HUMAN PAPILLOMA VIRUS VACCINES AND USES OF THE SAME FOR HPV ASSOCIATED DISEASES

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of U.S. Provisional Patent Application No. 63/118,222 filed on November 25, 2020, which is hereby incorporated by reference in its entirety.

REFERENCE TO SEQUENCE LISTING

[0002] The content of the electronically submitted sequence listing (Name: INX00511US-V1_SEQ-LIST_ST25; Size: 71,267 bytes; Date of Creation: November 24, 2020) filed with this application is incorporated herein by reference in its entirety.

FIELD OF THE DISCLOSURE

[0003] The present disclosure relates to innovative, multi-antigenic HPV molecular vaccines. More particularly, present disclosure relates to HPV molecular vaccines targeting HPV6 and HPV11 associated pathologies.

BACKGROUND OF THE DISCLOSURE

[0004] Globally, tens of millions of people are currently infected with human papillomavirus (HPV), and millions more become newly infected each year. Currently, vaccinations can protect against diseases caused by HPV when given to recommended age groups; however, there is a tremendous need for HPV vaccines with expanded coverage (e.g., against multiple and/or targeted subsets of HPV strains) and functionality against HPV-associated pathologies.

[0005] Recurrent respiratory papillomatosis (RRP) is rare disorder characterized by recurrent growth of wart-like benign papillomas in the respiratory tract. RRP occurs in both children and adults. Juvenile onset is almost always diagnosed by age ten and usually before the age of five. RRP is caused by human papillomavirus (HPV), in particular HPV6 and HPV11.

[0006] Genital warts are extremely common, and may also be known by other names such as anogenital warts or condylomata acuminate with between 500,000 to one million new cases diagnosed each year in the United States alone. HPV6 and HPV11 are responsible for about 90% of genital warts. The current treatment options are largely centered upon removal of the warts rather than elimination of the underlying viral infection. A wide range of therapies are presently in use, which are highly variable and can differ dramatically. However, no treatment has been 100% effective in eliminating warts and preventing them from coming back in all patients.

[0007] The current standard of treatment for RRP is surgical excision with adjuvant therapies as needed. However, the distinguishing aspect of RRP is the tendency for the growth to recur requiring repeated surgeries. The aggressiveness varies among patients - Some individuals may require surgery every few weeks while others may only require surgery twice a year or only a few times during their life.

[0008] Unfortunately, previously developed therapies have not been able to prevent multiple recurrence for RRP. There is a need to develop better treatment for RRP patients. The present specification describes an innovative immunotherapy approach, a novel therapeutic HPV vaccine aiming to induce anti-HPV6/11 immunity to treat and alter the course of RRP and other HPV6/11 associated disease such as genital warts.

[0009] Vaccines have traditionally been used as preventive measures against infectious diseases. Several HPV vaccines (GARDASIL and CERVARIX) have been developed and approved by the Food and Drug Administration (FDA) to protect women from cervical cancer, which is also associated with HPV infection of the same subtypes that cause RRP.

[0010] GARDASIL is a vaccine against HPV types 6, 11, 16 and 18 (RRP is caused by types 6 and 11), while CERVARIX is a vaccine against only HPV types 16 and 18. Gardasil-9 is targeted to protect against high-risk strains 31, 33, 45, 52, 58.

[0011] The prophylactic vaccines described above have been successful at preventing healthy people from acquiring HPV infection as well as previously infected patients from being re-infected; however, they are not able to treat or cure in patients who are already infected by HPV.

[0012] A number of therapeutic HPV vaccines have reached clinical trials, some show promising results. MVA E2, Recombinant Modified Vaccinia Ankara was reported to

show 90% lesion clearance in female patients and 100% lesion clearance in male patients. IN0-3106, a DNA vaccine consisting of plasmids encoding the E6 and E7 was reported to have clinical efficacy in a study of two patients with recurrent respiratory papillomatosis. Despite this apparent positive progress, however, there is a dearth of approved HPV therapeutic vaccines; indeed, as of this submission date, the patent applicants are not currently aware of any FDA-approved HPV therapeutic vaccine.

INCORPORATION BY REFERENCE

[0013] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

SUMMARY OF THE DISCLOSURE

[0014] Provided herein are non-naturally occurring polynucleotides (and polypeptides expressed therefrom) encoding non-naturally occurring polypeptides comprising immune response-inducing human papilloma virus (HPV) polypeptides. Also provided herein are non-naturally occurring, variably arrayed, configurations of HPV immune response-inducing polypeptides connected by various polypeptide linker sequences; thereby, comprising fusion proteins useful as vaccine antigens.

[0015] The invention includes, but is not limited to, compositions, methods of making compositions, and uses of non-naturally-occurring polynucleotides encoding non-naturally occurring polypeptides comprising antigens for vaccines directed against human papillomavirus (HPV), in particular HPV6 and HPV11; and, therapeutic approaches to treat pathological conditions caused by these particular HPV agents. In particular, the invention comprises polynucleotides, and polypeptides encoded by same, encoding multi-antigenic polypeptides derived from HPV6, HPV11 and HPV16 *and other polypeptide sequences* for use as vaccine components and for treatment of disorders associated with HPV infection. For example, an "indication" of the present invention is use of compositions described herein as therapeutic vaccines against HPV6 and/or HPV11 (HPV6/11) induced or associated diseases; such as, but not limited to, treatment of

recurrent respiratory papillomatosis (RRP), anogenital warts, and other HPV6/11-associated diseases.

[0016] The present invention includes unique and innovative configurations HPV polypeptide vaccine design approaches for HPV6/11 induced diseases. The invention includes a mix of design strategies; such as, use of full protein sequences, use of peptide "fragments", hybrid polypeptide constructs, introduction of amino acid substitutions, insertions, deletions and rearrangement of polypeptide and peptide HPV gene products (proteins). The five examples of design concepts (Designs No. 1-5) as described herein can be further modified and/or optimized into different variations as needed.

[0017] As indicated above, unique modifications of antigens described herein have been made by introducing point mutations, substitution mutations, and/or reordering of viral polypeptide sequences, in particular, to prevent oncogenic gene expression and/or disable essential viral functions (e.g. viral replication).

[0018] In certain embodiments, HPV early proteins E2 and E4 have been identified as novel antigens for HPV6/11 and incorporated them into designs described herein. In certain embodiments, the invention comprises the novel incorporation of four antigenic components from HPVs into one HPV vaccine design. In certain embodiments, the invention comprises the novel integration and combination of both high cancer risk and low cancer risk HPV genotype epitopes into one HPV antigen construct. The antigen designs described herein were carefully designed and selected for human *in vivo* antigenicity potential, *without* cross reactivity to endogenous human proteins, as determined via *in-silico* analysis.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] **Figure 1** is a depiction of the HPV6/11 antigen design work flow, including reference to tools and techniques utilized to design and evaluate the designs described herein.

[0020] **Figures 2 - 6** provide schematic, sequence and structural representations of the polypeptides and modules used therein for HPV6/11 Antigen Designs No. 1 through No. 5.

DETAILED DESCRIPTION OF THE DISCLOSURE

[0021] It is to be understood that the present disclosure is not limited to the particular embodiments described herein and as such can vary. Those of skill in the art will recognize that there are variations and modifications of the present disclosure, which are encompassed within its scope.

[0022] Rigid linker, cleavable linkers, HPV epitope linkers and HPV agonist enhancer linkers used to connect antigenic components as described herein and as shown in the Figures and Sequence Listing can be further configured in a variable number of other permutations.

[0023] Embodiments of the invention include use of a rigid linker peptide (EAAAK)**n**; *i.e.,* (SEQ ID NO:76)**n**; where "**n** "can vary from 1 to 10 (*e.g.*, EAAAK is n=1; EAAAKEAAAK is n=2, EAAAKEAAAKEAAAK is n=3; *etc.*) and which may be encoded by a variety of various nucleotide sequences (for example, but not limited to, SEQ ID Nos:77-81). It is also contemplated that the number of "E" (*aka*, Glu; Glutamic acid), "A" (*aka*, Ala; Alanine) and "K" (*aka*, Lys; Lysine) residues may be varied in number and patterns of repetition in generating rigid linker peptide sequences of similar composition.

[0024] Likewise, antigenic fragments as described herein and as shown in the Figures can be shuffled in various permutations both within and among each design.

[0025] Vaccine design as described herein and as shown in the figures can be utilized for treatment of recurrent respiratory papilloma pathologies and disorders as well as for treatment of genital, anogenital and other warts and other HPV-induced diseases, disorders and hyper-proliferative pathologies.

[0026] Constitutive or induced expression of HPV early (E) region proteins provide targets for an effective HPV6/11 vaccine.

[0027] HPV early region protein functions include the following: E1, E2 have functions in viral replication/transcription (*e.g.*, E2 regulates expression of E6 and E7; and, E1/E2 interaction is essential for viral replication); E4, E5 have increased expression during late stage of viral replication cycle (though HPV11 E4 is not essential for viral genome amplification); and, E6, E7 act co-operatively during replication (E6 is required for episomal genome maintenance, E7 expands compartment of epithelial cells active in DNA replication).

[0028] E2 is an excellent target, is expressed in infected cells, has been shown to generate a strong immune response and there is preclinical and clinical data using E2 as a sole target in a vaccine (plasmid DNA and MVA E2)

[0029] E4 is a unique target, is expressed in infected cells, enhanced expression late, while not essential for HPV11 replication may have a critical role in generation of the episomal state; no vaccine data available.

[0030] E6 and E7 are both are expressed in infected cells, but are absent on healthy cells, making them ideal antigenic targets, and there is clinical data using E6 and E7 as targets in vaccines.

[0031] All terms are intended to be understood as they would be understood by a person skilled in the art. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the disclosure pertains.

[0032] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

[0033] Although various features of the disclosure can be described in the context of a single embodiment, the features can also be provided separately or in any suitable combination. The following definitions supplement those in the art and are directed to the current application and are not to be imputed to any related or unrelated case, *e.g.*, to any commonly owned patent or application. Accordingly, the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

DEFINITIONS

[0034] In this application, the use of the singular includes the plural unless specifically stated otherwise. It must be noted that, as used in the specification, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

[0035] In this application, the use of "or" means "and/or" unless stated otherwise. The terms "and/or" and "any combination thereof" and their grammatical equivalents as used herein, can be used interchangeably. These terms can convey that any combination is specifically contemplated. Solely for illustrative purposes, the following phrases "A, B, and/or C" or "A, B, C, or any combination thereof" can mean "A individually; B individually; C individually; A and B; B and C; A and C; and A, B, and C." The term

"or" can be used conjunctively or disjunctively, unless the context specifically refers to a disjunctive use.

[0036] Furthermore, use of the term "including" as well as other forms, such as "include," "includes," and "included," is not limiting; *i.e.*, "including" does not mean "limited to".

[0037] Reference in the specification to "some embodiments," "an embodiment," "one embodiment" or "other embodiments" means that a particular feature, structure, or characteristic described in connection with the embodiments is included in at least some embodiments, but not necessarily all embodiments, of the present disclosures.

[0038] As used in this specification and claim(s), the words "comprising" (and any form of comprising, such as "comprise" and "comprises"), "having" (and any form of having, such as "have" and "has"), "including" (and any form of including, such as "includes" and "include") or "containing" (and any form of containing, such as "contains" and "contain") are inclusive or open-ended and do not exclude additional possible components, elements, or method steps. It is contemplated that any embodiment discussed in this specification can be implemented with respect to any method or composition of the disclosure, and vice versa. Furthermore, compositions of the present disclosure can be used to achieve methods of the present disclosure.

[0039] The term "about" in relation to a reference numerical value and its grammatical equivalents as used herein can include the numerical value itself and a range of values plus or minus 10% from that numerical value.

[0040] The term "about" or "approximately" means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, *i.e.*, the limitations of the measurement system. For example, "about" can mean within 1 or more than 1 standard deviation, per the practice in the art. Alternatively, "about" can mean a range of up to 20%, up to 10%, up to 5%, or up to 1% of a given value. In another example, the amount "about 10" includes 10 and any amounts from 9 to 11. In yet another example, the term "about" in relation to a reference numerical value can also include a range of values plus or minus 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1% from that value. Alternatively, particularly with respect to biological systems or processes, the term "about" can mean within an order of magnitude, preferably within 5-fold, and more preferably within 2-

fold, of a value. Where particular values are described in the application and claims, unless otherwise stated the term "about" meaning within an acceptable error range for the particular value should be assumed.

[0041] The term "isolated" and its grammatical equivalents as used herein refer to the removal of a nucleic acid from its natural environment. The term "purified" and its grammatical equivalents as used herein refer to a molecule or composition, whether removed from nature (including genomic DNA and mRNA) or synthesized (including cDNA) and/or amplified under laboratory conditions, that has been increased in purity, wherein "purity" is a relative term, not "absolute purity." It is to be understood, however, that nucleic acids and proteins can be formulated with diluents or adjuvants and still for practical purposes be isolated. For example, nucleic acids typically are mixed with an acceptable carrier or diluent when used for introduction into cells. The term "substantially purified" and its grammatical equivalents as used herein refer to a nucleic acid sequence, polypeptide, protein or other compound which is essentially free, *i.e.*, is more than about 50% free of, more than about 70% free of, more than about 90% free of, the polynucleotides, proteins, polypeptides and other molecules that the nucleic acid, polypeptide, protein or other compound is naturally associated with.

[0042] "Polynucleotide", "oligonucleotide", "polynucleotide construct", "gene", "gene construct", "heterologous gene" and their grammatical equivalents as used herein refer to a polymeric form of nucleotides or nucleic acids of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double and single stranded DNA, triplex DNA, as well as double and single stranded RNA. It also includes modified, for example, by methylation and/or by capping, and unmodified forms of the polynucleotide. The term is also meant to include molecules that include non-naturally occurring or synthetic nucleotides as well as nucleotide analogs. The nucleic acid sequences and vectors disclosed or contemplated herein can be introduced into a cell by, for example, transfection, transformation, or transduction.

[0043] "Polynucleotides encoding polypeptides" of the invention, and like or similar terms and phrases as used herein, include any polynucleotides which can be used to express (*i.e.,* encode) polypeptides of the invention. Indeed, it is well-known and understood by those of ordinary skill in the art how to construct any number of different

polynucleotide sequences which may encode the same polypeptide (*i.e.*, a polypeptide of identical amino acid sequence, although encoded by different sequences of codons (nucleotide triplets)). For example, one may use varying codons within a standard genetic code to generate a variety of polynucleotide sequences encoding a particular polypeptide. Moreover, one may also generate a variety of polynucleotide sequences encoding a particular polypeptide via use of codon selection as guided and selected via use of the "codon-use frequency" observed in a given organism; such as, by way of example, but not limited to, a *Homo sapiens* codon-use frequency table such as shown below.

[0044] Accordingly, it is also well-known and understood by those of ordinary skill in the art, amino acids are encoded by codons (nucleotide triplets) such as in the table below:

Table 1: Standard Genetic Code

Nucleotide position in codon - Amino Acid encoded					
First nucleotide	Second nucleotide				Third nucleotide
	U	C	A	G	
U	UUU - Phe	UCU - Ser	UAU - Tyr	UGU - Cys	U
	UUC - Phe	UCC - Ser	UAC - Tyr	UGC - Cys	C
	UUA - Leu	UCA - Ser	UAA - *	UGA - *	A
	UUG - Leu	UCG - Ser	UAG - *	UGG - Trp	G
C	CUU - Leu	CCU - Pro	CAU - His	CGU - Arg	U
	CUC - Leu	CCC - Pro	CAC - His	CGC - Arg	C
	CUA - Leu	CCA - Pro	CAA - Gln	CGA - Arg	A
	CUG - Leu	CCG - Pro	CAG - Gln	CGG - Arg	G
A	AUU - Ile	ACU - Thr	AAU - Asn	AGU - Ser	U
	AUC - Ile	ACC - Thr	AAC - Asn	AGC - Ser	C
	AUA - Ile	ACA - Thr	AAA - Lys	AGA - Arg	A
	AUG - Met	ACG - Thr	AAG - Lys	AGG - Arg	G
G	GUU - Val	GCU - Ala	GAU - Asp	GGU - Gly	U
	GUC - Val	GCC - Ala	GAC - Asp	GGC - Gly	C
	GUA - Val	GCA - Ala	GAA - Glu	GGA - Gly	A
	GUG - Val	GCG - Ala	GAG - Glu	GGG - Gly	G

(U indicating "Uracil" as incorporated into mRNA in place of T (Thymine) as found in DNA; both of which form complementary base-pairs with "A" (Adenine); "*" indicates stop codons.)

Table 2: Homo sapiens, table of codon-use frequency

[frequency per thousand nucleotide triplets in *H. sapiens*]

UUU [17.6]	UCU [15.2]	UAU [12.2]	UGU [10.6]
UUC [20.3]	UCC [17.7]	UAC [15.3]	UGC [12.6]
UUA [7.7]	UCA [12.2]	UAA [1.0]	UGA [1.6]
UUG [12.9]	UCG [4.4]	UAG [0.8]	UGG [13.2]
CUU [13.2]	CCU [17.5]	CAU [10.9]	CGU [4.5]
CUC [19.6]	CCC [19.8]	CAC [15.1]	CGC [10.4]
CUA [7.2]	CCA [16.9]	CAA [12.3]	CGA [6.2]
CUG [39.6]	CCG [6.9]	CAG [34.2]	CGG [11.4]
AUU [16.0]	ACU [13.1]	AAU [17.0]	AGU [12.1]
AUC [20.8]	ACC [18.9]	AAC [19.1]	AGC [19.5]
AUA [7.5]	ACA [15.1]	AAA [24.4]	AGA [12.2]
AUG [22.0]	ACG [6.1]	AAG [31.9]	AGG [12.0]
GUU [11.0]	GCU [18.4]	GAU [21.8]	GGU [10.8]
GUC [14.5]	GCC [27.7]	GAC [25.1]	GGC [22.2]
GUA [7.1]	GCA [15.8]	GAA [29.0]	GGA [16.5]
GUG [28.1]	GCG [7.4]	GAG [39.6]	GGG [16.5]

[0045] Furthermore, publicly available software resources are readily available for computer-generated "reverse-translation" (also known as, "back translation" of polypeptide sequences (*i.e.*, converting polypeptide sequences into nucleotide sequences encoding same). See, for example:

The EMBL-EBI search and sequence analysis tools APIs in 2019

Madeira, F., *et al., Nucleic Acids Res*, 47(W1), W636–W641 (2019)

DOI: 10.1093/nar/gkz268; EuropePMC: 30976793.

Using EMBL-EBI services via web interface and programmatically via web services

Madeira, F., *et al., Curr Protoc in Bioinformatics*, 66(1):e74 (2019)

DOI: 10.1002/cpbi.74; EuropePMC: 31039604.

Programmatic access to bioinformatics tools from EMBL-EBI update: 2017.

Chojnacki, S, *et al.*, *Nucleic Acids Res*. 2017 Jul 3;45(W1):W550-W553 (2017).

DOI: 10.1093/nar/gkx273. PMID: 28431173; PMCID: PMC5570243.

A new and updated resource for codon usage tables.

Athey, J., *et al., BMC Bioinformatics* 18:391 (2017)

DOI 10.1186/s12859-017-1793-7.

The MPI Bioinformatics Toolkit for protein sequence analysis

Biegert, A., *et al., Nucleic Acids Research*, Vol. 34, Iss. Suppl_2, pp. W335–W339 (2006)

An improved tool for molecular biology

Rawitch, A.B., *Science*; Washington Vol. 288, Iss. 5465, pp. 457-458 (2000).

EMBOSS: The European Molecular Biology Open Software Suite

Rice, P., *et al., Trends Genet.* 16(6)276-277 (2000)

DOI: 10.1016/S0168-9525(00)02024-2; PubMed: 10827456.

Microcomputer programs for back translation of protein to DNA sequences and analysis of ambiguous DNA sequences

Mount, D.W and Conrad, B. *Nucleic Acids Res.*, Vol. 12, Iss.1, Part 2, pp. 819–823 (1984).

Computer programs for handling nucleic acid sequences

Keller, C. *et al.*, Nucleic Acids Res., Vol. 12, Iss. 1, Part 1, pp. 379–386 (1984).

[0046] “Polypeptide”, “peptide” "polypeptide construct" and "peptide construct" and their grammatical equivalents as used herein refer to a polymer of amino acid residues. A “mature protein” is a protein which is full-length and which, optionally, includes glycosylation or other modifications typical for the protein in a given cellular environment. As disclosed herein, embodiments of the invention include HPV antigens/antigenic polypeptides, peptides, and mature proteins described herein and also polynucleotides (DNA or RNA) that encode the same. Polypeptides and proteins disclosed herein (including functional portions and functional variants thereof) can comprise synthetic amino acids in place of one or more naturally-occurring amino acids. Such synthetic amino acids are known in the art, and include, for example, aminocyclohexane carboxylic acid, norleucine, α-amino n-decanoic acid, homoserine, S-acetylaminomethyl-cysteine, trans-3- and trans-4-hydroxyproline, 4-aminophenylalanine, 4-nitrophenylalanine, 4-chlorophenylalanine, 4-carboxyphenylalanine, β-phenylserine β-

hydroxyphenylalanine, phenylglycine, α-naphthylalanine, cyclohexylalanine, cyclohexylglycine, indoline-2-carboxylic acid, 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, aminomalonic acid, aminomalonic acid monoamide, N'-benzyl-N'-methyl-lysine, N',N'-dibenzyl-lysine, 6-hydroxylysine, ornithine, α-aminocyclopentane carboxylic acid, α-aminocyclohexane carboxylic acid, α-aminocycloheptane carboxylic acid, α-(2-amino-2-norbornane)-carboxylic acid, α,γ-diaminobutyric acid, α,β-diaminopropionic acid, homophenylalanine, and α-tert-butylglycine.

[0047] Nucleic acids and/or nucleic acid sequences are "homologous" when they are derived, naturally or artificially, from a common ancestral nucleic acid or nucleic acid sequence. Proteins and/or protein sequences are "homologous" when their encoding DNAs are derived, naturally or artificially, from a common ancestral nucleic acid or nucleic acid sequence. The homologous molecules can be termed homologs. For example, any naturally occurring proteins, as described herein, can be modified by any available mutagenesis method. When expressed, this mutagenized nucleic acid encodes a polypeptide that is homologous to the protein encoded by the original nucleic acid. Homology is generally inferred from sequence identity between two or more nucleic acids or proteins (or sequences thereof). The precise percentage of identity between sequences that is useful in establishing homology varies with the nucleic acid and protein at issue, but as little as 25% sequence identity is routinely used to establish homology. Higher levels of sequence identity, *e.g.*, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or 99% or more can also be used to establish homology. Methods for determining sequence identity percentages (*e.g.*, BLASTP and BLASTN using default parameters) are described herein and are generally available.

[0048] The term "identical" and its grammatical equivalents as used herein or "sequence identity" in the context of two nucleic acid sequences or amino acid sequences of polypeptides refers to the residues in the two sequences which are the same when aligned for maximum correspondence over a specified comparison window. A "comparison window", as used herein, refers to a segment of at least about 20 contiguous positions, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence can be compared to a reference sequence of the same number of contiguous positions after the two sequences are aligned optimally. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison

can be conducted by the local homology algorithm of Smith and Waterman, *Adv. Appl. Math.*, 2:482 (1981); by the alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.*, 48:443 (1970); by the search for similarity method of Pearson and Lipman, *Proc. Nat. Acad. Sci U.S.A.*, 85:2444 (1988); by computerized implementations of these algorithms (including, but not limited to CLUSTAL in the PC/Gene program by Intelligentics, Mountain View Calif., GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wis., U.S.A.); the CLUSTAL program is well described by Higgins and Sharp, *Gene,* 73:237-244 (1988) and Higgins and Sharp, *CABIOS*, 5:151-153 (1989); Corpet et al., *Nucleic Acids Res.*, 16:10881-10890 (1988); Huang et al., *Computer Applications in the Biosciences*, 8:155-165 (1992); and Pearson et al., *Methods in Molecular Biology*, 24:307-331 (1994). Alignment is also often performed by inspection and manual alignment**.** In one class of embodiments, the polypeptides herein are at least 80%, 85%, 90%, 98% 99% or 100% identical to a reference polypeptide, or a fragment thereof, *e.g.*, as measured by BLASTP (or CLUSTAL, or any other available alignment software) using default parameters. Similarly, nucleic acids can also be described with reference to a starting nucleic acid, *e.g.*, they can be 50%, 60%, 70%, 75%, 80%, 85%, 90%, 98%, 99% or 100% identical to a reference nucleic acid or a fragment thereof, *e.g.*, as measured by BLASTN (or CLUSTAL, or any other available alignment software) using default parameters. When one molecule is said to have certain percentage of sequence identity with a larger molecule, it means that when the two molecules are optimally aligned, said percentage of residues in the smaller molecule finds a match residue in the larger molecule in accordance with the order by which the two molecules are optimally aligned.

[0049] The term "substantially identical" and its grammatical equivalents as applied to nucleic acid or amino acid sequences mean that a nucleic acid or amino acid sequence comprises a sequence that has at least 90% sequence identity or more, at least 95%, at least 98% and at least 99%, compared to a reference sequence using the programs described above, *e.g.*, BLAST, using standard parameters. For example, the BLASTN program (for nucleotide sequences) uses as defaults a word length (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a word length (W) of 3, an expectation

(E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1992)). Percentage of sequence identity is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window can comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. In embodiments, the substantial identity exists over a region of the sequences that is at least about 50 residues in length, over a region of at least about 100 residues, and in embodiments, the sequences are substantially identical over at least about 150 residues. In embodiments, the sequences are substantially identical over the entire length of the coding regions.

[0050] An "expression vector" or "vector" is any genetic element, *e.g.*, a plasmid, chromosome, virus, transposon, behaving either as an autonomous unit of polynucleotide replication within a cell. (*i.e.* capable of replication under its own control) or being rendered capable of replication by insertion into a host cell chromosome, having attached to it another polynucleotide segment, so as to bring about the replication and/or expression of the attached segment. Suitable vectors include, but are not limited to, plasmids, transposons, mammalian viruses, bacteriophages and cosmids. Vectors can contain polynucleotide sequences which are necessary to effect ligation or insertion of the vector into a desired host cell and to effect the expression of the attached segment. Such sequences differ depending on the host organism; they include promoter sequences to effect transcription, enhancer sequences to increase transcription, ribosomal binding site sequences and transcription and translation termination sequences. Alternatively, expression vectors can be capable of directly expressing nucleic acid sequence products encoded therein without ligation or integration of the vector into host cell DNA sequences. In some embodiments, the vector is an "episomal expression vector" or "episome," which is able to replicate in a host cell, and persists as an extrachromosomal

segment of DNA within the host cell in the presence of appropriate selective pressure (see, *e.g.*, Conese et al., *Gene Therapy, 11*:1735-1742 (2004)).

[0051] The term "adenovirus" and "adenoviral vector" as used herein, refers to an adenovirus that retains the ability to participate in the adenovirus life cycle and/or which has been physically inactivated by, for example, disruption (*e.g.*, sonication), denaturing (*e.g.*, using heat or solvents), or cross-linkage (*e.g.*, via formalin cross-linking). The "adenovirus life cycle" includes (1) virus binding and entry into cells, (2) transcription of the adenoviral genome and translation of adenovirus proteins, (3) replication of the adenoviral genome, and (4) viral particle assembly (see, *e.g.*, Fields Virology, 5th ed., Knipe et al. (eds.), Lippincott Williams & Wilkins, Philadelphia, PA (2006)). Adenoviruses, as used and described herein may also be rendered replication deficient (*i.e.*, do not retain ability to participate in the adenovirus life cycle) by deletion of one or more parts of the naturally occurring viral genome. "Adenoviruses" and "Adenoviral vector," as used and described herein, may include an adenovirus in which the adenoviral genome has been manipulated to accommodate a nucleic acid sequence that is non-native with respect to the adenoviral genome. Typically, an adenoviral vector is generated by introducing one or more mutations (*e.g.*, a deletion, insertion, or substitution) into the adenoviral genome of the adenovirus so as to accommodate the insertion of a non-native nucleic acid sequence, for example, for gene transfer, into the adenovirus.

[0052] The term "selectable marker gene" as used herein refers to a nucleic acid sequence that allows cells expressing the nucleic acid sequence to be specifically selected for or against, in the presence of a corresponding selective agent. Suitable selectable marker genes are known in the art and described in, *e.g.*, International Patent Application Publications WO 1992/08796 and WO 1994/28143; Wigler et al., *Proc. Natl. Acad. Sci. USA, 77*: 3567 (1980); O'Hare et al., *Proc. Natl. Acad. Sci. USA, 78*: 1527 (1981); Mulligan & Berg, *Proc. Natl. Acad. Sci. USA, 78*: 2072 (1981); *Colberre-Garapin et al., J. Mol. Biol., 150*:1 (1981); Santerre et al., *Gene, 30*: 147 (1984); Kent et al., *Science, 237*: 901-903 (1987); Wigler et al., *Cell, 11*: 223 (1977); Szybalska & Szybalski, *Proc. Natl. Acad. Sci. USA, 48*: 2026 (1962); Lowy et al., *Cell, 22*: 817 (1980); and U.S. Pat. Nos. 5,122,464 and 5,770,359.

[0053] The term "coding sequence" as used herein refers to a segment of a polynucleotide that codes for protein. The region or sequence is bounded nearer the 5'

end by a start codon and nearer the 3' end with a stop codon. Coding sequences can also be referred to as open reading frames.

[0054] The term "operably linked" as used herein refers to refers to the physical and/or functional linkage of a DNA segment to another DNA segment in such a way as to allow the segments to function in their intended manners. A DNA sequence encoding a gene product is operably linked to a regulatory sequence when it is linked to the regulatory sequence, such as, for example, promoters, enhancers and/or silencers, in a manner which allows modulation of transcription of the DNA sequence, directly or indirectly. For example, a DNA sequence is operably linked to a promoter when it is ligated to the promoter downstream with respect to the transcription initiation site of the promoter, in the correct reading frame with respect to the transcription initiation site and allows transcription elongation to proceed through the DNA sequence. An enhancer or silencer is operably linked to a DNA sequence coding for a gene product when it is ligated to the DNA sequence in such a manner as to increase or decrease, respectively, the transcription of the DNA sequence. Enhancers and silencers can be located upstream, downstream or embedded within the coding regions of the DNA sequence. A DNA for a signal sequence is operably linked to DNA coding for a polypeptide if the signal sequence is expressed as a pre-protein that participates in the secretion of the polypeptide. Linkage of DNA sequences to regulatory sequences is typically accomplished by ligation at suitable restriction sites or via adapters or linkers inserted in the sequence using restriction endonucleases known to one of skill in the art.

[0055] The terms "induce", "induction" and its grammatical equivalents as used herein refer to an increase in nucleic acid sequence transcription, promoter activity and/or expression brought about by a transcriptional regulator, relative to some basal level of transcription.

[0056] The term "transcriptional regulator" refers to a biochemical element that acts to prevent or inhibit the transcription of a promoter-driven DNA sequence under certain environmental conditions (*e.g.*, a repressor or nuclear inhibitory protein), or to permit or stimulate the transcription of the promoter-driven DNA sequence under certain environmental conditions (*e.g.*, an inducer or an enhancer).

[0057] The term "enhancer" as used herein, refers to a DNA sequence that increases transcription of, for example, a nucleic acid sequence to which it is operably linked.

Enhancers can be located many kilobases away from the coding region of the nucleic acid sequence and can mediate the binding of regulatory factors, patterns of DNA methylation, or changes in DNA structure. A large number of enhancers from a variety of different sources are well known in the art and are available as or within cloned polynucleotides (from, *e.g.*, depositories such as the ATCC as well as other commercial or individual sources). A number of polynucleotides comprising promoters (such as the commonly-used CMV promoter) also comprise enhancer sequences. Enhancers can be located upstream, within, or downstream of coding sequences. The term "Ig enhancers" refers to enhancer elements derived from enhancer regions mapped within the immunoglobulin (Ig) locus (such enhancers include for example, the heavy chain (mu) 5' enhancers, light chain (kappa) 5' enhancers, kappa and mu intronic enhancers, and 3' enhancers (see generally Paul W. E. (ed), Fundamental Immunology, 3rd Edition, Raven Press, New York (1993), pages 353-363; and U.S. Pat. No. 5,885,827).

[0058] The term "promoter" refers to a region of a polynucleotide that initiates transcription of a coding sequence. Promoters are located near the transcription start sites of genes, on the same strand and upstream on the DNA (towards the 5' region of the sense strand). Some promoters are constitutive as they are active in all circumstances in the cell, while others are regulated becoming active in response to specific stimuli, *e.g.*, an inducible promoter. The term "promoter activity" and its grammatical equivalents as used herein refer to the extent of expression of nucleotide sequence that is operably linked to the promoter whose activity is being measured. Promoter activity can be measured directly by determining the amount of RNA transcript produced, for example by Northern blot analysis or indirectly by determining the amount of product coded for by the linked nucleic acid sequence, such as a reporter nucleic acid sequence linked to the promoter.

[0059] "Inducible promoter" as used herein refers to a promoter which is induced into activity by the presence or absence of transcriptional regulators, *e.g.*, biotic or abiotic factors. Inducible promoters are useful because the expression of genes operably linked to them can be turned on or off at certain stages of development of an organism or in a particular tissue. Non-limiting examples of inducible promoters include alcohol-regulated promoters, tetracycline-regulated promoters, steroid-regulated promoters, metal-regulated promoters, pathogenesis-regulated promoters, temperature-regulated promoters and light-regulated promoters. The inducible promoter can be part of a gene

switch or genetic switch. The inducible promoter can be a gene switch ligand inducible promoter. In some cases, an inducible promoter can be a small molecule ligand-inducible two polypeptide ecdysone receptor-based gene switch. In some cases, a gene switch can be selected from ecdysone-based receptor components as described in, but without limitation to, any of the systems described in: International Patent Applications WO 2001/070816; WO 2002/029075; WO 2002/066613; WO 2002/066614; WO 2002/066612; WO 2002/066615; WO 2003/027266; WO 2003/027289; WO 2005/108617; WO 2009/045370; WO 2009/048560; WO 2010/042189; WO 2010/042189; WO 2011/119773; and WO 2012/122025; and U.S. Patent Nos. 7,091,038; 7,776,587; 7,807,417; 8,202,718; 8,105,825; 8,168,426; 7,531,326; 8,236,556; 8,598,409; 8,715,959; 7,601,508; 7,829,676; 7,919,269; 8,030,067; 7,563,879; 8,021,878; 8,497,093; 7,935,510; 8,076,454; 9,402,919; 9,493,540; 9,249,207; and 9,492,482, each of which is incorporated by reference in its entirety).

[0060] The term "gene switch" or "genetic switch" refers to the combination of a response element associated with a promoter, and for instance, an EcR based system, which, in the presence of one or more ligands, modulates the expression of a gene into which the response element and promoter are incorporated. Tightly regulated inducible gene expression systems or gene switches are useful for various applications such as gene therapy, large scale production of proteins in cells, cell based high throughput screening assays, functional genomics and regulation of traits in transgenic plants and animals. Such inducible gene expression systems can include ligand inducible heterologous gene expression systems.

[0061] "Sleeping Beauty (SB) Transposon System" refers a synthetic DNA transposon system for to introducing DNA sequences into the chromosomes of vertebrates. Some exemplary embodiments of the system are described, for example, in U.S. Pat. Nos. 6,489,458, 8,227,432, 9,228,180 and WO/2016/145146. The Sleeping Beauty transposon system is composed of a Sleeping Beauty (SB) transposase and a SB transposon. In embodiments, the Sleeping Beauty transposon system can include the SB11 transposon system, the SB100X transposon system, or the SB110 transposon system.

[0062] "Transposon" or "transposable element" (TE) is a vector DNA sequence that can change its position within the genome, sometimes creating or reversing mutations and altering the cell's genome size. Transposition often results in duplication of the TE.

Class I TEs are copied in two stages: first they are transcribed from DNA to RNA, and the RNA produced is then reverse transcribed to DNA. This copied DNA is then inserted at a new position into the genome. The reverse transcription step is catalyzed by a reverse transcriptase, which can be encoded by the TE itself. The characteristics of retrotransposons are similar to retroviruses, such as HIV. The cut-and-paste transposition mechanism of class II TEs does not involve an RNA intermediate. The transpositions are catalyzed by several transposase enzymes. Some transposases non-specifically bind to any target site in DNA, whereas others bind to specific DNA sequence targets. The transposase makes a staggered cut at the target site resulting in single-strand 5' or 3' DNA overhangs (sticky ends). This step cuts out the DNA transposon, which is then ligated into a new target site; this process involves activity of a DNA polymerase that fills in gaps and of a DNA ligase that closes the sugar-phosphate backbone. This results in duplication of the target site. The insertion sites of DNA transposons can be identified by short direct repeats which can be created by the staggered cut in the target DNA and filling in by DNA polymerase, followed by a series of inverted repeats important for the TE excision by transposase. Cut-and-paste TEs can be duplicated if their transposition takes place during S phase of the cell cycle when a donor site has already been replicated, but a target site has not yet been replicated. Transposition can be classified as either autonomous or non-autonomous in both Class I and Class II TEs. Autonomous TEs can move by themselves while non-autonomous TEs require the presence of another TE to move. This is often because non-autonomous TEs lack transposase (for class II) or reverse transcriptase (for class I).

[0063] "Transposase" refers an enzyme that binds to the end of a transposon and catalyzes the movement of the transposon to another part of the genome by a cut and paste mechanism or a replicative transposition mechanism.

[0064] "T cell" or "T lymphocyte" as used herein is a type of lymphocyte that plays a central role in cell-mediated immunity. They can be distinguished from other lymphocytes, such as B cells and natural killer cells (NK cells), by the presence of a T-cell receptor (TCR) on the cell surface.

[0065] "T helper cells" (T_H cells) assist other white blood cells in immunologic processes, including maturation of B cells into plasma cells and memory B cells, and activation of cytotoxic T cells and macrophages. These cells are also known as CD4+ T

cells because they express the CD4 glycoprotein on their surfaces. Helper T cells become activated when they are presented with peptide antigens by MHC class II molecules, which are expressed on the surface of antigen-presenting cells (APCs). Once activated, they divide rapidly and secrete small proteins called cytokines that regulate or assist in the active immune response. These cells can differentiate into one of several subtypes, including T_H1, T_H2, T_H3, T_H9, T_H17, T_H22 or T_{FH} (T follicular helper cells), which secrete different cytokines to facilitate different types of immune responses. Signaling from the APCs directs T cells into particular subtypes.

[0066] "Cytotoxic T cells" (TC cells, or CTLs) or "cytotoxic T lymphocytes" destroy virus-infected cells and tumor cells, and are also implicated in transplant rejection. These cells are also known as CD8+ T cells since they express the CD8 glycoprotein at their surfaces. These cells recognize their targets by binding to antigen associated with MHC class I molecules, which are present on the surface of all nucleated cells. Through IL-10, adenosine, and other molecules secreted by regulatory T cells, the CD8+ cells can be inactivated to an anergic state, which prevents autoimmune diseases.

[0067] "Memory T cells" are a subset of antigen-specific T cells that persist long-term after an infection has resolved. They quickly expand to large numbers of effector T cells upon re-exposure to their cognate antigen, thus providing the immune system with memory against past infections. Memory T cells comprise three subtypes: central memory T cells (T_{CM} cells) and two types of effector memory T cells (T_{EM} cells and T_{EMRA} cells). Memory cells can be either CD4+ or CD8+. Memory T cells typically express the cell surface proteins CD45RO, CD45RA and/or CCR7.

[0068] "Regulatory T cells" (Treg cells), formerly known as suppressor T cells, play a role in the maintenance of immunological tolerance. Their major role is to shut down T cell-mediated immunity toward the end of an immune reaction and to suppress autoreactive T cells that escaped the process of negative selection in the thymus.

[0069] "Natural killer T cells" (NKT cells – not to be confused with natural killer cells of the innate immune system) bridge the adaptive immune system with the innate immune system. Unlike conventional T cells that recognize peptide antigens presented by major histocompatibility complex (MHC) molecules, NKT cells recognize glycolipid antigen presented by a molecule called CD1d. Once activated, these cells can perform functions ascribed to both T helper (T_H) and cytotoxic T (TC) cells (*i.e.*, cytokine production and

release of cytolytic/cell killing molecules). They are also able to recognize and eliminate some tumor cells and cells infected with herpes viruses.

[0070] "Adoptive T cell transfer" refers to the isolation and *ex vivo* expansion of tumor specific T cells to achieve greater number of T cells than what could be obtained by vaccination alone or the patient's natural tumor response. The tumor specific T cells are then infused into patients with cancer in an attempt to give their immune system the ability to overwhelm remaining tumor via T cells which can attack and kill cancer. There are many forms of adoptive T cell therapy being used for cancer treatment; culturing tumor infiltrating lymphocytes or TIL, isolating and expanding one particular T cell or clone, and even using T cells that have been engineered to potently recognize and attack tumors.

[0071] "Antibody" as used herein refers to monoclonal or polyclonal antibodies. The term "monoclonal antibodies," as used herein, refers to antibodies that are produced by a single clone of B-cells and bind to the same epitope. In contrast, "polyclonal antibodies" refer to a population of antibodies that are produced by different B-cells and bind to different epitopes of the same antigen. A whole antibody typically consists of four polypeptides: two identical copies of a heavy (H) chain polypeptide and two identical copies of a light (L) chain polypeptide. Each of the heavy chains contains one N-terminal variable (VH) region and three C-terminal constant (CH1, CH2 and CH3) regions, and each light chain contains one N-terminal variable (VL) region and one C-terminal constant (CL) region. The variable regions of each pair of light and heavy chains form the antigen binding site of an antibody. The VH and VL regions have a similar general structure, with each region comprising four framework regions, whose sequences are relatively conserved. The framework regions are connected by three complementarity determining regions (CDRs). The three CDRs, known as CDR1, CDR2, and CDR3, form the "hypervariable region" of an antibody, which is responsible for antigen binding.

[0072] "Antibody like molecules" can be for example proteins that are members of the Ig-superfamily which are able to selectively bind a partner. MHC molecules and T cell receptors are such molecules. In one embodiment, the antibody-like molecule is an TCR. In one embodiment, the TCR has been modified to increase its MHC binding affinity.

[0073] The terms "fragment of an antibody," "antibody fragment," "functional fragment of an antibody," "antigen-binding portion" or its grammatical equivalents are used

interchangeably herein to mean one or more fragments or portions of an antibody that retain the ability to specifically bind to an antigen (see, generally, Holliger et al., *Nat. Biotech., 23(9)*:1126-1129 (2005)). The antibody fragment desirably comprises, for example, one or more CDRs, the variable region (or portions thereof), the constant region (or portions thereof), or combinations thereof. Non-limiting examples of antibody fragments include (i) a Fab fragment, which is a monovalent fragment consisting of the VL, VH, CL, and CH1 domains; (ii) a F(ab')2 fragment, which is a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the stalk region; (iii) a Fv fragment consisting of the VL and VH domains of a single arm of an antibody; (iv) a single chain Fv (scFv), which is a monovalent molecule consisting of the two domains of the Fv fragment (*i.e.*, VL and VH) joined by a synthetic linker which enables the two domains to be synthesized as a single polypeptide chain (see, *e.g.*, Bird et al., *Science, 242*: 423-426 (1988); Huston et al., *Proc. Natl. Acad. Sci. USA, 85*: 5879-5883 (1988); and Osbourn et al., *Nat. Biotechnol., 16*: 778 (1998)) and (v) a diabody, which is a dimer of polypeptide chains, wherein each polypeptide chain comprises a VH connected to a VL by a peptide linker that is too short to allow pairing between the VH and VL on the same polypeptide chain, thereby driving the pairing between the complementary domains on different VH-VL polypeptide chains to generate a dimeric molecule having two functional antigen binding sites. Antibody fragments are known in the art and are described in more detail in, *e.g.*, U.S. Patent 8,603,950.

[0074] "Antigen recognition moiety" or "antigen recognition domain" refers to a molecule or portion of a molecule that specifically binds to an antigen. In one embodiment, the antigen recognition moiety is an antibody, antibody like molecule or fragment thereof and the antigen is a tumor antigen.

[0075] The term "conservative amino acid substitution" or "conservative mutation" refers to the replacement of one amino acid by another amino acid with a common property. A functional way to define common properties between individual amino acids is to analyze the normalized frequencies of amino acid changes between corresponding proteins of homologous organisms (Schulz, G. E. and Schirmer, R. H., Principles of Protein Structure, Springer-Verlag, New York (1979)). According to such analyses, groups of amino acids can be defined where amino acids within a group exchange preferentially with each other, and therefore resemble each other most in their impact on the overall

protein structure (Schulz, G. E. and Schirmer, R. H., supra). Examples of conservative mutations include amino acid substitutions of amino acids within the sub-groups above, for example, lysine for arginine and vice versa such that a positive charge can be maintained; glutamic acid for aspartic acid and vice versa such that a negative charge can be maintained; serine for threonine such that a free –OH can be maintained; and glutamine for asparagine such that a free $–NH_2$ can be maintained. Alternatively, or additionally, the functional variants can comprise the amino acid sequence of the reference protein with at least one non-conservative amino acid substitution.

[0076] The term "non-conservative mutations" involve amino acid substitutions between different groups, for example, lysine for tryptophan, or phenylalanine for serine, etc. In this case, it is preferable for the non-conservative amino acid substitution to not interfere with, or inhibit the biological activity of, the functional variant. The non-conservative amino acid substitution can enhance the biological activity of the functional variant, such that the biological activity of the functional variant is increased as compared to the homologous parent protein.

[0077] The term "proliferative disease" as referred to herein refers to a unifying concept in which excessive proliferation of cells and/or turnover of cellular matrix contributes significantly to the pathogenesis of the disease, including cancer.

[0078] "Patient" or "subject" as used herein refers to a mammalian subject diagnosed with or suspected of having or developing a proliferative disorder such as cancer. In some embodiments, the term "patient" refers to a mammalian subject with a higher than average likelihood of developing a proliferative disorder such as cancer. Exemplary patients can be humans, apes, dogs, pigs, cattle, cats, horses, goats, sheep, rodents and other mammalians that can benefit from the therapies disclosed herein. Exemplary human patients can be male and/or female. "Patient in need thereof" or "subject in need thereof" is referred to herein as a patient diagnosed with or suspected of having a disease or disorder, for instance, but not restricted to human papilloma virus (HPV) infection.

[0079] "Administering" is referred to herein as providing one or more compositions described herein to a patient or a subject. By way of example and not limitation, composition administration, *e.g.*, injection, can be performed by intravenous (i.v.) injection, sub-cutaneous (s.c.) injection, intradermal (i.d.) injection, intraperitoneal (i.p.) injection, or intramuscular (i.m.) injection. One or more such routes can be employed.

Parenteral administration can be, for example, by bolus injection or by gradual perfusion over time. Alternatively, or concurrently, administration can be by the oral route. Additionally, administration can also be by surgical deposition, or positioning of a medical device. A pharmaceutical composition can comprise a composition of the invention as described herein, in combination with one or more pharmaceutically or physiologically acceptable carriers, diluents or excipients. Such compositions can comprise buffers such as neutral buffered saline, phosphate buffered saline and the like; carbohydrates such as glucose, mannose, sucrose or dextrans, mannitol; proteins; polypeptides or amino acids such as glycine; antioxidants; chelating agents such as EDTA or glutathione; adjuvants (*e.g.*, aluminum hydroxide); and preservatives.

[0080] As used herein, the term "treatment", "treating", or its grammatical equivalents refers to obtaining a desired pharmacologic and/or physiologic effect. In embodiments, the effect is therapeutic, *i.e.*, the effect partially or completely cures a disease and/or adverse symptom or pathological manifestation attributable to the disease. To this end, the inventive method comprises administering a therapeutically effective amount of a composition of the invention expressing the inventive nucleic acid sequence, or a vector comprising the inventive nucleic acid sequences.

[0081] The term "therapeutically effective amount", therapeutic amount", "immunologically effective amount", "anti-tumor effective amount", "tumor-inhibiting effective amount" or its grammatical equivalents refers to an amount effective, at dosages and for periods of time necessary, to achieve a desired therapeutic result. The therapeutically effective amount can vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of a composition described herein to elicit a desired response in one or more subjects.

[0082] Alternatively, the pharmacologic and/or physiologic effect of administration of one or more compositions described herein to a patient or a subject of can be "prophylactic," *i.e.*, the effect completely or partially prevents a disease or symptom thereof. A "prophylactically effective amount" refers to an amount effective, at dosages and for periods of time necessary, to achieve a desired prophylactic result (*e.g.*, prevention of disease or prevention of manifestation of a target pathology).

[0083] The term "in sequential order" or grammatical equivalents, in reference to polypeptide and amino acid sequences, refers to a N-terminal to C-terminal order of

sequences. The term “in sequential order” or grammatical equivalents, in reference to polynucleotide and nucleic acid sequences, refers to a 5’ (“five prime”) to 3’ (“three prime”) order of seqeunces.

HPV Molecular Vaccine

[0084] *Human papillomavirus* (HPV) is a group of more than 200 related viruses. Each HPV virus in this large group is given a number which is called its HPV type (or serotype). HPV is a small, non-enveloped deoxyribonucleic acid (DNA) virus that infects skin or mucosal cells. The circular, double-stranded viral genome is approximately 8-kb in length. The genome encodes for seven early proteins (E1 to E7) responsible for virus replication and two late proteins (L1 and L2), which are the viral structural proteins. Each gene has specific functions. At least 13 of more than 200 known HPV types can cause cancer of the cervix and are associated with other anogenital cancers and cancers of the head and neck. The two most common "high-risk" serotypes (HPV16 and HPV-18) cause approximately 70% of all cervical cancers. In HPV16 and HPV-18, two primary oncoproteins, E6 and E7, are constitutively expressed by HPV-associated tumors and are critical for the induction and maintenance of cellular transformation in HPV-infected cells. Recent evidence also suggests the E5 protein also impacts viral transformation. HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82 are considered carcinogenic. Two "low-risk" HPV 6 and 11 are known to cause genital warts, a common benign condition of the external genitalia that causes significant morbidity. HPV is highly transmissible, with peak incidence soon after the onset of sexual activity, and most persons acquire infection at some time in their lives.

[0085] Genital warts are extremely common, and may also be known by other names such as anogenital warts or condylomata acuminate with between 500,000 to one million new cases diagnosed each year in the United States alone. HPV6 and HPV11 are responsible for about 90% of genital warts. The current treatment options are largely centered upon removal of the warts rather than elimination of the underlying viral infection. A wide range of therapies are presently in use, which are highly variable and can differ dramatically. However, no treatment has been 100% effective in eliminating warts and preventing them from coming back in all patients.

[0086] The current standard of treatment for RRP is surgical excision with adjuvant therapies as needed. However, the distinguishing aspect of RRP is the tendency for the

growth to recur requiring repeated surgeries. The aggressiveness varies among patients - Some individuals may require surgery every few weeks while others may only require surgery twice a year or only a few times during their life.

[0087] Unfortunately, previously developed therapies have not been able to prevent multiple recurrence for RRP. There is a need to develop better treatment for RRP patients. The present specification describes an innovative immunotherapy approach, a novel therapeutic HPV vaccine aiming to induce anti-HPV6/11 immunity to treat and alter the course of RRP and other HPV6/11 associated disease such as genital warts.

[0088] Prophylactic vaccines composed of virus-like particles, only induce immunity to the capsid structure and not to the non-structural proteins responsible for cell transformation. Initial HPV studies in animal models showed that inoculation with species-specific papillomaviruses induced an immune response that conferred protection against homologous virus challenge. However, native papillomaviruses are not good substrates for vaccine development as they cannot be grown easily in tissue culture. Subsequent studies were initiated on the production of viral particles from expression of the structural proteins in heterologous expression systems, such as yeast or baculovirus vectors. Results showed that expression of L1 alone led to the production of virus-like particles (VLPs) which morphologically resemble the authentic HPV virions but contain no viral DNA. These VLPs are produced by self-assembly of the L1 protein when expressed in a heterologous cell substrate. In animal studies, VLPs were shown to protect against high dose experimental infection by homologous virus. HPV VLPs are highly immunogenic in mice or rabbits, and the resulting antibodies have been shown to be neutralizing and type restricted when tested in a pseudovirion neutralization assay. Immunization with denatured particles does not result in the production of neutralizing antibodies, or protect from experimental virus challenge, indicating that neutralizing epitopes are conformation dependent.

[0089] Provided herein are compositions, kits, and systems comprising methods of making HPV recombinant vaccines. The HPV recombinant vaccines (*e.g.*, HPV designs 1-5) in the present disclosure are engineered through protein engineering.

Delivery System

Gorilla Adenovirus Shuttle Vector

[0090] Certain aspects of the present disclosure are directed to a vector comprising a polynucleotide encoding a polypeptide comprising one or more immune response-inducing HPV polypeptides described herein. In certain embodiments, the vector is a viral vector. In particular embodiments, the vector is an adenoviral vector. Adenoviruses are generally associated with benign pathologies in humans, and the genomes of adenoviruses isolated from a variety of species, including humans, have been extensively studied. Adenovirus is a medium-sized (90-100 nm), non-enveloped icosahedral virus containing approximately 36 kb of double-stranded DNA. The adenovirus capsid mediates the key interactions of the early stages of the infection of a cell by the virus, and is required for packaging adenovirus genomes at the end of the adenovirus life cycle. The capsid comprises 252 capsomeres, which includes 240 hexons, 12 penton base proteins, and 12 fibers (Ginsberg et al., *Virology*, *28*: 782-83 (1966)). The hexon comprises three identical proteins, namely polypeptide II (Roberts et al., *Science*, *232*: 1148-51 (1986)). The penton base comprises five identical proteins and the fiber comprises three identical proteins. Proteins IIIa, VI, and IX are present in the adenoviral coat and are believed to stabilize the viral capsid (Stewart et al., *Cell*, *67*: 145-54 (1991), and Stewart et al., *EMBO J.*, *12*(7): 2589-99 (1993)). The expression of the capsid proteins, with the exception of pIX, is dependent on the adenovirus polymerase protein. Therefore, major components of an adenovirus particle are expressed from the genome only when the polymerase protein gene is present and expressed.

[0091] Several features of adenoviruses make them ideal vehicles for transferring genetic material to cells for therapeutic applications (*i.e.*, "gene therapy"), or for use as antigen delivery systems for vaccine applications. For example, adenoviruses can be produced in high titers (*e.g.*, about 10^{13} particle units (pu)), and can transfer genetic material to non-replicating and replicating cells. The adenoviral genome can be manipulated to carry a large amount of exogenous DNA (up to about 8 kb), and the adenoviral capsid can potentiate the transfer of even longer sequences (Curiel et al., *Hum. Gene Ther.*, *3*: 147-154 (1992)). Additionally, adenoviruses generally do not integrate into the host cell chromosome, but rather are maintained as a linear episome, thereby minimizing the likelihood that a recombinant adenovirus will interfere with normal cell function.

[0092] In some embodiments, the adenovirus described herein is isolated from a gorilla. There are four widely recognized gorilla subspecies within the two species of Eastern Gorilla (*Gorilla beringei*) and Western Gorilla (*Gorilla gorilla*). The Western Gorilla species includes the subspecies Western Lowland Gorilla (*Gorilla gorilla gorilla*) and Cross River Gorilla (*Gorilla gorilla diehli*). The Eastern Gorilla species includes the subspecies Mountain Gorilla (*Gorilla beringei beringei*) and Eastern Lowland Gorilla (*Gorilla beringei graueri*) (see, *e.g.*, Wilson and Reeder, eds., *Mammalian Species of the World*, 3rd ed., Johns Hopkins University Press, Baltimore, Maryland (2005)). In some embodiments, the adenovirus of the present disclosure is isolated from Mountain Gorilla (*Gorilla beringei beringei)*.

[0093] Various Gorilla adenoviruses or adenoviral vectors are described in International Patent Application Publications WO 2013/052832; WO 2013/052811; and WO 2013 052799, each of which is herein incorporated by reference in its entirety.

[0094] The adenovirus can be modified in the same manner as previously known adenoviruses to be used as an adenoviral vector, *e.g.*, a gene delivery vehicle. The adenovirus and adenoviral vector can be replication-competent, conditionally replication-competent, or replication-deficient.

[0095] A replication-competent adenovirus or adenoviral vector can replicate in typical host cells, *i.e.*, cells typically capable of being infected by an adenovirus. A replication-competent adenovirus or adenoviral vector can have one or more mutations as compared to the wild-type adenovirus (*e.g.*, one or more deletions, insertions, and/or substitutions) in the adenoviral genome that do not inhibit viral replication in host cells. For example, the adenovirus or adenoviral vector can have a partial or entire deletion of the adenoviral early region known as the E3 region, which is not essential for propagation of the adenovirus or adenoviral genome.

[0096] A conditionally-replicating adenovirus or adenoviral vector is an adenovirus or adenoviral vector that has been engineered to replicate under pre-determined conditions. For example, replication-essential gene functions, *e.g.*, gene functions encoded by the adenoviral early regions, can be operably linked to an inducible, repressible, or tissue-specific transcription control sequence, *e.g.*, promoter. In such an embodiment, replication requires the presence or absence of specific factors that interact with the

transcription control sequence. Conditionally-replicating adenoviral vectors are further described in U.S. Patent 5,998,205.

[0097] A replication-deficient adenovirus or adenoviral vector is an adenovirus or adenoviral vector that requires complementation of one or more gene functions or regions of the adenoviral genome that are required for replication, as a result of, for example, a deficiency in one or more replication-essential gene function or regions, such that the adenovirus or adenoviral vector does not replicate in typical host cells, especially those in a human to be infected by the adenovirus or adenoviral vector.

[0098] A deficiency in a gene function or genomic region, as used herein, is defined as a disruption (*e.g.*, deletion) of sufficient genetic material of the adenoviral genome to obliterate or impair the function of the gene (*e.g.*, such that the function of the gene product is reduced by at least about 2-fold, 5-fold, 10-fold, 20-fold, 30-fold, or 50-fold) whose nucleic acid sequence was disrupted (*e.g.*, deleted) in whole or in part. Deletion of an entire gene region often is not required for disruption of a replication-essential gene function. However, for the purpose of providing sufficient space in the adenoviral genome for one or more transgenes, removal of a majority of one or more gene regions can be desirable. While deletion of genetic material is preferred, mutation of genetic material by addition or substitution also is appropriate for disrupting gene function. Replication-essential gene functions are those gene functions that are required for adenovirus replication (*e.g.*, propagation) and are encoded by, for example, the adenoviral early regions (*e.g.*, the E1, E2, and E4 regions), late regions (*e.g.*, the L1, L2, L3, L4, and L5 regions), genes involved in viral packaging (*e.g.*, the IVa2 gene), and virus-associated RNAs (*e.g.*, VA-RNA-1 and/or VA-RNA-2).

[0099] Whether the adenovirus or adenoviral vector is replication-competent or replication-deficient, the adenovirus or adenoviral vector retains at least a portion of the adenoviral genome. The adenovirus or adenoviral vector can comprise any portion of the adenoviral genome, including protein coding and non-protein coding regions. Desirably, the adenovirus or adenoviral vector comprises at least one nucleic acid sequence that encodes an adenovirus protein. The adenovirus or adenoviral vector can comprise a nucleic acid sequence that encodes any suitable adenovirus protein, such as, for example, a protein encoded by any one of the early region genes (*i.e.*, E1A, E1B, E2A, E2B, E3,

and/or E4 regions), or a protein encoded by any one of the late region genes, which encode the virus structural proteins (*i.e.*, L1, L2, L3, L4, and L5 regions).

[0100] The adenovirus or adenoviral vector desirably comprises one or more nucleic acid sequences that encode the pIX protein, the DNA polymerase protein, the penton protein, the hexon protein, and/or the fiber protein. The adenovirus or adenoviral vector can comprise a full-length nucleic acid sequence that encodes a full-length amino acid sequence of an adenovirus protein. Alternatively, the adenovirus or adenoviral vector can comprise a portion of a full-length nucleic acid sequence that encodes a portion of a full-length amino acid sequence of an adenovirus protein.

[0101] A "portion" of a nucleic acid sequence comprises at least ten nucleotides (*e.g.*, about 10 to about 5000 nucleotides). Preferably, a "portion" of a nucleic acid sequence comprises 10 or more (*e.g.*, 15 or more, 20 or more, 25 or more, 30 or more, 35 or more, 40 or more, 45 or more, 50 or more, or 100 or more) nucleotides, but less than 5,000 (*e.g.*, 4900 or less, 4000 or less, 3000 or less, 2000 or less, 1000 or less, 800 or less, 500 or less, 300 or less, or 100 or less) nucleotides. Preferably, a portion of a nucleic acid sequence is about 10 to about 3500 nucleotides (*e.g.*, about 10, 20, 30, 50, 100, 300, 500, 700, 1000, 1500, 2000, 2500, or 3000 nucleotides), about 10 to about 1000 nucleotides (*e.g.*, about 25, 55, 125, 325, 525, 725, or 925 nucleotides), or about 10 to about 500 nucleotides (*e.g.*, about 15, 30, 40, 50, 60, 70, 80, 90, 150, 175, 250, 275, 350, 375, 450, 475, 480, 490, 495, or 499 nucleotides), or a range defined by any two of the foregoing values. More preferably, a "portion" of a nucleic acid sequence comprises no more than about 3200 nucleotides (*e.g.*, about 10 to about 3200 nucleotides, about 10 to about 3000 nucleotides, or about 30 to about 500 nucleotides, or a range defined by any two of the foregoing values).

[0102] A "portion" of an amino acid sequence comprises at least three amino acids (*e.g.*, about 3 to about 1,200 amino acids). Preferably, a "portion" of an amino acid sequence comprises 3 or more (*e.g.*, 5 or more, 10 or more, 15 or more, 20 or more, 25 or more, 30 or more, 40 or more, or 50 or more) amino acids, but less than 1,200 (*e.g.*, 1,000 or less, 800 or less, 700 or less, 600 or less, 500 or less, 400 or less, 300 or less, 200 or less, or 100 or less) amino acids. Preferably, a portion of an amino acid sequence is about 3 to about 500 amino acids (*e.g.*, about 10, 100, 200, 300, 400, or 500 amino acids), about 3 to about 300 amino acids (*e.g.*, about 20, 50, 75, 95, 150, 175, or 200 amino acids), or about

3 to about 100 amino acids (*e.g.*, about 15, 25, 35, 40, 45, 60, 65, 70, 80, 85, 90, 95, or 99 amino acids), or a range defined by any two of the foregoing values. More preferably, a "portion" of an amino acid sequence comprises no more than about 500 amino acids (*e.g.*, about 3 to about 400 amino acids, about 10 to about 250 amino acids, or about 50 to about 100 amino acids, or a range defined by any two of the foregoing values).

[0103] The adenovirus pIX protein is present in the adenovirus capsid, has been shown to strengthen hexon nonamer interactions, and is essential for the packaging of full-length genomes (see, *e.g.*, Boulanger et al., *J. Gen. Virol.*, *44*: 783-800 (1979); Horwitz M.S., "Adenoviridae and their replication" in *Virology*, 2nd ed., B.N. Fields et al. (eds.), Raven Press, Ltd., New York, pp. 1679-1721 (1990), Ghosh-Choudhury et al., *EMBO J.*, *6*: 1733-1739 (1987), and van Oostrum et al, *J. Virol.*, *56*: 439-448 (1985)). In addition to its contribution to adenovirus structure, pIX has also been shown to exhibit transcriptional properties, such as stimulation of adenovirus major late promoter (MLP) activity (see, *e.g.*, Lutz et al., *J. Virol.*, *71*(7): 5102-5109 (1997)).

[0104] The adenovirus DNA polymerase protein is essential for viral DNA replication both *in vitro* and *in vivo*. The polymerase co-purifies in a complex with the precursor (pTP) of the terminal protein (TP), which is covalently attached to the 5' ends of adenovirus DNA (Field et al., *J. Biol. Chem.*, *259*: 9487-9495 (1984)). Both the adenovirus DNA polymerase and pTP are encoded by the E2 region. The polymerase protein is required for the expression of all the structural proteins except for pIX. Without the gene sequence for polymerase protein, polymerase protein is not produced. As a result, the viral genome is not replicated, the Major Late Promoter is not activated, and the capsid proteins are not expressed.

[0105] The adenovirus hexon protein is the largest and most abundant protein in the adenovirus capsid. The hexon protein is essential for virus capsid assembly, determination of the icosahedral symmetry of the capsid (which in turn defines the limits on capsid volume and DNA packaging size), and integrity of the capsid. In addition, hexon is a primary target for modification in order to reduce neutralization of adenoviral vectors (see, *e.g.*, Gall et al., *J. Virol.*, 72: 10260-264 (1998), and Rux et al., *J. Virol.*, 77(17): 9553-9566 (2003)). The major structural features of the hexon protein are shared by adenoviruses across serotypes, but the hexon protein differs in size and immunological properties between serotypes (Jornvall et al., *J. Biol. Chem.*, *256*(12): 6181-6186 (1981)).

A comparison of 15 adenovirus hexon proteins revealed that the predominant antigenic and serotype-specific regions of the hexon appear to be in loops 1 and 2 (*i.e.*, LI or *l*1, and LII or *l*2, respectively), within which are seven discrete hypervariable regions (HVR1 to HVR7) varying in length and sequence between adenoviral serotypes (Crawford-Miksza et al., *J. Virol.*, *70*(3): 1836-1844 (1996)).

[0106] The adenovirus fiber protein is a homotrimer of the adenoviral polypeptide IV that has three domains: the tail, shaft, and knob. (Devaux et al., *J. Molec. Biol.*, *215*: 567-88 (1990), Yeh et al., *Virus Res.*, *33*: 179-98 (1991)). The fiber protein mediates primary viral binding to receptors on the cell surface via the knob and the shaft domains (Henry et al., *J. Virol.*, *68*(8): 5239-46 (1994)). The amino acid sequences for trimerization are located in the knob, which appears necessary for the amino terminus of the fiber (the tail) to properly associate with the penton base (Novelli et al., *Virology*, *185*: 365-76 (1991)). In addition to recognizing cell receptors and binding the penton base, the fiber contributes to serotype identity. Fiber proteins from different adenoviral serotypes differ considerably (see, *e.g.*, Green et al., *EMBO J.*, *2*: 1357-65 (1983), Chroboczek et al., *Virology*, *186*: 280-85 (1992), and Signas et al., *J. Virol.*, *53*: 672-78 (1985)). Thus, the fiber protein has multiple functions key to the life cycle of adenovirus.

[0107] The adenovirus penton base protein is located at the vertices of the icosahedral capsid and comprises five identical monomers. The penton base protein provides a structure for bridging the hexon proteins on multiple facets of the icosahedral capsid, and provides the essential interface for the fiber protein to be incorporated in the capsid. Each monomer of the penton base contains an RGD tripeptide motif (Neumann et al., *Gene*, *69*: 153-157 (1988)). The RGD tripeptide mediates binding to αv integrins and adenoviruses that have point mutations in the RGD sequence of the penton base are restricted in their ability to infect cells (Bai et al., *J. Virol.*, *67*: 5198-5205 (1993)). Thus, the penton base protein is essential for the architecture of the capsid and for maximum efficiency of virus-cell interaction.

[0108] Nucleic acid or amino acid sequence "identity," as described herein, can be determined by comparing a nucleic acid or amino acid sequence of interest to a reference nucleic acid or amino acid sequence. The numbers of nucleotides or amino acid residues that have been changed and/or modified (such as, *e.g.*, by point mutations, insertions, or deletions) in the reference sequence so as to result in the sequence of interest are counted.

The total number of such changes is subtracted from the total length of the sequence of interest, and the difference is divided by the length of the sequence of interest and expressed as a percentage. A number of mathematical algorithms for obtaining the optimal alignment and calculating identity between two or more sequences are known and incorporated into a number of available software programs. Examples of such programs include CLUSTAL-W, T-Coffee, and ALIGN (for alignment of nucleic acid and amino acid sequences), BLAST programs (*e.g.*, BLAST 2.1, BL2SEQ, and later versions thereof) and FASTA programs (*e.g.*, FASTA3x, FASTM, and SSEARCH) (for sequence alignment and sequence similarity searches). Sequence alignment algorithms also are disclosed in, for example, Altschul et al., *J. Molecular Biol.*, *215*(3): 403-410 (1990), Beigert et al., *Proc. Natl. Acad. Sci. USA*, *106*(10): 3770-3775 (2009), Durbin et al., eds., *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, UK (2009), Soding, *Bioinformatics*, *21*(7): 951-960 (2005), Altschul et al., *Nucleic Acids Res.*, *25*(17): 3389-3402 (1997), and Gusfield, *Algorithms on Strings, Trees and Sequences*, Cambridge University Press, Cambridge UK (1997)).

[0109] The adenovirus or adenoviral vector can comprise one, two, three, four, or all five of the aforementioned sequences alone or in any combination. In this respect, the adenovirus or adenoviral vector can comprise any combination of any two of the aforementioned sequences, any combination of any three of the aforementioned sequences, any combination of any four of the aforementioned sequences, or all five of the aforementioned sequences.

[0110] As discussed herein, the adenovirus or adenoviral vector can be replication-competent, conditionally-replicating, or replication-deficient. Preferably, the adenovirus or adenoviral vector is replication-deficient, such that the replication-deficient adenovirus or adenoviral vector requires complementation of at least one replication-essential gene function of one or more regions of the adenoviral genome for propagation (*e.g.*, to form adenoviral vector particles).

[0111] The replication-deficient adenovirus or adenoviral vector can be modified in any suitable manner to cause the deficiencies in the one or more replication-essential gene functions in one or more regions of the adenoviral genome for propagation. The complementation of the deficiencies in the one or more replication-essential gene

functions of one or more regions of the adenoviral genome refers to the use of exogenous means to provide the deficient replication-essential gene functions. Such complementation can be effected in any suitable manner, for example, by using complementing cells and/or exogenous DNA (*e.g.*, helper adenovirus) encoding the disrupted replication-essential gene functions.

[0112] The adenovirus or adenoviral vector can be deficient in one or more replication-essential gene functions of only the early regions (*i.e.*, E1-E4 regions) of the adenoviral genome, only the late regions (*i.e.*, L1-L5 regions) of the adenoviral genome, both the early and late regions of the adenoviral genome, or all adenoviral genes (*i.e.*, a high capacity adenovector (HC-Ad). See Morsy et al., *Proc. Natl. Acad. Sci. USA*, *95*: 965-976 (1998); Chen et al., *Proc. Natl. Acad. Sci. USA*, *94*: 1645-1650 (1997); and Kochanek et al., *Hum. Gene Ther.*, *10*: 2451-2459 (1999). Examples of replication-deficient adenoviral vectors are disclosed in U.S. Patents 5,837,511; 5,851,806; 5,994,106; 6,127,175; 6,482,616; and 7,195,896, and International Patent Application Publications WO 1994/028152, WO 1995/002697, WO 1995/016772, WO 1995/034671, WO 1996/022378, WO 1997/012986, WO 1997/021826, and WO 2003/022311.

[0113] The early regions of the adenoviral genome include the E1, E2, E3, and E4 regions. The E1 region comprises the E1A and E1B subregions, and one or more deficiencies in replication-essential gene functions in the E1 region can include one or more deficiencies in replication-essential gene functions in either or both of the E1A and E1B subregions, thereby requiring complementation of the E1A subregion and/or the E1B subregion of the adenoviral genome for the adenovirus or adenoviral vector to propagate (*e.g.*, to form adenoviral vector particles). The E2 region comprises the E2A and E2B subregions, and one or more deficiencies in replication-essential gene functions in the E2 region can include one or more deficiencies in replication-essential gene functions in either or both of the E2A and E2B subregions, thereby requiring complementation of the E2A subregion and/or the E2B subregion of the adenoviral genome for the adenovirus or adenoviral vector to propagate (*e.g.*, to form adenoviral vector particles).

[0114] The E3 region does not include any replication-essential gene functions, such that a deletion of the E3 region in part or in whole does not require complementation of any gene functions in the E3 region for the adenovirus or adenoviral vector to propagate (*e.g.*, to form adenoviral vector particles). In the context of the present disclosure, the E3

region is defined as the region that initiates with the open reading frame that encodes a protein with high homology to the 12.5K protein from the E3 region of human adenovirus 5 (NCBI reference sequence AP_000218) and ends with the open reading frame that encodes a protein with high homology to the 14.7K protein from the E3 region of human adenovirus 5 (NCBI reference sequence AP_000224.1). The E3 region can be deleted in whole or in part, or retained in whole or in part. The size of the deletion can be tailored so as to retain an adenovirus or adenoviral vector whose genome closely matches the optimum genome packaging size. A larger deletion will accommodate the insertion of larger heterologous nucleic acid sequences in the adenovirus or adenoviral genome. In one embodiment of the present disclosure, the L4 polyadenylation signal sequences, which reside in the E3 region, are retained.

[0115] The E4 region comprises multiple open reading frames (ORFs). An adenovirus or adenoviral vector with a deletion of all of the open reading frames of the E4 region except ORF6, and in some cases ORF3, does not require complementation of any gene functions in the E4 region for the adenovirus or adenoviral vector to propagate (*e.g.*, to form adenoviral vector particles). Conversely, an adenovirus or adenoviral vector with a disruption or deletion of ORF6, and in some cases ORF3, of the E4 region (*e.g.*, with a deficiency in a replication-essential gene function based in ORF6 and/or ORF3 of the E4 region), with or without a disruption or deletion of any of the other open reading frames of the E4 region or the native E4 promoter, polyadenylation sequence, and/or the right-side inverted terminal repeat (ITR), requires complementation of the E4 region (specifically, of ORF6 and/or ORF3 of the E4 region) for the adenovirus or adenoviral vector to propagate (*e.g.*, to form adenoviral vector particles). The late regions of the adenoviral genome include the L1, L2, L3, L4, and L5 regions. The adenovirus or adenoviral vector also can have a mutation in the major late promoter (MLP), as discussed in International Patent Application Publication WO 2000/000628, which can render the adenovirus or adenoviral vector replication-deficient if desired.

[0116] The one or more regions of the adenoviral genome that contain one or more deficiencies in replication-essential gene functions desirably are one or more early regions of the adenoviral genome, *i.e.*, the E1, E2, and/or E4 regions, optionally with the deletion in part or in whole of the E3 region.

[0117] The replication-deficient adenovirus or adenoviral vector also can have one or more mutations as compared to the wild-type adenovirus (*e.g.*, one or more deletions, insertions, and/or substitutions) in the adenoviral genome that do not inhibit viral replication in host cells. Thus, in addition to one or more deficiencies in replication-essential gene functions, the adenovirus or adenoviral vector can be deficient in other respects that are not replication-essential. For example, the adenovirus or adenoviral vector can have a partial or entire deletion of the adenoviral early region known as the E3 region, which is not essential for propagation of the adenovirus or adenoviral genome.

[0118] In one embodiment, the adenovirus or adenoviral vector is replication-deficient and requires, at most, complementation of the E1 region or the E4 region of the adenoviral genome, for propagation (*e.g.*, to form adenoviral vector particles). Thus, the replication-deficient adenovirus or adenoviral vector requires complementation of at least one replication-essential gene function of the E1A subregion and/or the E1B region of the adenoviral genome (denoted an E1-deficient adenoviral vector) or the E4 region of the adenoviral genome (denoted an E4-deficient adenoviral vector) for propagation (*e.g.*, to form adenoviral vector particles). The adenovirus or adenoviral vector can be deficient in at least one replication-essential gene function (desirably all replication-essential gene functions) of the E1 region of the adenoviral genome and at least one gene function of the nonessential E3 region of the adenoviral genome (denoted an E1/E3-deficient adenoviral vector). The adenovirus or adenoviral vector can be deficient in at least one replication-essential gene function (desirably all replication-essential gene functions) of the E4 region of the adenoviral genome and at least one gene function of the nonessential E3 region of the adenoviral genome (denoted an E3/E4-deficient adenoviral vector).

[0119] In one embodiment, the adenovirus or adenoviral vector is replication-deficient and requires, at most, complementation of the E2 region, preferably the E2A subregion, of the adenoviral genome, for propagation (*e.g.*, to form adenoviral vector particles). Thus, the replication-deficient adenovirus or adenoviral vector requires complementation of at least one replication-essential gene function of the E2A subregion of the adenoviral genome (denoted an E2A-deficient adenoviral vector) for propagation (*e.g.*, to form adenoviral vector particles). The adenovirus or adenoviral vector can be deficient in at least one replication-essential gene function (desirably all replication-essential gene functions) of the E2A region of the adenoviral genome and at least one gene function of

the nonessential E3 region of the adenoviral genome (denoted an E2A/E3-deficient adenoviral vector).

[0120] In one embodiment, the adenovirus or adenoviral vector is replication-deficient and requires, at most, complementation of the E1 and E4 regions of the adenoviral genome for propagation (*e.g.*, to form adenoviral vector particles). Thus, the replication-deficient adenovirus or adenoviral vector requires complementation of at least one replication-essential gene function of both the E1 and E4 regions of the adenoviral genome (denoted an E1/E4-deficient adenoviral vector) for propagation (*e.g.*, to form adenoviral vector particles). The adenovirus or adenoviral vector can be deficient in at least one replication-essential gene function (desirably all replication-essential gene functions) of the E1 region of the adenoviral genome, at least one replication-essential gene function of the E4 region of the adenoviral genome, and at least one gene function of the nonessential E3 region of the adenoviral genome (denoted an E1/E3/E4-deficient adenoviral vector). The adenovirus or adenoviral vector preferably requires, at most, complementation of the E1 region of the adenoviral genome for propagation, and does not require complementation of any other deficiency of the adenoviral genome for propagation. More preferably, the adenovirus or adenoviral vector requires, at most, complementation of the E1 and E4 regions of the adenoviral genome for propagation, and does not require complementation of any other deficiency of the adenoviral genome for propagation.

[0121] The adenovirus or adenoviral vector, when deficient in multiple replication-essential gene functions of the adenoviral genome (*e.g.*, an E1/E4-deficient adenoviral vector), can include a spacer sequence to provide viral growth in a complementing cell line similar to that achieved by adenoviruses or adenoviral vectors deficient in a single replication-essential gene function (*e.g.*, an E1-deficient adenoviral vector). The spacer sequence can contain any nucleotide sequence or sequences which are of a desired length, such as sequences at least about 15 base pairs (*e.g.*, between about 15 nucleotides and about 12,000 nucleotides), preferably about 100 nucleotides to about 10,000 nucleotides, more preferably about 500 nucleotides to about 8,000 nucleotides, even more preferably about 1,500 nucleotides to about 6,000 nucleotides, and most preferably about 2,000 to about 3,000 nucleotides in length, or a range defined by any two of the foregoing values. The spacer sequence can be coding or non-coding and native or non-native with respect to

the adenoviral genome, but does not restore the replication-essential function to the deficient region. The spacer also can contain an expression cassette. More preferably, the spacer comprises a polyadenylation sequence and/or a gene that is non-native with respect to the adenovirus or adenoviral vector. The use of a spacer in an adenoviral vector is further described in, for example, U.S. Patent 5,851,806 and International Patent Application Publication WO 1997/021826.

[0122] By removing all or part of the adenoviral genome, for example, the E1, E3, and E4 regions of the adenoviral genome, the resulting adenovirus or adenoviral vector is able to accept inserts of exogenous nucleic acid sequences while retaining the ability to be packaged into adenoviral capsids. An exogenous nucleic acid sequence can be inserted at any position in the adenoviral genome so long as insertion in the position allows for the formation of adenovirus or the adenoviral vector particle. The exogenous nucleic acid sequence preferably is positioned in the E1 region, the E3 region, or the E4 region of the adenoviral genome.

[0123] The replication-deficient adenovirus or adenoviral vector of the present disclosure can be produced in complementing cell lines that provide gene functions not present in the replication-deficient adenovirus or adenoviral vector, but required for viral propagation, at appropriate levels in order to generate high titers of viral vector stock. Such complementing cell lines are known and include, but are not limited to, 293 cells (described in, *e.g.*, Graham et al., *J. Gen. Virol.*, *36*: 59-72 (1977)), PER.C6 cells (described in, *e.g.*, International Patent Application Publication WO 1997/000326, and U.S. Patents 5,994,128 and 6,033,908), and 293-ORF6 cells (described in, *e.g.*, International Patent Application Publication WO 95/34671 and Brough et al., *J. Virol.*, *71*: 9206-9213 (1997)). Other suitable complementing cell lines to produce the replication-deficient adenovirus or adenoviral vector of the present disclosure include complementing cells that have been generated to propagate adenoviral vectors encoding transgenes whose expression inhibits viral growth in host cells (see, *e.g.*, U.S. Patent Application Publication No. 2008/0233650). Additional suitable complementing cells are described in, for example, U.S. Patents 6,677,156 and 6,682,929, and International Patent Application Publication WO 2003/020879. In some instances, the cellular genome need not comprise nucleic acid sequences, the gene products of which complement for all of the deficiencies of a replication-deficient adenoviral vector. One or more replication-

essential gene functions lacking in a replication-deficient adenoviral vector can be supplied by a helper virus, *e.g.*, an adenoviral vector that supplies in *trans* one or more essential gene functions required for replication of the replication-deficient adenovirus or adenoviral vector. Alternatively, the inventive adenovirus or adenoviral vector can comprise a non-native replication-essential gene that complements for the one or more replication-essential gene functions lacking in the inventive replication-deficient adenovirus or adenoviral vector. For example, an E1/E4-deficient adenoviral vector can be engineered to contain a nucleic acid sequence encoding E4 ORF 6 that is obtained or derived from a different adenovirus (*e.g.*, an adenovirus of a different serotype than the inventive adenovirus or adenoviral vector, or an adenovirus of a different species than the inventive adenovirus or adenoviral vector).

[0124] The adenovirus or adenoviral vector can further comprise a transgene. The term "transgene" is defined herein as a non-native nucleic acid sequence that is operably linked to appropriate regulatory elements (*e.g.*, a promoter), such that the non-native nucleic acid sequence can be expressed to produce a protein (*e.g.*, peptide or polypeptide). The regulatory elements (*e.g.*, promoter) can be native or non-native to the adenovirus or adenoviral vector.

[0125] A "non-native" nucleic acid sequence is any nucleic acid sequence (*e.g.*, DNA, RNA, or cDNA sequence) that is not a naturally occurring nucleic acid sequence of an adenovirus in a naturally occurring position. Thus, the non-native nucleic acid sequence can be naturally found in an adenovirus, but located at a non-native position within the adenoviral genome and/or operably linked to a non-native promoter. The terms "non-native nucleic acid sequence," "heterologous nucleic acid sequence," and "exogenous nucleic acid sequence" are synonymous and can be used interchangeably in the context of the present disclosure. The non-native nucleic acid sequence preferably is DNA and preferably encodes a protein (*i.e.*, one or more nucleic acid sequences encoding one or more proteins).

[0126] The non-native nucleic acid sequence can encode a therapeutic protein that can be used to prophylactically or therapeutically treat a mammal for a disease. Examples of suitable therapeutic proteins include cytokines, toxins, tumor suppressor proteins, growth factors, hormones, receptors, mitogens, immunoglobulins, neuropeptides, neurotransmitters, and enzymes. Alternatively, the non-native nucleic acid sequence can

encode an antigen of a pathogen (*e.g.*, a bacterium or a virus), and the adenovirus or adenoviral vector can be used as a vaccine.

Viral Based Delivery System

[0127] The present disclosure also provides delivery systems, such as viral-based systems, in which a nucleic acid described herein is inserted. Representative viral expression vectors include, but are not limited to, adeno-associated viral vectors, adenovirus-based vectors, lentivirus-based vectors, retroviral vectors, and herpes virus-based vectors. In an embodiment, the viral vector is a lentivirus vector. Vectors derived from retroviruses such as the lentivirus are suitable tools to achieve long-term gene transfer since they allow long-term, stable integration of a transgene and its propagation in daughter cells. Lentiviral vectors have the added advantage over vectors derived from onco-retroviruses such as murine leukemia viruses in that they can transduce non-proliferating cells, such as hepatocytes. They also have the added advantage of low immunogenicity. In an additional embodiment, the viral vector is an adeno-associated viral vector. In a further embodiment, the viral vector is a retroviral vector. In general, and in embodiments, a suitable vector contains an origin of replication functional in at least one organism, a promoter sequence, convenient restriction endonuclease sites, and one or more selectable markers.

[0128] Additional suitable vectors include integrating expression vectors, which can randomly integrate into the host cell's DNA, or can include a recombination site to enable the specific recombination between the expression vector and the host cell's chromosome. Such integrating expression vectors can utilize the endogenous expression control sequences of the host cell's chromosomes to effect expression of the desired protein. Examples of vectors that integrate in a site specific manner include, for example, components of the flp-in system from Invitrogen (Carlsbad, Calif.) (*e.g.*, pcDNATM5/FRT), or the cre-lox system, such as can be found in the pExchange-6 Core Vectors from Stratagene (La Jolla, Calif.). Examples of vectors that randomly integrate into host cell chromosomes include, for example, pcDNA3.1 (when introduced in the absence of T-antigen) from Invitrogen (Carlsbad, Calif.), and pCI or pFN10A (ACT) FLEXITM from Promega (Madison, Wis.). Additional promoter elements, *e.g.*, enhancers, regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 bp upstream of the start site, although a number of promoters have

recently been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In the thymidine kinase (tk) promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either cooperatively or independently to activate transcription.

[0129] One example of a suitable promoter is the immediate early cytomegalovirus (CMV) promoter sequence. This promoter sequence is a strong constitutive promoter sequence capable of driving high levels of expression of any polynucleotide sequence operatively linked thereto.

[0130] However, other constitutive promoter sequences can also be used, including, but not limited to the simian virus 40 (SV40) early promoter, mouse mammary tumor virus (MMTV), human immunodeficiency virus (HIV) long terminal repeat (LTR) promoter, MoMuLV promoter, an avian leukemia virus promoter, an Epstein-Barr virus immediate early promoter, a Rous sarcoma virus promoter, as well as human gene promoters such as, but not limited to, the actin promoter, the myosin promoter, the hemoglobin promoter, and the creatine kinase promoter. Further, the present disclosure should not be limited to the use of constitutive promoters. Inducible promoters are also contemplated as part of the present disclosure. The use of an inducible promoter provides a molecular switch capable of turning on expression of the polynucleotide sequence which it is operatively linked when such expression is desired, or turning off the expression when expression is not desired. Examples of inducible promoters include, but are not limited to a metallothionine promoter, a glucocorticoid promoter, a progesterone promoter, and a tetracycline promoter.

[0131] Methods of introducing and expressing genes into a cell are known in the art. In the context of an expression vector, the vector can be readily introduced into a host cell, *e.g.*, mammalian, bacterial, yeast, or insect cell by any method in the art. For example, the expression vector can be transferred into a host cell by physical, chemical, or biological means.

[0132] Physical methods for introducing a polynucleotide into a host cell include calcium phosphate precipitation, lipofection, particle bombardment, microinjection,

electroporation, and the like. Methods for producing cells comprising vectors and/or exogenous nucleic acids are well-known in the art. See, for example, Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York (2001)). In embodiments, a method for the introduction of a polynucleotide into a host cell is calcium phosphate transfection or polyethylenimine (PEI) Transfection.

[0133] Biological methods for introducing a polynucleotide of interest into a host cell include the use of DNA and RNA vectors. Viral vectors, and especially retroviral vectors, have become the most widely used method for inserting genes into mammalian, *e.g.*, human cells. Other viral vectors can be derived from lentivirus, poxviruses, herpes simplex virus I, adenoviruses and adeno-associated viruses, and the like. See, for example, U.S. Pat. Nos. 5,350,674 and 5,585,362.

Non-Viral Based Delivery System

[0134] Chemical means for introducing a polynucleotide into a host cell include colloidal dispersion systems, such as macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. An exemplary colloidal system for use as a delivery vehicle *in vitro* and *in vivo* is a liposome (*e.g.*, an artificial membrane vesicle).

[0135] The use of lipid formulations is contemplated for the introduction of the nucleic acids into a host cell (*in vitro*, *ex vivo* or *in vivo*). In another aspect, the nucleic acid can be associated with a lipid. The nucleic acid associated with a lipid can be encapsulated in the aqueous interior of a liposome, interspersed within the lipid bilayer of a liposome, attached to a liposome via a linking molecule that is associated with both the liposome and the oligonucleotide, entrapped in a liposome, complexed with a liposome, dispersed in a solution containing a lipid, mixed with a lipid, combined with a lipid, contained as a suspension in a lipid, contained or complexed with a micelle, or otherwise associated with a lipid. Lipid, lipid/DNA or lipid/expression vector associated compositions are not limited to any particular structure in solution. For example, they can be present in a bilayer structure, as micelles, or with a "collapsed" structure. They can also simply be interspersed in a solution, possibly forming aggregates that are not uniform in size or shape. Lipids are fatty substances which can be naturally occurring or synthetic lipids. For example, lipids include the fatty droplets that naturally occur in the cytoplasm as well

as the class of compounds which contain long-chain aliphatic hydrocarbons and their derivatives, such as fatty acids, alcohols, amines, amino alcohols, and aldehydes.

[0136] Lipids suitable for use can be obtained from commercial sources. For example, dimyristyl phosphatidylcholine ("DMPC") can be obtained from Sigma, St. Louis, Mo.; dicetyl phosphate ("DCP") can be obtained from K & K Laboratories (Plainview, N.Y.); cholesterol ("Choi") can be obtained from Calbiochem-Behring; dimyristyl phosphatidylglycerol ("DMPG") and other lipids can be obtained from Avanti Polar Lipids, Inc. (Birmingham, Ala.). Stock solutions of lipids in chloroform or chloroform/methanol can be stored at about -20o C. Chloroform is used as the only solvent since it is more readily evaporated than methanol. "Liposome" is a generic term encompassing a variety of single and multilamellar lipid vehicles formed by the generation of enclosed lipid bilayers or aggregates. Liposomes can be characterized as having vesicular structures with a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes have multiple lipid layers separated by aqueous medium. They form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers (Ghosh et al., *Glycobiology 5*: 505-10 (1991)). However, compositions that have different structures in solution than the normal vesicular structure are also encompassed. For example, the lipids can assume a micellar structure or merely exist as nonuniform aggregates of lipid molecules. Also contemplated are lipofectamine-nucleic acid complexes.

[0137] In some instances, polynucleotides encoding polypeptides can also be introduced into cells using non-viral based delivery systems, such as the “Sleeping Beauty (SB) Transposon System,” which refers a synthetic DNA transposon system for introducing DNA sequences into the chromosomes of vertebrates. Some exemplary embodiments of the system are described, for example, in U.S. Pat. Nos. 6,489,458 and 8,227,432. The Sleeping Beauty transposon system is composed of a Sleeping Beauty (SB) transposase and a SB transposon. In embodiments, the Sleeping Beauty transposon system can include the SB11 transposon system, the SB100X transposon system, or the SB110 transposon system.

[0138] DNA transposons translocate from one DNA site to another in a simple, cut-and-paste manner. Transposition is a precise process in which a defined DNA segment is excised from one DNA molecule and moved to another site in the same or different DNA molecule or genome. As do other Tc1/mariner-type transposases, SB transposase inserts a transposon into a TA dinucleotide base pair in a recipient DNA sequence. The insertion site can be elsewhere in the same DNA molecule, or in another DNA molecule (or chromosome). In mammalian genomes, including humans, there are approximately 200 million TA sites. The TA insertion site is duplicated in the process of transposon integration. This duplication of the TA sequence is a hallmark of transposition and used to ascertain the mechanism in some experiments. The transposase can be encoded either within the transposon or the transposase can be supplied by another source, for instance a DNA or mRNA source, in which case the transposon becomes a non-autonomous element. Non-autonomous transposons are most useful as genetic tools because after insertion they cannot independently continue to excise and re-insert. SB transposons envisaged to be used as non-viral vectors for introduction of genes into genomes of vertebrate animals and for gene therapy.

[0139] Regardless of the method used to introduce exogenous nucleic acids into a host cell or otherwise expose a cell to the inhibitor of the present disclosure, in order to confirm the presence of the recombinant DNA sequence in the host cell, a variety of assays can be performed. Such assays include, for example, molecular assays well known to those of skill in the art, such as Southern and Northern blotting, RT-PCR and PCR; "biochemical" assays, such as detecting the presence or absence of a particular peptide, *e.g.*, by immunological means (ELISAs and Western blots) or by assays described herein to identify agents falling within the scope of the present disclosure.

[0140] In embodiments, a modified effector cell described herein and other genetic elements are delivered to a cell using the SB11 transposon system, the SB100X transposon system, the SB110 transposon system, the piggyBac transposon system (see, *e.g.*, Wilson et al, "PiggyBac Transposon-mediated Gene Transfer in Human Cells," *Molecular Therapy 15*:139-145 (2007), incorporated herein by reference in its entirety) and/or the piggyBac transposon system (see, *e.g.*, Mitra et al., "Functional characterization of piggyBac from the bat Myotis lucifugus unveils an active mammalian DNA transposon," *Proc. Natl. Acad. Sci USA 110*:234-239 (2013). Additional

transposases and transposon systems are provided in U.S. Patent Nos.; 6,489,458; 6,613,752, 7,148,203; 7,985,739; 8,227,432; 9,228,180; U.S. Patent Pub. No. 2011/0117072; Mates et al., *Nat Genet, 41(6)*:753-61 (2009). doi: 10.1038/ng.343. Epub 2009 May 3, *Gene Ther., 18(9)*:849-56 (2011). doi: 10.1038/gt.2011.40. Epub 2011 Mar 31 and in Ivics et al., *Cell, 91(4)*:501-10, (1997), each of which is incorporated herein by reference in their entirety.

[0141] Additional suitable non-viral systems can include integrating expression vectors, which can randomly integrate into the host cell's DNA, or can include a recombination site to enable the specific recombination between the expression vector and the host cell's chromosome. Targeted integration of transgenes into predefined genetic loci is a desirable goal for many applications. First, a first recombination site for a site-specific recombinase is inserted at a genomic site, either at a random or at a predetermined location. Subsequently, the cells are transfected with a plasmid carrying the gene or DNA of interest and the second recombination site and a source for recombinase (expression plasmid, RNA, protein, or virus-expressing recombinase). Recombination between the first and second recombination sites leads to integration of plasmid DNA.

[0142] Such integrating expression vectors can utilize the endogenous expression control sequences of the host cell's chromosomes to effect expression of the desired protein. In some embodiments, targeted integration is promoted by the presence of sequences on the donor polynucleotide that are homologous to sequences flanking the integration site. For example, targeted integration using the donor polynucleotides described herein can be achieved following conventional transfection techniques, *e.g.* techniques used to create gene knockouts or knockins by homologous recombination. In other embodiments, targeted integration is promoted both by the presence of sequences on the donor polynucleotide that are homologous to sequences flanking the integration site, and by contacting the cells with donor polynucleotide in the presence of a site-specific recombinase. By a site-specific recombinase, or simply a recombinase, it is meant is a polypeptide that catalyzes conservative site-specific recombination between its compatible recombination sites. As used herein, a site-specific recombinase includes native polypeptides as well as derivatives, variants and/or fragments that retain activity, and native polynucleotides, derivatives, variants, and/or fragments that encode a recombinase that retains activity.

[0143] Also provided herein is a system for integrating heterologous genes in a host cell, said system comprising one or more gene expression cassettes. In some instances, the system includes a first gene expression cassette comprising a first polynucleotide encoding a first polypeptide construct. In other instances, the system can include a second gene expression cassette comprising a second polynucleotide encoding a second polypeptide construct. In yet other instances, the system can include a third gene expression cassette. In one embodiment, one of the gene expression cassettes can comprise a gene switch polynucleotide encoding one or more of: (i) a transactivation domain; (ii) nuclear receptor ligand binding domain; (iii) a DNA-binding domain; and (iv) ecdysone receptor binding domain. In another embodiment, the system further includes recombinant attachment sites; and a serine recombinase; such that upon contacting said host cell with at least said first gene expression cassette, in the presence of said serine recombinase, said heterologous genes are integrated in said host cell.

[0144] In some instances, the system further comprises a ligand; such that upon contacting said host cell, in the presence of said ligand, said heterologous gene are expressed in said host cell. In one instance, the system also includes recombinant attachment sites. In some instances, one recombination attachment site is a phage genomic recombination attachment site (attP) or a bacterial genomic recombination attachment site (attB). In one instance, the host cell is an eukaryotic cell. In another instance, the host cell is a human cell. In further instances, the host cell is a T cell or NK cell.

Promoters

[0145] "Promoter" refers to a region of a polynucleotide that initiates transcription of a coding sequence. Promoters are located near the transcription start sites of genes, on the same strand and upstream on the DNA (towards the 5' region of the sense strand). Some promoters are constitutive as they are active in all circumstances in the cell, while others are regulated becoming active in response to specific stimuli, *e.g.*, an inducible promoter. Yet other promoters are tissue specific or activated promoters, including but not limited to T-cell specific promoters.

[0146] The term "promoter activity" and its grammatical equivalents as used herein refer to the extent of expression of nucleotide sequence that is operably linked to the promoter whose activity is being measured. Promoter activity can be measured directly by determining the amount of RNA transcript produced, for example by Northern blot

analysis or indirectly by determining the amount of product coded for by the linked nucleic acid sequence, such as a reporter nucleic acid sequence linked to the promoter.

[0147] "Inducible promoter" as used herein refers to a promoter which is induced into activity by the presence or absence of transcriptional regulators, *e.g.*, biotic or abiotic factors. Inducible promoters are useful because the expression of genes operably linked to them can be turned on or off at certain stages of development of an organism or in a particular tissue. Examples of inducible promoters are alcohol-regulated promoters, tetracycline-regulated promoters, steroid-regulated promoters, metal-regulated promoters, pathogenesis-regulated promoters, temperature-regulated promoters and light-regulated promoters. In one embodiment, the inducible promoter is part of a genetic switch. The inducible promoter can be a gene switch ligand inducible promoter. In some cases, an inducible promoter can be a small molecule ligand-inducible two polypeptide ecdysone receptor-based gene switch, such as RHEOSWITCH® gene switch. In some cases, a gene switch can be selected from ecdysone-based receptor components as described in, but without limitation to, any of the systems described in: PCT/US2001/009050 (WO 2001/070816); U.S. Pat. Nos. 7,091,038; 7,776,587; 7,807,417; 8,202,718; PCT/US2001/030608 (WO 2002/029075); U.S. Pat. Nos. 8,105,825; 8,168,426; PCT/1J52002/005235 (WO 2002/066613); U.S. App. No. 10/468,200 (U.S. Pub. No. 20120167239); PCT/US2002/005706 (WO 2002/066614); U.S. Pat. Nos. 7,531,326; 8,236,556; 8,598,409; PCT/U52002/005090 (WO 2002/066612); U.S. Pat. No. 8,715,959 (U.S. Pub. No. 20060100416); PCT/US2002/005234 (WO 2003/027266); U.S. Pat. Nos. 7,601,508; 7,829,676; 7,919,269; 8,030,067; PCT/U52002/005708 (WO 2002/066615); U.S. App. No. 10/468,192 (U.S. Pub. No. 20110212528); PCT/US2002/005026 (WO 2003/027289); U.S. Pat. Nos. 7,563,879; 8,021,878; 8,497,093; PCT/US2005/015089 (WO 2005/108617); U.S. Pat. No. 7,935,510; 8,076,454; PCT/U52008/011270 (WO 2009/045370); U.S. App. No. 12/241,018 (U.S. Pub. No. 20090136465); PCT/US2008/011563 (WO 2009/048560); U.S. App. No. 12/247,738 (U.S. Pub. No. 20090123441); PCT/US2009/005510 (WO 2010/042189); U.S. App. No. 13/123,129 (U.S. Pub. No. 20110268766); PCT/US2011/029682 (WO 2011/119773); U.S. App. No. 13/636,473 (U.S. Pub. No. 20130195800); PCT/US2012/027515 (WO 2012/122025); and, U.S. Pat. No. 9,402,919 each of which is incorporated by reference in its entirety).

[0148] Provided herein are methods comprising administering to a subject at least one non-viral vector comprising a polynucleotide encoding a polypeptide sequence described herein comprising at least two functional proteins or portions thereof; at least one promoter; and at least one engineered recombination site; wherein said at least one promoter drives expression of said at least two functional proteins. In some cases, at least one promoter can be constitutive. In some cases, at least one promoter can be tissue-specific. In some cases, at least one promoter can be inducible. In some cases, an inducible promoter is a small molecule ligand-inducible two polypeptide ecdysone receptor-based gene switch. In other cases, a combination of promoters wherein at least one promoter can be inducible and at least one promoter can be activation specific can be utilized.

[0149] An inducible promoter utilizes a ligand for dose-regulated control of expression of said at least two genes. In some cases, a ligand can be selected from a group consisting of ecdysteroid, 9-cis-retinoic acid, synthetic analogs of retinoic acid, N,N'-diacylhydrazines, oxadiazolines, dibenzoylalkyl cyanohydrazines, N-alkyl-N,N'-diaroylhydrazines, N-acyl-N- alkylcarbonylhydrazines, N-aroyl-N-alkyl-N'-aroylhydrazines, arnidoketones, 3,5-di-tert-butyl- 4-hydroxy-N-isobutyl-benzamide, 8-O-acetylharpagide, oxysterols, 22(R) hydroxycholesterol, 24(S) hydroxycholesterol, 25-epoxycholesterol, T0901317, 5-alpha-6-alpha-epoxycholesterol-3- sulfate (ECHS), 7-ketocholesterol-3-sulfate, framesol, bile acids, 1,1-biphosphonate esters, juvenile hormone III, RG-115819 (3,5 -Dimethyl-benzoic acid N-(1-ethy1-2,2-dimethyl- propy1)-N'-(2-methyl-3-methoxy-benzoy1)-hydrazide-), RG-115932 ((R)-3,5-Dimethyl- benzoic acid N-(1-tert-butyl-buty1)-N'-(2-ethy1-3-methoxy-benzoy1)-hydrazide), and RG- 115830 (3,5 -Dimethyl-benzoic acid N-(1-tert-butyl-buty1)-N'-(2-ethy1-3-methoxy-benzoy1)- hydrazide), and any combination thereof.

[0150] In some embodiments, a promoter is an inducible promoter. In some embodiments, a promoter is a non-inducible promoter. In some cases, a promoter can be a tissue-specific promoter. Herein "tissue-specific" refers to regulated expression of a gene in a subset of tissues or cell types. In some cases, a tissue-specific promoter can be regulated spatially such that the promoter drives expression only in certain tissues or cell types of an organism. In other cases, a tissue-specific promoter can be regulated temporally such that the promoter drives expression in a cell type or tissue differently

across time, including during development of an organism. In some cases, a tissue-specific promoter is regulated both spatially and temporally. In certain embodiments, a tissue-specific promoter is activated in certain cell types either constitutively or intermittently at particular times or stages of the cell type. For example, a tissue-specific promoter can be a promoter that is activated when a specific cell such as a T cell or a NK cell is activated. T cells can be activated in a variety of ways, for example, when presented with peptide antigens by MHC class II molecules.

[0151] In one case, at least one promoter is an engineered promoter or variants thereof. As described herein, the promoter can incorporate minimal promoter sequences from IL-2 and one or more of the following: nuclear factor of activated T-cells (NFAT) response element(s); NFIL2D response element, NFkB/TCF response element, NF_AT/NFIL2B response element or NFIL2A/OCT response element. Examples of response elements are described in Mattila et al., *EMBO J. 9(13)*:4425-33 (1990); incorporated herein in its entirety.

[0152] Use of gene switch for ligand inducible control of IL-12 expression described herein can improve the safety profile of IL-12 by for example allowing for regulated expression and improving therapeutic index. However, a condition for ligand dose dependent expression of IL-12 using gene switch(es) is the presence or absence of activator ligand (*e.g.* veledimex). In certain embodiments, an additional conditional control for induction of IL-12 expression is contemplated. Gene switch components under the control of T cell activated specific promoters are provided. This results in conditional expression (*e.g.*,T cell activation) of gene switch components necessary for veledimex controlled expression of transgene(s) under control of a gene switch. In some embodiments, this results in preferential expression of cytokines such as IL-12 or IL-15 by tumor specific T cells when veledimex is present and T cells are activated. This can lead to increased localized levels of gene switch controlled transgene expression.

[0153] For example, T cell activation specific expression of gene switch components can be controlled by promoter comprising Nuclear Factor of Activated T-cells (NFAT) response element(s). NFAT transcription factors are key modulators of effector T-cell states. NFATs are early transcriptional checkpoint progressively driving exhaustion. NFATs are quickly activated in T cells following TCR stimulation and form a protein complex with AP-1 induced by appropriate co-stimulation signaling and regulate effector

genes and T-cell functions. NFAT response element(s) can be fused with other minimal promoter sequences (*e.g.* IL2 minimal promoter) to drive expression of transgenes in response to T cell activation.

[0154] Other examples of activation specific promoters include but are not limited to interleukin-2 (IL2) promoter and Programmed Death (PD)-1 (CD279) promoter. Gene switch components can also be conditionally expressed upon immune cell activation by fusing binding sites for other nuclear factors like NF-κB of proinflammatory signaling pathway to minimal promoter sequence (*e.g.* IL2).

[0155] In certain embodiments, the promoter can be any one or more of: IL-2 core promoter, IL-2 minimal promoter, IL-2 enhancer and promoter variant, $(NF\text{-}\kappa B)_1$-IL2 promoter variant, $(NF\text{-}\kappa B)_3$-IL2 promoter variant, $(NF\text{-}\kappa B)_6$-IL2 promoter variant, 1X NFAT response elements-IL2 promoter variant, 3X NFAT response elements-IL2 promoter variant , 6X NFAT response elements-IL2 promoter variant, human EEF1A1 promoter variant, human EEF1A1 promoter and enhancer, human UBC promoter and synthetic minimal promoter 1.

Gene Switch

[0156] Provided herein are gene switch polypeptides, polynucleotides encoding ligand-inducible gene switch polypeptides, and methods and systems incorporating these polypeptides and/or polynucleotides. In certain aspects, the present disclosure is directed to a polynucleotide comprising one or more polynucleotides encoding a gene switch system for inducible control of heterologous gene expression, wherein the heterologous gene expression is regulated by said gene switch system; and, wherein said heterologous gene comprises a polynucleotide encoding a polypeptide comprising one or more immune response-inducing human papilloma virus (HPV) polypeptides, disclosed herein.

[0157] The term “gene switch” refers to the combination of a response element associated with a promoter, and for instance, an EcR based system which, in the presence of one or more ligands, modulates the expression of a gene into which the response element and promoter are incorporated. Tightly regulated inducible gene expression systems or gene switches are useful for various applications such as gene therapy, large scale production of proteins in cells, cell based high throughput screening assays, functional genomics and regulation of traits in transgenic plants and animals. Such inducible gene expression systems can include ligand inducible heterologous gene expression systems.

[0158] An early version of EcR-based gene switch used *Drosophila melanogaster* EcR (DmEcR) and *Mus musculus* RXR (MmRXR) polypeptides and showed that these receptors in the presence of steroid, ponasteroneA, transactivate reporter genes in mammalian cell lines and transgenic mice (Christopherson et al., *Proc. Natl. Acad. Sci. USA* 89(14):6314-18 (1992); No et al., *Proc. Natl. Acad. Sci. USA* 93(8):3346-51 (1996)). Later, Suhr et al. (*Proc. Natl. Acad. Sci. USA* 95(14):7999-8004 (1998)) showed that non-steroidal ecdysone agonist, tebufenozide, induced high level of transactivation of reporter genes in mammalian cells through *Bombyx mori* EcR (BmEcR) in the absence of exogenous heterodimer partner.

[0159] International Patent Applications No. PCT/US97/05330 (WO 97/38117) and PCT/US99/08381 (WO99/58155) disclose methods for modulating the expression of an exogenous gene in which a DNA construct comprising the exogenous gene and an ecdysone response element is activated by a second DNA construct comprising an ecdysone receptor that, in the presence of a ligand therefor, and optionally in the presence of a receptor capable of acting as a silent partner, binds to the ecdysone response element to induce gene expression. In this example, the ecdysone receptor was isolated from *Drosophila melanogaster*. Typically, such systems require the presence of the silent partner, preferably retinoid X receptor (RXR), in order to provide optimum activation. In mammalian cells, insect ecdysone receptor (EcR) is capable of heterodimerizing with mammalian retinoid X receptor (RXR) and, thereby, be used to regulate expression of target genes or heterologous genes in a ligand dependent manner. International Patent Application No. PCT/US98/14215 (WO 99/02683) discloses that the ecdysone receptor isolated from the silk moth *Bombyx mori* is functional in mammalian systems without the need for an exogenous dimer partner.

[0160] U.S. Pat. No. 6,265,173 discloses that various members of the steroid/thyroid superfamily of receptors can combine with *Drosophila melanogaster* ultraspiracle receptor (USP) or fragments thereof comprising at least the dimerization domain of USP for use in a gene expression system. U.S. Pat. No. 5,880,333 discloses a *Drosophila melanogaster* EcR and ultraspiracle (USP) heterodimer system used in plants in which the transactivation domain and the DNA binding domain are positioned on two different hybrid proteins. In each of these cases, the transactivation domain and the DNA binding domain (either as native EcR as in International Patent Application No. PCT/US98/14215

or as modified EcR as in International Patent Application No. PCT/US97/05330) were incorporated into a single molecule and the other heterodimeric partners, either USP or RXR, were used in their native state.

[0161] International Patent Application No. PCT/US01/0905 discloses an ecdysone receptor-based inducible gene expression system in which the transactivation and DNA binding domains are separated from each other by placing them on two different proteins results in greatly reduced background activity in the absence of a ligand and significantly increased activity over background in the presence of a ligand. This two-hybrid system is a significantly improved inducible gene expression modulation system compared to the two systems disclosed in applications PCT/US97/05330 and PCT/US98/14215. The two-hybrid system is believed to exploit the ability of a pair of interacting proteins to bring the transcription activation domain into a more favorable position relative to the DNA binding domain such that when the DNA binding domain binds to the DNA binding site on the gene, the transactivation domain more effectively activates the promoter (see, for example, U.S. Pat. No. 5,283,173). The two-hybrid gene expression system comprises two gene expression cassettes; the first encoding a DNA binding domain fused to a nuclear receptor polypeptide, and the second encoding a transactivation domain fused to a different nuclear receptor polypeptide. In the presence of ligand, it is believed that a conformational change is induced which promotes interaction of the first polypeptide with the second polypeptide thereby resulting in dimerization of the DNA binding domain and the transactivation domain. Since the DNA binding and transactivation domains reside on two different molecules, the background activity in the absence of ligand is greatly reduced.

[0162] In some cases, an inducible promoter can be a small molecule ligand-inducible two polypeptide ecdysone receptor-based gene switch, such as Intrexon Corporation's RHEOSWITCH® gene switch. In some cases, a gene switch can be selected from ecdysone-based receptor components as described in, but without limitation to, any of the systems described in: PCT/US2001/009050 (WO 2001/070816); U.S. Pat. Nos. 7,091,038; 7,776,587; 7,807,417; 8,202,718; PCT/US2001/030608 (WO 2002/029075); U.S. Pat. Nos. 8,105,825; 8,168,426; PCT/1J52002/005235 (WO 2002/066613); U.S. App. No. 10/468,200 (U.S. Pub. No. 20120167239); PCT/US2002/005706 (WO 2002/066614); U.S. Pat. Nos. 7,531,326; 8,236,556; 8,598,409; PCT/U52002/005090

(WO 2002/066612); U.S. Pat. No. 8,715,959 (U.S. Pub. No. 20060100416); PCT/US2002/005234 (WO 2003/027266); U.S. Pat. Nos. 7,601,508; 7,829,676; 7,919,269; 8,030,067; PCT/U52002/005708 (WO 2002/066615); U.S. App. No. 10/468,192 (U.S. Pub. No. 20110212528); PCT/US2002/005026 (WO 2003/027289); U.S. Pat. Nos. 7,563,879; 8,021,878; 8,497,093; PCT/US2005/015089 (WO 2005/108617); U.S. Pat. No. 7,935,510; 8,076,454; PCT/U52008/011270 (WO 2009/045370); U.S. App. No. 12/241,018 (U.S. Pub. No. 20090136465); PCT/US2008/011563 (WO 2009/048560); U.S. App. No. 12/247,738 (U.S. Pub. No. 20090123441); PCT/US2009/005510 (WO 2010/042189); U.S. App. No. 13/123,129 (U.S. Pub. No. 20110268766); PCT/US2011/029682 (WO 2011/119773); U.S. App. No. 13/636,473 (U.S. Pub. No. 20130195800); PCT/US2012/027515 (WO 2012/122025); and, U.S. Pat. No. 9,402,919 each of which is incorporated by reference in its entirety.

[0163] Provided are systems for modulating the expression of a heterologous gene and an interleukin in a host cell, comprising polynucleotides expressing gene-switch polypeptides disclosed herein.

[0164] In some embodiments are systems for modulating the expression of a heterologous gene and a cytokine in a host cell, comprising a first gene expression cassette comprising a first polynucleotide encoding a first polypeptide; a second gene expression cassette comprising a second polynucleotide encoding a second polypeptide; and a ligand; wherein said first and second polypeptides comprise one or more of: (i) a transactivation domain; (ii) a DNA-binding domain; and (iii) a ligand binding domain; (iv) said heterologous gene; and (vi) said cytokine such that upon contacting said host cell with said first gene expression cassette and said second gene expression cassette in the presence of said ligand, said heterologous gene and said cytokine are expressed in said host cell. In some cases, the heterologous gene comprises an antigen binding polypeptide described herein. In some cases, the cytokine comprises at least one chemokine, interferon, interleukin, lymphokine, tumor necrosis factor, or variant or combination thereof. In some cases, the cytokine is an interleukin. In some cases the interleukin is at least one of IL12, IL2, IL15, IL21, and functional variants and fragments thereof. In some embodiments, the cytokines can be membrane bound or secreted. In other embodiments, the cytokines can be intracellular. The interleukin can comprise membrane bound IL-15 (mbIL-15) or a fusion of IL-15 and IL-15Rα. In some embodiments, a

mbIL-15 is a membrane-bound chimeric IL-15 which can be co-expressed with a modified effector cell described herein. In some embodiments, the mbIL-15 comprises a full-length IL-15 (*e.g.*, a native IL-15 polypeptide) or fragment or variant thereof, fused in frame with a full length IL-15Rα, functional fragment or variant thereof. In some cases, the IL-15 is indirectly linked to the IL-15Rα through a linker. In some instances, the mbIL-15 is as described in Hurton et al., "Tethered IL-15 augments antitumor activity and promotes a stem-cell memory subset in tumor-specific T cells," *Proc. Natl. Acad. Sci. USA 113(48)*:E7788-E7797 (2016). In another aspect, the interleukin can comprise IL-12. In some embodiments, the IL-12 is a single chain IL-12 (scIL-12), protease sensitive IL-12, destabilized IL-12, membrane bound IL-12, intercalated IL-12. In some instances, the IL-12 variants are as described in WO2015/095249, WO2016/048903, WO2017/062953, all of which is incorporated by reference in their entireties.

[0165] Provided herein are polynucleotides encoding gene switch polypeptides, wherein said gene switch polypeptides comprise: a) a first gene switch polypeptide comprising a DNA-binding domain fused to a nuclear receptor ligand binding domain, and b) a second gene switch polypeptide comprising a transactivation domain fused to a nuclear receptor ligand binding domain, wherein the first gene switch polypeptide and the second gene switch polypeptide are connected by a linker. In some cases, the linker can be a linker described herein, for instance GSG linker, furinlink, a 2A linker such as F/T2A, T2A, p2A, GSG-p2A, variants and derivatives thereof. In other instances, the linker can be an IRES.

[0166] In some cases, the DNA binding domain (DBD) comprises a DBD described herein, for instance at least one of GAL4 (GAL4 DBD), a LexA DBD, a transcription factor DBD, a steroid/thyroid hormone nuclear receptor superfamily member DBD, a bacterial LacZ DBD, and a yeast DBD. The transactivation domain can comprise a transactivation domain described herein, for instance one of a VP16 transactivation domain, a p53 transactivation domain and a B42 acidic activator transactivation domain. The Nuclear receptor ligand binding domain can comprise at least one of a ecdysone receptor (EcR), a ubiquitous receptor, an orphan receptor 1, a NER-1, a steroid hormone nuclear receptor 1, a retinoid X receptor interacting protein-15, a liver X receptor β, a steroid hormone receptor like protein, a liver X receptor, a liver X receptor α, a farnesoid X receptor, a receptor interacting protein 14, and a famesol receptor.

[0167] In some cases, the gene switch polypeptides connected by a polypeptide linker or ribosome-skipping sequence exhibit improved dose-dependent ligand-inducible control of gene expression compared to a ligand-inducible gene switch wherein the gene switch polypeptides are connected by non-coding sequences, such as an IRES. In some cases, the gene switch polypeptides connected by a 2A linker can exhibit improved dose-dependent ligand-inducible control of heterologous gene expression compared to a gene switch wherein said gene switch polypeptides are separated by an IRES.

[0168] In some embodiments, the gene switch comprises a VP16 transactivation domain. In one embodiment, the gene switch comprises at least one of an ecdysone receptor (EcR), a ubiquitous receptor, an orphan receptor 1, a NER-1, a steroid hormone nuclear receptor 1, a retinoid X receptor interacting protein-15, a liver X receptor β, a steroid hormone receptor like protein, a liver X receptor, a liver X receptor α, a farnesoid X receptor, a receptor interacting protein 14, and a famesol receptor. In another embodiment, a DNA-binding domain (DBD) of the gene switch comprises at least one of GAL4 (GAL4 DBD), a LexA DBD, a transcription factor DBD, a steroid/thyroid hormone nuclear receptor superfamily member DBD, a bacterial LacZ DBD, and a yeast DBD. In yet another case, the gene switch further comprises at least one of ultraspiracle protein (USP), retinoid receptor X (RXR), functional fragments and variants thereof wherein said functional fragments and variants are capable of binding to an EcR.

[0169] The polypeptides and polynucleotides as described herein can be expressed in an engineered cell. Herein an engineered cell is a cell which has been modified from its natural or endogenous state. An example of an engineered cell is a cell described herein which has been modified (*e.g.*, by transfection of a polynucleotide into the cell) to encode for example, gene switch polypeptides, gene of interest (GOI), cell tags, heterologous genes and any other polypeptides and polynucleotides described herein.

Ligands

[0170] In some embodiments, a ligand used for inducible gene switch regulation can be selected from any of, but without limitation to, following: N-[(1R)-1-(1,1-dimethylethyl)butyl]-N'-(2-ethyl-3-methoxybenzoyl)-3,5-dimethylbenzohydrazide (also referred to as veledimex), (2S,3R,5R,9R,10R,13R,14S,17R)-17- [(2S,3R)-3,6-dihydroxy-6-methylheptan- 2-yl]-2,3,14-trihydroxy-10,13-dimethyl- 2,3,4,5,9,11,12,15,16,17-decahydro- 1H-cyclopenta[a]phenanthren-6-one; N'-(3,5-Dimethylbenzoyl)-N'-[(3R)-2,2-

dimethyl-3-hexanyl]-2-ethyl-3-methoxybenzohydrazide; 5-Methyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(3,5-dimethyl-benzoyl)-N′-(1-ethyl-2,2-dimethyl-propyl)-hydrazide; 5-Methyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(3,5-dimethoxy-4-methyl-benzoyl)-N′-(1-ethyl-2,2-dimethyl-propyl)-hydrazide; 5-Methyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(1-tert-butyl-butyl)-N′-(3,5-dimethyl-benzoyl)-hydrazide; 5-Methyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(1-tert-butyl-butyl)-N′-(3,5-dimethoxy-4-methyl-benzoyl)-hydrazide; 5-Ethyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(3,5-dimethyl-benzoyl)-N′-(1-ethyl-2,2-dimethyl-propyl)-hydrazide; 5-Ethyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(3,5-dimethoxy-4-methyl-benzoyl)-N′-(1-ethyl-2,2-dimethyl-propyl)-hydrazide; 5-Ethyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(1-tert-butyl-butyl)-N′-(3,5-dimethyl-benzoyl)-hydrazide; 5-Ethyl-2,3-dihydro-benzo[1,4]dioxine-6-carboxylic acid N′-(1-tert-butyl-butyl)-N′-(3,5-dimethoxy-4-methyl-benzoyl)-hydrazide; 3,5-Dimethyl-benzoic acid N-(1-ethyl-2,2-dimethyl-propyl)-N′-(3-methoxy-2-methyl-benzoyl)-hydrazide; 3,5-Dimethoxy-4-methyl-benzoic acid N-(1-ethyl-2,2-dimethyl-propyl)-N′-(3-methoxy-2-methyl-benzoyl)-hydrazide; 3,5-Dimethyl-benzoic acid N-(1-tert-butyl-butyl)-N′-(3-methoxy-2-methyl-benzoyl)-hydrazide; 3,5-Dimethoxy-4-methyl-benzoic acid N-(1-tert-butyl-butyl)-N′-(3-methoxy-2-methyl-benzoyl)-hydrazide; 3,5-Dimethyl-benzoic acid N-(1-ethyl-2,2-dimethyl-propyl)-N′-(2-ethyl-3-methoxy-benzoyl)-hydrazide; 3,5-Dimethoxy-4-methyl-benzoic acid N-(1-ethyl-2,2-dimethyl-propyl)-N′-(2-ethyl-3-methoxy-benzoyl)-hydrazide; 3,5-Dimethyl-benzoic acid N-(1-tert-butyl-butyl)-N′-(2-ethyl-3-methoxy-benzoyl)-hydrazide; 3,5-Dimethoxy-4-methyl-benzoic acid N-(1-tert-butyl-butyl)-N′-(2-ethyl-3-methoxy-benzoyl)-hydrazide; 2-Methoxy-nicotinic acid N-(1-tert-butyl-pentyl)-N′-(4-ethyl-benzoyl)-hydrazide; 3,5-Dimethyl-benzoic acid N-(2,2-dimethyl-1-phenyl-propyl)-N′-(4-ethyl-benzoyl)-hydrazide; 3,5-Dimethyl-benzoic acid N-(1-tert-butyl-pentyl)-N′-(3-methoxy-2-methyl-benzoyl)-hydrazide; and 3,5-Dimethoxy-4-methyl-benzoic acid N-(1-tert-butyl-pentyl)-N′-(3-methoxy-2-methyl-benzoyl)-hydrazide.

[0171] In some cases, a ligand used for dose-regulated control of ecdysone receptor-based inducible gene switch can be selected from any of, but without limitation to, an ecdysteroid, such as ecdysone, 20-hydroxyecdysone, ponasterone A, muristerone A, and the like, 9-cis-retinoic acid, synthetic analogs of retinoic acid, N,N'-

diacylhydrazines such as those disclosed in U.S. Pat. Nos. 6,013,836; 5,117,057; 5,530,028; and 5,378,726 and U.S. Published Application Nos. 2005/0209283 and 2006/0020146; oxadiazolines as described in U.S. Published Application No. 2004/0171651; dibenzoylalkyl cyanohydrazines such as those disclosed in European Application No. 461,809; N-alkyl-N,N'-diaroylhydrazines such as those disclosed in U.S. Pat. No. 5,225,443; N-acyl-N-alkylcarbonylhydrazines such as those disclosed in European Application No. 234,994; N-aroyl-N-alkyl-N'-aroylhydrazines such as those described in U.S. Pat. No. 4,985,461; arnidoketones such as those described in U.S. Published Application No. 2004/0049037; each of which is incorporated herein by reference and other similar materials including 3,5-di-tert-butyl-4-hydroxy-N-isobutyl-benzamide, 8-O-acetylharpagide, oxysterols, 22(R) hydroxycholesterol, 24(S) hydroxycholesterol, 25-epoxycholesterol, T0901317, 5-alpha-6-alpha-epoxycholesterol-3-sulfate (ECHS), 7-ketocholesterol-3-sulfate, framesol, bile acids, 1,1-biphosphonate esters, juvenile hormone III, and the like. Examples of diacylhydrazine ligands useful in the present disclosure include RG-115819 (3,5 -Dimethyl-benzoic acid N-(1-ethy1-2,2-dimethyl-propy1)-N'-(2-methyl-3-methoxy-benzoy1)-hydrazide-), RG-115932 ((R)-3,5-Dimethyl-benzoic acid N-(1-tert-butyl-buty1)-N'-(2-ethy1-3-methoxy-benzoy1)-hydrazide), and RG-115830 (3,5-Dimethyl-benzoic acid N-(1-tert-butyl-buty1)-N'-(2-ethy1-3-methoxy-benzoy1)-hydrazide). See, *e.g.*, U.S. patent application Ser. No. 12/155,111, and PCT Appl. No. PCT/US2008/006757, both of which are incorporated herein by reference in their entireties.

Cytokines

[0172] In certain embodiments, HPV vaccine antigens provided herein are co-delivered and/or co-expressed (*e.g.*, as part of the same HPV antigen delivery vector or via a separate vector) along with other cytokines. In certain embodiments, HPV vaccine antigens Provided herein are polynucleotides encoding gene-switch polypeptides and a cytokine, or variant or derivative thereof, and methods and systems incorporating the same. Cytokine is a category of small proteins between about 5-20 kDa that are involved in cell signaling. In some instances, cytokines include chemokines, interferons, interleukins, colony-stimulating factors or tumor necrosis factors. In some embodiments, chemokines play a role as a chemoattractant to guide the migration of cells, and is classified into four subfamilies: CXC, CC, CX3C, and XC. Exemplary chemokines

include chemokines from the CC subfamily: CCL1, CCL2 (MCP-1), CCL3, CCL4, CCL5 (RANTES), CCL6, CCL7, CCL8, CCL9 (or CCL10), CCL11, CCL12, CCL13, CCL14, CCL15, CCL16, CCL17, CCL18, CCL19, CCL20, CCL21, CCL22, CCL23, CCL24, CCL25, CCL26, CCL27, and CCL28; the CXC subfamily: CXCL1, CXCL2, CXCL3, CXCL4, CXCL5, CXCL6, CXCL7, CXCL8, CXCL9, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL15, CXCL16, and CXCL17; the XC subfamily: XCL1 and XCL2; and the CX3C subfamily CX3CL1.

[0173] In certain embodiments, HPV vaccine antigens provided herein are co-delivered and/or co-expressed (*e.g.*, as part of the same HPV antigen delivery vector or via a separate vector) along with interferons. Interferons (IFNs) comprise interferon type I (*e.g.* IFN-α, IFN-β, IFN-ε, IFN-κ, and IFN-ω), interferon type II (*e.g.* IFN-γ), and interferon type III. In some embodiments, IFN-α is further classified into about 13 subtypes including IFNA1, IFNA2, IFNA4, IFNA5, IFNA6, IFNA7, IFNA8, IFNA10, IFNA13, IFNA14, IFNA16, IFNA17, and IFNA21.

[0174] In certain embodiments, HPV vaccine antigens provided herein are co-delivered and/or co-expressed (*e.g.*, as part of the same HPV antigen delivery vector or via a separate vector) along with other interleukins. Interleukins are expressed by leukocytes or white blood cells and they promote the development and differentiation of T and B lymphocytes and hematopoietic cells. Exemplary interleukines include IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8 (CXCL8), IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IL-35, and IL-36. In some embodiments, interleukins are IL-2, IL-12, IL-15, IL-21 or a fusion of IL-15 and IL-15α.

[0175] In some aspects, the interleukin can comprise IL-12. In some embodiments, the IL-12 is a single chain IL-12 (scIL-12), protease sensitive IL-12, destabilized IL-12, membrane bound IL-12, intercalated IL-12. In some instances, the IL-12 variants are as described in WO2015/095249, WO2016/048903, WO2017/062953, all of which is incorporated by reference in their entireties.

[0176] In some embodiments, an interleukin comprises mbIL-15. In some embodiments, a mbIL-15 is a membrane-bound chimeric IL-15 which can be co-expressed with a modified effector cell described herein. In some embodiments, the mbIL-15 comprises a full-length IL-15 (*e.g.*, a native IL-15 polypeptide) or fragment or variant thereof, fused in

frame with a full length IL-15Rα, functional fragment or variant thereof. In some cases, the IL-15 is indirectly linked to the IL-15Rα through a linker. In some instances, the mbIL-15 is as described in Hurton et al., "Tethered IL-15 augments antitumor activity and promotes a stem-cell memory subset in tumor-specific T cells," PNAS 2016.

[0177] In certain embodiments, HPV vaccine antigens provided herein are co-delivered and/or co-expressed (*e.g.*, as part of the same HPV antigen delivery vector or via a separate vector) along with tumor necrosis factors. Tumor necrosis factors (TNFs) are a group of cytokines that modulate apoptosis. In some instances, there are about 19 members within the TNF family, including, not limited to, TNFα, lymphotoxin-alpha (LT-alpha), lymphotoxin-beta (LT-beta), T cell antigen gp39 (CD40L), CD27L, CD30L, FASL, 4-1BBL, OX40L, and TNF-related apoptosis inducing ligand (TRAIL).

[0178] In certain embodiments, HPV vaccine antigens provided herein are co-delivered and/or co-expressed (*e.g.*, as part of the same HPV antigen delivery vector or via a separate vector) along with colony stimulating factors. Colony-stimulating factors (CSFs) are secreted glycoproteins that interact with receptor proteins on the surface of hemopoietic stem cells, which subsequently modulates cell proliferation and differentiation into specific kind of blood cells. In some instances, a CSF comprises macrophage colony-stimulating factor, granulocyte macrophage colony-stimulating factor (GM-CSF), granulocyte colony-stimulating factor (G-CSF) or promegapoietin.

[0179] In some embodiments, the cytokine is a membrane-bound cytokine, which is co-expressed with a chimeric antigen receptor described herein. In some embodiments, one or more methods described herein further comprise administration of a cytokine. In some instances, the cytokine comprises a chemokine, an interferon, an interleukin, a colony-stimulating factor or a tumor necrosis factor. In some instances, one or more methods described herein further comprise administration of a cytokine selected from a chemokine, an interferon, an interleukin, a colony-stimulating factor or a tumor necrosis factor. In some instances, one or more methods described herein further comprise administration of a cytokine selected from IL2, IL7, IL12, IL15, a fusion of IL-15 and IL-15Rα, IL21, IFNγ or TNF-α.

Interleukin-12

[0180] In particular embodiments, HPV vaccine antigens provided herein are delivered to, and/or are expressed in a subject, in conjunction with delivery and/or expression of

Interleukin-12 cytokine (*e.g.*, IL-12 polypeptides are expressed from the same HPV vaccine antigen expression vector or are expressed from a separate vector in conjunction with HPV vaccine antigen delivery or expression). In certain embodiments, in conjunction with HPV vaccine antigen delivery or expression, expression of IL-12 in a subject is controlled by constitutive or inducible regulation of expression. In a preferred embodiment, in conjunction with HPV vaccine antigen delivery or expression, expression of IL-12 in a subject is controlled by inducible regulation of expression (*also referred to as,* inducibly regulated expression of IL-12). Interleukin 12 (IL-12) is an interleukin that is naturally produced by dendritic cells, macrophages, neutrophils, and human B-lymphoblastoid cells (NC-37) in response to antigenic stimulation. IL-12 is composed of a bundle of four alpha helices. It is a heterodimeric cytokine encoded by two separate genes, IL-12A (p35) and IL-12B (p40). The active heterodimer (referred to as p70), and a homodimer of p40 are formed following protein synthesis. IL-12 is the master regulator of the immune system. IL-12 promotes immune response by activating NK cells and T cells (**FIG. 11**).

[0181] Provided herein are compositions, kits, and system comprising and methods of making HPV recombinant vaccines. Also provided herein are polynucleotides encoding gene-switch polypeptides and IL-12 or variant or derivative thereof, and methods and systems incorporating the same.

Linkers

[0182] Also disclosed are constructs comprising a linker to facilitate the expression and functionality of the polynucleotides and polypeptides described herein. In some embodiments, a polynucleotide linker can be utilized in a polynucleotide described herein. A polynucleotide linker can be a double-stranded segment of DNA containing desired restriction sites that can be added to create end structures that are compatible with a vector comprising a polynucleotide described herein. In some cases, a polynucleotide linker can be useful for modifying vectors comprising polynucleotides described herein. For example, a vector modification comprising a polynucleotide linker can be a change in a multiple cloning site, or the addition of a poly-histidine tail. Polynucleotide linkers can also be used to adapt the ends of blunt insert DNA for cloning into a vector cleaved with a restriction enzyme with cohesive end termini. The use of polynucleotide linkers can be more efficient than a blunt ligation into a vector and can provide a method of releasing an

insert from a vector in downstream applications. In some cases an insert can be a polynucleotide sequence encoding polypeptides useful for therapeutic applications. In some cases, a linker can be a cleavable linker.

[0183] A polynucleotide linker can be an oligomer. A polynucleotide linker can be a DNA double strand, single strand, or a combination thereof. In some cases, a linker can be RNA. A polynucleotide linker can be ligated into a vector comprising a polynucleotide described herein by a T4 ligase in some cases. To facilitate a ligation an excess of polynucleotide linkers can be added to a composition comprising an insert and a vector. In some cases, an insert and vector are pre-treated before a linker is introduced. For example, pre-treatment with a methylase can prevent unwanted cleavage of insert DNA.

[0184] In certain embodiments, two or more polypeptides encoded by a polynucleotide described herein can be separated by an intervening sequence encoding an intervening linker polypeptide. Herein the term "intervening linker polypeptide" referring to an amino acid sequence separating two or more polypeptides encoded by a polynucleotide is distinguished from the term "peptide linker" which refers to the sequence of amino acids which is optionally included in a polypeptide construct disclosed herein to connect the transmembrane domain to the cell surface polypeptide (*e.g.*, comprising a truncated variant of a natural polypeptide). In certain cases, the intervening linker is a cleavage-susceptible intervening linker polypeptide. In some embodiments, the linker is a cleavable or ribosome skipping linker. In some embodiments, the cleavable linker or ribosome skipping linker sequence is selected from the group consisting of 2A, GSG-2A, GSG linker, SGSG linker, furinlink variants and derivatives thereof. In some embodiments, the 2A linker is a p2A linker, a T2A linker, F2A linker or E2A linker. In some embodiments, polypeptides of interest are expressed as fusion proteins linked by a cleavage-susceptible intervening linker polypeptide. In certain embodiments, cleavage-susceptible intervening linker polypeptide(s) can be any one or more of: F/T2A, T2A, p2A, 2A, GSG-p2A, GSG linker, and furinlink variants. Linkers (polynucleotide and polypeptide sequences) as disclosed in PCT/US2016/061668 (WO2017083750) published 18-May-2017 are incorporated by reference herein.

[0185] In certain embodiments, the linker polypeptide comprises disclosed in the table below:

Table 3: Linker amino acid sequences.

SEQ ID NO:	Linker Type	Amino Acid Sequence
34, 76	RIGID LINKER	EAAAK$_{(n)}$ (Where n = sequence repeating any multiple number of times such as 1, 2, 3, 4, 5 times, *etc*…)
40	HPV6 E6 Epitope/Linker	TAEIYSYAYKQLKVL
42	HPV11 E6 Epitope/Linker	TAEIYAYAYKNLKVV
45	HPV11 E7 Epitope/Linker	HCYEQLEDSSEDEVD
52	HPV16 E7 agonist	RTLEDLLMGV
48	HPV16 E6 agonist1	KLPQLCTEV
50	HPV16 E6 agonist2	QLYNKPLCDV
57	Flex linker1	VSQTSKLTR

[0186] In some cases, a viral 2A sequence can be used. 2A elements can be shorter than IRES, having from 5 to 100 base pairs. In some cases, a 2A sequence can have 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, or 100 nucleotides in length. 2A linked genes can be expressed in one single open reading frame and "self-cleavage" can occur co-translationally between the last two amino acids, GP, at the C-terminus of the 2A polypeptide, giving rise to equal amounts of co-expressed proteins.

[0187] A viral 2A sequence can be about 20 amino acids. In some cases, a viral 2A sequence can contain a consensus motif Asp-Val/Ile-Glu-X-Asn-Pro-Gly-Pro. A consensus motif sequence can act co-translationally. For example, formation of a normal peptide bond between a glycine and proline residue can be prevented, which can result in ribosomal skipping and cleavage of a nascent polypeptide. This effect can produce multiple genes at equimolar levels.

[0188] A 2A peptide can allow translation of multiple proteins in a single open reading frame into a polypeptide that can be subsequently cleaved into individual polypeptide through a ribosome-skipping mechanism (Funston et al., *J. Gen. Virol. 89(Pt 2)*:389-96 (2008)). In some embodiments, a 2A sequence can include: F/T2A, T2A, p2A, 2A, T2A, E2A, F2A, and BmCPV2A, BmIFV2A, and any combination thereof.

[0189] In some cases, a vector can comprise an IRES sequence and a 2A linker sequence. In other cases, expression of multiple genes linked with 2A peptides can be facilitated by a spacer sequence (GSG) ahead of the 2A peptides. In some cases, constructs can combine a spacers, linkers, adaptors, promoters, or combinations thereof. For example, a linker can have a spacer (SGSG or GSG or Whitlow linker) and furin linker (R-A-K-R) cleavage site with different 2A peptides. A spacer can be an I-Ceui. In some cases, a linker can be engineered. For example, a linker can be designed to comprise chemical characteristics such as hydrophobicity. In some cases, at least two linker sequences can produce the same protein. In other cases, multiple linkers can be used in a vector. For example, genes of interest can be separated by at least two linkers.

[0190] In certain embodiments, two or more polypeptides encoded by a polynucleotide described herein can be separated by an intervening sequence encoding a linker polypeptide. In certain cases, the linker is a cleavage-susceptible linker. In some embodiments, polypeptides of interest are expressed as fusion proteins linked by a cleavage-susceptible linker polypeptide. In certain embodiments, cleavage-susceptible linker polypeptide(s) can be any one or two of: Furinlink, fmdv, p2a, GSG-p2a, and/or fp2a described below. In some cases, a linker is APVKQGSG.

[0191] In certain cases, a linker polypeptide can comprise an amino acid sequence "RAKR". In certain cases, a Furin linker polypeptide can be encoded by a polynucleotide sequence polynucleotide sequence comprising "CGTGCAAAGCGT" or "AGAGCTAAGAGG".

[0192] In some embodiments, a linker can be utilized in a polynucleotide described herein. A linker can be a flexible linker, a rigid linker, an *in vivo* cleavable linker, or any combination thereof. In some cases, a linker can link functional domains together (as in flexible and rigid linkers) or releasing free functional domain *in vivo* as in *in vivo* cleavable linkers.

[0193] Linkers can improve biological activity, increase expression yield, and achieving desirable pharmacokinetic profiles. A linker can also comprise hydrazone, peptide, disulfide, or thioesther.

[0194] In some cases, a linker sequence described herein can include a flexible linker. Flexible linkers can be applied when a joined domain requires a certain degree of movement or interaction. Flexible linkers can be composed of small, non-polar (*e.g.*, Gly)

or polar (*e.g.*, Ser or Thr) amino acids. A flexible linker can have sequences consisting primarily of stretches of Gly and Ser residues ("GS" linker). An example of a flexible linker can have the sequence of (Gly-Gly-Gly-Gly-Ser)n . By adjusting the copy number "n", the length of this exemplary GS linker can be optimized to achieve appropriate separation of functional domains, or to maintain necessary inter-domain interactions. Besides GS linkers, other flexible linkers can be utilized for recombinant fusion proteins. In some cases, flexible linkers can also be rich in small or polar amino acids such as Gly and Ser, but can contain additional amino acids such as Thr and Ala to maintain flexibility. In other cases, polar amino acids such as Lys and Glu can be used to improve solubility.

[0195] Flexible linkers included in linker sequences described herein, can be rich in small or polar amino acids such as Gly and Ser to provide good flexibility and solubility. Flexible linkers can be suitable choices when certain movements or interactions are desired for fusion protein domains. In addition, although flexible linkers cannot have rigid structures, they can serve as a passive linker to keep a distance between functional domains. The length of flexible linkers can be adjusted to allow for proper folding or to achieve optimal biological activity of the fusion proteins.

[0196] A linker described herein can further include a rigid linker in some cases. A rigid linker can be utilized to maintain a fixed distance between domains of a polypeptide. Examples of rigid linkers can be: Alpha helix-forming linkers, Pro-rich sequence, (XP)n, X-Pro backbone, A(EAAAK)nA (n = 2-5), to name a few. Rigid linkers can exhibit relatively stiff structures by adopting α-helical structures or by containing multiple Pro residues in some cases.

[0197] A linker described herein can be cleavable in some cases. In other cases a linker is not cleavable. Linkers that are not cleavable can covalently join functional domains together to act as one molecule throughout an *in vivo* processes or an *ex vivo* process. A linker can also be cleavable *in vivo*. A cleavable linker can be introduced to release free functional domains *in vivo*. A cleavable linker can be cleaved by the presence of reducing reagents, proteases, to name a few. For example, a reduction of a disulfide bond can be utilized to produce a cleavable linker. In the case of a disulfide linker, a cleavage event through disulfide exchange with a thiol, such as glutathione, could produce a cleavage. In other cases, an *in vivo* cleavage of a linker in a recombinant fusion protein

can also be carried out by proteases that can be expressed *in vivo* under pathological conditions (*e.g.* cancer or inflammation), in specific cells or tissues, or constrained within certain cellular compartments. In some cases, a cleavable linker can allow for targeted cleavage. For example, the specificity of many proteases can offer slower cleavage of a linker in constrained compartments. A cleavable linker can also comprise hydrazone, peptides, disulfide, or thioesther. For example, a hydrazone can confer serum stability. In other cases, a hydrazone can allow for cleavage in an acidic compartment. An acidic compartment can have a pH up to 7. A linker can also include a thioether. A thioether can be nonreducible A thioether can be designed for intracellular proteolytic degradation.

[0198] In some cases, genes can be separated by 2, 3, 4, 5, 6, 7, 8, 9, or up to 10 linkers.

[0199] A linker can be an engineered linker. Methods of designing linkers can be computational. In some cases, computational methods can include graphic techniques. Computation methods can be used to search for suitable peptides from libraries of three-dimensional peptide structures derived from databases. For example, a Brookhaven Protein Data Bank (PDB) can be used to span the distance in space between selected amino acids of a linker.

[0200] In some embodiments are polynucleotides encoding a polypeptide construct comprising a furin polypeptide and a 2A polypeptide, wherein the furin polypeptide and the 2A polypeptide are connected by a polypeptide linker comprising at least three hydrophobic amino acids. In some cases, at least three hydrophobic amino acids are selected from the list consisting of glycine (Gly)(G), alanine (Ala)(A), valine (Val)(V), leucine (Leu)(L), isoleucine (Ile)(I), proline (Pro)(P), phenylalanine (Phe)(F), methionine (Met)(M), tryptophan (Trp)(W). In some cases, a polypeptide linker can also include one or more GS linker sequences, for instance (GS)n, (SG)n, (GSG)n, and (SGSG)n, wherein n can be any number from zero to fifteen.

[0201] Provided are methods of obtaining an improved expression of a polypeptide construct comprising: providing a polynucleotide encoding said polypeptide construct comprising a first functional polypeptide and a second functional polypeptide, wherein said first functional polypeptide and second functional polypeptide are connected by a linker polypeptide comprising a sequence with at least 60% identity to the sequence APVKQ; and expressing said polynucleotide in a host cell, wherein said expressing results in an improved expression of the polypeptide construct as compared to a

corresponding polypeptide construct that does not have a linker polypeptide comprising a sequence with at least 60% identity to the sequence APVKQ.

IRES Elements

[0202] Also disclosed herein are constructs comprising an IRES element to facilitate the expression and functionality of the polynucleotides and polypeptides described herein. The term "internal ribosome entry site (IRES)" as used herein can be intended to mean internal ribosomal entry site. In a vector comprising an IRES sequence, a first gene can be translated by a cap-dependent, ribosome scanning, mechanism with its own 5′-UTR, whereas translation of a subsequent gene can be accomplished by direct recruitment of a ribosome to an IRES in a cap-independent manner. An IRES sequence can allow eukaryotic ribosomes to bind and begin translation without binding to a 5′ capped end. An IRES sequence can allow expression of multiple genes from one transcript (Mountford and Smith, *Trends Genet. 11(5)*:179-84 (1995)).

[0203] The term "CAP" or "cap" as used herein refers to a modified nucleotide, generally a 7-methyl guanosine, linked 3′ to 5′ (7meG-ppp-G), to the 5′ end of a eukaryotic mRNA, that serves as a required element in the normal translation initiation pathway during expression of protein from that mRNA.

[0204] In certain cases, an IRES region can be derived from a virus, such as picornavirus, encephalomyocarditis virus, hepatitis C virus IRES sequence. In other cases, an IRES sequence can be derived from an encephalomyocarditis virus. The term "EMCV" or "encephalomyocarditis virus" as used herein refers to any member isolate or strain of the encephalomyocarditis virus species of the genus of the family Picornaviridae. Examples are: EMCV-R (Rueckert) strain virus, Columbia-SK virus. In some cases, a cellular IRES element, such as eukaryotic initiation factor 4G, immunoglobulin heavy chain binding protein, c-myc proto-oncogene, vascular endothelial growth factor, fibroblast growth factor-1 IRES, or any combination or modification thereof can be used. In some cases, a cellular IRES can have increased gene expression when compared to a viral IRES.

[0205] An IRES sequence of viral, cellular or a combination thereof can be utilized in a vector. An IRES can be from encephalomyocarditis (EMCV) or poliovirus (PV). In some cases, an IRES element is selected from a group consisting of Poliovirus (PV), Encephalomyelitis virus (EMCV), Foot-and-mouth disease virus (FMDV), Porcine teschovirus-1 (PTV-1), Aichivirus (AiV), Seneca Valley virus (SVV), Hepatitis C virus

(HCV), Classical swine fever virus (CSFV), Human immunodeficiency virus-2 (HIV-2), Human immunodeficiency virus-1 (HIV-1), Moloney murine leukemia virus (MoMLV), Feline immunodeficiency virus (FIV), Mouse mammary tumor virus (MMTV), Human cytomegalovirus latency (pUL138), Epstein-Barr virus (EBNA-1), Herpes virus Marek's disease (MDV RLORF9), SV40 polycistronic 19S (SV40 19S), Rhopalosiphum padi virus (RhPV), Cricket paralysis virus (CrPV), Ectropis obliqua picorna-like virus (EoPV), Plautia stali intestine virus (PSIV), Triatoma virus (TrV), Bee paralysis dicistrovirus (IAPV, KBV), Black currant reversion virus (BRV), Pelargonium flower break virus (PFBV), Hibiscus chlorotic ringspot virus (HCRSV), Crucifer-infecting tobamovirus (CrTMV), Potato leaf roll polerovirus (PLRV), Tobacco etch virus (TEV), Giardiavirus (GLV), Leishmania RNA virus-1 (LRV-1), and combinations or modifications thereof. In some cases, an IRES is selected from a group consisting of Apaf-1, XIAP, HIAP2/c-IAP1, DAP5, Bcl-2, c-myc, CAT-1, INR, Differentiation LEF-1, PDGF2, HIF-1a, VEGF, FGF2, BiP, BAG-1, CIRP, p53, SHMT1, PITSLREp58, CDK1, Rpr, hid, hsp70, grim, skl, Antennapedia, dFoxO, dInR, Adh-Adhr, HSP101, ADH, URE-2,GPR1, NCE102, YMR181a, MSN1, BOI1, FLO8, GIC1, and any combination or modification thereof. When an IRES element is included between two open reading frames (ORFs), initiation of translation can occur by a canonical 5′- m7GpppN cap-dependent mechanism in a first ORF and a cap-independent mechanism in a second ORF downstream of the IRES element.

[0206] In some cases, genes can be linked by an internal ribosomal entry site (IRES). An IRES can allow simultaneous expression of multiple genes. For example, an IRES sequence can permit production of multiple proteins from a single mRNA transcript. A ribosome can bind to an IRES in a 5′-cap independent manner and initiate translation.

[0207] In some cases, an IRES sequence can be or can be about 500 base pairs. An IRES sequence can be from 300 base pairs to 1000 base pairs. For example, an IRES can be 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, or 1000 base pairs long.

[0208] In some cases, expression of a downstream gene within a vector comprising an IRES sequence can be reduced. For example, a gene following an IRES sequence can have reduced expression over a gene preceding an IRES sequence. Reduced expression can be from 1% to 99% reduction over a preceding gene.

Methods of Regulating Expression

[0209] In one embodiment, a method of regulating the expression of a heterologous gene in an engineered cell is provided. Polynucleotides encoding for gene switch polypeptides for ligand inducible control of a heterologous gene expression, an antigen binding polypeptide and a heterologous gene is provided. In some instances, the polynucleotides are in one or more gene expression cassettes as depicted in any one of Figures 1 through 16. In another instance, the polynucleotides are incorporated into an engineered cell via viral or non-viral vectors. Viral vectors can include lentiviral vectors, retroviral vectors or adenoviral vectors. Non-viral vectors can include Sleeping Beauty transposons. In other instances, the polynucleotides are incorporated into an engineered cell via recombinases or gene editing techniques. Examples of recombinases are serine recombinases as described herein. Examples of gene editing techniques can include CRISPR or Argonaute systems. Herein a "CRISPR gene editing system" of "CRISPR system" refers to any RNA-guided Cas protein-mediated process for targeting a change in DNA sequence to a specific region of a genome. Herein "Argonaute gene editing system" refers to any single-stranded DNA guided Argonaute endonuclease-mediated process for targeting a change in DNA sequence to a specific region of a genome.

Pharmaceutical Compostiions and Dosage

[0210] The present disclosure provides a composition comprising the adenovirus or adenoviral vector described herein and a carrier therefor (*e.g.*, a pharmaceutically acceptable carrier). The composition desirably is a physiologically acceptable (*e.g.*, pharmaceutically acceptable) composition, which comprises a carrier, preferably a physiologically (*e.g.*, pharmaceutically) acceptable carrier, and the adenovirus or adenoviral vector. Any suitable carrier can be used within the context of the present disclosure, and such carriers are well known in the art. The choice of carrier will be determined, in part, by the particular use of the composition (*e.g.*, administration to an animal) and the particular method used to administer the composition. Ideally, in the context of replication-deficient adenoviral vectors, the pharmaceutical composition preferably is free of replication-competent adenovirus. The pharmaceutical composition optionally can be sterile.

[0211] Suitable compositions include aqueous and non-aqueous isotonic sterile solutions, which can contain anti-oxidants, buffers, and bacteriostats, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The composition can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials, and can be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example, water, immediately prior to use. Extemporaneous solutions and suspensions can be prepared from sterile powders, granules, and tablets. Preferably, the carrier is a buffered saline solution. More preferably, the adenovirus or adenoviral vector is part of a composition formulated to protect the adenovirus or adenoviral vector from damage prior to administration. For example, the composition can be formulated to reduce loss of the adenovirus or adenoviral vector on devices used to prepare, store, or administer the adenovirus or adenoviral vector, such as glassware, syringes, or needles. The composition can be formulated to decrease the light sensitivity and/or temperature sensitivity of the adenovirus or adenoviral vector. To this end, the composition preferably comprises a pharmaceutically acceptable liquid carrier, such as, for example, those described above, and a stabilizing agent selected from the group consisting of polysorbate 80, L-arginine, polyvinylpyrrolidone, trehalose, and combinations thereof. Use of such a composition will extend the shelf life of the adenovirus or adenoviral vector, and facilitate its administration. Formulations for adenovirus or adenoviral vector-containing compositions are further described in, for example, U.S. Patent 6,225,289, U.S. Patent 6,514,943, and International Patent Application Publication WO 2000/034444.

[0212] The composition also can be formulated to enhance transduction efficiency. In addition, one of ordinary skill in the art will appreciate that the adenovirus or adenoviral vector can be present in a composition with other therapeutic or biologically-active agents. For example, factors that control inflammation, such as ibuprofen or steroids, can be part of the composition to reduce swelling and inflammation associated with *in vivo* administration of the adenovirus or adenoviral vector. If the adenovirus or adenoviral vector is used to deliver an antigen-encoding nucleic acid sequence to a host, immune system stimulators or adjuvants, *e.g.*, interleukins, lipopolysaccharide, or double-stranded RNA, can be administered to enhance or modify any immune response to the antigen.

Antibiotics, *i.e.*, microbicides and fungicides, can be present to treat existing infection and/or reduce the risk of future infection, such as infection associated with gene transfer procedures.

[0213] In some embodiments, disclosed herein are compositions comprising a polynucleotide or polypeptide disclosed herein for administration in a subject. In some instances, are modified effector cell compositions encoding a polynucleotide or polypeptide disclosed herein, and optionally containing a cytokine and/or an additional therapeutic agent. In some instances, also included herein are vectors encoding gene-switch polypeptides for regulating expression of a chimeric antigen receptor for modification of an effector cell.

[0214] In some instances, pharmaceutical compositions of a modified effector cell or a vector encoding gene-switch polypeptides and a chimeric antigen receptor are formulated in a conventional manner using one or more physiologically acceptable carriers including excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen. A summary of pharmaceutical compositions described herein is found, for example, in Remington: The Science and Practice of Pharmacy, Nineteenth Ed (Easton, Pa.: Mack Publishing Company, 1995); Hoover, John E., Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pennsylvania 1975; Liberman, H.A. and Lachman, L., Eds., Pharmaceutical Dosage Forms, Marcel Decker, New York, N.Y., 1980; and Pharmaceutical Dosage Forms and Drug Delivery Systems, Seventh Ed. (Lippincott Williams & Wilkins1999).

[0215] Pharmaceutical compositions are optionally manufactured in a conventional manner, such as, by way of example only, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or compression processes.

[0216] In certain embodiments, compositions can also include one or more pH adjusting agents or buffering agents, including acids such as acetic, boric, citric, lactic, phosphoric and hydrochloric acids; bases such as sodium hydroxide, sodium phosphate, sodium borate, sodium citrate, sodium acetate, sodium lactate and tris-hydroxymethylaminomethane; and buffers such as citrate/dextrose, sodium bicarbonate

and ammonium chloride. Such acids, bases and buffers are included in an amount required to maintain pH of the composition in an acceptable range.

[0217] In other embodiments, compositions can also include one or more salts in an amount required to bring osmolality of the composition into an acceptable range. Such salts include those having sodium, potassium or ammonium cations and chloride, citrate, ascorbate, borate, phosphate, bicarbonate, sulfate, thiosulfate or bisulfite anions; suitable salts include sodium chloride, potassium chloride, sodium thiosulfate, sodium bisulfite and ammonium sulfate.

[0218] The pharmaceutical compositions described herein are administered by any suitable administration route, including but not limited to, oral, parenteral (*e.g.*, intravenous, subcutaneous, intramuscular, intracerebral, intracerebroventricular, intra-articular, intraperitoneal, or intracranial), intranasal, buccal, sublingual, or rectal administration routes. In some instances, the pharmaceutical composition is formulated for parenteral (*e.g.*, intravenous, subcutaneous, intramuscular, intracerebral, intracerebroventricular, intra-articular, intraperitoneal, or intracranial) administration.

[0219] The pharmaceutical compositions described herein are formulated into any suitable dosage form, including but not limited to, aqueous oral dispersions, liquids, gels, syrups, elixirs, slurries, suspensions and the like, for oral ingestion by an individual to be treated, solid oral dosage forms, aerosols, controlled release formulations, fast melt formulations, effervescent formulations, lyophilized formulations, tablets, powders, pills, dragees, capsules, delayed release formulations, extended release formulations, pulsatile release formulations, multiparticulate formulations, and mixed immediate release and controlled release formulations. In some embodiments, the pharmaceutical compositions are formulated into capsules. In some embodiments, the pharmaceutical compositions are formulated into solutions (for example, for IV administration). In some cases, the pharmaceutical composition is formulated as an infusion. In some cases, the pharmaceutical composition is formulated as an injection.

[0220] The pharmaceutical solid dosage forms described herein optionally include a compound described herein and one or more pharmaceutically acceptable additives such as a compatible carrier, binder, filling agent, suspending agent, flavoring agent, sweetening agent, disintegrating agent, dispersing agent, surfactant, lubricant, colorant,

diluent, solubilizer, moistening agent, plasticizer, stabilizer, penetration enhancer, wetting agent, anti-foaming agent, antioxidant, preservative, or one or more combination thereof.

[0221] In still other aspects, using standard coating procedures, such as those described in Remington's Pharmaceutical Sciences, 20th Edition (2000), a film coating is provided around the compositions. In some embodiments, the compositions are formulated into particles (for example for administration by capsule) and some or all of the particles are coated. In some embodiments, the compositions are formulated into particles (for example for administration by capsule) and some or all of the particles are microencapsulated. In some embodiments, the compositions are formulated into particles (for example for administration by capsule) and some or all of the particles are not microencapsulated and are uncoated.

[0222] In certain embodiments, compositions provided herein can also include one or more preservatives to inhibit microbial activity. Suitable preservatives include mercury-containing substances such as merfen and thiomersal; stabilized chlorine dioxide; and quaternary ammonium compounds such as benzalkonium chloride, cetyltrimethylammonium bromide and cetylpyridinium chloride.

[0223] "Antifoaming agents" reduce foaming during processing which can result in coagulation of aqueous dispersions, bubbles in the finished film, or generally impair processing. Exemplary anti-foaming agents include silicon emulsions or sorbitan sesquoleate.

[0224] "Antioxidants" include, for example, butylated hydroxytoluene (BHT), sodium ascorbate, ascorbic acid, sodium metabisulfite and tocopherol. In certain embodiments, antioxidants enhance chemical stability where required.

[0225] Formulations described herein can benefit from antioxidants, metal chelating agents, thiol containing compounds and other general stabilizing agents. Examples of such stabilizing agents, include, but are not limited to: (a) about 0.5% to about 2% w/v glycerol, (b) about 0.1% to about 1% w/v methionine, (c) about 0.1% to about 2% w/v monothioglycerol, (d) about 1 mM to about 10 mM EDTA, (e) about 0.01% to about 2% w/v ascorbic acid, (f) 0.003% to about 0.02% w/v polysorbate 80, (g) 0.001% to about 0.05% w/v. polysorbate 20, (h) arginine, (i) heparin, (j) dextran sulfate, (k) cyclodextrins, (l) pentosan polysulfate and other heparinoids, (m) divalent cations such as magnesium and zinc; or (n) combinations thereof.

[0226] "Binders" impart cohesive qualities and include, *e.g.*, alginic acid and salts thereof; cellulose derivatives such as carboxymethylcellulose, methylcellulose (*e.g.*, Methocel®), hydroxypropylmethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose (*e.g.*, Klucel®), ethylcellulose (*e.g.*, Ethocel®), and microcrystalline cellulose (*e.g.*, Avicel®); microcrystalline dextrose; amylose; magnesium aluminum silicate; polysaccharide acids; bentonites; gelatin; polyvinylpyrrolidone/vinyl acetate copolymer; crospovidone; povidone; starch; pregelatinized starch; tragacanth, dextrin, a sugar, such as sucrose (*e.g.*, Dipac®), glucose, dextrose, molasses, mannitol, sorbitol, xylitol (*e.g.*, Xylitab®), and lactose; a natural or synthetic gum such as acacia, tragacanth, ghatti gum, mucilage of isapol husks, polyvinylpyrrolidone (*e.g.*, Polyvidone® CL, Kollidon® CL, Polyplasdone® XL-10), larch arabogalactan, Veegum®, polyethylene glycol, waxes, sodium alginate, and the like.

[0227] A "carrier" or "carrier materials" include any commonly used excipients in pharmaceutics and should be selected on the basis of compatibility with compounds disclosed herein, such as, compounds of ibrutinib and An anticancer agent, and the release profile properties of the desired dosage form. Exemplary carrier materials include, *e.g.*, binders, suspending agents, disintegration agents, filling agents, surfactants, solubilizers, stabilizers, lubricants, wetting agents, diluents, and the like. "Pharmaceutically compatible carrier materials" can include, but are not limited to, acacia, gelatin, colloidal silicon dioxide, calcium glycerophosphate, calcium lactate, maltodextrin, glycerine, magnesium silicate, polyvinylpyrrollidone (PVP), cholesterol, cholesterol esters, sodium caseinate, soy lecithin, taurocholic acid, phosphotidylcholine, sodium chloride, tricalcium phosphate, dipotassium phosphate, cellulose and cellulose conjugates, sugars sodium stearoyl lactylate, carrageenan, monoglyceride, diglyceride, pregelatinized starch, and the like. See, *e.g.*, Remington: The Science and Practice of Pharmacy, Nineteenth Ed (Easton, Pa.: Mack Publishing Company, 1995); Hoover, John E., Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pennsylvania 1975; Liberman, H.A. and Lachman, L., Eds., Pharmaceutical Dosage Forms, Marcel Decker, New York, N.Y., 1980; and Pharmaceutical Dosage Forms and Drug Delivery Systems, Seventh Ed. (Lippincott Williams & Wilkins1999).

[0228] "Dispersing agents," and/or "viscosity modulating agents" include materials that control the diffusion and homogeneity of a drug through liquid media or a granulation method or blend method. In some embodiments, these agents also facilitate the effectiveness of a coating or eroding matrix. Exemplary diffusion facilitators/dispersing agents include, *e.g.*, hydrophilic polymers, electrolytes, Tween ® 60 or 80, PEG, polyvinylpyrrolidone (PVP; commercially known as Plasdone®), and the carbohydrate-based dispersing agents such as, for example, hydroxypropyl celluloses (*e.g.*, HPC, HPC-SL, and HPC-L), hydroxypropyl methylcelluloses (*e.g.*, HPMC K100, HPMC K4M, HPMC K15M, and HPMC K100M), carboxymethylcellulose sodium, methylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose phthalate, hydroxypropylmethylcellulose acetate stearate (HPMCAS), noncrystalline cellulose, magnesium aluminum silicate, triethanolamine, polyvinyl alcohol (PVA), vinyl pyrrolidone/vinyl acetate copolymer (S630), 4-(1,1,3,3-tetramethylbutyl)-phenol polymer with ethylene oxide and formaldehyde (also known as tyloxapol), poloxamers (*e.g.*, Pluronics F68®, F88®, and F108®, which are block copolymers of ethylene oxide and propylene oxide); and poloxamines (*e.g.*, Tetronic 908®, also known as Poloxamine 908®, which is a tetrafunctional block copolymer derived from sequential addition of propylene oxide and ethylene oxide to ethylenediamine (BASF Corporation, Parsippany, N.J.)), polyvinylpyrrolidone K12, polyvinylpyrrolidone K17, polyvinylpyrrolidone K25, or polyvinylpyrrolidone K30, polyvinylpyrrolidone/vinyl acetate copolymer (S-630), polyethylene glycol, *e.g.*, the polyethylene glycol can have a molecular weight of about 300 to about 6000, or about 3350 to about 4000, or about 7000 to about 5400, sodium carboxymethylcellulose, methylcellulose, polysorbate-80, sodium alginate, gums, such as, *e.g.*, gum tragacanth and gum acacia, guar gum, xanthans, including xanthan gum, sugars, cellulosics, such as, *e.g.*, sodium carboxymethylcellulose, methylcellulose, sodium carboxymethylcellulose, polysorbate-80, sodium alginate, polyethoxylated sorbitan monolaurate, polyethoxylated sorbitan monolaurate, povidone, carbomers, polyvinyl alcohol (PVA), alginates, chitosans and combinations thereof. Plasticizers such as cellulose or triethyl cellulose can also be used as dispersing agents. Dispersing agents particularly useful in liposomal dispersions and self-emulsifying dispersions are dimyristoyl phosphatidyl choline, natural phosphatidyl choline from eggs, natural phosphatidyl glycerol from eggs, cholesterol and isopropyl myristate.

[0229] Combinations of one or more erosion facilitator with one or more diffusion facilitator can also be used in the present compositions.

[0230] The term "diluent" refers to chemical compounds that are used to dilute the compound of interest prior to delivery. Diluents can also be used to stabilize compounds because they can provide a more stable environment. Salts dissolved in buffered solutions (which also can provide pH control or maintenance) are utilized as diluents in the art, including, but not limited to a phosphate buffered saline solution. In certain embodiments, diluents increase bulk of the composition to facilitate compression or create sufficient bulk for homogenous blend for capsule filling. Such compounds include *e.g.*, lactose, starch, mannitol, sorbitol, dextrose, microcrystalline cellulose such as Avicel®; dibasic calcium phosphate, dicalcium phosphate dihydrate; tricalcium phosphate, calcium phosphate; anhydrous lactose, spray-dried lactose; pregelatinized starch, compressible sugar, such as Di-Pac® (Amstar); mannitol, hydroxypropylmethylcellulose, hydroxypropylmethylcellulose acetate stearate, sucrose-based diluents, confectioner's sugar; monobasic calcium sulfate monohydrate, calcium sulfate dihydrate; calcium lactate trihydrate, dextrates; hydrolyzed cereal solids, amylose; powdered cellulose, calcium carbonate; glycine, kaolin; mannitol, sodium chloride; inositol, bentonite, and the like.

[0231] "Filling agents" include compounds such as lactose, calcium carbonate, calcium phosphate, dibasic calcium phosphate, calcium sulfate, microcrystalline cellulose, cellulose powder, dextrose, dextrates, dextran, starches, pregelatinized starch, sucrose, xylitol, lactitol, mannitol, sorbitol, sodium chloride, polyethylene glycol, and the like.

[0232] "Lubricants" and "glidants" are compounds that prevent, reduce or inhibit adhesion or friction of materials. Exemplary lubricants include, *e.g.*, stearic acid, calcium hydroxide, talc, sodium stearyl fumerate, a hydrocarbon such as mineral oil, or hydrogenated vegetable oil such as hydrogenated soybean oil (Sterotex®), higher fatty acids and their alkali-metal and alkaline earth metal salts, such as aluminum, calcium, magnesium, zinc, stearic acid, sodium stearates, glycerol, talc, waxes, Stearowet®, boric acid, sodium benzoate, sodium acetate, sodium chloride, leucine, a polyethylene glycol (*e.g.*, PEG-4000) or a methoxypolyethylene glycol such as Carbowax™, sodium oleate, sodium benzoate, glyceryl behenate, polyethylene glycol, magnesium or sodium lauryl sulfate, colloidal silica such as Syloid™, Cab-O-Sil®, a starch such as corn starch, silicone oil, a surfactant, and the like.

[0233] "Plasticizers" are compounds used to soften the microencapsulation material or film coatings to make them less brittle. Suitable plasticizers include, *e.g.*, polyethylene glycols such as PEG 300, PEG 400, PEG 600, PEG 1450, PEG 3350, and PEG 800, stearic acid, propylene glycol, oleic acid, triethyl cellulose and triacetin. In some embodiments, plasticizers can also function as dispersing agents or wetting agents.

[0234] "Solubilizers" include compounds such as triacetin, triethylcitrate, ethyl oleate, ethyl caprylate, sodium lauryl sulfate, sodium doccusate, vitamin E TPGS, dimethylacetamide, N-methylpyrrolidone, N-hydroxyethylpyrrolidone, polyvinylpyrrolidone, hydroxypropylmethyl cellulose, hydroxypropyl cyclodextrins, ethanol, n-butanol, isopropyl alcohol, cholesterol, bile salts, polyethylene glycol 200-600, glycofurol, transcutol, propylene glycol, and dimethyl isosorbide and the like.

[0235] "Stabilizers" include compounds such as any antioxidation agents, buffers, acids, preservatives and the like.

[0236] "Suspending agents" include compounds such as polyvinylpyrrolidone, *e.g.*, polyvinylpyrrolidone K12, polyvinylpyrrolidone K17, polyvinylpyrrolidone K25, or polyvinylpyrrolidone K30, vinyl pyrrolidone/vinyl acetate copolymer (S630), polyethylene glycol, *e.g.*, the polyethylene glycol can have a molecular weight of about 300 to about 6000, or about 3350 to about 4000, or about 7000 to about 5400, sodium carboxymethylcellulose, methylcellulose, hydroxypropylmethylcellulose, hydroxymethylcellulose acetate stearate, polysorbate-80, hydroxyethylcellulose, sodium alginate, gums, such as, *e.g.*, gum tragacanth and gum acacia, guar gum, xanthans, including xanthan gum, sugars, cellulosics, such as, *e.g.*, sodium carboxymethylcellulose, methylcellulose, sodium carboxymethylcellulose, hydroxypropylmethylcellulose, hydroxyethylcellulose, polysorbate-80, sodium alginate, polyethoxylated sorbitan monolaurate, polyethoxylated sorbitan monolaurate, povidone and the like.

[0237] "Surfactants" include compounds such as sodium lauryl sulfate, sodium docusate, Tween 60 or 80, triacetin, vitamin E TPGS, sorbitan monooleate, polyoxyethylene sorbitan monooleate, polysorbates, polaxomers, bile salts, glyceryl monostearate, copolymers of ethylene oxide and propylene oxide, *e.g.*, Pluronic® (BASF), and the like. Some other surfactants include polyoxyethylene fatty acid glycerides and vegetable oils, *e.g.*, polyoxyethylene (60) hydrogenated castor oil; and polyoxyethylene alkylethers and

alkylphenyl ethers, *e.g.*, octoxynol 10, octoxynol 40. In some embodiments, surfactants can be included to enhance physical stability or for other purposes.

[0238] "Viscosity enhancing agents" include, *e.g.*, methyl cellulose, xanthan gum, carboxymethyl cellulose, hydroxypropyl cellulose, hydroxypropylmethyl cellulose, hydroxypropylmethyl cellulose acetate stearate, hydroxypropylmethyl cellulose phthalate, carbomer, polyvinyl alcohol, alginates, acacia, chitosans and combinations thereof.

[0239] "Wetting agents" include compounds such as oleic acid, glyceryl monostearate, sorbitan monooleate, sorbitan monolaurate, triethanolamine oleate, polyoxyethylene sorbitan monooleate, polyoxyethylene sorbitan monolaurate, sodium docusate, sodium oleate, sodium lauryl sulfate, sodium doccusate, triacetin, Tween 80, vitamin E TPGS, ammonium salts and the like.

Kits/Article of Manufacture

[0240] Disclosed herein, in certain embodiments, are kits and articles of manufacture for use with one or more methods described herein. Such kits include a carrier, package, or container that is compartmentalized to receive one or more containers such as vials, tubes, and the like, each of the container(s) comprising one of the separate elements to be used in a method described herein. Suitable containers include, for example, bottles, vials, syringes, and test tubes. In one embodiment, the containers are formed from a variety of materials such as glass or plastic.

[0241] The articles of manufacture provided herein contain packaging materials. Examples of pharmaceutical packaging materials include, but are not limited to, blister packs, bottles, tubes, bags, containers, bottles, and any packaging material suitable for a selected formulation and intended mode of administration and treatment.

[0242] A kit typically includes labels listing contents and/or instructions for use, and package inserts with instructions for use. A set of instructions will also typically be included.

[0243] In some embodiments, a label is on or associated with the container. In one embodiment, a label is on a container when letters, numbers or other characters forming the label are attached, molded or etched into the container itself; a label is associated with a container when it is present within a receptacle or carrier that also holds the container, *e.g.*, as a package insert. In one embodiment, a label is used to indicate that the contents

are to be used for a specific therapeutic application. The label also indicates directions for use of the contents, such as in the methods described herein.

Antigenicity Bioinformatics Workflow for HPV Vaccine Designs

Identification of new HPV antigenic components

[0244] Based on the important role that HPV E2 and E4 genetic components play in HPV essential functions, the location of the corresponding proteins, as well as *in-silico* prediction, E2- and E4-derived antigens were identified for an HPV therapeutic vaccine. *See*, FIGURE 1. Non-oncogenic and viral inactivation genetic modifications were also applied to eliminate viral and oncogenic biological activity from HPV proteins, such as in HPV E2 and E6 proteins. Genetic manipulation was also applied to achieve production of protein sequences reordered in such manner as to retain immunogenic features of the peptides, but eliminate oncogenic and viral amplification functions of E7 and E4, respectively.

[0245] As such, innovative aspects of the five designs exemplified in this specification include: (1) use of gene constructs encoding fusion proteins comprising four or more different HPV proteins; (2) combining amino acid point mutations and overlapping polypeptide sequence shuffling techniques to inactivate oncogenic and essential viral functions; (3) incorporation of HPV proteins comprising multiple antigenic components from HPV proteins which are highly expressed in host infected cells; (4) first known hybrid antigen designs; (5) combining epitopes from high cancer risk and low cancer risk HPV strains; (6) use of mixed and regularly repeating linkers; (7) use of rigid linkers to stabilize polypeptide subunits and prevent undesirable intra-molecular interactions; (8) use of cleavable linkers between epitopes; (9) dual use of linker sequence to provide both protein-protein linker (-) function as well as antigens and epitopes, *per se* (*i.e.*, antigenicity conferred by the linker sequences).

Table 4. *In-silico* Evaluation

Design No.	Antigenicity		Allergenicity	Host Cross-Reactivity
	Virus	Tumor		
1	0.5288	0.6134	Non-allergenic	No hits
2	0.5252	0.5814	Non-allergenic	No hits

3	0.5587	0.6322	Non-allergenic	No hits
4	0.5451	0.5809	Non-allergenic	No hits
5	0.4964	0.5997	Non-allergenic	No hits

[0246] Antigenicity is the capacity to stimulate the production of antibodies or cell-mediated immune responses. The antigenicity of the final design sequences was predicted by the Vaxjen software which is alignment-independent models for antigen recognition based on main chemical properties of amino acid sequences. The results indicate that the five antigen sequences are antigenic. *See*, Table 4 (Antigenicity/Virus & Tumor).

[0247] Allergens are small antigens that commonly provoke an antibody response. Allergenicity, whether the antigen is an allergen or non-allergen was predicted by ALLERTOP, a bioinformatics-based allergen prediction software with machine learning methods for classification. It includes logistic regression (LR), decision tree (DT), naïve Bayes (NB), random forest (RF), multilayer perceptron (MLP) and k nearest neighbors (kNN). The results indicate that the five antigen sequences are non-allergeneic. *See*, Table 4 (Allergenicity).

[0248] Cross-reactivity or invocation of autoimmune side effects in various tissues has important safety implications in adoptive immunotherapy. Sequence homology analyses were performed to assess if those novel antigens have cross-reactivity with human proteome with blast search, basic local alignment search tool. No host cross reactivity was identified in these five antigen sequences. *See*, Table 4 (Host Cross-Reactivity).

Software/Tools:

[0249] Software tools utilized in performance of the designs described herein include:

ALLERTOP (*See*, *AllerTOP v.2--a server for in silico prediction of allergens*; J Mol Model. 2014 Jun;20(6):2278. doi: 10.1007/s00894-014-2278-5. Epub 2014 May 31.)

ANN (*See*, *Reliable prediction of T-cell epitopes using neural networks with novel sequence representations*; Protein Sci. 2003 May;12(5):1007-17.)

BLAST (Basic Local Alignment Search Tool; NCBI, National Center for Biotechnology Information, U.S. National Library of Medicine 8600 Rockville Pike, Bethesda MD, 20894 USA.)

CLUSTALW2 (EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK. +44 (0)1223 49 44 44.)

GENEIOUS V11.1.5 (*See*, geneious.com/biopharma/)

IEDB CONSENSUS (*See*, *A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus*; Nat Biotechnol. 2006 Jul;24(7):817-9. Epub 2006 Jun 11..)

NETMHCPAN 4.0 (*See*, *Gapped sequence alignment using artificial neural networks: application to the MHC class I system*; Bioinformatics. 2016 Feb 15;32(4):511-7. doi: 10.1093/bioinformatics/btv639. Epub 2015 Oct 29.)

PHYRE2 (*See*, The Phyre2 web portal for protein modeling, prediction and analysis; 07 May 2015; nature.com/articles/nprot.2015.053; doi.org/10.1038/nprot.2015.053.)

PYMOL MOLECULAR GRAPHICS SYSTEM V2.1.1 (*See*, pymol.org/2/; sourceforge.net/projects/pymol/support)

VAXJEN (*See*, *VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines*; BMC Bioinformatics. 2007 Jan 5;8:4..)

Table 5. Nomenclature for Antigen Designs and Component Parts

Assigned Component Symbol	**SEQ ID NO** & Sequence Type	**Descriptive name**	**Sequence type**
A	1 AA	HPV6 E2*	Amino acid
	2 DNA	HPV6 E2*	Nucleic acid
B	3 AA	HPV6 E4 1-56	Amino acid
	4 DNA	HPV6 E4 1-56	Nucleic acid
C	5 AA	HPV6 E7 70-98	Amino acid
	6 DNA	HPV6 E7 70-98	Nucleic acid
D	7 AA	HPV6 E4 36-90	Amino acid
	8 DNA	HPV6 E4 36-90	Nucleic acid

E	9 AA	HPV6 E7 1-58	Amino acid
	10 DNA	HPV6 E7 1-58	Nucleic acid
F	11 AA	HPV6 E6*	Amino acid
	12 DNA	HPV6 E6*	Nucleic acid
G	13 AA	HPV6 E2* 1-368 Del_206-306	Amino acid
	14 DNA	HPV6 E2* 1-368 Del_206-306	Nucleic acid
H	15 AA	HPV6 E4 1-70 Del_21-50	Amino acid
	16 DNA	HPV6 E4 1-70 Del_21-50	Nucleic acid
I	17 AA	HPV6 E4 51-70	Amino acid
	18 DNA	HPV6 E4 51-70	Nucleic acid
	19 DNA	HPV6 E4 51-70	Nucleic acid
J	20 AA	HPV6 E6* 1-120 Del C-term	Amino acid
	21 DNA	HPV6 E6* 1-120 Del C-term	Nucleic acid
K	22 AA	HPV6 E7 1-98 Del_40-75	Amino acid
	23 DNA	HPV6 E7 1-98 Del_40-75	Nucleic acid

L	24 AA	HPV6 E2* 1-205	Amino acid
	25 DNA	HPV6 E2* 1-205	Nucleic acid
M	26 AA	HPV6 E4 1-20	Amino acid
	27 DNA	HPV6 E4 1-20	Nucleic acid
N	28 AA	HPV6 E2* 307-368	Amino acid
	29 DNA	HPV6 E2* 307-368	Nucleic acid
O	30 AA	HPV6 E4 Epitope-No.1	Amino acid
	31 DNA	HPV6 E4 Epitope-No.1	Nucleic acid
P	32 AA	HPV6 E4 Epitope-No.2	Amino acid
	33 DNA	HPV6 E4 Epitope-No.2	Nucleic acid
L1	34 AA	Rigid linker	Amino acid
	35 DNA	Rigid linker	Nucleic acid
	36 DNA	Rigid linker	Nucleic acid
	37 DNA	Rigid linker	Nucleic acid
	38 DNA	Rigid linker	Nucleic acid
	39 DNA	Rigid linker	Nucleic acid

L2	40 AA	HPV6 E6 epitope linker	Amino acid
	41 DNA	HPV6 E6 epitope linker	Nucleic acid
L3	42 AA	HPV11 E6 epitope linker	Amino acid
	43 DNA	HPV11 E6 epitope linker	Nucleic acid
	44 DNA	HPV11 E6 epitope linker	Nucleic acid
L4	45 AA	HPV11 E7 epitope linker	Amino acid
	46 DNA	HPV11 E7 epitope linker	Nucleic acid
	47 DNA	HPV11 E7 epitope linker	Nucleic acid
L5	48 AA	HPV16 E6 Agonist Enhancer No.1	Amino acid
	49 DNA	HPV16 E6 Agonist Enhancer No.1	Nucleic acid
L6	50 AA	HPV16 E6 Agonist Enhancer No.2	Amino acid
	51 DNA	HPV16 E6 Agonist Enhancer No.2	Nucleic acid
L7	52 AA	HPV16 E7 agonist enhancer	Amino acid
	53 DNA	HPV16 E7 agonist enhancer	Nucleic acid
	54 DNA	HPV16 E7 agonist enhancer	Nucleic acid
	55 DNA	HPV16 E7 agonist enhancer	Nucleic acid
	56 DNA	HPV16 E7 agonist enhancer	Nucleic acid

L8	57 AA	Flex linker	Amino acid
	58 DNA	Flex linker	Nucleic acid
	59 DNA	Flex linker	Nucleic acid
	60 DNA	Flex linker	Nucleic acid
	61 DNA	Flex linker	Nucleic acid
	62 DNA	Flex linker	Nucleic acid
	63 DNA	Flex linker	Nucleic acid
	64 DNA	Flex linker	Nucleic acid
	65 DNA	Flex linker	Nucleic acid
D1	66 AA	Design No. 1	Amino acid
	67 DNA	Design No. 1	Nucleic acid
D2	68 AA	Design No. 2	Amino acid
	69 DNA	Design No. 2	Nucleic acid
D3	70 AA	Design No. 3	Amino acid
	71 DNA	Design No. 3	Nucleic acid
D4	72 AA	Design No. 4	Amino acid

	73 DNA	Design No. 4	Nucleic acid
D5	74 AA	Design No. 5	Amino acid
	75 DNA	Design No. 5	Nucleic acid
L9	76 AA	Rigid linker polypeptide	Amino acid
	77 DNA	Rigid linker DNA	Nucleic acid
	78 DNA	Rigid linker DNA	Nucleic acid
	79 DNA	Rigid linker DNA	Nucleic acid
	80 DNA	Rigid linker DNA	Nucleic acid
	81 DNA	Rigid linker DNA	Nucleic acid

Additional Embodiments of the Invention

[0250] Embodiments ("**E**") of the invention include, but are not limited to:

E1. A polynucleotide encoding a non-naturally occurring polypeptide comprising at least one epitope derived from a low cancer risk human papilloma virus (HPV) protein and at least one epitope derived from a high cancer risk HPV protein.

E2. A polynucleotide encoding a non-naturally occurring polypeptide comprising at least one epitope derived from human papilloma virus 6 (HPV6) and at least one epitope derived from human papilloma virus 11 (HPV11).

E3. The polynucleotide of **E2**, wherein the non-naturally occurring polypeptide further comprises at least one epitope derived from human papilloma virus 16 (HPV16).

E4. A polynucleotide encoding a non-naturally occurring polypeptide comprising HPV epitopes derived from any four of:

an HPV6 E2 protein;
an HPV6 E4 protein;
an HPV6 E6 protein;
an HPV6 E7 protein;
an HPV11 E6 protein; and/or
an HPV11 E7 protein;

wherein the polypeptide comprises at least one HPV6 epitope and at least one HPV11 epitope.

E5. The polynucleotide of **E4**, wherein the non-naturally occurring polypeptide further comprises at least one epitope derived from HPV16 proteins.

E6. A polynucleotide encoding a non-naturally occurring polypeptide comprising HPV epitopes derived from any three of:

an HPV6 E2 protein;
an HPV6 E4 protein;
an HPV6 E6 protein; or
an HPV6 E7 protein;

wherein the HPV epitopes are linked by amino acid sequences not naturally encoded by an HPV genome.

E7. A polynucleotide encoding a non-naturally occurring polypeptide comprising HPV epitopes derived from each of:

an HPV6 E2 protein;
an HPV6 E4 protein;
an HPV6 E6 protein; and
an HPV6 E7 protein,

wherein the HPV epitopes are linked by amino acid sequences not naturally encoded by an HPV genome.

E8. The polynucleotide of **E6** or **E7**, wherein an amino acid sequence linking at least two of the HPV epitopes comprises the sequence of SEQ ID NO: 34.

E9. A polynucleotide encoding a non-naturally occurring polypeptide, wherein the non-naturally occurring polypeptide comprises HPV epitopes derived from:

a) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, and an HPV11 E7 protein; or,

b) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, an HPV11 E7 protein; an HPV16 E6 protein, and an HPV16 E7 protein.

E10. The polynucleotide of **E9**, further comprising one or more polynucleotide sequences encoding one or more polypeptide linkers, wherein the one or more linkers comprise an amino acid sequence of:

a) SEQ ID NO: 45;
b) SEQ ID NO: 42;
c) SEQ ID NO: 40;
d) SEQ ID NO: 34;
e) SEQ ID NO: 52;
f) SEQ ID NO: 48;
g) SEQ ID NO: 50; and/or,
h) SEQ ID NO: 57.

E11. A polynucleotide encoding a non-naturally occurring polypeptide, wherein the non-naturally occurring polypeptide comprises HPV epitopes derived from:

a) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, and an HPV11 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 45, SEQ ID NO: 42, or SEQ ID NO: 40;

b) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, and an HPV11 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 34;

c) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV11 E6 protein, an HPV11 E7 protein, an HPV16 E6 protein; an HPV16 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 52, SEQ ID NO: 48, or SEQ ID NO: 50; or,

d) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV11 E6 protein, an HPV11 E7 protein, an HPV16 E6 protein; and an HPV16 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 57.

E12. A polynucleotide comprising a sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 98% identical or 100% identical to:

a) SEQ ID NO: 67;

b) SEQ ID NO: 69;

c) SEQ ID NO: 71;

d) SEQ ID NO: 73; or,

e) SEQ ID NO: 75.

E13. A polynucleotide comprising a sequence encoding a non-naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to:

a) SEQ ID NO: 66;

b) SEQ ID NO: 68;

c) SEQ ID NO: 70;

d) SEQ ID NO: 72; or,

e) SEQ ID NO: 74.

E14. The polynucleotide of **E13**, wherein the amino acid sequence is at least 85% identical, at least 90% identical, at least 95% identical, at least 98% identical, or is 100% identical to any one of a), b), c), d) or e).

E15. The polynucleotide in any one of **E1** to **E14**, wherein one or more of the HPV epitopes comprise amino acid substitutions, amino acid insertions, amino sequence deletions, amino acid sequence separations or amino acid sequence rearrangements as compared to the corresponding wild type HPV sequences, wherein said amino acid sequence substitutions, insertions, deletions, separations, or rearrangements function to eliminate naturally-occurring HPV oncogenic biologic activity or essential viral replication functions.

E16. The polynucleotide in any one of **E1** to **E15**, wherein said polynucleotide is useful in generating an *in vivo* immune response to HPV6 or HPV11 antigens in a mammal, or wherein said polynucleotide is useful in generating an *in vivo* immune responses to HPV6 *and* HPV11 antigens in a mammal.

E17. A polypeptide encoded by the polynucleotide in any one of **E1** to **E16**.

E18. A vector encoding the polynucleotide in any one of **E1** to **E16.**

E19. The vector of **E18**, wherein said vector is a plasmid, a viral, or a non-viral vector.

E20. The vector of **E19**, wherein said viral vector is an adenoviral vector.

E21. The vector of **E20**, wherein said adenoviral vector is a non-human primate viral vector.

E22. A composition or vaccine comprising a pharmaceutically acceptable carrier and the polynucleotide of any one of **E1** to **E16**, the polypeptide of **E17**, or the vector of **E18** to **E21**.

E23. A method of generating an HPV6 and HPV11 composition, vaccine or polynucleotide encoding a non-naturally occurring polypeptide for use in a composition or vaccine, the method comprising engineering HPV polynucleotide sequences to encode derivatives of naturally occurring HPV polypeptides, wherein at least one of the viral polypeptides has oncogenic biologic activity or an essential viral replication function, wherein a combination of one or more amino acid sequence substitutions, deletions, insertions, separations, or rearrangements are expressed via the composition, vaccine or encoded by the polynucleotide to generate derivative viral polypeptides, wherein HPV oncogenic biologic activity or essential HPV replication function is eliminated from the derivative viral polypeptides, but wherein host immune-responsiveness and/or immunogenic activity against the derivative viral polypeptides is retained or enhanced compared to the corresponding naturally occurring viral polypeptides.

E24. The method of **E23**, wherein the polynucleotide or derivative viral polypeptide is designed utilizing steps comprising identification of viral antigens combined with any two or more of: *in silico* prediction of *in vivo* antigen epitope recognition, *in silico* prediction of *in vivo* allergen recognition, human proteome cross-reactivity sequence homology analysis, amino acid sequence physiochemical property analysis, three-dimensional (3D) structure analysis, identification of oncogenic and/or viral function inactivation mutations, deletions, substitutions and/or rearrangements.

E25. The method of **E24**, further comprising *in vitro* or *in vivo* evaluation of immunogenicity, immune-protective effects, oncogenic, or viral function of the derivative viral polypeptides.

E26. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more components of Table 5.

E27. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, or P of Table 5.

E28. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any one or more of the components L1, L2, L3, L4, L5, L6, L7, L8 or L9 of Table 5.

E29. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components A, B, C, D, E, or F of Table 5.

E30. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components A, B, C, D, E, F, L1, L2, L3, or L4 of Table 5.

E31. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components A, B, C, D, E, F, or L1 of Table 5.

E32. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components A, B, C, D, E, F, L2, L3, or L4 of Table 5.

E33. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components A, B, C, D, E, F and L1 of Table 5.

E34. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components A, B, C, D, E, F, L2, L3 and L4 of Table 5.

E35. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components G, H, I, J, or K of Table 5.

E36. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components G, H, I, J, K, L1, L2, or L4 of Table 5.

E37. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences components G, H, I, J and K of Table 5.

E38. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components G, H, I, J, K, L1, L2 and L4 of Table 5.

E39. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components I, L, M, or N of Table 5.

E40. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components I, L, M, N, L2, L4, L5, L6 or L7 of Table 5.

E41. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components I, L, M, and N of Table 5.

E42. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components I, L, M, N, L2, L4, L5, L6 and L7 of Table 5.

E43. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components N, O, or P of Table 5.

E44. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more of the components N, O, P, L2, L3, L4, L5, L6, L7 or L8 of Table 5.

E45. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components N, O and P of Table 5.

E46. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components N, O, P, L2, L3, L4, L5, L6, L7 and L8 of Table 5.

E47. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components A, B, C, D, E, F, and L1 of Table 5, in any order.

E48. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises, in sequential order, sequence components A, L1, B, L1, C, L1, D, L1, E, L1, and F of Table 5.

E49. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components A, B, C, D, E, F, L2, L3 and L4 of Table 5, in any order.

E50. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises, in sequential order, sequence components A, L4, B, L3, C, L2, D, L3, E, L4 and F of Table 5.

E51. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components G, H, I, J, K, L1, L2 and L4 of Table 5, in any order.

E52. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises, in sequential order, sequence components G, L1, L2, L1, H, L1, L4, L1, I, L1, J, L1 and K of Table 5.

E53. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components I, L, M, N, L2, L4, L5, L6 and L7 of Table 5, in any order.

E54. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises, in sequential order, sequence components L, L7, L2, L7, M, L5, L4, L6, I, L7, L2, L7 and N of Table 5.

E55. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of components N, O, P, L2, L3, L4, L5, L6, L7 and L8 of Table 5, in any order.

E56. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises, in sequential order, sequence components N, L8, L2, L8, O, L8, P, L8, L4, L8, L3, L8, L5, L8, L6, L8 and L7 of Table 5.

E57. A polynucleotide or polypeptide in any one of **E26** to **E56**, wherein any one or more L1 sequence components are substituted with L9 sequence components of Table 5.

E58. A polypeptide comprising any two or more of the sequences of: AA Subset No. 1, AA Subset No. 2, AA Subset No. 3, AA Subset No. 4, or AA Subset No. 5, wherein the sequences are in any order, or are in sequential order, as shown below.

E59. A polypeptide comprising the sequences of: AA Subset No. 1, AA Subset No. 2, AA Subset No. 3, AA Subset No. 4, or AA Subset No. 5, wherein the sequences are in any order, or are in sequential order, as shown below.

E60. A polynucleotide comprising a nucleic acid sequence encoding a polypeptide comprising any two or more of the sequences of: AA Subset No. 1, AA Subset No. 2, AA Subset No. 3, AA Subset No. 4, or AA Subset No. 5, wherein the sequences are in any order, or are in sequential order, as shown below.

E61. A polynucleotide comprising a nucleic acid sequence encoding a polypeptide comprising the sequences of: AA Subset No. 1, AA Subset No. 2, AA Subset No. 3, AA Subset No. 4, or AA Subset No. 5, wherein the sequences are in any order, or are in sequential order, as shown below.

AA Subset No. 1:

SEQ ID NO: 1 (HPV6 E2*);
SEQ ID NO: 34 (rigid linker);
SEQ ID NO: 3 (HPV6 E4 1-56);
SEQ ID NO: 34 (rigid linker);
SEQ ID NO: 5 (HPV6 E7 70-98);
SEQ ID NO: 34 (rigid linker);
SEQ ID NO: 7 (HPV6 E4 36-90);

SEQ ID NO: 34 (rigid linker);

SEQ ID NO: 9 (HPV6 E7 1-58);

SEQ ID NO: 34 (rigid linker); and/or,

SEQ ID NO: 11 (HPV6 E6*).

AA Subset No. 2:

SEQ ID NO: 1 (HPV6 E2*);

SEQ ID NO: 45 (HPV11 E7 epitope linker);

SEQ ID NO: 3 (HPV6 E4);

SEQ ID NO: 42 (HPV11 E6 linker);

SEQ ID NO: 5 (HPV6 E7);

SEQ ID NO: 40 (HPV6 E6 linker);

SEQ ID NO: 7 (HPV6 E4);

SEQ ID NO: 42 (HPV11 E6 linker);

SEQ ID NO: 9 (HPV6 E7);

SEQ ID NO: 45 (HPV11 E7 epitope linker); and/or,

SEQ ID NO: 11 (HPV6 E6*).

AA Subset No. 3:

SEQ ID NO: 13 (HPV6 E2* 1-368 (Del 206-306));

SEQ ID NO: 34 (rigid linker);

SEQ ID NO: 42 (HPV11 E6 epitope “linker”);

SEQ ID NO: 34 (rigid linker);

SEQ ID NO: 15 (HPV6 E4 1-70 (Del 21-50));

SEQ ID NO: 34 (rigid linker);

SEQ ID NO: 45 (HPV11 E7 epitope “linker”);

SEQ ID NO: 34 (rigid linker);

SEQ ID NO: 17 (HPV6 E4 51-70);

SEQ ID NO: 34 (rigid linker);

SEQ ID NO: 20 (HPV6 E6* 1-120 (Del C-term));

SEQ ID NO: 34 (rigid linker); and/or,

SEQ ID NO: 22 (HPV6 E7 1-98 (Del 40-75)).

AA Subset No. 4:

SEQ ID NO: 24 (HPV6 E2* 1-205);

SEQ ID NO: 52 (HPV16 E7 agonist enhancer linker);

SEQ ID NO: 42 (HPV11 E6 epitope linker);

SEQ ID NO: 52 (HPV16 E7 agonist enhancer linker);

SEQ ID NO: 26 (HPV6 E4 1-20);

SEQ ID NO: 48 (HPV16 E6 agonist enhancer No.1 linker);

SEQ ID NO: 45 (HPV11 E7 epitope linker);

SEQ ID NO: 50 (HPV16 E6 agonist enhancer No.2 linker);

SEQ ID NO: 17 (HPV6 E4 51-70);

SEQ ID NO: 52 (HPV16 E7 agonist enhancer linker);

SEQ ID NO: 40 (HPV6 E6 epitope linker);

SEQ ID NO: 52 (HPV16 E7 agonist enhancer linker); and/or,

SEQ ID NO: 28 (HPV6 E2* 307-368).

AA Subset No. 5:

SEQ ID NO: 28 (HPV6 E2*);

SEQ ID NO: 57 (Flex linker);

SEQ ID NO: 40 (HPV6 E6 epitope linker);

SEQ ID NO: 57 (Flex linker);

SEQ ID NO: 30 (HPV6 E4 Epitope-No.1);

SEQ ID NO: 57 (Flex linker);

SEQ ID NO: 32 (HPV6 E4 Epitope -No.2);

SEQ ID NO: 57 (Flex linker);

SEQ ID NO: 45 (HPV11 E7 epitope linker);

SEQ ID NO: 57 (Flex linker);

SEQ ID NO: 42 (HPV11 E6 epitope linker);

SEQ ID NO: 57 (Flex linker);

SEQ ID NO: 48 (HPV16 E6 Agonist Enhancer No.1);

SEQ ID NO: 57 (Flex linker);

SEQ ID NO: 50 (HPV16 E6 Agonist Enhancer No.2);

SEQ ID NO: 57 (Flex linker); and/or,

SEQ ID NO: 52 (HPV16 E7 agonist enhancer).

E62. A polynucleotide comprising any two or more nucleic acid sequences of: DNA Subset No. 1, DNA Subset No. 2, DNA Subset No. 3, DNA Subset No. 4, or DNA Subset No. 5, wherein the sequences are in any order, or are in sequential order, as shown below.

E63. A polynucleotide comprising nucleic acid sequences of: DNA Subset No. 1, DNA Subset No. 2, DNA Subset No. 3, DNA Subset No. 4, or DNA Subset No. 5, wherein the sequences are in any order, or are in sequential order, as shown below.

DNA Subset No. 1:

SEQ ID NO: 2 (HPV6 E2*);

SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);

SEQ ID NO: 4 (HPV6 E4 1-56);

SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);

SEQ ID NO: 6 (HPV6 E7 70-98);

SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);

SEQ ID NO: 8 (HPV6 E4 36-90);

SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);

SEQ ID NO: 10 (HPV6 E7 1-58);

SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker); and/or,

SEQ ID NO: 12 (HPV6 E6*).

DNA Subset No. 2:

SEQ ID NO: 2 (HPV6 E2*);

SEQ ID NO: 46 or 47 (HPV11 E7 epitope linker);

SEQ ID NO: 4 (HPV6 E4);

SEQ ID NO: 43 or 44 (HPV11 E6 linker);

SEQ ID NO: 6 (HPV6 E7);

SEQ ID NO: 41 (HPV6 E6 linker);

SEQ ID NO: 8 (HPV6 E4);

SEQ ID NO: 43 or 44 (HPV11 E6 linker);

SEQ ID NO: 10 (HPV6 E7 1-58);

SEQ ID NO: 46 or 47 (HPV11 E7 epitope linker); and/or,

SEQ ID NO: 12 (HPV6 E6*).

DNA Subset No. 3:

SEQ ID NO: 14 (HPV6 E2* 1-368 (Del 206-306));
SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);
SEQ ID NO: 43 or 44 (HPV11 E6 epitope linker);
SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);
SEQ ID NO: 16 (HPV6 E4 1-70 (Del 21-50));
SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);
SEQ ID NO: 46 or 47 (HPV11 E7 epitope linker);
SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);
SEQ ID NO: 18 or 19 (HPV6 E4 51-70);
SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker);
SEQ ID NO: 21 (HPV6 E6* 1-120 (Del C-term));
SEQ ID NO: 35, 36, 37, 38 or 39 (rigid linker); and/or,
SEQ ID NO: 23 (HPV6 E7 1-98 (Del 40-75)).

DNA Subset No. 4:

SEQ ID NO: 25 (HPV6 E2* 1-205);
SEQ ID NO: 53, 54, 55, or 56 (HPV16 E7 agonist enhancer linker);
SEQ ID NO: 43 or 44 (HPV11 E6 epitope linker);
SEQ ID NO: 53, 54, 55, or 56 (HPV16 E7 agonist enhancer linker);
SEQ ID NO: 27 (HPV6 E4 1-20);
SEQ ID NO: 49 (HPV16 E6 Agonist Enhancer No.1 linker);
SEQ ID NO: 46 or 47 (HPV11 E7 epitope linker);
SEQ ID NO: 51 (HPV16 E6 Agonist Enhancer No.2 linker);
SEQ ID NO: 18 or 19 (HPV6 E4 51-70);
SEQ ID NO: 53, 54, 55, or 56 (HPV16 E7 agonist enhancer linker);
SEQ ID NO: 41 (HPV6 E6 epitope linker);
SEQ ID NO: 53, 54, 55, or 56 (HPV16 E7 agonist enhancer linker); and/or,
SEQ ID NO: 29 (HPV6 E2* 307-368).

DNA Subset No. 5:

SEQ ID NO: 29 (HPV6 E2*);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker);

SEQ ID NO: 41 (HPV6 E6 epitope linker);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker);

SEQ ID NO: 31 (HPV6 E4 Epitope-No.1);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker);

SEQ ID NO: 33 (HPV6 E4 Epitope-No.2);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker);

SEQ ID NO: 46 or 47 (HPV11 E7 epitope linker);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker);

SEQ ID NO: 43 or 44 (HPV11 E6 epitope linker);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker);

SEQ ID NO: 49 (HPV16 E6 Agonist Enhancer No.1);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker);

SEQ ID NO: 51 (HPV16 E6 Agonist Enhancer No.2);

SEQ ID NO: 58, 59, 60, 61, 62, 63, 64 or 65 (Flex linker); and/or,

SEQ ID NO: 53, 54, 55, or 56 (HPV16 E7 agonist enhancer).

E64. A vector encoding the polynucleotide in any one of **E1** to **E16** and **E26** to **E63.**

E65. The vector of **E64**, wherein the vector is a plasmid, a viral, or a non-viral vector.

E66. The vector of **E65**, wherein the viral vector is an adenoviral vector.

E67. The vector of **E66**, wherein the adenoviral vector is a non-human primate viral vector.

E68. A composition or vaccine comprising a pharmaceutically acceptable carrier and a polynucleotide in any one of **E1** to **E16** and **E26** to **E63**, a polypeptide in any one of **E17** and **E26** to **E59**, or a vector in any one of **E18 to E21** and **E64** to **E67**.

E69. The polynucleotide, polypeptide, vector, composition or vaccine in any one of **E1** to **E22** and **E26** to **E68**, wherein the polynucleotide, polypeptide, vector, composition or vaccine is capable of inducing an anti-HPV immune response.

E70. Use of a polynucleotide, polypeptide, vector, composition or vaccine in **E69** for treatment of an HPV-associated disease or disorder.

E72. The use of a polynucleotide, polypeptide, vector, composition or vaccine in **E70**, wherein the disease or disorder is an HPV6, HPV11 or HPV16-associated disease or disorder.

E72. The use of a polynucleotide, polypeptide, vector, composition or vaccine in **E70**, wherein the disease or disorder is Recurrent Respiratory Papillomatosis (RRP).

References

[0251] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. In the event that any disclosure made in sources incorporated by reference conflict with disclosure in the present application, the disclosure of the present application shall be deemed to control and supersede such sources.

Abbate, E. A., Berger, J. M., & Botchan, M. R. (2004). The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2. *Genes & Development*, 18(16), 1981–1996. doi:10.1101/gad.1220104.

Ahn, J., Peng, S., Hung, C. F., Roden, R., Wu, T. C., & Best, S. R. (2017). Immunologic responses to a novel DNA vaccine targeting human papillomavirus-11 E6E7. *The Laryngoscope*, 127(12), 2713–2720. doi:10.1002/lary.26737.

Anderson, Karen S., *et al*. (2015). Biologic predictors of serologic responses to HPV in oropharyngeal cancer: The HOTSPOT study. *Oral Oncology*, 51(8): 751–758. doi: 10.1016/j.oraloncology.2015.05.007

Brokaw, J. L., Blanco, M., & McBride, A. A. (1996). Amino acids critical for the functions of the bovine papillomavirus type 1 E2 transactivator. Journal of Virology, 70(1), 23–29.

Graham, S. (2017). The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clinical Science*, 131(17), 2201-2221. doi: 10.1042/CS20160786.

Holm, A., Nagaeva, O., Nagaev, I., Loizou, C., Laurell, G., Mincheva-Nilsson, L., Olofsson, K. (2017). Lymphocyte profile and cytokine mRNA expression in peripheral blood mononuclear cells of patients with recurrent respiratory papillomatosis suggest dysregulated cytokine mRNA response and impaired cytotoxic capacity. *Immunity, Inflammation and Disease*, 5(4), 541–550. doi:10.1002/iid3.188.

Oh, ST, Longworth, MS, & Laimins, LA. (2004) Roles of the E6 and E7 proteins in the life cycle of low-risk human papillomavirus type 11. *Journal of Virology*, 78(5), 2620–2626. doi:10.1128/jvi.78.5.2620-2626.2004.

Seedat, R., Combrinck, C., Bester, P., Lee, J., & Burt, F. (2016). Determination of the complete genome and functional analysis of HPV6 isolate VBD19/10 from a patient with aggressive recurrent respiratory papillomatosis. *Epidemiology and Infection*, 144(10), 2128-2135. doi:10.1017/S0950268816000388.

Shin, T. H., Pankhong, P., Yan, J., Khan, A. S., Sardesai, N. Y., & Weiner, D. B. (2012). Induction of robust cellular immunity against HPV6 and HPV11 in mice by DNA vaccine encoding for E6/E7 antigen. *Human Vaccines & Immunotherapeutics*, 8(4), 470–478. doi:10.4161/hv.19180.

Trimble, C. L., Morrow, M. P., Kraynyak, K. A., Shen, X., Dallas, M., Yan, J., Bagarazzi, M. L. (2015). Safety, efficacy, and immunogenicity of VGX-3100, a therapeutic synthetic DNA vaccine targeting human papillomavirus 16 and 18 E6 and E7 proteins for cervical intraepithelial neoplasia 2/3: a randomised, double-blind, placebo-controlled phase 2b trial. *Lancet* (London, England), 386(10008), 2078–2088. doi:10.1016/S0140-6736(15)00239-1.

Wang, Y., Coulombe, R., Cameron, DR., Thauvette, L., Massariol, M., Amon, LM., Fink, D., Titolo, S., Welchner, E., Yoakim, C., Arachambault, J. & White, P.W. (2004). Crystal structure of the E2 domain of human papillomavirus type 11 bound to a protein interaction inhibitor. *Journal of Biological Chemistry*. 279(8), 6976-6985. doi:10.1074/jbc.M311376200.

Yiu, Y., Fayson, S., Smith, H., & Matrka, L. (2019). Implementation of Routine HPV Vaccination in the Management of Recurrent Respiratory Papillomatosis. *Annals of Otology, Rhinology & Laryngology*, 128(4), 309–315.

Claims

1. A polynucleotide encoding a non-naturally occurring polypeptide comprising at least one epitope derived from human papilloma virus 6 (HPV6) and at least one epitope derived from human papilloma virus 11 (HPV11).

2. The polynucleotide of claim 1, wherein the non-naturally occurring polypeptide comprises at least one epitope derived from human papilloma virus 16 (HPV16).

3. The polynucleotide of claim 1, wherein the non-naturally occurring polypeptide comprises HPV epitopes derived from any four of:
   an HPV6 E2 protein;
   an HPV6 E4 protein;
   an HPV6 E6 protein;
   an HPV6 E7 protein;
   an HPV11 E6 protein; and/or
   an HPV11 E7 protein;
   wherein the polypeptide comprises at least one HPV6 epitope and at least one HPV11 epitope.

4. The polynucleotide of claim 1, wherein the non-naturally occurring polypeptide further comprises at least one epitope derived from HPV16 proteins.

5. The polynucleotide of claim 1, wherein the non-naturally occurring polypeptide, wherein the non-naturally occurring polypeptide comprises HPV epitopes derived from:
   a) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, and an HPV11 E7 protein; or,
   b) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, an HPV11 E7 protein; an HPV16 E6 protein, and an HPV16 E7 protein.

6. The polynucleotide of claim 5, further comprising one or more polynucleotide sequences encoding one or more polypeptide linkers, wherein the one or more linkers comprise an amino acid sequence of:

a) SEQ ID NO: 34;

b) SEQ ID NO: 40;

c) SEQ ID NO: 42;

d) SEQ ID NO: 45;

e) SEQ ID NO: 48;

f) SEQ ID NO: 50;

g) SEQ ID NO: 52; and/or,

h) SEQ ID NO: 57.

7. A polynucleotide encoding a non-naturally occurring polypeptide, wherein the non-naturally occurring polypeptide comprises HPV epitopes derived from:

a) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, and an HPV11 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 45, SEQ ID NO: 42, or SEQ ID NO: 40;

b) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV6 E7 protein, an HPV11 E6 protein, and an HPV11 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 34;

c) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV11 E6 protein, an HPV11 E7 protein, an HPV16 E6 protein; an HPV16 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 52, SEQ ID NO: 48, or SEQ ID NO: 50; or,

d) an HPV6 E2 protein, an HPV6 E4 protein, an HPV6 E6 protein, an HPV11 E6 protein, an HPV11 E7 protein, an HPV16 E6 protein; and an HPV16 E7 protein, wherein the polypeptide further comprises at least one linker sequence of SEQ ID NO: 57.

8. A polynucleotide comprising a sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 98% identical or 100% identical to:

a) SEQ ID NO: 67;

b) SEQ ID NO: 69;

c) SEQ ID NO: 71;

d) SEQ ID NO: 73; or,

e) SEQ ID NO: 75.

9. A polynucleotide comprising a sequence encoding a non-naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to:

a) SEQ ID NO: 66;

b) SEQ ID NO: 68;

c) SEQ ID NO: 70;

d) SEQ ID NO: 72; or,

e) SEQ ID NO: 74.

10. The polynucleotide of claim 9, wherein the amino acid sequence is at least 85% identical, at least 90% identical, at least 95% identical, at least 98% identical, or is 100% identical to any one of a), b), c), d) or e).

11. A polypeptide encoded by the polynucleotide of claim 1.

12. A method of generating an HPV6 and HPV11 composition, vaccine or polynucleotide encoding a non-naturally occurring polypeptide for use in a composition or vaccine, the method comprising engineering HPV polynucleotide sequences to encode derivatives of naturally occurring HPV polypeptides, wherein at least one of the viral polypeptides has oncogenic biologic activity or an essential viral replication function, wherein a combination of one or more amino acid sequence substitutions, deletions, insertions, separations, or rearrangements are expressed via the composition, vaccine or encoded by the polynucleotide to generate derivative viral polypeptides, wherein HPV oncogenic biologic activity or essential HPV replication function is eliminated from the derivative viral polypeptides, but wherein host immune-responsiveness and/or immunogenic activity against the derivative viral polypeptides is retained or enhanced compared to the corresponding naturally occurring viral polypeptides.

13. The method of claim 12, wherein the polynucleotide or derivative viral polypeptide is designed utilizing steps comprising identification of viral antigens combined with any two or more of: *in silico* prediction of *in vivo* antigen epitope recognition, *in silico* prediction of *in vivo* allergen recognition, human proteome cross-reactivity sequence homology analysis, amino acid sequence physiochemical property analysis, three-dimensional (3D) structure analysis, identification of oncogenic and/or viral function inactivation mutations, deletions, substitutions and/or rearrangements.

14. A non-naturally occurring polypeptide or polynucleotide, wherein the polypeptide or polynucleotide comprises sequences of any two or more components of Table 5.

15. The non-naturally occurring polypeptide or polynucleotide of claim 14, wherein the polypeptide or polynucleotide comprises sequences of components A, B, C, D, E, F and L1 of Table 5.

16. The non-naturally occurring polypeptide or polynucleotide of claim 14, wherein the polypeptide or polynucleotide comprises sequences of components A, B, C, D, E, F, L2, L3 and L4 of Table 5.

17. The non-naturally occurring polypeptide or polynucleotide of claim 14, wherein the polypeptide or polynucleotide comprises sequences of components G, H, I, J, K, L1, L2 and L4 of Table 5.

18. The non-naturally occurring polypeptide or polynucleotide of claim 14, wherein the polypeptide or polynucleotide comprises sequences of components I, L, M, N, L2, L4, L5, L6 and L7 of Table 5.

19. The non-naturally occurring polypeptide or polynucleotide of claim 14, wherein the polypeptide or polynucleotide comprises sequences of components N, O, P, L2, L3, L4, L5, L6, L7 and L8 of Table 5.

20. A vector encoding the polynucleotide of claim 1.

21. The vector of claim 20, wherein the vector is a plasmid, a viral, or a non-viral vector.

22. The vector of claim 21, wherein the viral vector is an adenoviral vector.

23. The vector of claim 22, wherein the adenoviral vector is a non-human primate viral vector.

24. A composition or vaccine comprising a pharmaceutically acceptable carrier and the polynucleotide of claim 1.

25. Use of the composition or vaccine of claim 24, for treatment of an HPV-associated disease or disorder.

26. Use of the composition or vaccine of claim 25, wherein the disease or disorder is an HPV6, HPV11 or HPV16-associated disease or disorder.

27. Use of the composition or vaccine of claim 25, wherein the disease or disorder is Recurrent Respiratory Papillomatosis (RRP).