administration of a viral vaccine commences. Similarly, multiple bacterial or yeast vaccine compositions may be administered prior to a viral vaccine composition. In the same way, the DNA, bacterial, and/or yeast vaccine may be given once and the viral vaccine may be given multiple times. Therefore, it should be noted that in a prime/boost regimen, the prime vaccination may use a modality that is different from the boost vaccination (*e.g.*, DNA, bacterial, or yeast vaccination as prime, viral vaccination as boost).

[00127] It should further be recognized that the administrations of the specific modalities will be spaced apart by several days to allow an immune response to develop. Most typically, the first administration will be spaced apart from the second administration by at least two days, more typically by at least four days, even more typically by at least one week, and most typically by two weeks or even longer. Moreover, it should be noted that the immune system of the patient may also be pre-conditioned using immune stimulatory cytokines (*e.g.*, IL-2,

IL-15, IL-21) or cytokine analogs (*e.g.*, ALT-803), using checkpoint inhibitors or Treg or M2 macrophage inhibitors to reduce immune suppression, or using cytokines that support an immune response and generation of immune memory (*e.g.*, IL-12).

[00128] As used herein, the term "administering" a pharmaceutical composition or drug refers to both direct and indirect administration of the pharmaceutical composition or drug, wherein direct administration of the pharmaceutical composition or drug is typically performed by a health care professional (*e.g.*, physician, nurse, etc.), and wherein indirect administration includes a step of providing or making available the pharmaceutical composition or drug to the health care professional for direct administration (*e.g.*, via injection, infusion, oral delivery, topical delivery, etc.). Most preferably, the recombinant virus is administered via subcutaneous or subdermal injection. However, in other contemplated aspects, administration may also be intravenous injection. Alternatively, or additionally, antigen presenting cells may be isolated or grown from cells of the patient, infected *in vitro*, and then transfused to the patient. Therefore, it should be appreciated that contemplated systems and methods can be considered a complete drug discovery system (*e.g.*, drug discovery, treatment protocol, validation, etc.) for highly personalized cancer treatment. As will also be appreciated, contemplated treatments may be repeated over time, particularly where new neoepitopes have developed (*e.g.*, as a result of clonal expansion).

[00129] Moreover, it is noted that additional treatment regimens may be implemented to assist contemplated methods and compositions. Such additional treatment regimens will preferably be performed to increase 'visibility' of a tumor to the immune system. For example, to trigger overexpression or transcription of stress signals, it is contemplated that chemotherapy and/or radiation may be employed using at a low-dose regimen, preferably in a metronomic fashion. For example, it is generally preferred that such treatment will use doses effective to affect at least one of protein expression, cell division, and cell cycle, preferably to induce apoptosis or at least to induce or increase the expression of stress-related genes (and particularly NKG2D ligands). Thus, in further contemplated aspects, such treatment will include low dose treatment using one or more chemotherapeutic agents. Most typically, low dose treatments will be at exposures that are equal or less than 70%, equal or less than 50%, equal or less than 40%, equal or less than 30%, equal or less than 20% , equal or less than 10%, or equal or less than 5% of the LD_{50} or IC_{50} for the chemotherapeutic agent. Additionally, where advantageous, such low-dose regimen may be performed in a

metronomic manner as described, for example, in US 7758891, US 7771751, US 7780984, US 7981445, and US 8034375.

[00130] With respect to the particular drug used in low-dose regimens, it is contemplated that all chemotherapeutic agents are deemed suitable. Among other suitable drugs, kinase inhibitors, receptor agonists and antagonists, anti-metabolic, cytostatic and cytotoxic drugs are all contemplated herein. However, particularly preferred agents include those identified to interfere or inhibit a component of a pathway that drives growth or development of the tumor. Suitable drugs can be identified using pathway analysis on omics data as described in, for example, WO 2011/139345 and WO 2013/062505. Most notably, so achieved expression of stress-related genes in the tumor cells will result in surface presentation of NKG2D, NKP30, NKP44, and/or NKP46 ligands, which in turn activate NK cells to specifically destroy the tumor cells. Thus, it should be appreciated that low-dose chemotherapy may be employed as a trigger in tumor cells to express and display stress related proteins, which in turn will trigger NK-cell activation and/or NK-cell mediated tumor cell killing. Additionally, NK-cell mediated killing will be associated with release of intracellular tumor specific antigens, which is thought to further enhance the immune response.

Examples

[00131] Exemplary Sequence Arrangements

[00132] Neoepitope sequences were determined *in silico* by location-guided synchronous alignment of tumor and normal samples as, for example, disclosed in US 2012/0059670 and US 2012/0066001 using BAM files and BAM servers. Specifically, DNA analysis of the tumor was from the B16-F10 mouse melanoma line and matched normal was blood from C57bl/6 parental mouse DNA. The results were filtered for expression by RNA sequencing of this tumor cell line. Neoepitopes that were found expressed were further analyzed for binding affinity towards murine MHC-I (here: Kb) and MHC-II (here: I-Ab). Selected binders (with affinity of equal or less than 200 nM) were further analyzed after a further step of dbSNP filtering using positional permutations of all neoepitopes that were then processed through a weight matrix and neural network prediction to generate a score representing the likelihood of presence and/or strength of hydrophobic sequences or signal peptides. The best scoring arrangement (lowest likelihood of hydrophobic sequences or signal peptides) for the polytope (not shown) was used for further experiments. Neoepitopes were prioritized by

detection in RNAseq or other quantitative system that yielded expression strength for a specific gene harboring the neoepitope mutation.

[00133] **Table 1** shows exemplary neoepitopes that were expressed as determined by RNAseq along with gene name and mutated amino acid and position of the mutated amino acid. The neoepitope listed with * was discarded after dbSNP filtering as that neoepitope occurred as variant Rs71257443 in 28% of the population.

Table 1

Gene	Position	Neoepitope-a	Neoepitope-b
VIPR2	V73M	GETVT**M**PCP	
LILRB3	T187N	VGPV**N**PSHR*	
FCRL1	R286C	GLGAQ**C**SEA	
FAT4	S1613L	RK**L**TTELTI	PERRK**L**TTE
PIEZO2	T2356M	MDWVW**M**DTT	VW**M**DTTLSL
SIGLEC14	A292T	GK**T**LNPSQT	REGK**T**LNPS
SIGLEC1	D1143N	VR**N**ATSYRC	NVTVR**N**ATS
SLC4A11	Q678P	FAMAQ**I**P**S**L	AQ**I**P**S**LSLR

[00134] **Table 2** shows further examples of neoepitopes in which the position of the mutated amino acid was changed, and shows further alternate sequences for MHC-I presentation (9-mer) and MHC-II presentation (15-mer). The neoepitope sequence for MHC-II presentation was back-translated to the corresponding nucleic acid sequence, which is also shown in Table 2.

Table 2

Gene	Change	Neoepitope-a	Neoepitope-b	Extended 15 mer	Nucleotide Sequence
SLC4A11	Q678P	FAMAQ**I**P**S**L	AQ**I**P**S**LSLR	PFAMAQ**I**P**S**LSLRAV	CCCTTCGCCATGGCCC AGATCCCCAGCCTGA GCCTGAGGGCCGTG
SIGLEC1	D1143N	VR**N**ATSYRC	NVTVR**N**ATS	LPNVTVR**N**ATSYRCG	CTGCCCAACGTGACC GTGAGGAACGCCACC AGCTACAGGTGCGGC
SIGLEC14	A292T	GK**T**LNPSQT	REGK**T**LNPS	SWFREGK**T**LNPSQTS	AGCTGGTTCAGGGAG GGCAAGACCCTGAAC CCCAGCCAGACCAGC
PIEZO2	T2356M	MDWVW**M**DTT	VW**M**DTTLSL	AVMDWVW**M**DTTLSLS	GCCGTGATGGACTGG GTGTGGATGGACACC ACCCTGAGCCTGAGC
FAT4	S1613L	RK**L**TTELTI	PERRK**L**TTE	LGPERRK**L**TTELTII	CTGGGCCCCGAGAGG AGGAAGCTGACCACC GAGCTGACCATCATC
FCRL1	R286C	GLGAQ**C**SEA		NNGLGAQ**C**SEAVTLN	AACAACGGCCTGGGC GCCCAGTGCAGCGAG GCCGTGACCCTGAAC
VIPR2	V73M	GETVT**M**PCP		NVGETVT**M**PCPKVFS	AACGTGGGCGAGACC GTGACCATGCCCTGC CCCAAGGTGTTCAGC

FLRT2	R346W	EQV**W**GMAVR		CQGPEQV**W**GMAVREL	TGCCAGGGCCCCGAG CAGGTGTGGGGCATG GCCGTGAGGGAGCT G

[00135] Sequence Trafficking

[00136] Model cancer: Murine B16-F10 melanoma (derived from C57/Bl6 mouse) was used tumors were screened in a tumor versus normal manner as described above, and expressed mutant epitopes were identified in the B16F10 melanoma cell line. Candidate neoepitopes were further filtered as described above using sequencing data analysis and binding analysis to murine MHC I (H2-Kb, H2-Dd) and MHC II (I-Ab). Nine distinct polytope constructs were then prepared for testing various trafficking schemes, and each construct was prepared as the corresponding recombinant nucleic acid under the control of a CMV promoter. Each construct was cloned into an AdV5 expression vector that had deleted E1 and E2b genes, and the resulting recombinant virus was then used for transfection of mice as is further discussed below.

[00137] More specifically, three polytope constructs included MHC I binding neoepitopes for MHC-I presentation and were therefore targeted to the cytoplasmic compartment. While one construct had an unmodified N-terminus, another construct had an N-terminal non-cleavable ubiquitin, and yet another construct had an N-terminal cleavable ubiquitin. Ubiquitination was used to target the proteasome in the cytosol. Three further polytope constructs included MHC I binding neoepitopes for MHC-II presentation and were therefore targeted to the lysosomal/endosomal compartments compartment. While one construct had lysosomal targeting sequence, another construct had a recycling endosomal targeting sequence, and yet another construct had a sorting endosomal targeting sequence. Three additional polytope constructs included MHC II binding neoepitopes for MHC-II presentation and were also targeted to the lysosomal/endosomal compartments compartment. Once more, one construct had lysosomal targeting sequence, another construct had a recycling endosomal targeting sequence, and yet another construct had a sorting endosomal targeting sequence. These nine constructs had sequence arrangements as follows.

[00138] In the following exemplary sequences, for MHC-I presentation, ubiquitin (cleavable and non-cleavable) were used for proteasome targeting, while the CD1b leader peptide was used as an export leader peptide for trafficking the polypeptide out of the cytosol for all MHC-II directed sequences. LAMP1-TM/cytoplasmic tail was used as a lysosomal

targeting sequence, while LAMP1-TM/CD1a tail was used as a recycling endosomes targeting sequence, and LAMP1-TM/CD1c tail was used as a sorting endosomes targeting domain .

[00139] It should further be noted that various internal controls were also used in the above polypeptides to account for expression and presentation. More specifically, the SIINFEKL peptide was used as an internal control for a MHC I restricted (Kb) peptide epitope, while the Esat6 peptide was used as an internal control for a secreted protein for MHC II presentation. FLAG-tag was used as an internal control epitope for detection of expression, and cMYC used as an internal control Tag for simple protein detection.

[00140] Exemplary constructs for MHC-I epitopes directed to MHC-I presentation (traffic through proteasome, cytoplasmic targeting):

[00141] Polyepitope only: 12aa-A^m-12aa-AAAA-12aa-B^m-12aa-AAAA-(12aa-X^m-12aa-AAAA)$_n$-SIINFEKL-AAAA-Esat6-cMYC. **Figure 5A** exemplarily depicts the polypeptide structure of such arrangement where the SIINFEKL motif is underlined, the Esat6 motif is in italics, and where the cMY motif is in bold type font. The nucleotide sequence for Figure 5A is SEQ ID NO: 1, and the polypeptide sequence for Figure 5A is SEQ ID NO: 4.

[00142] Polyepitope and cleavable ubiquitin GGR N-terminus: Ubiquitin-GGR-12aa-A^m-12aa-AAAA-12aa-B^m-12aa-AAAA-(12aa-X^m-12aa-AAAA)$_n$-SIINFEKL-AAAA-Esat6-cMYC. **Figure 5B** exemplarily depicts the polypeptide structure of such arrangement where the cleavable ubiquitin moiety is italics and underlined, the SIINFEKL motif is underlined, the Esat6 motif is in italics, and where the cMY motif is in bold type font. The nucleotide sequence for Figure 5B is SEQ ID NO: 2, and the polypeptide sequence for Figure 5B is SEQ ID NO: 5.

[00143] Polyepitope and non-cleavable ubiquitin G N-terminus: Ubiquitin-G-12aa-A^m-12aa-AAAA-12aa-B^m-12aa-AAAA-(12aa-X^m-12aa-AAAA)$_n$-SIINFEKL-AAAA-Esat6-cMYC. **Figure 5C** exemplarily depicts the polypeptide structure of such arrangement where the cleavable ubiquitin moiety is italics and underlined, the SIINFEKL motif is underlined, the Esat6 motif is in italics, and where the cMY motif is in bold type font. The nucleotide sequence for Figure 5C is SEQ ID NO: 3, and the polypeptide sequence for Figure 5C is SEQ ID NO: 6.

[00144] Exemplary constructs for MHC-I epitopes directed to MHC-II presentation (export from cytoplasm, traffic through endo/lysosome):

[00145] Lysosomal targeting of Kb epitope peptides: (CD1b leader peptide)-20aa-A^m-20aa-GPGPG-20aa B^m-20aa-GPGPG-(20aa-X^m-20aa-GPGPG-)$_n$-Esat6-FlagTag-LAMP1TM/ cytoplasmic tail. **Figure 6A** exemplarily depicts the polypeptide structure of such arrangement where the CD1b leader peptide moiety is bold, the Esat6 motif is in underlined, the Flag-tag motif is italics, and where the LAMP1TM/ cytoplasmic tail is in bold/underline type font. The nucleotide sequence for Figure 6A is SEQ ID NO: 7, and the polypeptide sequence for Figure 6A is SEQ ID NO: 10.

[00146] Recycling lysosome targeting of Kb epitope peptides: (CD1b leader peptide)-20aa-A^m-20aa-GPGPG-20aa B^m-20aa-GPGPG-(20aa-X^m-20aa-GPGPG-)$_n$-Esat6-FlagTag-LAMP1TM/CD1a tail. **Figure 6B** exemplarily depicts the polypeptide structure of such arrangement where the CD1b leader peptide moiety is bold, the Esat6 motif is in underlined, the Flag-tag motif is italics, and where the LAMP1TM motif is in bold/underline type font, and the CD1a targeting motif is underline/italics. The nucleotide sequence for Figure 6B is SEQ ID NO: 8, and the polypeptide sequence for Figure 6B is SEQ ID NO: 11.

[00147] Sorting endosome targeting of Kb epitope peptides: (CD1b leader peptide)-20aa-A^m-20aa-GPGPG-20aa B^m-20aa-GPGPG-(20aa-X^m-20aa-GPGPG-)$_n$-Esat6-FlagTag-LAMP1TM/CD1c tail. **Figure 6C** exemplarily depicts the polypeptide structure of such arrangement where the CD1b leader peptide moiety is bold, the Esat6 motif is in underlined, the Flag-tag motif is italics, and where the LAMP1TM motif is in bold/underline type font, and the CD1c targeting motif is underline/italics. The nucleotide sequence for Figure 6C is SEQ ID NO: 9, and the polypeptide sequence for Figure 6C is SEQ ID NO: 12.

[00148] Exemplary constructs for MHC-II epitopes directed to MHC-II presentation (export from cytoplasm, traffic through endo/lysosome):

[00149] Lysosomal targeting of IAb epitope peptides: (CD1b leader peptide)-20aa-A^m-20aa-GPGPG-20aa B^m-20aa-GPGPG-(20aa-X^m-20aa-GPGPG-)$_n$-Esat6-FlagTag-LAMP1TM/ cytoplasmic tail. **Figure 7A** exemplarily depicts the polypeptide structure of such arrangement where the CD1b leader peptide moiety is bold, the SIINFEKL and Esat6 motifs are underlined, the Flag-tag motif is italics, and where the LAMP1TM/cytoplasmic tail is in

bold/underline type font. The nucleotide sequence for Figure 7A is SEQ ID NO: 13, and the polypeptide sequence for Figure 7A is SEQ ID NO: 16.

[00150] Recycling lysosome targeting of IAb epitope peptides: (CD1b leader peptide)-20aa-A^m-20aa-GPGPG-20aa B^m-20aa-GPGPG-(20aa-X^m-20aa-GPGPG-)$_n$-Esat6-FlagTag-LAMP1TM/CD1a tail. **Figure 7B** exemplarily depicts the polypeptide structure of such arrangement where the CD1b leader peptide moiety is bold, the SIINFEKL and Esat6 motifs are underlined, the Flag-tag motif is italics, where the LAMP1TM motif is in bold/underline type font, and where the CD1a tail is in bold/italics. The nucleotide sequence for Figure 7B is SEQ ID NO: 14, and the polypeptide sequence for Figure 7B is SEQ ID NO: 17.

[00151] Sorting endosome targeting of IAb epitope peptides: (CD1b leader peptide)-20aa-A^m-20aa-GPGPG-20aa B^m-20aa-GPGPG-(20aa-X^m-20aa-GPGPG-)$_n$-Esat6-FlagTag-LAMP1TM/CD1c tail. **Figure 7C** exemplarily depicts the polypeptide structure of such arrangement where the CD1b leader peptide moiety is bold, the SIINFEKL and Esat6 motifs are underlined, the Flag-tag motif is italics, where the LAMP1TM motif is in bold/underline type font, and where the CD1c tail is in bold/italics. The nucleotide sequence for Figure 7C is SEQ ID NO: 15, and the polypeptide sequence for Figure 7C is SEQ ID NO: 18.

[00152] *In vivo* vaccination

[00153] **Figure 8** depicts an exemplary *in vivo* vaccination experiment where nine groups of C57bl/6 mice were immunized with nine distinct recombinant Ad5 viruses comprising the sequence arrangements substantially as described above. Immunization followed a biweekly schedule of administration with distinct routes for separate groups of animals (subcutaneous and intravenous) as is schematically shown in Figure 8. Tumor challenge with B16-F10 melanoma cells was at day 42, followed by administration of an M2 macrophage suppressive drug (RP 182) and IL-15 superagonist (Alt-803). **Figures 9A-9C** depict exemplary results for subcutaneous administration, while **Figures 10A-10C** depict exemplary results for intravenous administration.

[00154] Notably, subcutaneous injection of adenovirus encoding Class I polytopes directed to the cytoplasm and MHC-I presentation did not provide a significant immune protection, regardless of the presence or absence of ubiquitination as can be taken from Figure 9A. On the other hand, where the Class I polytopes were directed to the endosomal and lysosomal compartments for processing and presentation via MHC-II, some protective immunity was

observed for direction to the recycling endosomal compartment and lysosomal compartment as is evident from Figure 9B. Even stronger immune protection was observed when Class II polytopes were directed to the endosomal and lysosomal compartments for processing and presentation via MHC-II. Here, the strongest protection was observed for lysosomal and sorting endosomal compartments as is shown in Figure 9C.

[00155] When immunization was performed with the same viral constructs, albeit via intravenous injection, protective effect of neoepitope vaccination was observed for Class I neoepitopes directed to the cytoplasm where the polytope included cleavable ubiquitin, and some protective effect was observed where the polytope included non-cleavable ubiquitin as can be seen from Figure 10A. Notably, when the Class I polytopes were directed to the endosomal and lysosomal compartment, stronger protective effect was observed in all vaccinations as is shown in Figure 10B. Moreover, strong protective effect was observed when Class II polytopes were directed to the endosomal and lysosomal compartments for processing and presentation via MHC-II. Here, the strongest protection was observed for recycling and sorting endosomal compartments as is shown in the graph of Figure 10C.

[00156] Comparison of Routes and Vectors

[00157] In still further experiments, the inventors further investigated if the actual route of administration and type of expression vector had an effect on the therapeutic efficacy. To that end, a substantially identical mouse model as described above was employed and the polytope was constructed from neoepitopes from B16F10 melanoma cells. Figures 11A and 11B show exemplary results for the experiments where the expression vector was an adenovirus. As can be readily seen, subcutaneous administration of an adenoviral expression system encoding MHC II targeted polytope conferred a significant immune protective effect for all sub-cellular locations while a null vector failed to provide immune protective effect as can be taken from **Figure 11A**. Notably, when the same vector constructs were tested using intravenous administration, immune protective effect was less pronounced as can be taken from **Figure 11B**.

[00158] Conversely, where the expression vector was a plasmid (here: shuttle vector for generating the adenoviral expression vector) targeting MHC II presentation as above, subcutaneous administration of the plasmid conferred a notable immune protective effect versus null vector as is shown in **Figure 11C**. Moreover, and unexpectedly, where the same

plasmid was administered by intravenous injection, a substantial immune protective effect was observed, even for the null vector as can be taken from **Figure 11D**. While not wishing to be bound by any particular theory or hypothesis, the inventors therefore contemplate that the immune protective effect may be at least in part due to an innate and an adaptive immune response.

[00159] **Figure 11E** illustrates a comparison between use of different routes using empty plasmid ('null') versus empty viral ('null') expression vector. As can be readily seen from the results, subcutaneous administration did not confer immune protective effect, while the strongest immune protective effect was observed by intravenous administration of null plasmid.

[00160] **Figure 12** depicts the data shown in **Table 3** below where the type of expression vector is shown as "Type", and the route of administration indicated as "Route". The particular MHC targeting and targeting of sub-cellular location is shown under "Nucleic Acid", while tumor volume is indicated for the dates measured after implantation of B16-F10 melanoma cells.

Table 3

#	Type	Route	Nucleic Acid	Day 7	Day 10	Day 14	Day 17	Day 21	Day 24
1	AdV	Sub Q	Null AdV	29.49791667	57.92242	260.014833	688.7629	1246.774	2119.323
2	AdV	Sub Q	K^b-Cyto AdV	23.97941667	39.50383	187.24075	529.0868	1715.039	2993.372
3	AdV	Sub Q	K^b-UBQ (cleavable) AdV	26.4185	50.95125	253.255	685.7033	1882.06	3773.93
4	AdV	Sub Q	I-A^b-CD1a AdV	16.135	16.195	26.6793333	60.27433	157.3007	341.9303
5	AdV	Sub Q	I-A^b-CD1c AdV	15.75	17.37667	29.32725	59.65917	257.8108	385.198
6	AdV	Sub Q	I-A^b-LAMP1 AdV	17.64666667	18.82333	21.9275	50.01192	116.2528	310.7113
7	AdV	IV	Null AdV	29.06025	60.5985	219.399833	531.3525	990.4138	2527.871
8	AdV	IV	K^b-Cyto AdV	36.07041667	48.72017	113.244	408.8533	1372.974	2887.699
9	AdV	IV	K^b-UBQ (cleavable) AdV	26.62333333	65.407	202.246667	438.495	1565.227	3128.09
10	AdV	IV	I-A^b-CD1a AdV	22.08333333	44.61333	114.910333	267.6967	767.7163	1018.658
11	AdV	IV	I-A^b-CD1c AdV	28.1535	34.5925	62.64375	169.7857	542.2878	1140.391
12	AdV	IV	I-A^b-LAMP1 AdV	29.82733333	70.32883	183.320333	344.7302	648.9571	1033.862
13	Plasmid	Sub Q	Null/pShuttle plasmid	38.22775	76.19925	263.548875	620.3451	837.5773	1473.021
14	Plasmid	Sub Q	K^b-Cyto Plasmid	31.7225	81.33738	277.45375	724.6457	1107.705	2506.988
15	Plasmid	Sub Q	K^b-UBQ (cleavable) Plasmid	19.42	40.41075	181.246	454.0527	951.4715	1995.16
16	Plasmid	Sub Q	I-A^b-CD1a Plasmid	25.595	32.52	70.540625	179.315	522.0373	1340.285
17	Plasmid	Sub Q	I-A^b-CD1c Plasmid	15.19	28.58125	149.4485	358.5129	933.2926	1681.392
18	Plasmid	Sub Q	I-A^b-LAMP1 Plasmid	19.6025	32.69625	80.381875	160.4238	542.5311	1493.744
19	Plasmid	IV	Null/pShuttle Plasmid	21.3825	42.272	116.60725	324.2005	659.2577	1528.49
20	Plasmid	IV	K^b-Cyto Plasmid	14.155	30.05125	106.206	319.6288	562.9468	1113.58
21	Plasmid	IV	K^b-UBQ (cleavable) Plasmid	20.4275	45.167	198.383625	486.2856	1111.359	2619.425
22	Plasmid	IV	I-A^b-CD1a Plasmid	16.45	22.293	61.65325	151.3369	554.821	1612.671
23	Plasmid	IV	I-A^b-CD1c Plasmid	11.725	13.7225	27.265125	73.65725	184.9038	447.9333
24	Plasmid	IV	I-A^b-LAMP1 Plasmid	15.6875	19.625	70.805875	259.7303	422.639	1223.331

[00161] As can be readily seen from the data presented here, targeting MHC-II matched neoepitopes of a polytope and targeting the polytope towards MHC-II presentation via CD1c, LAMP1, and CD1a was significantly effective when administered intravenously in plasmid

form, and subcutaneously in adenovirus form. Notably, and as also reflected in the data, targeting MHC-I was significantly less effective to provide immune protection.

[00162] The recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (*e.g.*, "such as") provided with respect to certain embodiments herein is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the invention.

[00163] It should be apparent to those skilled in the art that many more modifications besides those already described are possible without departing from the inventive concepts herein. The inventive subject matter, therefore, is not to be restricted except in the scope of the appended claims. Moreover, in interpreting both the specification and the claims, all terms should be interpreted in the broadest possible manner consistent with the context. In particular, the terms "comprises" and "comprising" should be interpreted as referring to elements, components, or steps in a non-exclusive manner, indicating that the referenced elements, components, or steps may be present, or utilized, or combined with other elements, components, or steps that are not expressly referenced. Where the specification claims refers to at least one of something selected from the group consisting of A, B, C .... and N, the text should be interpreted as requiring only one element from the group, not A plus N, or B plus N, etc.