

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

21 September 2023 (21.09.2023)



WIPO | PCT



(10) International Publication Number

WO 2023/174998 A1

(51) International Patent Classification:

A61K 39/00 (2006.01)

C07K 16/28 (2006.01)

A61K 39/12 (2006.01)

(21) International Application Number:

PCT/EP2023/056591

(22) International Filing Date:

15 March 2023 (15.03.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

22162533.8

16 March 2022 (16.03.2022)

EP

22204783.9

31 October 2022 (31.10.2022)

EP

(71) Applicant: **BOEHRINGER INGELHEIM INTERNATIONAL GMBH** [DE/DE]; Binger Strasse 173, 55216 Ingelheim am Rhein (DE).

(72) Inventors: **BELNOUE, Elodie**; Amal Therapeutics, 64, Avenue de la Roseaie, 1205 Geneve (CH). **DEROUAZI, Madiha**; Amal Therapeutics, 64, Avenue de la Roseaie, 1205 Geneve (CH). **ADAM, Paul**; Boehringer Ingelheim International GmbH, Global Patents, Binger Strasse 173, 55216 Ingelheim am Rhein (DE). **HOFMANN, Irmgard Maria Rita**; Boehringer Ingelheim International GmbH, Global Patents, Binger Strasse 173, 55216 Ingelheim am Rhein (DE). **LEONHARDT, Ralf**; Boehringer Ingelheim International GmbH, Global Patents, Binger Strasse 173, 55216 Ingelheim am Rhein (DE). **LUKOWSKI, Samuel**; Boehringer Ingelheim International GmbH, Global Patents, Binger Strasse 173, 55216 Ingelheim am Rhein (DE). **NOLDEN, Tobias**; Boehringer Ingelheim International GmbH, Global Patents, Binger Straße 173, 55216 Ingelheim am Rhein (DE). **TRAPANI, Francesca**; Boehringer Ingelheim International GmbH, Global Patents, Binger Strasse 173, 55216 Ingelheim am Rhein (DE).

(74) Agent: **GRAF VON STOSCH PATENTANWALTSGESELLSCHAFT MBH** et al.; Prinzregentenstr. 22, 80538 Munich (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,

TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

Published:

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))

(54) Title: TUMOR ANTIGENS, COMPOUNDS COMPRISING THE TUMOR ANTIGENS KRAS, TPX2 OR AURKA AND USES THEREOF

(57) Abstract: The present invention relates to tumor antigens encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of different mRNAs. Compositions and peptides comprising such tumor antigens and a virus encoding such tumor antigens are provided. The present invention also relates to the use of such compositions, peptides and viruses in the treatment of cancer.

WO 2023/174998 A1

5 **TUMOR ANTIGENS, COMPOUNDS COMPRISING THE TUMOR ANTIGENS KRAS,
TPX2 OR AURKA AND USES THEREOF**

10 The present invention relates to the field of vaccination and immunotherapy, in particular to cancer immunotherapy. More specifically, the present invention relates to tumor antigens encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of different mRNAs. Compositions and peptides comprising such tumor antigens and a virus encoding such tumor antigens are provided. The present invention also relates to the use of such compositions, peptides and viruses in the treatment of cancer.

15 The immune system can recognize and to some extent eliminate tumor cells, however, this anti-tumor response is often of low amplitude and inefficient. Boosting this weak anti-tumor response with therapeutic vaccination has been a long sought goal for cancer therapy. Modulating the immune system to enhance immune responses has thus become a promising therapeutic approach in oncology as it can be combined with standard of care treatments.

20 Tumor cells commonly express several antigens, such as tumor-associated antigens (TAAs), viral antigens (oncovirus) or mutation-derived antigens (neoantigens). Various TAAs, viral antigens or neoantigens expressed in cancer cells have been identified and utilized as targets for cancer vaccines. One approach to elicit tumor-specific immune responses is the peptide-based cancer vaccination involving administration of TAAs, viral antigens or neoantigen-derived to treat cancer according to the nature of the tumor.

30 While tumor associated antigens are usually also present on normal cells, tumor-specific antigens are present only on tumor cells. Tumor-specific antigens include mutation-derived antigens as well as antigens derived from cryptic open reading frames and/or frameshift mutations. Such antigens that can be recognized by the immune system as foreign antigens and, therefore, drive an anti-tumor immune response in animal tumor-models and cancer

patients. However, neoantigens with tumor-specific mutations often require vaccines tailored to individual patients, which is cumbersome and cost-intensive. Recently, so-called “dark antigens” received increasing interest, which are peptides expressed specifically in tumors, which are encoded in sequence regions previously believed to be “non-coding”, such as 5’
5 UTRs of mRNAs.

In view thereof, it is the object of the present invention to provide novel tumor antigens, which are specifically expressed in tumors of many patients. Such tumor antigens provide tumor specificity, but do not require cumbersome and cost-intensive personalized
10 approaches. It is also an object of the present invention to provide vaccines, in particular vaccines which can be used in a heterologous prime-boost regimen, combining such tumor antigens with further advantageous features, such as a multi-antigenic domain. This is especially relevant as each human being has a different set of MHC molecules and presentation of multiple peptides is allowing representation of the best (highest affinity)
15 peptides in the respective HLA.

This object is achieved by means of the subject-matter set out below and in the appended claims.

20 Although the present invention is described in detail below, it is to be understood that this invention is not limited to the particular methodologies, protocols and reagents described herein as these may vary. It is also to be understood that the terminology used herein is not intended to limit the scope of the present invention which will be limited only by the appended claims. Unless defined otherwise, all technical and scientific terms used herein
25 have the same meanings as commonly understood by one of ordinary skill in the art.

In the following, the elements of the present invention will be described. These elements are listed with specific embodiments, however, it should be understood that they may be combined in any manner and in any number to create additional embodiments. The variously
30 described examples and preferred embodiments should not be construed to limit the present invention to only the explicitly described embodiments. This description should be understood to support and encompass embodiments which combine the explicitly described

embodiments with any number of the disclosed and/or preferred elements. Furthermore, any permutations and combinations of all described elements in this application should be considered disclosed by the description of the present application unless the context indicates otherwise.

5

Throughout this specification and the claims which follow, unless the context requires otherwise, the term "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated member, integer or step but not the exclusion of any other non-stated member, integer or step. The term "consist of" is a particular
10 embodiment of the term "comprise", wherein any other non-stated member, integer or step is excluded. In the context of the present invention, the term "comprise" encompasses the term "consist of". The term "comprising" thus encompasses "including" as well as "consisting" *e.g.*, a composition "comprising" X may consist exclusively of X or may include something additional *e.g.*, X + Y.

15

The terms "a" and "an" and "the" and similar reference used in the context of describing the invention (especially in the context of the claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. Recitation of ranges of values herein is merely intended to serve as a shorthand method of
20 referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually recited herein. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the invention.

25

The word "substantially" does not exclude "completely" *e.g.*, a composition which is "substantially free" from Y may be completely free from Y. Where necessary, the word "substantially" may be omitted from the definition of the invention.

The term "about" in relation to a numerical value x means $x \pm 10\%$.

30

Throughout the present specification, the terms "peptide", "polypeptide", "protein" and variations of these terms are used interchangeably. These terms refer to peptides,

oligopeptides, or proteins including fusion proteins, which comprise at least two amino acids joined to each other, preferably by a ("normal") peptide bond. Alternatively, the amino acids may be joined to each other by a modified peptide bond, such as for example in the cases of isosteric peptides. Accordingly, as used herein, the term "peptide" refers to shorter (oligo)peptides as well as to longer (polypeptides) and to proteins, independently of their length. "Classical" peptides, which are typically composed of amino acids selected from the 20 amino acids defined by the genetic code, linked to each other by a normal peptide bond, are preferred. However, a peptide can also comprise amino acids other than the 20 amino acids defined by the genetic code in addition to these amino acids, or it can be composed of amino acids other than the 20 amino acids defined by the genetic code. In some embodiments, a peptide can be composed of amino acids modified by natural processes, such as post-translational maturation processes, or by chemical processes, which are well known to a person skilled in the art. Modifications can appear anywhere in the peptide: in the peptide skeleton, in the amino acid chain or even at the carboxy- or amino-terminal end of the peptide. Accordingly, the terms "peptide", "polypeptide", "protein" also include modified peptides, polypeptides and proteins. For example, a peptide modification can include acetylation, acylation, ADP-ribosylation, amidation, covalent fixation of a nucleotide or of a nucleotide derivative, covalent fixation of a lipid or of a lipidic derivative, the covalent fixation of a phosphatidylinositol, covalent or non-covalent cross-linking, cyclization, disulfide bond formation, demethylation, glycosylation including pegylation, hydroxylation, iodization, methylation, myristoylation, oxidation, proteolytic processes, phosphorylation, prenylation, racemization, seneloylation, sulfatation, amino acid addition such as arginylation or ubiquitination.

As used herein (i.e. throughout the present application), the term "sequence variant" refers to any alteration in a reference sequence. The term "sequence variant" includes nucleotide sequence variants and amino acid sequence variants. Preferably, a reference sequence is any of the sequences listed in the "Table of Sequences and SEQ ID Numbers" (Sequence listing), i.e. SEQ ID NO: 1 to SEQ ID NO: 51. In particular, a sequence variant shares (over the whole length of the sequence) at least 70% or at least 75%, preferably at least 80% or at least 85%, more preferably at least 90% or at least 95%, even more preferably at least 97% or at least 98%, particularly preferably at least 99% sequence identity with a reference sequence.

Sequence identity may be calculated as described below. In particular, a sequence variant usually preserves the specific function of the reference sequence. In some embodiments, an amino acid sequence variant has an altered sequence in which one or more (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) of the amino acids in the reference sequence is deleted or substituted, or one or more (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) amino acids are inserted into or added to the sequence of the reference amino acid sequence. As a result of the alterations, the amino acid sequence variant has an amino acid sequence which is at least 70% or at least 75%, preferably at least 80% or at least 85%, more preferably at least 90% or at least 95%, even more preferably at least 97% or at least 98%, particularly preferably at least 99% identical to the reference sequence. For example, variant sequences which are at least 90% identical have no more than 10 alterations, i.e., any combination of deletions, insertions or substitutions, per 100 amino acids of the reference sequence. The same, of course, also applies similarly to nucleic acid sequences.

For (amino acid or nucleic acid) sequences without exact correspondence, a "% identity" of a first sequence may be determined with respect to a second sequence. In general, these two sequences to be compared may be aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may then be determined over the whole length of each of the sequences being compared (so-called global alignment), that is particularly suitable for sequences of the same or similar length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length. Methods for comparing the identity and homology of two or more sequences are well known in the art. The percentage to which two sequences are identical can e.g., be determined using a mathematical algorithm. A preferred, but not limiting, example of a mathematical algorithm which can be used is the algorithm of Karlin *et al.* (1993), PNAS USA, 90:5873-5877. Such an algorithm is integrated in the BLAST family of programs, e.g., BLAST or NBLAST program (see also Altschul *et al.*, 1990, J. Mol. Biol. 215, 403-410 or Altschul *et al.* (1997), Nucleic Acids Res, 25:3389-3402), accessible through the home page of the NCBI at world wide web site ncbi.nlm.nih.gov and FASTA (Pearson (1990), Methods Enzymol. 183, 63-98; Pearson and Lipman (1988), Proc. Natl. Acad. Sci. U. S. A 85, 2444-2448.).

In general, substitutions for one or more amino acids present in the referenced amino acid sequence are preferably made conservatively. Examples of conservative substitutions include substitution of one aliphatic residue for another, such as Ile, Val, Leu, or Ala for one another, or substitutions of one polar residue for another, such as between Lys and Arg; Glu and Asp; 5 or Gln and Asn. Other such conservative substitutions, for example, substitutions of entire regions having similar hydrophobicity properties, are well known (*Kyte and Doolittle, 1982, J. Mol. Biol. 157(1):105- 132*). Substitutions of one or more L-amino acids with one or more D-amino acids are to be considered as conservative substitutions in the context of the present invention. Exemplary amino acid substitutions are presented in Table 1 below:

10

Original residues	Examples of substitutions
Ala (A)	Val, Leu, Ile, Gly
Arg (R)	His, Lys
Asn (N)	Gln
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Pro, Ala
His (H)	Lys, Arg
Ile (I)	Leu, Val, Met, Ala, Phe
Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, His
Met (M)	Leu, Ile, Phe
Phe (F)	Leu, Val, Ile, Tyr, Trp, Met
Pro (P)	Ala, Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr, Phe
Tyr (Y)	Trp, Phe
Original residues	Examples of substitutions
Val (V)	Ile, Met, Leu, Phe, Ala

(Table 1)

As used herein, an “antigen” is any structural substance which serves as a target for the receptors of an adaptive immune response, in particular as a target for antibodies, T cell receptors, and/or B cell receptors. Preferably, the antigen is a peptide (including polypeptides and proteins). An “epitope”, also known as “antigenic determinant”, is the part (or fragment) of an antigen that is recognized by the immune system, in particular by antibodies, T cell receptors, and/or B cell receptors. Thus, one antigen has at least one epitope, i.e., a single antigen has one or more epitopes. In the context of the present invention, the term “epitope” is mainly used to designate T cell epitopes, which are presented on the surface of an antigen-presenting cell, where they are bound to Major Histocompatibility Complex (MHC). For peptide antigens, T cell epitopes presented by MHC class I molecules are typically, but not exclusively, peptides between 8 and 11 amino acids in length, whereas MHC class II molecules present longer peptides, generally, but not exclusively, between 12 and 25 amino acids in length.

The terms “CD4⁺ epitope” or “CD4⁺-restricted epitope”, as used herein, designate an epitope recognized by a CD4⁺ T cell, said epitope in particular consisting of an antigen fragment lying in the groove of a MHC class II molecule. A single CD4⁺ epitope preferably consists of about 12-25 amino acids. It can also consist of, for example, about 8-25 amino acids or about 6-100 amino acids.

The terms “CD8⁺ epitope” or “CD8⁺-restricted epitope”, as used herein, designate an epitope recognized by a CD8⁺ T cell, said epitope in particular consisting of an antigen fragment lying in the groove of a MHC class I molecule. A single CD8⁺ epitope preferably consists of about 8-11 amino acids. It can also consist of, for example, about 8-15 amino acids or about 6-100 amino acids.

As used herein, a “fragment” of an antigen has usually a minimum length of 8 amino acids, i.e. the fragment comprises at least 8 consecutive amino acids of the antigen, preferably at least 10 consecutive amino acids of the antigen, more preferably at least 15 consecutive amino acids of the antigen, even more preferably at least 20 consecutive amino acids of the antigen, still more preferably at least 25 consecutive amino acids of the antigen and most preferably at least 30 consecutive amino acids of the antigen. An antigen fragment usually

comprises one or more epitopes. Accordingly, the fragment of the antigen is typically immunogenic. A „sequence variant“ of an antigen (or a fragment thereof) has usually an (amino acid) sequence which is at least 70% or at least 75%, preferably at least 80% or at least 85%, more preferably at least 90% or at least 95%, even more preferably at least 97% or at least 98%, particularly preferably at least 99% identical to the reference sequence. A “functional” sequence variant means in the context of an antigen/antigen fragment, that the function of the epitope(s), e.g., comprised by the antigen (fragment), is not impaired or abolished. In other words, a “functional” sequence variant of an antigen (fragment) is immunogenic, preferably it has substantially the same immunogenicity as the reference antigen (fragment). In some embodiments, the amino acid sequence of one or more epitope(s), e.g., comprised in the tumor antigen (fragment), is not mutated and, thus, identical to a (naturally occurring) reference epitope sequence.

It is understood that the skilled person usually selects the antigen, or the fragment thereof, in view of the disease to be treated. Accordingly, the antigen or antigenic epitope is usually associated with (or related to) the disease to be treated. A large number of antigens is known in the art in the context of specific diseases. For example, to treat a tumor/cancer, the skilled person selects a tumor antigen (or a fragment thereof), which is useful for the specific type of tumor/cancer. In some embodiments, the patient may be tested/screened for specific antigens (e.g., by using an isolated sample) to identify whether or not the cancer/tumor expresses the specific antigen.

As used herein, a “tumor antigen” is an antigen produced by cancer/tumor cells. Tumor antigens include tumor-associated antigens and tumor-specific antigens. Tumor-associated (also tumor-related) antigens (TAAs) are antigens, which are expressed in both, cancer/tumor cells and normal cells. Tumor-specific antigens (TSAs), in contrast, are antigens, which are expressed specifically by cancer/tumor cells, but not by normal cells. Examples of TSA include neoantigens, which were not present in the body before the cancer/tumor developed and are, thus, neoantigens are “new” to the immune system. Neoantigens are often due to somatic mutations. Suitable epitopes of tumor antigens can be retrieved, for example, from cancer/tumor epitope databases, e.g., as described in Vigneron et al. 2013, Cancer Immunol. 13:15; URL: <http://www.cancerimmunity.org/peptide/>, or from the database

"Tantigen" (TANTIGEN version 1.0, Dec 1, 2009; developed by Bioinformatics Core at Cancer Vaccine Center, Dana-Farber Cancer Institute; URL: <http://cvc.dfci.harvard.edu/tadb/>).

- 5 The term "pharmaceutical composition" as used herein refers in particular to preparations which are in such a form as to permit biological activity of the active ingredient(s) to be unequivocally effective and which contain no additional component which would be toxic to subjects to which the said formulation would be administered. In some embodiments, the (pharmaceutical) composition does not contain a further active component (e.g., "active"
- 10 regarding cancer treatment) in addition to the active components described herein below.

As used herein, the term "vaccine" refers to a biological preparation that provides innate and/or adaptive immunity, typically to a particular disease, preferably cancer. Thus, a vaccine supports in particular an innate and/or an adaptive immune response of the immune system

15 of a subject to be treated. For example, the antigens or the multi-antigenic domain as described herein typically leads to or supports an adaptive immune response in the patient to be treated, and the TLR peptide agonist as described herein may lead to or support an innate immune response.

- 20 The term "disease" as used herein is intended to be generally synonymous, and is used interchangeably with, the terms "disorder" and "condition" (as in medical condition), in that all reflect an abnormal condition of the human or animal body or of one of its parts that impairs normal functioning, is typically manifested by distinguishing signs and symptoms, and causes the human or animal to have a reduced duration or quality of life.

25 As used herein, reference to "treatment" of a subject or patient is intended to include prevention, prophylaxis, attenuation, amelioration and therapy. The terms "subject" or "patient" are used interchangeably herein to mean all mammals including humans. Examples of subjects include humans, cows, dogs, cats, horses, goats, sheep, pigs, and rabbits.

30 Preferably, the subject or patient is a human.

Composition

In a first aspect the present invention provides a composition comprising

- (i) at least one tumor antigen, or a fragment or sequence variant thereof,
5 wherein the tumor antigen is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2);
- (ii) a nucleic acid encoding said tumor antigen, or the fragment or sequence variant thereof;
- 10 (iii) an antigen-presenting cell (APC) containing (i) or (ii); or
- (iv) a T-cell expressing either a T-cell receptor or a CAR T-cell receptor targeting said tumor antigen.

The present inventors identified upstream open reading frames (ORFs), i.e., ORFs encoded in
15 the 5'-UTR (5'-untranslated region) of tumor antigens, by analyzing selected Ribo-seq datasets, for example, from the RPFdb v.2.0 database (Wang et al, 2019, Nucleic Acids Res. 47, D230-D234). Thereby, the inventors have surprisingly identified peptides encoded in a 5'-upstream open reading frame (uORF) within the 5'-UTR of KRAS (Kirsten rat sarcoma viral oncogene homolog) mRNAs, peptides encoded in a 5'-upstream open reading frame (uORF)
20 within the 5'-UTR of TPX2 (Targeting protein for Xklp2) mRNAs and peptides encoded in a 5'-upstream open reading frame (uORF) within the 5'-UTR of of AURKA (Aurora A kinase) mRNAs as tumor antigens. As demonstrated in the appended examples, such KRAS, TPX2 and AURKA transcripts were found to be strongly expressed in tumors in contrast to normal healthy tissue. The present inventors also found immunogenic epitopes within the KRAS-
25 uORF1, TPX2-uORF1 and AURKA-uORF2 peptides and that epitopes derived from KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 peptides were presented by human dendritic cells in the context of MHC class I and/or MHC class II. In summary, these findings support the role of the KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 peptides as novel tumor-specific antigens useful in cancer immunotherapy.

30

Accordingly, the present invention provides a composition comprising at least one tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of

the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2); or a fragment or sequence variant thereof. The tumor antigen is typically a peptide tumor antigen. These tumor antigens (i.e., on the peptide level) are also referred to herein as "KRAS-uORF1", "TPX2-uORF1" and "AURKA-uORF2", respectively. The person
 5 skilled in the art would tell immediately from each context whether they are referred to a nucleotide or to a peptide.

In some embodiments, the composition comprises

- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-
 10 upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1),
- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-
 upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-uORF1), and/or
- 15 - a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-
 upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2).

In some embodiments, the composition comprises the tumor antigen encoded in the 5'-
 20 upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 1 or 2. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%,
 25 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
 30 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 2. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

In some embodiments, the composition comprises the tumor antigen encoded in the 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-uORF1) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 3 or 4. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 3. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

In some embodiments, the composition comprises the tumor antigen encoded in the 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 5. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

In addition to the tumor antigen(s) described above, the composition may further comprise at least one tumor antigen selected from the group consisting of CEACAM5 (or a fragment or sequence variant thereof as defined above), DUOXA2 (or a fragment or sequence variant thereof as defined above), and KRAS (or a fragment or sequence variant thereof as defined above). In contrast to the tumor antigens encoded in the 5'-upstream open reading frame

(uORF) within a 5' UTR as described above, the additional antigens CEACAM5, DUOXA2 and KRAS, as referred to herein, are tumor antigens, which are encoded in the ("classical") coding sequence (CDS) of an mRNA (i.e., not in an "untranslated" region/UTR).

5 In some embodiments, the composition comprises a fragment of CEACAM5 (Carcinoembryonic antigen-related cell adhesion molecule 5), which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 6. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence
10 identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 6. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

15 In some embodiments, the composition comprises a fragment of CEACAM5, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 7. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70%
20 sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 7. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

25 In some embodiments, the composition comprises a fragment of CEACAM5, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 8. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70%
30 sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 8. The functionality as tumor antigen is preferably maintained

in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

5 In some embodiments, the fragments of CEACAM5 as described above may be linked to each other (in particular as fusion peptide/protein). Accordingly, the composition may comprise a (fusion) peptide comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 9. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 10 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

15 In some embodiments, the composition comprises a fragment of DUOXA2 (Dual oxidase maturation factor 2), which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 10. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 20 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

25 In some embodiments, the composition comprises a fragment of KRAS. As KRAS mutations are common in cancer, the fragment of KRAS preferably includes a mutation. The KRAS mutation may occur, for example at G12, G13 or Q61 of KRAS. Accordingly, the fragment of KRAS may include positions G12, G13 and/or Q61 of KRAS, preferably wherein at least one of the amino acid residues at these positions is substituted. Preferably, the fragment of KRAS 30 includes a substitution at G12. Non-limiting examples of such substitutions include G12D, G12V, G12C, G12A and G12R. Preferably, the fragment of KRAS includes position G12 of KRAS with a G12D or G12V substitution. Accordingly, KRAS, or the fragment thereof, is

preferably KRAS-G12D, or a fragment thereof, or KRAS-G12V, or a fragment thereof (wherein the fragment includes the G12 position of KRAS and the respective substitution).

In some embodiments, the composition comprises a fragment of KRAS, which may comprise
5 (or consist of) an amino acid sequence according to SEQ ID NO: 11 or 12. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino
10 acid sequence of SEQ ID NO: 11. The composition may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12.
15 The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope).

Preferably, the composition comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and
20 AURKA-uORF2. In some embodiments, the composition comprises (exactly or at least) two different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the composition comprises (exactly or at least) three different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and
25 AURKA-uORF2. In some embodiments, the composition comprises (exactly or at least) four different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the composition comprises (exactly or at least) five different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and
30 AURKA-uORF2. In some embodiments, the composition comprises all of the six different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. A multi-antigenic vaccine will (i) avoid

outgrowth of antigen-loss variants, (ii) target different tumor cells within a heterogeneous tumor mass and (iii) circumvent patient-to-patient tumor variability.

Preferably, the composition comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS-G12D, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the composition comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS-G12V, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the composition comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS-G12D, KRAS-G12V, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2.

Preferably, the tumor antigen selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2, or the fragment or sequence variant thereof, comprises a CD4⁺ and/or a CD8⁺ epitope.

In some embodiments, the composition according to the present invention comprises at least one CD4⁺ epitope and at least one CD8⁺ epitope, which may be included in the same or different tumor antigens (or fragments or variants thereof). T_h cells (CD4⁺ T cells) play a central role in the anti-tumor immune response both in DC (dendritic cell) licensing and in the recruitment and maintenance of CTLs (CD8⁺ T cells, cytotoxic T lymphocytes) at the tumor site. Therefore, a composition according to the present invention comprising at least one CD4⁺ epitope and at least one CD8⁺ epitope provides an integrated immune response allowing simultaneous priming of CTLs and T_h cells and is thus preferable to immunity against only CD8⁺ or only CD4⁺ epitopes.

Preferably, the composition comprises (i) different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2; and (ii) different CD4⁺ epitopes and different CD8⁺ epitopes. Such a composition induces multi-epitopic CD8⁺ CTLs and CD4⁺ T_h cells to function synergistically to counter tumor cells and promote efficient anti-tumor immunity. T_h cells are also involved in the maintenance of long-lasting cellular immunity that was monitored after vaccination.

Such a composition induces polyclonal, multi-epitopic immune responses and poly-functional CD8⁺ and CD4⁺ T cells, and thus efficacious anti-tumor activity.

In some embodiments, the composition comprises the tumor antigen comprises at least one
5 of amino acid sequence, wherein the amino acid sequence is

- an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11;
- 10 - an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12;
- an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof
15 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1, or an amino acid sequence according to SEQ ID NO: 2, or a sequence variant thereof having at least
20 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 2;
- an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%
25 sequence identity to the amino acid sequence of SEQ ID NO: 3, or an amino acid sequence according to SEQ ID NO: 4, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4;
- 30 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%,

86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5;

- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9; and/or

- an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10.

In some embodiments, the composition comprises

- a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11;

- a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12;

- a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1, or a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 2, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 2;

- a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 3, or a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 4, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4;
- a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5;
- a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9; and/or
- a tumor antigen comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10.

It is understood that the tumor antigens (or the fragments or variants thereof) are preferably comprised in the composition as peptides, as described above. Alternatively (or additionally), the composition may comprise a nucleic acid encoding the tumor antigen (or the fragment or variant thereof), as described above.

Nucleic acids preferably comprise single stranded, double stranded or partially double stranded nucleic acids, preferably selected from genomic DNA, cDNA, RNA, siRNA,

antisense DNA, antisense RNA, ribozyme, complimentary RNA/DNA sequences with or without expression elements, a mini-gene, gene fragments, regulatory elements, promoters, and combinations thereof. Further preferred examples of nucleic acid (molecules) and/or polynucleotides include, e.g., a recombinant polynucleotide, a vector, an oligonucleotide, an RNA molecule such as an rRNA, an mRNA, an miRNA, an siRNA, or a tRNA, or a DNA molecule as described above. It is thus preferred that the nucleic acid (molecule) is a DNA molecule or an RNA molecule; preferably selected from genomic DNA; cDNA; siRNA; rRNA; mRNA; antisense DNA; antisense RNA; ribozyme; complimentary RNA and/or DNA sequences; RNA and/or DNA sequences with or without expression elements, regulatory elements, and/or promoters; a vector; and combinations thereof.

In some embodiments, the composition comprises

- a nucleic acid encoding KRAS-uORF1, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 37, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 37;
- a nucleic acid encoding TPX2-uORF1, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 38, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 38;
- a nucleic acid encoding AURKA-uORF2, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 39, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 39;
- a nucleic acid encoding a fragment of CEACAM5, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 40, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 40;

- a nucleic acid encoding a fragment of CEACAM5, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 41, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 41;
- a nucleic acid encoding a fragment of CEACAM5, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 42, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 42;
- a nucleic acid encoding a fragment of DUOXA2, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 43, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 43;
- a nucleic acid encoding a fragment of KRAS-G12D, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 44, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 44; and/or
- a nucleic acid encoding a fragment of KRAS-G12V, the nucleic acid preferably comprising a polynucleotide sequence according to SEQ ID NO: 45, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 45.

The composition may comprise one or more nucleic acid molecules encoding different tumor antigens (or the fragments or variants thereof), as described above, e.g. 2, 3, 4, 5 or all 6 tumor antigens (or the fragments or variants thereof) as described above. The different tumor antigens (or the fragments or variants thereof), as described above, may be encoded by the same nucleic acid molecule (e.g., in a polycistronic manner) or in different nucleic acid molecules.

Alternatively (or additionally), the composition may comprise an antigen-presenting cell (APC) containing (i) the tumor antigen (or the fragment or variant thereof) as described above or (ii) the nucleic acid, as described above, encoding the tumor antigen (or the fragment or variant thereof). In some embodiments, the antigen-presenting cell (APC) may be a dendritic cell (DC).

In some embodiments, the APC may be loaded with the tumor antigen (or the fragment or variant thereof) as described above, such that the APC can present one or more epitopes of the tumor antigen (or the fragment or variant thereof).

The composition may comprise different antigen-presenting cell(s), containing (i) different tumor antigens (or the fragments or variants thereof); or (ii) different nucleic acid molecules as described above. However, a single APC may also contain (i) different tumor antigens (or the fragments or variants thereof); or (ii) different nucleic acid molecules as described above.

Alternatively (or additionally), the composition may comprise a T-cell expressing either a T-cell receptor or a CAR T-cell receptor targeting the tumor antigens (or the fragments or variants thereof), as described above. In some embodiments, the composition may comprise different T-cells, each expressing a different T-cell receptor or CAR T-cell receptor targeting a different tumor antigen (or the fragments or variants thereof), as described above.

In general, the composition may be a pharmaceutical composition and/or a vaccine. In particular, such a composition is preferably a (pharmaceutical) composition which optionally comprises a pharmaceutically acceptable carrier and/or vehicle, or any excipient, buffer, stabilizer or other materials well known to those skilled in the art.

As a further ingredient, the (pharmaceutical) composition may in particular comprise a pharmaceutically acceptable carrier and/or vehicle. In the context of the present invention, a pharmaceutically acceptable carrier typically includes the liquid or non-liquid basis of the (pharmaceutical) composition. If the (pharmaceutical) composition is provided in liquid form, the carrier will typically be pyrogen-free water; isotonic saline or buffered (aqueous) solutions, e.g., phosphate, citrate etc. buffered solutions. Particularly for injection of the

(pharmaceutical) composition, water or preferably a buffer, more preferably an aqueous buffer, may be used, containing a sodium salt, preferably at least 30 mM of a sodium salt, a calcium salt, preferably at least 0.05 mM of a calcium salt, and optionally a potassium salt, preferably at least 1 mM of a potassium salt. According to a preferred embodiment, the sodium, calcium and, optionally, potassium salts may occur in the form of their halogenides, e.g., chlorides, iodides, or bromides, in the form of their hydroxides, carbonates, hydrogen carbonates, or sulfates, etc. Without being limited thereto, examples of sodium salts include e.g. NaCl, NaI, NaBr, Na₂CO₃, NaHCO₃, Na₂SO₄, examples of the optional potassium salts include e.g. KCl, KI, KBr, K₂CO₃, KHCO₃, K₂SO₄, and examples of calcium salts include e.g. CaCl₂, CaI₂, CaBr₂, CaCO₃, CaSO₄, Ca(OH)₂. Furthermore, organic anions of the aforementioned cations may be contained in the buffer. According to a more preferred embodiment, the buffer suitable for injection purposes as defined above, may contain salts selected from sodium chloride (NaCl), calcium chloride (CaCl₂) and optionally potassium chloride (KCl), wherein further anions may be present additional to the chlorides. CaCl₂ can also be replaced by another salt like KCl. Typically, the salts in the injection buffer are present in a concentration of at least 30 mM sodium chloride (NaCl), at least 1 mM potassium chloride (KCl) and at least 0.05 mM calcium chloride (CaCl₂). The injection buffer may be hypertonic, isotonic or hypotonic with reference to the specific reference medium, i.e., the buffer may have a higher, identical or lower salt content with reference to the specific reference medium, wherein preferably such concentrations of the afore mentioned salts may be used, which do not lead to damage of cells due to osmosis or other concentration effects. Reference media are e.g., liquids occurring in "*in vivo*" methods, such as blood, lymph, cytosolic liquids, or other body liquids, or e.g. liquids, which may be used as reference media in "*in vitro*" methods, such as common buffers or liquids. Such common buffers or liquids are known to a skilled person. Saline (0.9% NaCl) and Ringer-Lactate solution are particularly preferred as a liquid basis. In some embodiments, the (pharmaceutical) composition further comprises arginine, such as L-arginine.

In this context, prescription of treatment, e.g., decisions on dosage etc. when using the above (pharmaceutical) composition is typically within the responsibility of general practitioners and other medical doctors, and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other

factors known to practitioners. Accordingly, the (pharmaceutical) composition typically comprises a therapeutically effective amount of the active components (the tumor antigen(s)). The (pharmaceutical) composition may be used for human and also for veterinary medical purposes, preferably for human medical purposes, as a (pharmaceutical) composition in
5 general or as a vaccine.

Examples of suitable adjuvants and/or immunomodulatory materials in the context of the present invention include MPL® (Corixa), aluminum-based minerals including aluminum compounds (generically called Alum), ASO1-4, MF59, CalciumPhosphate, Liposomes,
10 Iscom, polyinosinic:polycytidylic acid (polyIC), including its stabilized form poly-ICLC (Hiltonol), CpG oligodeoxynucleotides, Granulocyte-macrophage colony-stimulating factor (GM-CSF), lipopolysaccharide (LPS), Montanide, polylactide co-glycolide (PLG), Flagellin, Soap Bark tree saponins (QS21), amino alkyl glucosamide compounds (e.g. RC529), two component antibacterial peptides with synthetic oligodeoxynucleotides (e.g. IC31),
15 Imiquimod, Resiquimod, Immunostimulatory sequences (ISS), monophosphoryl lipid A (MPLA), and Fibroblast-stimulating lipopeptide (FSL1).

Further materials as well as formulation processing techniques and the like, which are useful in the context of compositions, in particular pharmaceutical compositions and vaccines, or
20 in the context of their preparation are known to the skilled artisan.

Peptide

- 25 In a further aspect the present invention provides a peptide comprising:
- a cell penetrating peptide;
 - a multi-antigenic domain comprising at least one tumor antigen, or a fragment or sequence variant thereof,
wherein the tumor antigen is encoded in a 5'-upstream open reading frame (uORF)
30 within the 5' UTR of the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2); and
 - a TLR peptide agonist.

In the peptide of the present invention, the components a) – c) are covalently linked, usually with a peptide bond. Accordingly, said peptide is a recombinant peptide (not occurring in nature) or a “fusion” peptide (wherein the components a) – c) are “fused” to each other).

5

Such a peptide provides simultaneous (i) stimulation of multi-epitopic cytotoxic T cell-mediated immunity, (ii) induction of T_h cells and (iii) promotion of immunological memory. Thereby, the peptide according to the present invention provides a potent vaccine, in particular having improved anti-tumor activity.

10

Multi-antigenic domain

The peptide of the invention comprises a multi-antigenic domain. As used herein the term “multi-antigenic domain” refers to a domain, such as a peptide, comprising at least two (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or more) distinct antigens or fragments of at least two (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or more) distinct antigens. In some embodiments, the multi-antigenic domain comprises three or more different antigens, or fragments thereof, in particular 3, 4, 5, 6, 7, 8, 9, 10 or more different antigens or fragments thereof. Preferably, the “multi-antigenic domain” comprises (fragments of) at least two (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or more) distinct antigens, wherein each fragment or antigen comprises at least one antigenic epitope. More preferably, the “multi-antigenic domain” comprises (fragments of) at least two to eight distinct antigens, wherein each fragment/antigen comprises at least one antigenic epitope. Even more preferably, the “multi-antigenic domain” comprises (fragments of) five or six distinct antigens, wherein each fragment comprises at least one antigenic epitope.

25

In particular, the different (fragments of the) antigens are positioned consecutively in the multi-antigenic domain. In some embodiments, the different (fragments of the) antigens are linked to each other for example by a peptide spacer or linker (e.g., a GS-linker). The spacer or linker is usually neither component a), i.e., the cell penetrating peptide, nor component c), i.e. the TLR peptide agonist. In other embodiments, the different (fragments of the) antigens are directly linked to each other, i.e., without spacer or linker.

30

In some embodiments, each antigen, or fragment thereof, in the multi-antigenic domain comprises at least one CD4⁺ epitope and/or at least one CD8⁺ epitope, which may be included in the same or different tumor antigens (or fragments or variants thereof). T_h cells (CD4⁺ T cells) play a central role in the anti-tumor immune response both in DC (dendritic cell) licensing and in the recruitment and maintenance of CTLs (CD8⁺ T cells, cytotoxic T lymphocytes) at the tumor site. Therefore, a multi-antigenic domain comprising at least one CD4⁺ epitope and at least one CD8⁺ epitope provides an integrated immune response allowing simultaneous priming of CTLs and T_h cells and is thus preferable to immunity against only CD8⁺ or only CD4⁺ epitopes.

The multi-antigenic domain of the peptide comprises at least one tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2); or a fragment or sequence variant thereof as defined above. The tumor antigen is typically a peptide tumor antigen. These tumor antigens (i.e., on the peptide level) are also referred to herein as "KRAS-uORF1", "TPX2-uORF1" and "AURKA-uORF2", respectively. The person skilled in the art would tell immediately from each context whether they are referred to a nucleotide and/or to a peptide.

In some embodiments, the multi-antigenic domain comprises

- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1),
- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-uORF1) and/or
- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2).

In some embodiments, the multi-antigenic domain comprises the tumor antigen encoded in the 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 1 or 2. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 2. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 1 or 2, or a sequence variant thereof.

In some embodiments, the multi-antigenic domain comprises the tumor antigen encoded in the 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (uORF-TPX2) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 3 or 4. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 3. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at

least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 3 or 4, or a sequence variant thereof.

In some embodiments, the multi-antigenic domain comprises the tumor antigen encoded in
5 the 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 5. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%,
10 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 5, or a sequence
15 variant thereof.

In addition to the tumor antigen(s) described above, the multi-antigenic domain may further comprise at least one tumor antigen selected from the group consisting of CEACAM5 (or a fragment or sequence variant thereof as defined above), DUOXA2 (or a fragment or sequence
20 variant thereof as defined above), and KRAS (or a fragment or sequence variant thereof as defined above).

In some embodiments, the multi-antigenic domain comprises (a fragment of) CEACAM5 (Carcinoembryonic antigen-related cell adhesion molecule 5), which may comprise (or
25 consist of) an amino acid sequence according to SEQ ID NO: 6. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino
30 acid sequence of SEQ ID NO: 6. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at

least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 6, or a sequence variant thereof.

In some embodiments, the multi-antigenic domain comprises (a fragment of) CEACAM5, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 7. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 7. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 7, or a sequence variant thereof.

In some embodiments, the multi-antigenic domain comprises (a fragment of) CEACAM5, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 8. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 8. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 8, or a sequence variant thereof.

In some embodiments, the fragments of CEACAM5 as described above may be linked to each other (in particular as fusion peptide/protein). Accordingly, the multi-antigenic domain may comprise a (fusion) peptide comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 9. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids

or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant
5 (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof.

In some embodiments, the multi-antigenic domain comprises (a fragment of) DUOXA2 (Dual
10 oxidase maturation factor 2), which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 10. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
15 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof.

20

In some embodiments, the multi-antigenic domain comprises (a fragment of) KRAS. As KRAS mutations are common in cancer, the (fragment of) KRAS preferably includes a mutation. The KRAS mutation may occur, for example at G12, G13 or Q61 of KRAS. Accordingly, the fragment of KRAS may include positions G12, G13 and/or Q61 of KRAS, preferably wherein
25 at least one of the amino acid residues at these positions is substituted. Preferably, the fragment of KRAS includes a substitution at G12. Non-limiting examples of such substitutions include G12D, G12V, G12C, G12A and G12R. Preferably, the fragment of KRAS includes position G12 of KRAS with a G12D or G12V substitution. Accordingly, KRAS, or the fragment thereof, is preferably KRAS-G12D, or a fragment thereof, or KRAS-G12V, or a fragment
30 thereof (wherein the fragment includes the G12 position of KRAS and the respective substitution).

In some embodiments, the multi-antigenic domain comprises a fragment of KRAS, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 11 or 12. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant
5 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence
10 identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino
15 acid sequence according to SEQ ID NO: 11 or 12, or a sequence variant thereof.

Accordingly, the multi-antigenic domain may comprise at least one amino acid sequence selected from the group consisting of:

- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof
20 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9;
- an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof
25 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12;
- an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof
30 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11;
- an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof
having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%,

86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1, or an amino acid sequence according to SEQ ID NO: 2, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 2;

- an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 3, or an amino acid sequence according to SEQ ID NO: 4, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4;

- an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5; or
- an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10.

Preferably, the multi-antigenic domain comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) two different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) three different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) four different tumor antigens, or fragments or variants thereof,

selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) five different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises all of the six different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. A multi-antigenic vaccine will (i) avoid outgrowth of antigen-loss variants, (ii) target different tumor cells within a heterogeneous tumor mass and (iii) circumvent patient-to-patient tumor variability.

Preferably, the multi-antigenic domain comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS-G12D, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the composition comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS-G12V, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the composition comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS-G12D, KRAS-G12V, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2.

Preferably, the tumor antigen selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2, or the fragment or sequence variant thereof, comprises a CD4⁺ and/or a CD8⁺ epitope. Preferably, the multi-antigenic domain comprises (i) different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2; and (ii) different CD4⁺ epitopes and different CD8⁺ epitopes. A peptide comprising such a multi-antigenic domain induces multi-epitopic CD8⁺ CTLs and CD4⁺ T_h cells to function synergistically to counter tumor cells and promote efficient anti-tumor immunity. T_h cells are also involved in the maintenance of long-lasting cellular immunity that was monitored after vaccination. A peptide comprising such a multi-antigenic domain induces polyclonal, multi-epitopic immune responses and poly-functional CD8⁺ and CD4⁺ T cells, and thus efficacious anti-tumor activity.

In some embodiments, the multi-antigenic domain comprises, preferably in N- to C-terminal direction:

- CEACAM5 or a fragment or sequence variant thereof,
 - KRAS or a fragment or sequence variant thereof,
 - 5 - KRAS-uORF1, or a fragment or sequence variant thereof,
 - TPX2-uORF1, or a fragment or sequence variant thereof,
 - AURKA-uORF2, or a fragment or sequence variant thereof, and
 - DUOXA2 or a fragment or sequence variant thereof.
- 10 Preferably, the multi-antigenic domain comprises, preferably in N- to C-terminal direction:
- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9;
 - 15 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11, and/or an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least
 - 20 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12;
 - an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%,
 - 25 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1, or an amino acid sequence according to SEQ ID NO: 2, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence
 - 30 identity to the amino acid sequence of SEQ ID NO: 2;
 - an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%,

86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 3, or an amino acid sequence according to SEQ ID NO: 4, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4;

- an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5; and
- an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10.

More preferably, the multi-antigenic domain comprises (or consists of) an amino acid sequence according to SEQ ID NO: 13, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 13. It is also more preferred that the multi-antigenic domain comprises (or consists of) an amino acid sequence according to SEQ ID NO: 14, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 14.

Cell penetrating peptide

The cell penetrating peptide (CPP) allows for efficient delivery, i.e., transport and loading, in particular of the multi-antigenic domain, into the antigen presenting cells (APCs), in particular into the dendritic cells (DCs) and thus to the dendritic cells' antigen processing machinery.

The term "cell penetrating peptide" ("CPP") is generally used to designate short peptides that are able to transport different types of cargo molecules across plasma membrane, and, thus, facilitate cellular uptake of various molecular cargoes (from nanosize particles to small chemical molecules and large fragments of DNA). "Cellular internalization" of the cargo molecule linked to the cell penetrating peptide generally means transport of the cargo molecule across the plasma membrane and thus entry of the cargo molecule into the cell. Depending on the particular case, the cargo molecule can, then, be released in the cytoplasm, directed to an intracellular organelle, or further presented at the cell surface. Cell penetrating ability, or internalization, of the cell penetrating peptide or of the peptide (comprising said cell penetrating peptide) according to the invention can be checked by standard methods known to one skilled in the art, including flow cytometry or fluorescence microscopy of live and fixed cells, immunocytochemistry of cells transduced with said peptide, and Western blot.

Usually, cell penetrating peptides (CPPs) have a length of 8 to 50 residues. Cell penetrating peptides typically have an amino acid composition that either contains a high relative abundance of positively charged amino acids such as lysine or arginine or have a sequence that contains an alternating pattern of polar/charged amino acids and non-polar, hydrophobic amino acids. These two types of structures are referred to as polycationic or amphipathic, respectively. Cell-Penetrating peptides are of different sizes, amino acid sequences, and charges, but all CPPs have a common characteristic that is the ability to translocate the plasma membrane and facilitate the delivery of various molecular cargoes to the cytoplasm or to an organelle of a cell. Cell-penetrating peptides have found numerous applications in medicine as drug delivery agents in the treatment of different diseases including cancer and virus inhibitors, as well as contrast agents for cell labeling and imaging. Various CPPs, which can be used as cell penetrating peptide, i.e., as component a), in the peptide according to the present invention, are also disclosed in the review: Milletti, F., Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov Today* 17 (15-16): 850-60, 2012.

Preferably, the cell penetrating peptide comprised in the peptide of the present invention has an amino acid sequence comprising or consisting of an amino acid sequence according to

SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 22, or SEQ ID NO: 23, or is a sequence variant thereof as described above.

5 In some embodiments, the cell penetrating peptide according to the invention has an amino acid sequence comprising or consisting of SEQ ID NO: 15, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 15.

10 In some embodiments, the cell penetrating peptide according to the invention has an amino acid sequence comprising or consisting of SEQ ID NO: 21, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 21.

15 In some embodiments, the cell penetrating peptide according to the invention has an amino acid sequence comprising or consisting of SEQ ID NO: 22, or sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 22.

20 In some embodiments, the cell penetrating peptide according to the invention has an amino acid sequence comprising or consisting of SEQ ID NO: 23, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 23.

25 These cell penetrating peptides are derived from the "ZEBRA" protein of the Epstein-Barr virus (EBV). "ZEBRA" (also known as Zta, Z, EB1, or BZLF1) generally refers to the basic-leucine zipper (bZIP) transcriptional activator of the Epstein-Barr virus (EBV). Without being bound to any theory, it is assumed that these cell penetrating peptides promote both, MHC class I and II restricted presentation of cargo antigens to CD8⁺ and CD4⁺ T cells, respectively.
30 Accordingly, such a CPP can deliver multi-epitopic peptides to dendritic cells (DCs), and subsequently promote CTL and Th cell activation and anti-tumor function. Such a CPP can

thus efficiently deliver the peptide according to the present invention to antigen presenting cells (APCs) and lead to multi-epitopic MHC class I and II restricted presentation.

Preferably, the cell penetrating peptide has an amino acid sequence comprising or consisting
5 of an amino acid sequence according to SEQ ID NO: 15, or is a sequence variant thereof as described above.

It will be understood by one skilled in the art that the primary amino acid sequence of the cell penetrating peptide may further be post-translationally modified, such as by glycosylation
10 or phosphorylation, without departing from the invention.

In certain embodiments, the cell penetrating peptide optionally further comprises, in addition to its amino acid sequence as described above, any one of, or any combination of:

- 15 (i) a nuclear localization signal (NLS). Such signals are well known to the skilled person and are described in Nair et al. (2003, Nucleic Acids Res. 31(1): 397–399); and/or
- (ii) a targeting peptide, including tumor homing peptides such as those described in Kapoor et al. (2012, PLoS ONE 7(4): e35187).

20 Preferably, the cell penetrating peptide is (directly) linked to multi-antigenic domain and facilitates the cellular internalization of the multi-antigenic domain.

TLR peptide agonist

25 In the peptide according to the present invention, the TLR peptide agonist allows an increased targeting of the vaccine towards dendritic cells along with self-adjuvancity. Physical linkage of a TLR peptide agonist to the CPP and the at least one antigen or antigenic epitope in the peptide according to the present invention provides an enhanced immune response by
30 simultaneous stimulation of antigen presenting cells, in particular dendritic cells, that internalize, metabolize and display antigen(s).

As used herein, a "TLR peptide agonist" is an agonist of a Toll-like receptor (TLR), i.e., it binds to a TLR and activates the TLR, in particular to produce a biological response. Moreover, the TLR peptide agonist is a peptide as defined above. Preferably, the TLR peptide agonist comprises from 10 to 150 amino acids, more preferably from 15 to 130 amino acids, even
5 more preferably from 20 to 120 amino acids, particularly preferably from 25 to 110 amino acids, and most preferably from 30 to 100 amino acids.

Toll like receptors (TLRs) are transmembrane proteins that are characterized by extracellular, transmembrane, and cytosolic domains. The extracellular domains containing leucine-rich
10 repeats (LRRs) with horseshoe-like shapes are involved in recognition of common molecular patterns derived from diverse microbes. Toll like receptors include TLRs1 - 10. Compounds capable of activating TLR receptors and modifications and derivatives thereof are well documented in the art. TLR1 may be activated by bacterial lipoproteins and acetylated forms thereof, TLR2 may in addition be activated by Gram positive bacterial glycolipids, LPS, LP A,
15 LTA, fimbriae, outer membrane proteins, heat shock proteins from bacteria or from the host, and Mycobacterial lipoarabinomannans. TLR3 may be activated by dsRNA, in particular of viral origin, or by the chemical compound poly(LC). TLR4 may be activated by Gram negative LPS, LTA, Heat shock proteins from the host or from bacterial origin, viral coat or envelope proteins, taxol or derivatives thereof, hyaluronan containing oligosaccharides and
20 fibronectins. TLR5 may be activated with bacterial flagellae or flagellin. TLR6 may be activated by mycobacterial lipoproteins and group B streptococcus heat labile soluble factor (GBS- F) or staphylococcus modulins. TLR7 may be activated by imidazoquinolines. TLR9 may be activated by unmethylated CpG DNA or chromatin - IgG complexes.

25 Preferably, the TLR peptide agonist comprised by the peptide according to the present invention is an agonist of TLR1, 2, 4, 5, 6, and/or 10. TLRs are expressed either on the cell surface (TLR1, 2, 4, 5, 6, and 10) or on membranes of intracellular organelles, such as endosomes (TLR3, 4, 7, 8, and 9). The natural ligands for the endosomal receptors turned out to be nucleic acid-based molecules (except for TLR4). The cell surface-expressed TLR1, 2, 4,
30 5, 6, and 10 recognize molecular patterns of extracellular microbes (Monie, T. P., Bryant, C. E., et al. 2009: Activating immunity: Lessons from the TLRs and NLRs. Trends Biochem. Sci.

34(11), 553–561). TLRs are expressed on several cell types but virtually all TLRs are expressed on DCs allowing these specialized cells to sense all possible pathogens and danger signals.

However, TLR2, 4, and 5 are constitutively expressed at the surface of DCs. Accordingly, the
5 TLR peptide agonist comprised by the peptide according to the present invention is more preferably a peptide agonist of TLR2, TLR4 and/or TLR5. Even more preferably, the TLR peptide agonist is a TLR2 peptide agonist and/or a TLR4 peptide agonist.

TLR2 can detect a wide variety of ligands derived from bacteria, viruses, parasites, and fungi.
10 TLR2 interacts with a broad and structurally diverse range of ligands, including molecules expressed by microbes and fungi.

A preferred TLR2 peptide agonist is annexin II or an immunomodulatory fragment thereof (having TLR agonist functionality), which is described in detail in WO 2012/048190 A1 and
15 US patent application 13/0331546, in particular a TLR2 peptide agonist comprising an amino acid sequence according to SEQ ID NO: 7 of WO 2012/048190 A1 or fragments or variants thereof are preferred.

A TLR2 peptide agonist comprising or consisting of an amino acid sequence according to
20 SEQ ID NO: 16 or 24; or a sequence variant thereof (which is, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 16, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%,
25 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 24) is preferred as component c), i.e. as the TLR peptide agonist, comprised by the peptide. Particularly preferably, the TLR peptide agonist has an amino acid sequence according to SEQ ID NO: 16.

30 A diversity of ligands interact with TLR4, including Monophosphoryl Lipid A from *Salmonella minnesota* R595 (MPLA), lipopolysaccharides (LPS), mannans (*Candida albicans*), glycoinositolphospholipids (*Trypanosoma*), viral envelope proteins (RSV and MMTV) and

endogenous antigens including fibrinogen and heat-shock proteins. Preferred TLR4 peptide agonists correspond to motifs that bind to TLR4, in particular (i) peptides mimicking the natural LPS ligand (RS01: Gln-Glu-Ile-Asn-Ser-Ser- Tyr and RS09: Ala-Pro-Pro-His-Ala-Leu-Ser) and (ii) Fibronectin derived peptides. The cellular glycoprotein Fibronectin (FN) has multiple isoforms generated from a single gene by alternative splicing of three exons. One of these isoforms is the extra domain A (EDA), which interacts with TLR4. Accordingly, suitable TLR peptide agonists comprise a fibronectin EDA domain or a fragment or variant thereof. Such suitable fibronectin EDA domains or a fragments or variants thereof are disclosed in EP 1 913 954 B1, EP 2 476 440 A1, US 2009/0220532 A1, and WO 2011/101332 A1.

In some embodiments, the peptide according to the present invention comprises

- a) a cell penetrating peptide having an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 22, or SEQ ID NO: 23; or sequence variants thereof having at least 70% sequence identity;
- b) a multi-antigenic domain comprising:
 - a fragment of CEACAM5 having a minimum length of 8 amino acids,
 - a fragment of DUOXA2 having a minimum length of 8 amino acids,
 - a fragment of KRAS having a minimum length of 8 amino acids,
 - KRAS-uORF1, or a fragment thereof having a minimum length of 8 amino acids,
 - TPX2-uORF1, or a fragment thereof having a minimum length of 8 amino acids, and/or
 - AURKA-uORF2, or a fragment thereof having a minimum length of 8 amino acids; and
- c) a TLR peptide agonist, which is a TLR2 peptide agonist and/or a TLR4 peptide agonist.

The different components a), b) and c) of the peptide may be directly or indirectly linked (e.g., by a peptide bond). In other words, two components may directly adjoin or they may be linked by an additional component of the peptide, e.g. a peptide spacer or a linker. Preferably, a peptidic spacer consists of about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids, more preferably of about 1, 2, 3, 4, or 5 amino acids. The amino acid sequence of the peptidic spacer may be identical to that of the N-terminal or C-terminal flanking region of any of the components a),

b), or c). Alternatively a peptidic spacer can consist of non-natural amino acid sequences such as an amino acid sequence resulting from conservative amino acid substitutions of said natural flanking regions or sequences of known cleavage sites for proteases. In some embodiments, the peptidic spacer does not contain any Cys (C) residues. In some
 5 embodiments the linker sequence contains at least 20%, more preferably at least 40% and even more preferably at least 50% Gly or β -alanine residues. Appropriate linker sequences can be easily selected and prepared by a person skilled in the art. They may be composed of D and/or L amino acids.

10 Preferably, all three components a), b), and c) are linked via main-chain/main-chain linkage, thus resulting in a main chain of the peptide, which comprises the main chain of the cell penetrating peptide, the main chain of the multi-antigenic domain and the main chain of the TLR peptide agonist. In other words, the main chain of the cell penetrating peptide, the main chain of the multi-antigenic domain and the main chain of the TLR peptide agonist constitute
 15 the main chain of the peptide, optionally together with further components, for example linker(s) or spacer(s). Accordingly, the following arrangements of the components a), b), and c) are preferred, wherein said preferred arrangements are shown below in N-terminal $\rightarrow$ C-terminal direction of the main chain of the peptide and wherein all three components a), b), and c) are linked via main-chain/main-chain linkage and may be optionally linked by a linker,
 20 a spacer or another additional component:

- (α) component a) (cell penetrating peptide) – component b) (multi-antigenic domain) – component c) (TLR peptide agonist);
- (β) component c) (TLR peptide agonist) – component a) (cell penetrating peptide) – component b) (multi-antigenic domain);
- 25 (γ) component a) (cell penetrating peptide) – component c) (TLR peptide agonist) – component b) (multi-antigenic domain);
- (δ) component c) (TLR peptide agonist) – component b) (multi-antigenic domain) – component a) (cell penetrating peptide);
- (ϵ) component b) (multi-antigenic domain) – component a) (cell penetrating peptide) –
 30 component c) (TLR peptide agonist); or

(ζ) component b) (multi-antigenic domain) – component c) (TLR peptide agonist) – component a) (cell penetrating peptide).

Preferably, the multi-antigenic domain is positioned C-terminally of the cell penetrating peptide. More preferably, the components a), b) and c) are positioned in N-terminal to C-terminal direction of the main chain of said peptide in the order:

(α) component a) – component b) – component c); or

(β) component c) – component a) – component b),

wherein the components may be linked by a further component, in particular by a linker or a spacer.

A preferred exemplified peptide of the invention comprises, preferably in N-terminal to C-terminal direction:

a) a cell penetrating peptide having an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 15, or sequence variants thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 15;

b) a multi-antigenic domain comprising

- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9;

- an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11, and/or an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%,

94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12;

5 - an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1;

10 - an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 3;

15 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5; and

20 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10; and

25 c) a TLR peptide agonist having an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 16, or sequence variants thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 16.

30 In some embodiments, the peptide of the invention may comprise or consist of an amino acid sequence according to SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 46, or SEQ ID NO: 47; or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%,

97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 46, or SEQ ID NO: 47. For example, the peptide of the invention may comprise or consist of an amino acid sequence according to SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 46, or SEQ ID NO: 47.

- 5 Preferably, the peptide of the invention may comprise or consist of an amino acid sequence according to SEQ ID NO: 17 or SEQ ID NO: 18; or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 17 or SEQ ID NO: 18. For example, the peptide
10 of the invention may comprise or consist of an amino acid sequence according to SEQ ID NO: 17 or 18.

- The present invention also provides a (pharmaceutical) composition comprising the peptide of the invention as described above. The (pharmaceutical) composition may comprise a
15 pharmaceutically acceptable carrier and/or vehicle, or any excipient, buffer, stabilizer or other materials well known to those skilled in the art. Regarding such further ingredients, i.e. carriers, vehicles, excipients, buffers, stabilizers, adjuvants and the like, the detailed description of the composition according to the present invention above applies accordingly to the (pharmaceutical) composition comprising the peptide of the invention. The
20 (pharmaceutical) composition typically comprises a therapeutically effective amount of the active components (the peptide). The (pharmaceutical) composition may be used for human and also for veterinary medical purposes, preferably for human medical purposes, as a (pharmaceutical) composition in general or as a vaccine.

- 25 The present invention also provides a nucleic acid (molecule) comprising a polynucleotide encoding the peptide of the invention as described above. Nucleic acids preferably comprise single stranded, double stranded or partially double stranded nucleic acids. Preferably, the nucleic acid (molecule) is selected from genomic DNA, cDNA, RNA, siRNA, mRNA, antisense DNA, antisense RNA, ribozyme, complimentary RNA/DNA sequences. It may (or
30 may not) contain expression elements, a mini-gene, gene fragments, regulatory elements, promoters, and combinations thereof. Further preferred examples of nucleic acid (molecules) and/or polynucleotides include, e.g., a recombinant polynucleotide, a vector, an

oligonucleotide, an RNA molecule such as an rRNA, an mRNA, an miRNA, an siRNA, or a tRNA, or a DNA molecule. It is thus preferred that the nucleic acid (molecule) is a DNA molecule or an RNA molecule; preferably selected from genomic DNA; cDNA; siRNA; rRNA; mRNA; antisense DNA; antisense RNA; ribozyme; complimentary RNA and/or DNA sequences; RNA and/or DNA sequences with or without expression elements, regulatory elements, and/or promoters; a vector; and combinations thereof. In some embodiments, the nucleic acid (molecule) may be a vector, e.g. an expression vector for expression of the peptide of the invention.

Vesicular stomatitis virus (VSV)

In a further aspect the present invention provides a recombinant vesicular stomatitis virus (VSV) encoding a multi-antigenic domain comprising at least one tumor antigen, or a fragment or sequence variant thereof, wherein the tumor antigen is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2).

Vesicular stomatitis viruses (VSV) belong to rhabdoviruses (family rhabdoviridae) and are negative strand RNA-viruses. The genome of VSV is on a single molecule of negative-sense RNA molecule that encodes five major proteins: G protein (G), large protein (L), phosphoprotein (P), matrix protein (M) and nucleoprotein (N). In some embodiments, the vesicular stomatitis virus is Vesicular stomatitis Indiana virus (VSIV) or Vesicular stomatitis New Jersey virus (VSNJV).

Multi-antigenic domain

The VSV of the invention is a recombinant VSV, which comprises a multi-antigenic domain defined as above, in the context of the peptide. Accordingly, it refers to a domain, such as a peptide, comprising at least two (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or more) distinct antigens or fragments of at least two (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or more) distinct antigens. In some

embodiments, the multi-antigenic domain comprises three or more different antigens, or fragments thereof, in particular 3, 4, 5, 6, 7, 8, 9, 10 or more different antigens or fragments thereof. Preferably, the “multi-antigenic domain” comprises (fragments of) at least two (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or more) distinct antigens, wherein each fragment or antigen comprises at least one antigenic epitope. More preferably, the “multi-antigenic domain” comprises (fragments of) at least two to eight distinct antigens, wherein each fragment/antigen comprises at least one antigenic epitope. Even more preferably, the “multi-antigenic domain” comprises (fragments of) five or six distinct antigens, wherein each fragment comprises at least one antigenic epitope.

In particular, the different (fragments of the) antigens are positioned consecutively in the multi-antigenic domain. In some embodiments, the different (fragments of the) antigens are linked to each other for example by a peptide spacer or linker (e.g., a GS-linker). In other embodiments, the different (fragments of the) antigens are directly linked to each other, i.e., without spacer or linker.

In some embodiments, each antigen, or fragment thereof, in the multi-antigenic domain comprises at least one CD4⁺ epitope and/or at least one CD8⁺ epitope, which may be included in the same or different tumor antigens (or fragments or variants thereof). T_h cells (CD4⁺ T cells) play a central role in the anti-tumor immune response both in DC (dendritic cell) licensing and in the recruitment and maintenance of CTLs (CD8⁺ T cells, cytotoxic T lymphocytes) at the tumor site. Therefore, a multi-antigenic domain comprising at least one CD4⁺ epitope and at least one CD8⁺ epitope provides an integrated immune response allowing simultaneous priming of CTLs and T_h cells and is thus preferable to immunity against only CD8⁺ or only CD4⁺ epitopes.

The multi-antigenic domain of the VSV comprises at least one tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2); or a fragment or sequence variant thereof as defined above. The tumor antigen is typically a peptide tumor antigen. These tumor antigens (i.e. on the peptide level) are also referred to herein as “KRAS-uORF1”, “TPX2-uORF1” and “AURKA-uORF2”, respectively. The person

skilled in the art would tell immediately from each context whether they are referred to a nucleotide and/or to a peptide.

In some embodiments, the multi-antigenic domain comprises

- 5 - a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1),
- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-
- 10 uORF1) and/or
- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2).

15 In some embodiments, the multi-antigenic domain comprises the tumor antigen encoded in the 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 1 or 2. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence

20 variant having at least 70% sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 1 or 2, or a sequence variant thereof.

25 In some embodiments, the multi-antigenic domain comprises the tumor antigen encoded in the 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (uORF-TPX2) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 3 or 4. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence

30 variant having at least 70% sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant

contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 3 or 4, or a sequence variant thereof.

5 In some embodiments, the multi-antigenic domain comprises the tumor antigen encoded in the 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2) comprises (or consists of) an amino acid sequence according to SEQ ID NO: 5. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity. The functionality as tumor antigen is preferably
10 maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof.

15 In addition to the tumor antigen(s) described above, the multi-antigenic domain may further comprise at least one tumor antigen selected from the group consisting of CEACAM5 (or a fragment or sequence variant thereof as defined above), DUOXA2 (or a fragment or sequence variant thereof as defined above), and KRAS (or a fragment or sequence variant thereof as defined above).

20 In some embodiments, the multi-antigenic domain comprises (a fragment of) CEACAM5 (Carcinoembryonic antigen-related cell adhesion molecule 5), which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 6. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined above, i.e., a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70%
25 sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 6, or a sequence variant thereof.

30 In some embodiments, the multi-antigenic domain comprises (a fragment of) CEACAM5, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 7. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined

above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises
5 an amino acid sequence according to SEQ ID NO: 7, or a sequence variant thereof.

In some embodiments, the multi-antigenic domain comprises (a fragment of) CEACAM5, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 8. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined
10 above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 8, or a sequence variant thereof.

15 In some embodiments, the fragments of CEACAM5 as described above may be linked to each other (in particular as fusion peptide/protein). Accordingly, the multi-antigenic domain may comprise a (fusion) peptide comprising (or consisting of) an amino acid sequence according to SEQ ID NO: 9. The multi-antigenic domain may also comprise a fragment or sequence
20 variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 9, or a sequence
25 variant thereof.

In some embodiments, the multi-antigenic domain comprises (a fragment of) DUOXA2 (Dual oxidase maturation factor 2), which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 10. The multi-antigenic domain may also comprise a fragment or
30 sequence variant thereof as defined above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that

the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof.

5 In some embodiments, the multi-antigenic domain comprises (a fragment of) KRAS. As KRAS mutations are common in cancer, the (fragment of) KRAS preferably includes a mutation. The KRAS mutation may occur, for example at G12, G13 or Q61 of KRAS. Accordingly, the fragment of KRAS may include positions G12, G13 and/or Q61 of KRAS, preferably wherein at least one of the amino acid residues at these positions is substituted. Preferably, the
10 fragment of KRAS includes a substitution at G12. Non-limiting examples of such substitutions include G12D, G12V, G12C, G12A and G12R. Preferably, the fragment of KRAS includes position G12 of KRAS with a G12D or G12V substitution. Accordingly, KRAS, or the fragment thereof, is preferably KRAS-G12D, or a fragment thereof, or KRAS-G12V, or a fragment thereof (wherein the fragment includes the G12 position of KRAS and the respective
15 substitution).

In some embodiments, the multi-antigenic domain comprises a fragment of KRAS, which may comprise (or consist of) an amino acid sequence according to SEQ ID NO: 11 or 12. The multi-antigenic domain may also comprise a fragment or sequence variant thereof as defined
20 above, i.e. a fragment having a minimum length of 8 amino acids or a sequence variant having at least 70% sequence identity. The functionality as tumor antigen is preferably maintained in the fragment or sequence variant (e.g., in that the fragment or sequence variant contains at least one epitope). Accordingly, the multi-antigenic domain preferably comprises an amino acid sequence according to SEQ ID NO: 11 or 12, or a sequence variant thereof.

25 Accordingly, the multi-antigenic domain may comprise at least one amino acid sequence selected from the group consisting of:

- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%,
30 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9;

- an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12;
- 5 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11;
- an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof
- 10 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1, or an amino acid sequence according to SEQ ID NO: 2, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%,
- 15 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 2;
- an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%
- 20 sequence identity to the amino acid sequence of SEQ ID NO: 3, or an amino acid sequence according to SEQ ID NO: 4, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4;
- 25 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5; or
- an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof
- 30 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10.

Preferably, the multi-antigenic domain comprises different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) two different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) three different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) four different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises (exactly or at least) five different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain comprises all of the six different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. A multi-antigenic vaccine will (i) avoid outgrowth of antigen-loss variants, (ii) target different tumor cells within a heterogeneous tumor mass and (iii) circumvent patient-to-patient tumor variability.

Preferably, the tumor antigen selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2, or the fragment or sequence variant thereof, comprises a CD4⁺ and/or a CD8⁺ epitope. Preferably, the multi-antigenic domain comprises (i) different tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2; and (ii) different CD4⁺ epitopes and different CD8⁺ epitopes.

In some embodiments, the multi-antigenic domain comprises, preferably in N- to C-terminal direction:

- CEACAM5 or a fragment or sequence variant thereof,
- KRAS or a fragment or sequence variant thereof,
- KRAS-uORF1, or a fragment or sequence variant thereof,

- TPX2-uORF1, or a fragment or sequence variant thereof,
- AURKA-uORF2, or a fragment or sequence variant thereof, and
- DUOXA2 or a fragment or sequence variant thereof.

5 Preferably, the multi-antigenic domain comprises, preferably in N- to C-terminal direction:

- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 9;
- 10 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 11, and/or an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least
- 15 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 12;
- an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%,
- 20 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 1, or an amino acid sequence according to SEQ ID NO: 2, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence
- 25 identity to the amino acid sequence of SEQ ID NO: 2;
- an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 2, or an amino acid
- 30 sequence according to SEQ ID NO: 4, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%,

88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 4;

- an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 5; and

- an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 10.

More preferably, the multi-antigenic domain comprises (or consists of) an amino acid sequence according to SEQ ID NO: 19 or 48, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 19, or 48. For example, the multi-antigenic domain comprises (or consists of) an amino acid sequence according to SEQ ID NO: 19 or 48, preferably SEQ ID No: 19.

VSV: further features

In some embodiments, the vesicular stomatitis virus (VSV) is an oncolytic vesicular stomatitis virus (VSV). As used herein, an "oncolytic" virus refers to a virus that preferentially infects and kills cancer cells. It is assumed that the infected cancer cells, which are destroyed by oncolysis, release new infectious virus particles or virions to targeting the "remaining" cancer/tumor. Oncolytic activity of the recombinant rhabdovirus of the invention may be tested in different assay systems known to the skilled artisan (an exemplary *in vitro* assay is described by Muik et al., Cancer Res., 74(13), 3567-78, 2014). It is to be understood that an oncolytic VSV may infect and lyse only specific types of cancer cells. Also, the oncolytic effect may vary depending on the type of cancer cells.

In some embodiments, the vesicular stomatitis virus (VSV) is replication-competent. The terms “replication competent” or “replication competent virus” as used herein refer to a virus which contains all the information within its genome to allow it to replicate within a cell. For example, replication competence of the recombinant vesicular stomatitis virus of the invention may be assessed according to the methods disclosed in Tani et al. JOURNAL OF VIROLOGY, Aug. 2007, p. 8601–8612; or Garbutt et al. JOURNAL OF VIROLOGY, May 2004, p. 5458–5465. Preferably, the VSV is capable of replicating specifically within cancer cells. Preferably, the VSV specifically replicates in tumor cells, which have lost the ability to mount and respond to anti-viral innate immune responses (e.g., type-I IFN signaling). Viral replication in tumor cells leads to the cell death, and is assumed to result in the release of tumor associated antigens, local inflammation and the induction of anti-tumor immunity. On the other hand, abortive replication is preferred in “healthy cells”, such that the VSV may be rapidly excluded from normal tissues.

Preferably, in the genome of the VSV the gene coding for the glycoprotein G is replaced by the gene coding for the glycoprotein GP of Lymphocyte choriomeningitis virus (LCMV), wherein the glycoprotein GP of LCMV preferably comprises the amino acid sequence according to SEQ ID NO: 25, or a functional sequence variant thereof which is at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical thereto. Accordingly, a vesicular stomatitis virus (VSV) and, in particular, the VSV-GP (recombinant VSV with GP of LCMV, e.g. as disclosed in WO2010/040526) is preferred. Advantageous properties of the VSV-GP include one or more of the following: very potent and fast killer (<8h); oncolytic virus; systemic application possible; significantly reduced neurotropism with abolished neurotoxicity; it reproduces lytically; strong activation of innate immunity; about 3kb space for immunomodulatory cargos and antigens; recombinant with an arenavirus glycoprotein from the lymphocytic-choriomeningitis-virus (LCMV); favorable safety features in terms of reduced neurotoxicity and less sensitive to neutralizing antibody responses and complement destruction as compared to the wild type VSV (VSV-G); specifically replicates in tumor cells, which have lost the ability to mount and respond to anti-viral innate immune responses (e.g. type-I IFN signaling); abortive replication in “healthy cells” so is rapidly excluded from normal tissues;

viral replication in tumor cells leads to the cell death, and assumed to result in the release of tumor associated antigens, local inflammation and the induction of anti-tumor immunity.

In some embodiments, the vesicular stomatitis virus (VSV) exhibits the following features:

- 5 - it encodes in its genome the multi-antigenic domain as defined above,
- it encodes in its genome a vesicular stomatitis virus nucleoprotein (N), large protein (L), phosphoprotein (P), and matrix protein (M),
- the gene coding for the glycoprotein G of the vesicular stomatitis virus is replaced by the gene coding for the glycoprotein GP of lymphocyte choriomeningitis virus (LCMV),
- 10 and/or
- the glycoprotein G of the vesicular stomatitis virus is replaced by the glycoprotein GP of LCMV.

Preferably, the vesicular stomatitis virus (VSV) exhibits the following features:

- 15 - it encodes in its genome a vesicular stomatitis virus nucleoprotein (N) comprising an amino acid as set forth in SEQ ID NO: 26, or a sequence variant thereof having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity;
- it encodes in its genome a vesicular stomatitis virus phosphoprotein (P) comprising an amino acid as set forth in SEQ ID NO: 27, or a sequence variant thereof having at least
- 20 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity;
- it encodes in its genome a vesicular stomatitis virus large protein (L) comprising an amino acid as set forth in SEQ ID NO: 28, or a sequence variant thereof having at least
- 25 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity;
- it encodes in its genome a vesicular stomatitis virus matrix protein (M) comprising an amino acid as set forth in SEQ ID NO: 29, or a sequence variant thereof having at least
- 30 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity;
- it encodes in its genome a vesicular stomatitis virus the multi-antigenic domain as defined above,

- the gene coding for the glycoprotein G of the vesicular stomatitis virus is replaced by the gene coding for the glycoprotein GP of lymphocyte choriomeningitis virus (LCMV), and/or
- the glycoprotein G is replaced by the glycoprotein GP of LCMV.

5

More preferably, the vesicular stomatitis virus (VSV) according to any one of claims 43 – 85, and wherein it encodes in its genome

- a phosphoprotein (P) comprising the amino acid sequence according to SEQ ID NO: 27, or a sequence variant thereof having at least 70% sequence identity,
- 10 - a nucleoprotein (N) comprising the amino acid sequence according to SEQ ID NO: 26, or a sequence variant thereof having at least 70% sequence identity,
- a matrix protein (M) comprising the amino acid sequence according to SEQ ID NO: 29, or a sequence variant thereof having at least 70% sequence identity,
- a large protein (L) comprising the amino acid sequence according to SEQ ID NO: 28,
- 15 or a sequence variant thereof having at least 70% sequence identity,
- a glycoprotein (GP) comprising the amino acid sequence according to SEQ ID NO: 25, or a sequence variant thereof having at least 70% sequence identity more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid
- 20 sequence of SEQ ID NO: 25, and
- a multi-antigenic domain which comprises the amino acid sequence according to SEQ ID NO: 19, or a sequence variant thereof having at least 70% sequence identity more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid
- 25 sequence of SEQ ID NO: 19, or the amino acid sequence according to SEQ ID NO: 48, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 48.

30

Accordingly, a vesicular stomatitis virus (VSV) comprising an RNA genome, which comprises (or consists of) an RNA sequence according to SEQ ID NO: 30, or a sequence variant thereof

as defined above (e.g., a sequence variant having at least 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity), is particularly preferred. In some embodiments, the vesicular stomatitis virus (VSV) comprises an RNA genome, which comprises (or consists of) an RNA sequence, which corresponds to the cDNA sequence according to SEQ ID NO: 49, or to a sequence variant thereof as defined above (e.g., a sequence variant having at least 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity).

- 10 The present invention also provides a (pharmaceutical) composition comprising the VSV of the invention as described above. The (pharmaceutical) composition may further comprise a pharmaceutically acceptable carrier and/or vehicle, or any excipient, buffer, stabilizer or other materials well known to those skilled in the art. Regarding such further ingredients, i.e. carriers, vehicles, excipients, buffers, stabilizers, adjuvants and the like, the detailed description of the composition according to the present invention above applies accordingly to the (pharmaceutical) composition comprising the VSV of the invention. The (pharmaceutical) composition typically comprises a therapeutically effective amount of the active components (the VSV). The (pharmaceutical) composition may be used for human and also for veterinary medical purposes, preferably for human medical purposes, as a (pharmaceutical) composition in general or as a vaccine.

Vaccine, kit and combination

- 25 In a further aspect, the present invention also provides a vaccine comprising
- (i) the composition according to the present invention as described above;
 - (ii) the peptide according to the present invention as described above; or
 - (iii) the vesicular stomatitis virus (VSV) according to the present invention as described above.

30

In general, a vaccine induces, supports or enhances an (innate and/or an adaptive) immune response of the immune system of a subject or patient to be treated. For example, the antigens,

in particular the multi-antigenic domain, as described herein, typically lead to or support an adaptive immune response in the patient to be treated. The TLR peptide agonist of the peptide as described herein may lead to or support an innate immune response.

5 The vaccine may also comprise a pharmaceutically acceptable carrier, adjuvant, and/or vehicle as defined above for the (pharmaceutical) composition. In the specific context of the vaccine, the choice of a pharmaceutically acceptable carrier is determined in principle by the manner in which the vaccine is administered. The vaccine can be administered, for example, systemically or locally as described above. More preferably, vaccines may be
10 administered by an intravenous, intratumoral, intradermal, subcutaneous, or intramuscular route. The vaccine is therefore preferably formulated in liquid (or sometimes in solid) form. The suitable amount of the vaccine to be administered can be determined by routine experiments with animal models. Such models include, without implying any limitation, rabbit, sheep, mouse, rat, dog and non-human primate models. Preferred unit dose forms for
15 injection include sterile solutions of water, physiological saline or mixtures thereof. The pH of such solutions should be adjusted to about 7.4. Suitable carriers for injection include hydrogels, devices for controlled or delayed release, polylactic acid and collagen matrices.

The vaccine can additionally contain one or more auxiliary substances in order to further
20 increase its immunogenicity. In some embodiments, a synergistic action of the composition containing the antigen(s), the peptide and/or the VSV as described above and of an auxiliary substance, which may be optionally contained in the vaccine, may be achieved. Depending on the various types of auxiliary substances, various mechanisms can come into consideration in this respect. For example, compounds that permit the maturation of dendritic cells (DCs),
25 for example lipopolysaccharides or TNF-alpha, form a first class of suitable auxiliary substances. In general, it is possible to use as auxiliary substance any agent that influences the immune system in the manner of a "danger signal" (LPS, GP96, etc.) or cytokines, such as GM-CSF, which allow an immune response produced by the composition containing the antigen(s), the peptide and/or the VSV as described above to be enhanced and/or influenced
30 in a targeted manner. Particularly preferred auxiliary substances are cytokines, such as monokines, lymphokines, interleukins or chemokines, that further promote the innate immune response, such as IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, 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, IFN-alpha, IFN-beta, IFN-gamma, GM-CSF, G-CSF, M-CSF, LT-beta or TNF-alpha, growth factors, such as hGH.

5 Preferably, the vaccine comprises

- (i) the peptide according to the present invention as described above; and
- (ii) the vesicular stomatitis virus (VSV) according to the present invention as described above.

10 In a further aspect, the present invention also provides a kit comprising

- (i) the peptide according to the present invention as described above; and
- (ii) the vesicular stomatitis virus (VSV) according to the present invention as described above.

15 In a further aspect, the present invention also provides a combination comprising

- (i) the peptide according to the present invention as described above; and
- (ii) the vesicular stomatitis virus (VSV) according to the present invention as described above.

20 In general, in the vaccine, the kit and the combination of the invention, the different components (i) and (ii) may be provided (a) in the same composition or in the same container; or (b) in a (spatially) separate manner (e.g., in different compositions or different containers). Preferably, in the vaccine, the kit and the combination of the invention, (i) the peptide and (ii) the VSV are provided in a separate manner, e.g., in separate containers or separate
25 compositions as described above. Separate provision of the different components (i) and (ii) enables separate administration of (i) the peptide and (ii) the VSV, e.g., via distinct routes, in distinct compositions and/or at different times/at a different schedule. In particular, separate provision of the different components (i) and (ii) enables the use of (i) the peptide and (ii) the VSV in a heterologous prime-boost regimen.

30

In the vaccine, the kit or the combination it is preferred that the multi-antigenic domain encoded in the (genome of the) vesicular stomatitis virus (VSV) comprises a (tumor) antigen,

or a fragment or sequence variant thereof, which is also contained in the multi-antigenic domain of the peptide. It is also preferred that the multi-antigenic domain encoded in the (genome of the) vesicular stomatitis virus (VSV) comprises the amino acid sequences of each of the (tumor) antigens, fragment or sequence variant thereof, which is also contained in the multi-antigenic domain of the peptide. It is also preferred that the multi-antigenic domain of the peptide comprises an antigen, or a fragment thereof, which is contained in the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV). It is further preferred that the multi-antigenic domain of the peptide comprises the amino acid sequences of each of the (tumor) antigens, fragment or sequence variant thereof, which is contained in the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV).

This enables an advantageous heterologous prime-boost regimen. The principle of the heterologous prime-boost technology is to force the immune system to focus its response on a specific target antigen by avoiding an immune response against the antigen carrier or delivery system after sequential administration of the same antigen carrier or delivery system when used in homologous prime-boost regimens. In heterologous prime boost regimens the administration of the first immunogen primes cytotoxic T lymphocytes (CTLs) specific for the target antigen, however, priming also occurs for the antigen carrier or delivery system. By administering an unrelated second antigen carrier or delivery system, such as for example a viral vector during the "boost" phase, the immune system is faced with a large number of new antigens. As the second antigen carrier or delivery system also encodes/delivers the target antigen for which primed cells already exist, a strong memory response is raised by the immune system, expanding previously primed CTLs, which are specific for the target antigen.

Accordingly, the vaccine, the kit or the combination of the invention advantageously provide a target antigen in different contexts, namely, (i) contained in the peptide of the invention; and (ii) encoded in the VSV of the invention, which enables the use in a heterologous prime-boost regimen.

The corresponding antigen (tumor) antigen, or the fragment or sequence variant thereof, which is contained in the multi-antigenic domain of the peptide and in the multi-antigenic domain encoded in the VSV may be any antigen (or fragment or variant thereof) as described

herein. Preferably, corresponding antigen (tumor) antigen, or the fragment or sequence variant thereof, is selected from the group consisting of:

- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1),
- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-uORF1),
- a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2),
- CEACAM5, or a fragment or sequence variant thereof as defined above,
- DUOXA2, or a fragment or sequence variant thereof as defined above, and
- KRAS, or a fragment or sequence variant thereof as defined above.

Preferably, the KRAS (or the fragment or sequence variant thereof) is a mutant KRAS as described above. More preferably, the KRAS (or the fragment or sequence variant thereof) is KRAS-G12D (or a fragment or sequence variant thereof) or KRAS-G12V (or a fragment or sequence variant thereof).

In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1).

In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises a tumor antigen, or a fragment or sequence variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-uORF1).

In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises a tumor antigen, or a fragment or sequence

variant thereof, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2).

5 In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises CEACAM5, or a fragment or sequence variant thereof as defined above.

10 In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises DUOXA2, or a fragment or sequence variant thereof as defined above.

15 In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises KRAS, or a fragment or sequence variant thereof as defined above. Preferably, the KRAS (or the fragment or sequence variant thereof) is a mutant KRAS as described above. More preferably, the KRAS (or the fragment or sequence variant thereof) is KRAS-G12D (or a fragment or sequence variant thereof) or KRAS-G12V (or a fragment or sequence variant thereof).

20 Preferably, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises at least one amino acid sequence selected from the group consisting of:

- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity;
- an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity;
- an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity;
- an amino acid sequence according to SEQ ID NO: 1 or 2, or a sequence variant thereof having at least 70% sequence identity;
- 30 - an amino acid sequence according to SEQ ID NO: 3 or 4, or a sequence variant thereof having at least 70% sequence identity;

- an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity; or
- an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity.

5

More preferably, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises different (i.e., more than one) corresponding tumor antigens, or fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises (exactly or at least) two corresponding tumor antigens, or corresponding fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises (exactly or at least) three corresponding tumor antigens, or corresponding fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises (exactly or at least) four corresponding tumor antigens, or corresponding fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises (exactly or at least) five corresponding tumor antigens, or corresponding fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2. In some embodiments, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV each comprises all of the six different corresponding tumor antigens, or corresponding fragments or variants thereof, selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2.

Preferably, the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) comprises the amino acid sequences of each of the antigens, or fragments or sequence variants thereof, which are contained in the multi-antigenic domain of the peptide. In some embodiments, the multi-antigenic domain encoded in the genome of the vesicular

30

stomatitis virus (VSV) includes one or more additional amino acid sequences, preferably of antigen(s), or fragment(s) or sequence variant(s) thereof, which are not included in the multi-antigenic domain of the peptide. In other embodiments, the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) consists of the amino acid sequences of
5 each of the antigens, or fragments or sequence variants thereof, which are contained in the multi-antigenic domain of the peptide.

In some embodiments, the multi-antigenic domain of the peptide comprises the amino acid sequences of each of the antigens, or fragments thereof, of the multi-antigenic domain which
10 is encoded in the genome of the vesicular stomatitis virus (VSV). In some embodiments, the multi-antigenic domain of the peptide includes one or more additional amino acid sequences, preferably of antigen(s), or fragment(s) or sequence variant(s) thereof, which are not included in the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV). In other embodiments, the multi-antigenic domain of the peptide consists of the amino acid
15 sequences of each of the antigens, or fragments or sequence variants thereof, which are contained in the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV).

Preferably, the multi-antigenic domain encoded in the genome of the vesicular stomatitis
20 virus (VSV) may comprise both (i) KRAS-G12D, or a fragment or sequence variant thereof; and (ii) KRAS-G12V, or a fragment or sequence variant thereof, while the multi-antigenic domain of the peptide may comprise either (i) KRAS-G12D, or a fragment or sequence variant thereof; or (ii) KRAS-G12V, or a fragment or sequence variant thereof. Accordingly, the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) may comprise
25 both (i) an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity; and (ii) an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity, while the multi-antigenic domain of the peptide may comprise either (i) an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity; or
30 (ii) an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity.

In some embodiments, the multi-antigenic domain of the peptide may comprise both (i) KRAS-G12D, or a fragment or sequence variant thereof; and (ii) KRAS-G12V, or a fragment or sequence variant thereof, while the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) may comprise either (i) KRAS-G12D, or a fragment or sequence variant thereof; or (ii) KRAS-G12V, or a fragment or sequence variant thereof. Accordingly, the multi-antigenic domain of the peptide may comprise both (i) an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity; and (ii) an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity, while the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) may comprise either (i) an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity; or (ii) an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity.

More preferably, in the vaccine, the kit or the combination of the invention,

- the multi-antigenic domain of the peptide comprises an amino acid sequence according to SEQ ID NO: 13, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 13, or an amino acid sequence according to SEQ ID NO: 14, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 14, and
- the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) comprises the amino acid sequence according to SEQ ID NO: 19, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 19 or the amino acid sequence according to SEQ ID NO: 48, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%,

86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 48.

Still more preferably, in the vaccine, the kit or the combination of the invention,

- 5 - the peptide comprises an amino acid sequence according to SEQ ID NO: 17, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 17, an amino acid sequence according to SEQ ID NO: 18, or a sequence variant thereof
- 10 having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 18, an amino acid sequence according to SEQ ID NO: 46, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%,
- 15 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of SEQ ID NO: 46, or an amino acid sequence according to SEQ ID NO: 47, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity
- 20 to the amino acid sequence of SEQ ID NO: 47, and
- the vesicular stomatitis virus (VSV) comprises an RNA genome comprising an RNA sequence according to SEQ ID NO: 30, or a sequence variant thereof having at least 70% sequence identity, more preferably one of at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence
- 25 identity to the amino acid sequence of SEQ ID NO: 30.

In some embodiments of the vaccine, the kit or the combination of the invention, the vesicular stomatitis virus (VSV) comprises an RNA genome, which comprises (or consists of) an RNA sequence, which corresponds to the cDNA sequence according to SEQ ID NO: 49, or to a sequence variant thereof as defined above (e.g., a sequence variant having at least 75%, 76%,

30 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity).

It is also preferred that the vaccine, the kit or the combination of the invention further comprises

(iii) an inhibitor of the PD-1/PD-L1 pathway.

- 5 While, in general, the inhibitor of the PD-1/PD-L1 pathway may be provided (a) in the same composition or in the same container as any one of components (i) and (ii) (the peptide and the VSV); or (b) in a (spatially) separate manner (e.g. in different compositions or different containers), it is preferred in the vaccine, the kit and the combination of the invention, that the inhibitor of the PD-1/PD-L1 pathway is provided in a separate manner (separate from (i)
- 10 the peptide and (ii) the VSV), e.g. in a separate container or in a separate composition. As used herein, the term "inhibitor" includes reduction, decrease, blocking and inhibition of the PD-1/PD-L1 pathway, including antagonists (and inverse agonists) of the PD-1/PD-L1 pathway.
- 15 In general, the PD-1/PD-L1 pathway is well-known in the art and described, for example, in Han Y, Liu D, Li L. PD-1/PD-L1 pathway: current researches in cancer. Am J Cancer Res. 2020 Mar 1;10(3):727-742. PD-1 (Programmed Cell Death Protein 1) and its ligand PD-L1 (Programmed Cell Death Ligand 1) are well-known for their role in cancer immune escape.
- 20 Non-limiting examples of inhibitors of the PD-1/PD-L1 pathway include pembrolizumab (anti-PD-1 antibody); nivolumab (anti-PD-1 antibody); pidilizumab (anti-PD-1 antibody); cemiplimab (anti-PD-1 antibody); PDR-001 (anti-PD-1 antibody); atezolizumab (anti-PD-L1 antibody); avelumab (anti-PD-L1 antibody); durvalumab (anti-PD-L1 antibody); and PD1-1, PD1-2 and PD1-3, as described herein (anti-PD-1 antibodies). Accordingly, the inhibitor of
- 25 the PD-1/PD-L1 pathway may be selected from the group consisting of pembrolizumab; nivolumab; pidilizumab; cemiplimab; PDR-001; atezolizumab; avelumab; durvalumab, ezabenlimab, an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 31 and a light chain comprising the amino acid sequence of SEQ ID NO: 32 (PD1-1); an antibody comprising a heavy chain comprising the amino acid sequence of SEQ
- 30 ID NO: 33 and a light chain comprising the amino acid sequence of SEQ ID NO: 34 (PD1-2); and an antibody comprising a heavy chain comprising the amino acid sequence of SEQ

ID NO: 35 and a light chain comprising the amino acid sequence of SEQ ID NO: 36 (PD1-3).

Further PD-1 antagonists disclosed by Li et al. (supra), or known to be in clinical trials, such as AMP-224, MEDI0680 (AMP-514), BMS-936559, JS001-PD-1, SHR-1210, BMS-936559, TSR-042, JNJ-63723283, MEDI4736, MPDL3280A, may be used as alternative or in addition to the above mentioned antagonists. The INNs as used herein are meant to also encompass all biosimilar antibodies having the same, or substantially the same, amino acid sequences as the originator antibody, including but not limited to those biosimilar antibodies authorized under 42 USC §262 subsection (k) in the US and equivalent regulations in other jurisdictions.

Antibodies PD1-1, PD1-2 and PD1-3 are antibodies comprising the following (full length) heavy chain and (full length) light chain sequences:

- PD1-1: a heavy chain comprising the amino acid sequence of SEQ ID NO: 31 and a light chain comprising the amino acid sequence of SEQ ID NO: 32;
- PD1-2: a heavy chain comprising the amino acid sequence of SEQ ID NO: 33 and a light chain comprising the amino acid sequence of SEQ ID NO: 34;
- PD1-3: a heavy chain comprising the amino acid sequence of SEQ ID NO: 35 and a light chain comprising the amino acid sequence of SEQ ID NO: 36.

20

Preferably, the inhibitor of the PD-1/PD-L1 pathway may be selected from the group consisting of ezabenzimab; an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 31 and a light chain comprising the amino acid sequence of SEQ ID NO: 32 (PD1-1); an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 33 and a light chain comprising the amino acid sequence of SEQ ID NO: 34 (PD1-2); and an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 35 and a light chain comprising the amino acid sequence of SEQ ID NO: 36 (PD1-3).

Medical treatment and uses

The composition according to the present invention as described above, the peptide according to the present invention as described above, the vesicular stomatitis virus (VSV) according to the present invention as described above, the kit according to the present invention as described above, the vaccine according to the present invention as described above, or the combination according to the present invention as described above may be for use in medicine. Preferably, the composition according to the present invention as described above, the peptide according to the present invention as described above, the vesicular stomatitis virus (VSV) according to the present invention as described above, the kit according to the present invention as described above, the vaccine according to the present invention as described above, or the combination according to the present invention as described above are for use in the treatment of cancer.

Accordingly, the present invention also provides the use of the composition according to the present invention as described above, the peptide according to the present invention as described above, the vesicular stomatitis virus (VSV) according to the present invention as described above, the kit according to the present invention as described above, the vaccine according to the present invention as described above, or the combination according to the present invention as described above for the manufacture of a medicament for the treatment of cancer.

In a further aspect, the present invention provides a method for ameliorating, treating, or reducing (the risk of occurrence) of a cancer, or for inducing or enhancing an anti-tumor response, the method comprising administration of (an effective amount of) the composition according to the present invention as described above, the peptide according to the present invention as described above, the vesicular stomatitis virus (VSV) according to the present invention as described above, the kit according to the present invention as described above, the vaccine according to the present invention as described above, or the combination according to the present invention as described above to a subject in need thereof.

Preferably, the cancer (as referred to in the above methods and uses) is a cancer of the gastrointestinal tract (GI). Non-limiting examples of GI cancers include anal cancer; appendix cancer; cholangiocarcinoma/bile duct cancer, in particular extrahepatic bile duct cancer; gastrointestinal carcinoid tumor; colorectal cancer, in particular colon cancer, rectal cancer and metastatic colorectal cancer; esophageal cancer; gallbladder cancer; gastric (stomach) cancer; gastrointestinal stromal tumor (GIST), and pancreatic cancer, such as pancreatic ductal adenocarcinoma. Accordingly, the cancer is preferably selected from the group consisting of anal cancer; appendix cancer; cholangiocarcinoma/bile duct cancer, in particular extrahepatic bile duct cancer; gastrointestinal carcinoid tumor; colorectal cancer, in particular colon cancer, rectal cancer and metastatic colorectal cancer; esophageal cancer; gallbladder cancer; gastric (stomach) cancer; gastrointestinal stromal tumor (GIST), and pancreatic cancer, such as pancreatic ductal adenocarcinoma. More preferably, the cancer is selected from the group consisting of colon cancer, rectal cancer, colorectal cancer, metastatic colorectal cancer, pancreatic cancer, and pancreatic ductal adenocarcinoma.

For the envisaged medical treatment, in particular for the treatment of cancer (e.g., as described above), the peptide of the present invention as described above is preferably combined with the vesicular stomatitis virus (VSV) of the present invention as described above. Accordingly, the peptide for use as described above may be administered in combination with the vesicular stomatitis virus (VSV) of the present invention as described above. Moreover, the vesicular stomatitis virus (VSV) for use as described above may be administered in combination with the peptide of the present invention as described above. In such a combined use, the multi-antigenic domain of the peptide and the multi-antigenic domain encoded in the VSV are preferably adapted to each other as described in detail above, in the context of the vaccine, the kit and the combination of the invention comprising the peptide and the VSV.

In particular, a "combination" or "combined" use of the peptide and the VSV as described herein usually means that the treatment with the peptide as described herein is combined with the treatment with the VSV as described herein. In such a combined use of the peptide and the VSV of the present invention as described above (i.e., also for the respective vaccine, kit and combination of the invention as described above), each of the peptide and the

vesicular stomatitis virus (VSV) are usually administered at least once. Preferably, the peptide and the vesicular stomatitis virus (VSV) are administered consecutively (not simultaneously). However, even if one component (the peptide or the VSV) is not administered, e.g., at the same day as the other component, their treatment schedules are usually intertwined. Thereby, an advantageous heterologous prime-boost regimen, as described above, can be achieved. In particular, one component, e.g., the peptide, may be administered first (as "prime"), while the other component, e.g., the VSV, may be administered later (as "boost"); e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more days or weeks after the "prime". Thereby, the interval between prime and boost is usually selected such, that a strong immune response can be found. Preferably the peptide is administered prior to the administration of the vesicular stomatitis virus (VSV). Accordingly, the peptide may function as "prime" and the VSV as "boost" in the heterologous prime-boost regimen.

In some embodiments, the peptide is administered at least twice, preferably prior to and subsequent to the administration of the VSV. Accordingly, the peptide and the vesicular stomatitis virus (VSV) are administered in the order K-V-K, wherein "K" refers to a (single) administration of the peptide and "V" refers to a (single) administration of the VSV. In some embodiments, the peptide is administered repeatedly. For example, the peptide and the vesicular stomatitis virus (VSV) may be administered in the order K-V-K, K-V-K-K, K-V-K-K-K, or K-V-K-K-K-K (with "K" referring to the peptide and "V" to the VSV). In some embodiments, the treatment schedule comprises a single administration of the vesicular stomatitis virus (VSV), i.e., the VSV is administered only once.

In some embodiments, the peptide and the VSV are administered 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21 days apart from each other, preferably from about 4 days to about 24 days apart from each other, more preferably from about 10 days to about 18 days apart from each other, even more preferably from about 12 days to about 16 days apart from each other. In some embodiments, the peptide is administered repeatedly, preferably once per every 2 to 6 weeks, more preferably once per every 3 to 5 weeks, even more preferably once every 4 weeks.

In some embodiments, the dose of the peptide may be from about 0.5 nmol to about 10 nmol. In some embodiments, the dose of the VSV may be from about 10^6 TCID₅₀ to about 10^{11} TCID₅₀. Preferably, the dose of the VSV may be from about 10^7 TCID₅₀ to about 10^9 TCID₅₀, more preferably about 10^8 TCID₅₀ to about 10^9 TCID₅₀, or about 10^7 TCID₅₀ to about 10^8 TCID₅₀, more preferably, from about 1×10^7 TCID₅₀ to about 1×10^8 TCID₅₀.

In general, the peptide and the vesicular stomatitis virus (VSV) may be administered via the same route or different routes of administration. Preferably, the route of administration is selected from intravenous, subcutaneous, and intramuscular administration. Preferably, the peptide and the VSV are administered via distinct routes of administration. More preferably, the peptide is administered subcutaneously, and the vesicular stomatitis virus (VSV) are administered intravenously, or intratumorally, preferably, intravenously.

The (combined) use of the peptide and the VSV (also in the vaccine, kit and combination as described herein) may further comprise administration of an inhibitor of the PD-1/PD-L1 pathway, such as the inhibitors of the PD-1/PD-L1 pathway described above. Accordingly, the inhibitor of the PD-1/PD-L1 pathway may be selected from the group consisting of pembrolizumab; nivolumab; pidilizumab; cemiplimab; PDR-001; atezolizumab; avelumab; durvalumab, ezabenlimab, an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 31 and a light chain comprising the amino acid sequence of SEQ ID NO: 32 (PD1-1); an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 33 and a light chain comprising the amino acid sequence of SEQ ID NO: 34 (PD1-2); and an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 35 and a light chain comprising the amino acid sequence of SEQ ID NO: 36 (PD1-3).

The inhibitor of the PD-1/PD-L1 pathway is administered concomitantly, sequentially or alternately with the peptide or the vesicular stomatitis virus (VSV). In some embodiments, the inhibitor of the PD-1/PD-L1 pathway is administered at the same day and/or alternately with the peptide.

BRIEF DESCRIPTION OF THE FIGURES

In the following a brief description of the appended figures will be given. The figures are intended to illustrate the present invention in more detail. However, they are not intended to limit the subject matter of the invention in any way.

Figure 1 shows for Example 1 KRAS-uORF1-, TPX2-uORF1- and AURKA-uORF2-encoding transcript expression as measured by BaseScope on tissue arrays (HalioDX) and tumor whole sections (A). Probes corresponding to the KRAS-uORF1-, TPX2-uORF1- and AURKA-uORF2-antigen sequences were employed. Bar diagrams display the quantification of the experiments through an automated scoring algorithm (HALO-Indica lab). Note that CRC-adjacent healthy tissue is characterized by substantially higher expression of the transcripts of interest than normal healthy colon tissue. (B) RNA-Seq-based expression of KRAS, TPX2, and AURKA transcripts in healthy normal colon/pancreas versus tumor-adjacent colon/pancreas tissue (*TCGA/GTEX*). mRNA expression of KRAS, TPX2 and AURKA data were extracted from the following public databases: the Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx). (C) Translation of KRAS-uORF1-, TPX2-uORF1- and AURKA-uORF2-tumor antigen ORFs in human pancreatic (PDAC), gastric (GC), and colorectal (CRC) tumors or tumor-adjacent healthy tissue. Translation was measured by RiboSeq. The number of *in-frame* reads across the respective ORFs was normalized for ORF length (assuming translation is driven by the innermost possible start codon) and sample library size (*i.e.*, the sum of all trimmed reads in the sample mapping to the human genome after removal of rRNA- and tRNA-derived reads). An unpaired two-tailed t-test was performed to assess statistical significance (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

Figure 2 shows for Example 2 (A) Expression of DUOXA2 and CEACAM5 transcripts in various GI tract tumors as well as in health tissue (TCGA/GTEx). (B) CEACAM5 protein expression as measured by immunohistochemistry staining on tissue

arrays (HalioDX). The bar diagram (upper panel) displays the quantification of the experiment. (C) DUOXA2 transcript expression as measured by RNAScope on tissue arrays (HalioDX). The bar diagram (left panel) displays the quantification of the experiment.

5

Figure 3 shows for Example 3 (A) Co-expression on the mRNA level of mutant KRAS-G12D/V with CEACAM5 and DUOXA2 in pancreatic (PDAC) and colorectal (CRC) cancer (based on various RNASeq-based datasets including TCGA). (B) Immunogenicity of epitopes derived from KRAS-uORF1, TPX2-uORF1, AURKA-uORF2, KRAS-G12D/V as well as for the DUOXA2. *In vitro* priming was performed employing the indicated peptides (*e.g.*, KLVVVGADGV) in the context of the given HLA alleles (*e.g.*, HLA-A*0201) to stimulate specific T cell responses among healthy donor PBMCs. Cultures thereby enriched for peptide-specific T cells were subsequently assayed on various tumor cell line targets pulsed with the respective peptides (specific response). In parallel, the same tumor cell lines were pulsed either with an irrelevant control peptide or were not pulsed at all (unspecific background response). T cell activation was measured by IFN γ ELISpot. In panel (B) each spot corresponds to a different tumor cell line or moDC target. Displayed is the ratio of the specific response to the unspecific background response. A ratio above 1 indicates a specific T cell response. A negative control (*i.e.*, a peptide epitope that did not elicit a specific T cell response) is included in the panel.

10

15

20

25

30

Figure 4 shows for Example 4 the presentation of ATP150- and ATP152-derived epitopes by human moDCs. moDCs were incubated with either ATP150 or ATP152 to allow processing of the vaccine by the endogenous antigen presentation machinery. Cross-presented antigen was subsequently analyzed by mass spectrometry. In the panel, detected peptide epitopes were mapped onto the ATP150/ATP152 backbone to visualize the region from which they are derived. Each line corresponds to one of the 10 PBMCs' donors. Note that the 10 donors differ from each other by their HLA haplotype and are, hence, expected to present different peptide epitopes.

- Figure 5 shows for Example 5 the immunogenicity after ATP150/VSV-GP154 heterologous prime-boost vaccination. IFN γ -producing CEACAM5- and DUOXA2-specific T cells by ELISpot (A) Outline of the prime-boost protocol. Mice [n=5] were either both primed (d0) and boosted (d14) with 10 nmol ATP150 (s.c.) or primed (d0) with 10 nmol ATP150 (s.c.) and boosted (d14) with 10⁷ TCID₅₀ VSV-GP154 (i.v.). (B) On day 21, splenocytes were harvested and pulsed with overlapping 11mer peptides corresponding to the CEACAM5 and DUOXA2 regions incorporated in the vaccine. *** p $\leq$ 0.001; **** p $\leq$ 0.0001.
- Figure 6 shows for Example 6 (A) Tumor growth (mean $\pm$ SEM) and (B) survival of mice injected with 1x10⁵ TC-1 cells s.c. and immunized with 2 nmol the peptide construct (multi-antigenic domain "Mad25") s.c. 7 days later or with 1x10⁷ TCID₅₀ VSV-GP-HPV i.v. 14 days post tumor implantation. Additional doses of K and V were administered as indicated by dotted lines and tumor growth was monitored (n=7).
- Figure 7 shows for Example 7 the vaccination schedule (A) and detection of HPV-specific CD8 T cells (B) after heterologous prime-boost and anti-PD-1 in TC-1 tumor-bearing mice. Mice were bled at days 14, 21, 35, 42, 56, and multimer staining detecting HPV-specific CD8 T cells was performed and analyzed by flow cytometry.
- Figure 8 shows for Example 7 the efficacy of heterologous prime-boost vaccination and anti-PD-1 in TC-1 tumor-bearing mice. Mean tumor volume and survival curve of C57BL/6 mice (n=7 mice/group) implanted s.c. with TC-1 cells. Mice were vaccinated with the peptide construct (multi-antigenic domain "Mad25") (K, 2 nmol/mouse; s.c.) and VSV-GP-HPV (V; 10⁷ TCID₅₀/mouse; i.v.). Anti-PD-1 (200ug) was administrated i.v. Numbers of tumor-free mice are indicated.

Figure 9 shows for Example 8 the efficacy of heterologous prime-boost vaccination comparing three different doses of VSV-GP-HPV in tumor-bearing mice. Tumor volumes (upper/left panels) and survival curves (lower/right panel) of C57BL/6 mice (n=7 mice/group) implanted s.c. with TC-1 cells are shown. Mice were vaccinated at d7, d14 and d28 (Vac) with either the peptide construct (multi-antigenic domain "Mad25") (K) or VSV-GP-HPV (V) at 3 different doses, 1×10^7 , 3×10^7 or 5×10^7 TCID₅₀. Median survival is indicated on the graph (ms). *, p<0.05; **, p<0.01; ***, p<0.001.

EXAMPLES

Example 1: Tumor antigens encoded within the 5' UTR of KRAS mRNA (KRAS-uORF1), TPX2 mRNA (TPX2-uORF1) and AURKA mRNA (AURKA-uORF2)

5

Novel tumor antigens, which are encoded in open reading frames (ORFs) in the 5' UTR of KRAS mRNA (KRAS-uORF1), TPX2 mRNA (TPX2-uORF1) and AURKA mRNA (AURKA-uORF2) were identified. The tumor antigens (peptides) are also referred to herein as "KRAS-uORF1 peptide" or "KRAS-uORF1", "TPX2-uORF1 peptide" or "TPX2-uORF1", and "AURKA-uORF2 peptide" or "AURKA-uORF2", respectively. The respective tumor antigens have the amino acid sequences according to SEQ ID NO: 1 (KRAS-uORF1 peptide), SEQ ID NO: 3 (TPX2-uORF1 peptide) and SEQ ID NO: NO 5 (AURKA-uORF2 peptide).

10

15

Because uORFs are often characterized by unusual, non-canonical translation start sites, and KRAS-uORF1, TPX2-uORF1, and AURKA-uORF2 do not contain a conventional start codon, it was first verified that they could be translated in human cancer cell lines. To this end, the antigens were fused N-terminally to dTomato and the expression of the resulting fluorescent fusion proteins was confirmed by flow cytometry.

20

To examine tumor expression of KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 on RNA level, BaseScope detection of relevant transcripts was used. To this end, probes corresponding only to the KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 coding sequences themselves were employed. This allows to distinguish between those mRNA transcripts of a gene locus that encode the tumor antigen of interest and those, which do not (*e.g.*, according to Ensembl, the AURKA gene locus encompasses 13 alternatively spliced transcripts with seven of these 13 transcripts encoding the full-length AURKA kinase, but only three encoding AURKA-uORF2).

25

30

In a first experiment transcription, i.e., the generation of relevant KRAS, TPX2, and AURKA mRNA transcripts, was assessed in (i) pancreatic and colorectal tumors, (ii) tumor-adjacent 'healthy' tissue; and (iii) healthy tissue from non-diseased organs. To this end, the stainings were performed on LEICA Bond Max©. The slides were evaluated for RNA quality by using

positive control probe (housekeeping gene PPIB). After QC tests, the slides were tested using BaseScope probe: Basescope® 2.5 LS Probe BA-Hs-Aurka-Tu-1-1zz-stC1 Lot: 21175A (ACD#1084338-C1), Basescope® LS Probe BA-Hs-KRAS-1zz-stC1 Lot: 21175A (ACD#1084358) according to manufacturer's instruction. The slides were counterstained by hematoxylin and mounted using VectaMount mounting medium. A whole slide image was scanned at 40 times objective magnification using Nanozoomer S360© Digital Pathology Scanning System (Hamamatsu). The number of cancer cells with ≥ 1 red color signal (positive cancer cells) were evaluated as positive. The percentage of positive cancer cells was calculated with the aid of HALO software (Indica Lab) by a certified pathologist.

10

As the KRAS-uORF1, TPX2-uORF1, and AURKA-uORF2 peptide tumor antigens lack any known domains, it appears unlikely that they stably fold and accumulate in the cell. Without being bound to any theory, it is assumed that they undergo rapid proteasomal degradation, which may be advantageous, as this may lead to the production of antigenic epitopes. However, the expected ultra-low steady-state protein levels of these entities render them essentially undetectable with antibodies. To nevertheless assess their translation (*i.e.*, their protein synthesis) in human pancreatic (PDAC), gastric (GC), and colorectal (CRC) tumors (or tumor-adjacent healthy tissue), we resorted to the RiboSeq technique (Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* 324, 218–223). Briefly, at least 200 mg of fresh frozen tumor tissues as well as corresponding tumor-adjacent tissues derived from PDAC [$n = 12$; only five of which with tumor-adjacent material], GC [$n = 5$], and CRC [$n = 44$] were purchased from a commercial provider (Indivumed GmbH, Hamburg) for RiboSeq library preparation, sequencing (NovaSeq6000; ≥ 200 M reads/library), and bioinformatic analysis at a partner CRO (TB-Seq Inc, South San Francisco). Adapters were trimmed using Cutadapt before aligning them to the human genome assembly GRCh38.p13 using STAR. It was confirmed that ribosomal footprints matched the expected lengths spectrum. 28-36mers were considered for further analysis. Dataset quality and proper 3-nt periodicity was validated by assessing metagene read distributions around start/stop codons. To quantify AURKA-uORF2, TPX2-uORF1, and KRAS-uORF1 translation, the numbers of relevant in-frame reads (multiplied by 10^7) were normalized for ORF length (under the assumption that translation is driven by the innermost RiboSeq-supported start codon) and

30

sample library size (i.e., the sum of all trimmed, mappable reads post removal of rRNA- and tRNA-derived sequences) to allow comparison across samples.

Results are shown in Figure 1 with KRAS-uORF1-, TPX2-uORF1- and AURKA-uORF2-
5 encoding transcript expression (A), RNA-Seq-based expression of KRAS, TPX2, and AURKA transcripts in healthy normal colon/pancreas versus tumor-adjacent colon/pancreas tissue (B) and translation of the uORFs of interest in human pancreatic (PDAC), gastric (GC), and colorectal (CRC) tumors or tumor-adjacent healthy tissue (C).

10 These data demonstrated strong expression of relevant KRAS, TPX2, and AURKA transcripts in pancreatic and colorectal tumors (Figure 1A), while normal healthy tissue from colon, pancreas and other organs showed only low signal above background (Figure 1A). This underscores that the encoded tumor antigens are likely targetable in these cancers. Tumor-adjacent 'healthy' tissue - unlike truly healthy tissue from non-diseased organs - displayed an
15 intermediate transcript signal. This trend is also seen in RNASeq-based TCGA/GTEX datasets (Figure 1B) and points to transcriptional dysregulation spreading from the tumor site into the immediately neighboring areas. Without being bound to any theory, it is thus assumed that tumor-adjacent healthy tissue might not always be fully representative of truly healthy tissue.

20 Figure 1C shows that KRAS-uORF1, TPX2-uORF1, and AURKA-uORF2 peptide tumor antigens are expressed in tumors, albeit to various degrees in different indications including pancreatic, gastric, and colorectal cancer. In general, low or no synthesis was observed in healthy tissue immediately bordering the tumor (Figure 1C). Without being bound to any theory, it is assumed that the KRAS-uORF1, TPX2-uORF1, and AURKA-uORF2 translation is
25 not detected in most tumor-adjacent healthy tissues, even though BaseScope (Figure 1A) and RNASeq (Figure 1B) suggest the expression of corresponding mRNA transcripts, because of tumor-selective up-regulation of upstream ORF translation, which may be linked to the process of tumorigenesis itself. Thus, for KRAS-uORF1, TPX2-uORF1, and AURKA-uORF2, cancer-specific translation may add an additional layer of tumor specificity on top of largely
30 cancer-specific transcription.

Example 2: Expression profiles of additional tumor antigens CEACAM5 and DUOXA2

To select additional tumor antigens of interest, mRNA expression of CEACAM5 and DUOXA2 data were extracted from the following public databases: the Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx). Based on the tumor-selective expression according to public RNASeq (Figure 2A), CEACAM5 and DUOXA2 were selected.

Subsequently, the expression profile of CEACAM5 and DUOXA2 was experimentally validated in human tumor as well as non-tumor GI tract tissue. To measure CEACAM5 expression, Formalin-Fixed and Paraffin Embedded (FFPE) tissue arrays were purchased from a commercial provider (Indivumed GmbH, Hamburg). After QC tests, CEACAM5 staining was performed by ImmunoHistoChemistry on using CEACAM5-specific antibody (Dako M707229-2). The slides were counterstained by hematoxylin and mounted using VectaMount mounting medium. A whole slide image was scanned at 40 times objective magnification using Nanozoomer S360© Digital Pathology Scanning System (Hamamatsu). The number of cancer cells with ≥ 1 red color signal (positive cancer cells) were evaluated as positive. The percentage of positive cancer cells was calculated with the aid of HALO software (Indica Lab) by a certified pathologist.

Further, DUOXA2 expression was examined in GI tract tissue. To this end, Formalin-Fixed and Paraffin Embedded (FFPE) tissue arrays were purchased from a commercial provider (Indivumed GmbH, Hamburg). The slides were evaluated for RNA quality by using positive control probe (housekeeping gene PPIB). After QC tests, DUOXA2 transcript expression was performed by using RNAscope probe for DUOXA2 (Bio-Techne HD-RM-000619). according to manufacturer's instruction. The slides were counterstained by hematoxylin and mounted using VectaMount mounting medium. A whole slide image was scanned at 40 times objective magnification using Nanozoomer S360© Digital Pathology Scanning System (Hamamatsu). The number of cancer cells with ≥ 1 red color signal (positive cancer cells) were evaluated as positive. The percentage of positive cancer cells was calculated with the aid of HALO software (Indica Lab) by a certified pathologist.

Results are shown in Figure 2. Specifically, CEACAM5 was detected by immunohistochemistry in 80-100% of pancreatic and colorectal cancers, but not in completely healthy or tumor-adjacent healthy tissue (Figure 2B). DUOXA2 expression was solidly detected in 37-74% of pancreatic and colorectal cancers, while expression in
5 corresponding healthy tissues was rare (*e.g.*, only one out of 16 pancreatic cancers) (Figure 2C).

Example 3: Immunogenicity of KRAS-uORF1, TPX2-uORF1, AURKA-uORF2, CEACAM5,
10 DUOXA2 and KRAS-G12D/V antigens

In addition to the tumor antigens of Examples 1 and 2 above, a further tumor antigen of interest, namely KRAS, in particular KRAS-G12D/V, was included in the further analysis. Oncogenic driver mutations KRAS-G12D/V are generally not found outside tumors.

15 In a first step, co-expression on the mRNA level of mutant KRAS-G12D/V with CEACAM5 and DUOXA2 in pancreatic (PDAC) and colorectal (CRC) cancer (based on various RNASeq-based datasets including TCGA).

20 In addition, MHC class I epitopes for KRAS-uORF1, TPX2-uORF1, AURKA-uORF2, as well as for DUOXA2 were predicted using the NetMHCpan 4.1 prediction algorithm in order to identify peptides that bind to specific HLA class I alleles. A subset of these peptides and corresponding HLA allele-expressing donor PBMCs were then used for *in vitro* priming experiments to determine immunogenicity. To this end, dendritic cells were isolated from
25 these PBMCs and pulsed with selected peptides, before they were used for the stimulation of autologous CD8 T cells, whose response was measured via IFN γ ELISpot assays.

Next, the immunogenicity of KRAS-G12D/V was analyzed using two independent healthy peripheral blood mononuclear cells (PBMC) donors. To this end, CD8⁺ T cells were isolated
30 from healthy donor peripheral blood mononuclear cells (PBMCs) by magnetic cell sorting using CD8 MicroBeads and columns (Miltenyi Biotec, Bergisch Gladbach, Germany). Monocytes were isolated from CD8⁺ depleted PBMCs by plastic adherence and subsequently

treated with cytokines for generation and maturation of dendritic cells (DCs, antigen presenting cells).). Selection of donors occurred due to their expression of specific HLA alleles that are predicted for presentation of the respective peptide by DCs. CD8+ T cells were primed at least 3 times by peptide-loaded DCs before their specific immune response against a peptide was tested. A priming cycle included a co-culture of CD8+ T cells and peptide-loaded DCs for 7-10 days. After 3 priming cycles, CD8 T cells were plated in an ELISpot assay (Immunospot, Shaker Heights, OH 44122) for assessing their capacity to secrete IFN- γ after overnight stimulation with tumor cell lines or peptide-loaded DCs. Spots were counted using ELISPOT reader.

Results are shown in Figure 3. As shown in Figure 3A, co-expression analysis on the mRNA level of mutant KRAS-G12D/V with CEACAM5 and DUOXA2 in pancreatic (PDAC) and colorectal (CRC) cancer revealed that 59% of KRAS-G12D/V-mutant PDAC/CRC tumors express both CEACAM5 and DUOXA2 transcripts at > 5 TPM. Figure 3B shows the immunogenicity of epitopes derived from KRAS-uORF1, TPX2-uORF1, AURKA-uORF2, KRAS-G12D/V as well as for the DUOXA2. All of the tested peptides were found to be immunogenic (Figure 3B). This suggests that cognate T cells will likely be available for activation in cancer patients.

Example 4: Design and evaluation of peptides comprising a cell-penetrating peptide, a multi-antigenic domain and a TLR agonist

Next, peptides comprising (i) a cell-penetrating peptide, (ii) a multi-antigenic domain and (iii) a TLR agonist were designed, essentially as described in WO2016/146260. To this end, two different multi-antigenic domains were designed, which comprise either (immunogenic) fragments of KRAS-uORF1, TPX2-uORF1, AURKA-uORF2, KRAS-G12D, CEACAM5 and DUOXA2; or (immunogenic) fragments of KRAS-uORF1, TPX2-uORF1, AURKA-uORF2, KRAS-G12V, CEACAM5 and DUOXA2.

The multi-antigenic domain of both constructs comprises the following antigenic fragments of CEACAM5:

CEACAM5-1

NRTLTLFNVTRNDARAYVSGIQNSVSVANRSDPVTLDVL

[SEQ ID NO: 6]

5

CEACAM5-2

PDSSYLSGANLNLSCHSAS

[SEQ ID NO: 7]

10 CEACAM5-3

PQYSWRINGIPQQHTQVLFIKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSA

[SEQ ID NO: 8]

15 Based on the three different fragments, a CEACAM5 fusion construct of the following
sequence was designed for the multi-antigenic domain of both constructs:

NRTLTLFNVTRNDARAYVSGIQNSVSVANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYSWR

INGIPQQHTQVLFIKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSA

[SEQ ID NO: 9]

20

In addition, the following antigens or antigenic fragments were included in the multi-antigenic
domain of both constructs:

DUOXA2:

25 TGVLSLFLGGAVVSLQYVRPSALRTLDDQSAKD

[SEQ ID NO: 10]

KRAS-uORF1:

VAAARPVLPAPAI SDWERARRRH

30 [SEQ ID NO: 1]

TPX2-uORF1:

VGRKEGAARARVSLLPEFGTAEVHLPAPSAVRAARPGF

[SEQ ID NO: 3]

AURKA-uORF2:

5 GGDKGRLVGVAERQVPCRFLRP

[SEQ ID NO: 5]

The multi-antigenic domain of both constructs differs in that one comprises a fragment of KRAS-G12D (SEQ ID NO: 11), while the other comprises a fragment of KRAS-G12V (SEQ ID

10 NO: 12):

KRAS-G12D:

TEYKLVVVGADGVGKSALTIQLIQ

(SEQ ID NO: 11)

15

KRAS-G12V:

TEYKLVVVGAVGVGKSALTIQLIQ

(SEQ ID NO: 12)

20 The two different peptide constructs are referred to as "ATP150" (comprising KRAS-G12D) and "ATP152" (comprising KRAS-G12V). They comprise the following multi-antigenic domains:

multi-antigenic domain ATP150:

25 NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYSWR
INGIPQQHTQVLFIKITPNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSATEYKLVVV
GADGVGKSALTIQLIQVAAARPVLPAPASDWERARRRHVGRKEGAARARVSLLPEFGTAEVH
LPAPSAVRAARPGFGGDKGRLVGVAERQVPCRFLRPTGVLSLFLGGAVVSLQYVRPSALRTLLD
QSAKD

30 [SEQ ID NO: 13]

multi-antigenic domain ATP152:

NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYSWR
 INGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSATEYKLVVV
 GAVGVGKSALTIQLIQVAAARPVLPAPAISDWERARRRHVGRKEGAARARVSLLPEFGTAEVHL
 PPSAVRAARPGFGGDKGRLVGVAERQVPCRFLRPTGVLSLFLGGAVVSLQYVRPSALRTLLD

5 QSAKD

[SEQ ID NO: 14]

As described above, the peptide constructs ATP150 and ATP152 comprise, in addition to the multi-antigenic domain, a cell penetrating peptide. The cell penetrating peptide included in

10 ATP150 and ATP152 has the following sequence:

KRYKNRVASRKSRKFKQLLQHYREVAAAKSSENDRLRLLK

[SEQ ID NO: 15]

In addition, the peptide constructs ATP150 and ATP152 further comprise a TLR agonist:

15 STVHEILSKLSLEGDHSTPPSAYGSVKPYTNFDAE

[SEQ ID NO: 16]

In summary, the peptide constructs ATP150 and ATP152 exhibit the following sequences:

20 ATP150

MKRYKNRVASRKSRKFKQLLQHYREVAAAKSSENDRLRLLKNRTLTLFNVTRNDARAYVSGI
 QNSVSANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYSWRINGIPQQHTQVLFIAKITPNN
 NGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSATEYKLVVVGADGVGKSALTIQLIQVAAA
 RPVLPAPAISDWERARRRHVGRKEGAARARVSLLPEFGTAEVHLPAPSAVRAARPGFGGDKGR
 25 LVGVAERQVPCRFLRPTGVLSLFLGGAVVSLQYVRPSALRTLLDQSAKDSTVHEILSKLSLEGDH
 STPPSAYGSVKPYTNFDAE

[SEQ ID NO: 17]

ATP152

30 MKRYKNRVASRKSRKFKQLLQHYREVAAAKSSENDRLRLLKNRTLTLFNVTRNDARAYVSGI
 QNSVSANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYSWRINGIPQQHTQVLFIAKITPNN
 NGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSATEYKLVVVGAVGVGKSALTIQLIQVAAAR

PVLPAPAISDWERARRRHVGRKEGAARARVSLPEFGTAEVHLPAPSAVRAARPGFGGDKGRL
VGVAERQVPCRFLRPTGVLSLFLGGAVVSLQYVRPSALRTLDDQSAKDSTVHEILSKLSLEGDHS
TPPSAYGSKPYTNFDAE
[SEQ ID NO: 18]

5

Next, the immunogenicity of the antigens/antigenic fragments contained in the multi-antigenic domain of ATP150 and ATP152 was tested in monocyte-derived human dendritic cells (moDCs). To this end, cells were extracted from fresh human Leukapheresis products named Leukopaks supplied by Tissue Solution (Scotland). The HLA typing of the 15 donors was done by the Immunology and Transplant Unit and National Reference Laboratory for Histocompatibility (LNRH) of the Geneva University Hospital. Peripheral blood mononuclear cells (PBMCs) from leukopaks samples were prepared after density gradient centrifugation on Ficoll-Paque and monocytes were enriched by cold aggregation. Remaining T lymphocytes were eliminated using rosette formation with sheep red blood cells and density gradient centrifugation on Ficoll-Paque. Monocytes were then plated with complete RPMI medium supplemented with GM-CSF and IL-4 during 7 days. Cells were then concentrated and Interferon alpha 2a was added together with 600nM of ATP150 or of ATP152. Cells were incubated overnight at 37°C and 5% CO₂. Cells were then scrapped, numerated. and centrifuged. The supernatant was aspirated and the cell pellet was frozen at -20°C before ligandome analysis.

The cell pellet was lysed with a solubilization buffer and homogenized by pulsed sonification. After debris elimination and sterile filtration, HLA class I and class II molecules were isolated from the soluble fraction using standard immunoaffinity purification. The HLA ligands were then eluted by acid elution and isolated by ultracentrifugation, desalted and pre-concentrated prior to LC-MS/MS analysis.

Peptide samples were separated by nanoflow high performance liquid chromatography (RSLCnano, Thermo Fisher Scientific) using a 50 µm x 25 cm PepMap rapid separation liquid chromatography column (Thermo Fisher Scientific) and a gradient ranging from 2.4% to 32.0% acetonitrile over the course of 90 min. Eluting peptides were analyzed in an online-coupled LTQ Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) using a

top speed collision-induced dissociation (CID) fragmentation method for HLA class I or a higher-energy collisional dissociation (HCD) fragmentation method for HLA class II. For label-free quantification (LFQ) of the relative HLA ligand abundances between ATP150 (or ATP152)-loaded and non-loaded activated samples, the injected peptide amounts of paired
5 samples calculated from the average precursor ion intensities determined in dose-finding runs were normalized and LC-MS/MS analysis was performed in five technical replicates for each sample.

The raw files were processed using Proteome Discoverer 1.4 (Thermo Fisher Scientific). The
10 SEQUEST HT search engine (University of Washington) was used to search the human proteome as comprised in the Swiss-Prot database combined with the construct ATP150 (or ATP152). After data processing specific filters were applied. For class I the data were filtered with an FDR of 5%, peptide lengths of 8-12 amino acids and search engine rank 1. For class
15 II the data were filtered with an FDR of 1%, peptide lengths of 8-25 amino acids and search engine rank 1. Protein inference was disabled, allowing for multiple protein annotations of peptides. To determine HLA class I binding peptides, the NetMHCpan-3.0 algorithm with a percentile binding rank below 2% and the SYFPEITHI algorithm with a threshold of 50% of the maximum score were used.

20 Results are shown in Figure 4. These data show that monocyte-derived human dendritic cells (moDCs), which were incubated with ATP150 or ATP152, presented MHC class I and/or MHC class II epitopes derived from all incorporated antigens, further validating that endogenous processing in professional antigen-presenting cells can drive the display of meaningful antigens to patients' CD8 and CD4 T cells. CEACAM5 and DUOXA2 epitopes are
25 presented by MHC class I in human GI tract tumors *in vivo*.

The Ligandome analysis study confirms the correct delivery of the ATP150 and ATP152 into human moDCs cells, the translation of the multi-antigenic domain of ATP150 and ATP152, and the multi-epitopic and multi-allelic processing and presentation of the all the antigens of
30 the multi-antigenic domain of ATP150 and ATP152.

Example 5: CEACAM5- and DUOXA2-specific T cell response of ATP150/VSV-GP154 heterologous prime-boost

Next, antigen-specific T-cell responses were assessed in a heterologous prime-boost vaccination scheme, wherein ATP150 (as described in Example 4 above) was combined with a recombinant vesicular stomatitis virus (VSV) encoding corresponding antigens/antigenic fragments (referred to herein as "VSV-GP154") as described below.

The multi-antigenic domain encoded in the VSV comprises the antigens/antigenic fragments according to SEQ ID NOs 1, 3 and 5 – 12. The multi-antigenic domain of VSV-GP154 has the following amino acid sequence:

MNRTLTLFNVTRNDARAYVSGIQNSVSANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYS
WRINGIPQQHTQVLFIKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSATEYKLV
VVGADGVGKSALTILQTEYKLVVVGAVGVGKSALTILQVAAARPVLPAPAISDWERARRR
HVGKRKEGAARARVSLPEFGTAEVHLPAPSAVRAARPGFGGDKGRLVGVAERQVPCRFLRPTG
VLSLFLGGAVVSLQYVRPSALRTLLDQSAKD
[SEQ ID NO: 19]

Specifically, VSV-GP154 has a sequence according to SEQ ID NO: 30 (RNA sequence of VSV-GP154, complement without reverse). The corresponding cDNA sequence is provided in SEQ ID NO: 49.

To test antigen-specific T-cell responses were assessed in a heterologous prime-boost vaccination scheme, peripheral immune response at day 21 (1 week after VSV-GP154 boost) was investigated to determine CEACAM5 and DUOXA2-specific T cells by enzyme-linked immunospot (ELISpot) in mice. To this end, C57BL/6 mice (6 and 10 weeks old) were injected with 2 cycles of vaccination: a first s.c. vaccination with 10 nmoles of ATP150; a boost with 10 nmoles of ATP150 s.c. or 10^7 TCID₅₀ of VSV-GP154 i.v. at day 14. At day 21, spleen were harvested. Spleen cells were isolated and assessed in ELISpot assay. ELISpot assay was performed using the Murine IFN γ ELISpot Diaclone kit (Ref 862.031.015S) according to manufacturer instructions. Splenocytes were isolated with Ficoll Paque Plus, GE Healthcare,

ref 171440.02 and plated in a concentration gradient from 2×10^5 – 1.0×10^6 cells. Cells were pulsed with overlapping 11mer peptides corresponding to the CEACAM5 and DUOXA2 regions incorporated in the ATP150 and VSV-GP154. Spots were counted using ELISPOT reader.

5

Figure 5 shows the prime-boost protocol (A) and the results (B). Boost with VSV-GP154 promotes potent increase of CEACAM5 and DUOXA2-specific IFN γ -producing T cells as compared to homologous prime-boost.

10

Example 6: Heterologous prime-boost vaccination in a mouse tumor model

To assess treatment effects like tumor volume and survival rates in a heterologous prime-boost vaccination regimen with a peptide construct and recombinant VSV, a mouse tumor model (TC-1 tumor model) was used. To this end, the multi-antigenic domain of the peptide construct and the multi-antigenic domain of the VSV needed to be adapted to the model tumor. The multi-antigenic domains as described in above examples 4 and 5 were designed for use in humans, while animal tumor models require vaccination with antigens/antigenic fragments corresponding to the model tumor.

20

To this end, the multi-antigenic domain described in above examples 4 and 5 was replaced with "multi-antigenic domain 25" (Mad-25; SEQ ID NO: 20), which contains both CD8 and CD4 H-2b epitopes from E7 HPV. Accordingly, the peptide construct of the present example comprises the multi-antigenic domain Mad25. VSV-GP-HPV encodes the attenuated E6/E7 fusion construct (Cassetti et al., 2004, Vaccine 22(3-4): 520-527) in addition to wild-type E2.

25

Briefly, TC-1 cells were provided by T.C. Wu (Johns Hopkins University, Maryland, US) and cultured in complete RPMI 1640 with 0.4 mg/ml geneticin. For tumor implantation, mice were injected subcutaneously with 1×10^5 TC-1 cells in the back. Seven days later, mice were immunized with 2 nmol of the peptide construct (multi-antigenic domain "Mad25"; in Figure 6 "K") prime s.c. or with 1×10^7 TCID $_{50}$ VSV-GP-HPV i.v. (in Figure 6 "V"). 14 days post tumor implantation mice received 1×10^7 TCID $_{50}$ VSV-GP-HPV i.v.. Additional doses of the peptide

30

construct "K" and the virus "V" were administered as indicated by dotted lines in Fig. 6A. For monitoring tumor growth, tumor diameter was measured 2-3 times per week using a caliper and volume was calculated using the formula: $0.4 \times \text{length} \times \text{width}^2$. Mice were sacrificed when tumor-size reached the size specified by the respective institutional veterinary authorities or tumors showed signs of ulcerations.

Results are shown in Figure 6. Virus treatment ("V") alone resulted in delayed tumor growth (Fig. 6A). However, virus treatment ("boost") after the "prime" with the peptide construct led additionally to a complete remission in all tumors; even in large tumors (Fig. 6B). This strongly indicates that priming with the peptide (K) can induce tumor regression of sizeable tumors treated with virus two weeks after grafting. These data suggest that the peptide (K) prime lays the immunological foundation for the strong tumor remission that follows the virus (V) boost in the TC-1 tumor model.

15

Example 7: Immunogenicity and efficacy of heterologous prime-boost in combination with anti-PD-1 in tumor-bearing mice

To assess immunogenicity and efficacy of combinatory treatment with heterologous prime-boost vaccination and a checkpoint inhibitor (anti-PD-1), the TC-1 mouse tumor model was used, essentially as described in Example 6 above. As described in Example 6, TC-1 cells were implanted and mice received the peptide construct "K" (with the multi-antigenic domain Mad25) and the VSV-GP-HPV "V", as described in Example 6. Mice were also administered 200 µg i.v. of PD-1 antibody (clone RMP1-14, BioXcell, Lebanon, New Hampshire, US) at day 7, 15, 28 and 49 according to the schedule shown in Figure 7A. To assess circulating anti-HPV CD8 T cells, mice were bled at days 14, 21, 35, 42, 56, and multimer staining detecting HPV-specific CD8 T cells was performed and analyzed by flow cytometry.

Results are shown in Figures 7 and 8. As shown in Figure 7B, circulating levels of HPV-specific CD8 T cells were increased similarly in the KVKK regimen, independently whether mice received anti-PD-1 treatment or not. As shown in Figure 8, therapeutic vaccination starting at day 7 after tumor cell implantation mediated a strong anti-tumor effect in the TC-1

tumor model in both groups treated with KVKK. Increased survival was observed when mice received additionally anti-PD-1 injection at each vaccination, with 4 mice out of 7 remaining tumor-free in the triple combination group (Figure 8B).

5

Example 8: Effects of different VSV doses in heterologous prime-boost vaccination

To assess the efficacy of different doses of VSV-GP-HPV "V" in heterologous prime-boost vaccination, the TC-1 tumor model, as described in examples 6 and 7 above, was used. For
10 tumor implantation, mice were injected subcutaneously with 1×10^5 TC-1 cells in the back. Seven days later, mice were immunized with 2 nmol of the peptide construct ("K") prime s.c.. VSV-GP-HPV boost was administrated at 3 different doses: 1×10^7 , 3×10^7 or 5×10^7 TCID₅₀ at day 14 after tumor cell implantation. Fourteen days later, mice received 2 nmol of peptide construct ("K").

15

Results are shown in Figure 9. The anti-tumor effect observed previously is highlighted by a similar transient remission in the three K/V/K groups as observed by the strong decreased of the tumor volume after the first boost (VSV-GP-HPV injection). Although the group at VSV-GP-HPV 5×10^7 TCID₅₀ shows a slightly enhanced survival, no major differences are observed
20 in the 3 treated groups displaying the different doses of VSV-GP-HPV.

Example 9: Effects of plurality of immunogenic epitopes including the fragments of KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2

25

To investigate the effects of a plurality of immunogenic epitopes including the fragments of KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 in the peptide according to the present invention (ATP150), a comparative construct lacking the fragments of KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 ("ATP132") was prepared.

30

The multi-antigenic domain of ATP132 has to following sequence:

NRTLTLFNVTRNDARAYVSGIQNSVNSANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYSWR
 INGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSATEYKLVVV
 GADGVGKSALTIQLIQ
 [SEQ ID NO: 50]

5

In addition to said multi-antigenic domain, ATP132 comprises the cell penetrating peptide according to SEQ ID NO: 15 and the TLR agonist according to SEQ ID NO: 16.

In summary, the peptide construct ATP132 exhibits the following sequence:

10 MKRYKNRVASRKSRAKFKQLLQHYREVAAAKSENDRLRLLKNRTLTLFNVTRNDARAYVSGI
 QNSVNSANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQYSWRINGIPQQHTQVLFIAKITPNN
 NGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSATEYKLVVVGADGVGKSALTIQLIQSTVHE
 ILSKLSLEGDHSTPPSAYGSVKPYTNFDAE
 [SEQ ID NO: 51]

15

The immunogenicity of the antigens/antigenic fragments contained in the multi-antigenic domain of ATP150 and ATP132 was compared in monocyte-derived human dendritic cells (moDCs). To this end, cells were extracted from fresh human Leukapheresis products named Leukopaks supplied by Tissue Solution (Scotland). The HLA typing of the donors was done
 20 by the Immunology and Transplant Unit and National Reference Laboratory for Histocompatibility (LNRH) of the Geneva University Hospital. Peripheral blood mononuclear cells (PBMCs) from leukopaks samples were prepared after density gradient centrifugation on Ficoll-Paque and monocytes were enriched by cold aggregation. Remaining T lymphocytes were eliminated using rosette formation with sheep red blood cells and density gradient
 25 centrifugation on Ficoll-Paque. Monocytes were then plated with complete RPMI medium supplemented with GM-CSF and IL-4 during 7 days. Cells were then concentrated and Interferon alpha 2a was added together with 600 nM of ATP150 or 600 nM of ATP132. Cells were incubated overnight at 37°C and 5% CO₂. Cells were then scrapped, numerated. and centrifuged. The supernatant was aspirated, and the cell pellet was frozen at -20°C before
 30 ligandome analysis.

The cell pellet was lysed with a solubilization buffer and homogenized by pulsed sonification. After debris elimination and sterile filtration, HLA class I and class II molecules were isolated from the soluble fraction using standard immunoaffinity purification. The HLA ligands were then eluted by acid elution and isolated by ultracentrifugation, desalted and pre-concentrated prior to LC-MS/MS analysis.

Peptide samples were separated by nanoflow high performance liquid chromatography (RSLCnano, Thermo Fisher Scientific) using a 50 μ m x 25 cm PepMap rapid separation liquid chromatography column (Thermo Fisher Scientific) and a gradient ranging from 2.4% to 32.0% acetonitrile over the course of 90 min. Eluting peptides were analyzed in an online-coupled LTQ Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) using a top speed collision-induced dissociation (CID) fragmentation method for HLA class I or a higher-energy collisional dissociation (HCD) fragmentation method for HLA class II. For label-free quantification (LFQ) of the relative HLA ligand abundances between ATP150- or ATP132-loaded and non-loaded activated samples, the injected peptide amounts of paired samples calculated from the average precursor ion intensities determined in dose-finding runs were normalized and LC-MS/MS analysis was performed in five technical replicates for each sample.

The raw files were processed using Proteome Discoverer 1.4 (Thermo Fisher Scientific). The SEQUEST HT search engine (University of Washington) was used to search the human proteome as comprised in the Swiss-Prot database combined with the construct ATP150 or ATP132. After data processing specific filters were applied. For class I the data were filtered with an FDR of 5%, peptide lengths of 8-12 amino acids and search engine rank 1. For class II the data were filtered with an FDR of 1%, peptide lengths of 8-25 amino acids and search engine rank 1. Protein inference was disabled, allowing for multiple protein annotations of peptides. To determine HLA class I binding peptides, the NetMHCpan-3.0 algorithm with a percentile binding rank below 2% and the SYFPEITHI algorithm with a threshold of 50% of the maximum score were used.

The results are depicted in Table 2 below:

# peptide/donor	ATP132	ATP150
MHC I	1.2	2.0
MHC II	22.8	79.3

(Table 2)

In Table 2, the average number of detected peptide epitopes for MHC class I and MHC class II per donor is shown.

5

These data show that monocyte-derived human dendritic cells (moDCs), which were incubated with a peptide according to the present invention comprising fragments of KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 (ATP150) presented more MHC class I and/or MHC class II epitopes, which are derived from all incorporated antigens, than moDCs incubated with the comparative peptide lacking the fragments of KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2 (ATP132). These data further validate that a higher number of immunogenic fragments result in a superior endogenous processing in professional antigen-presenting cells which can drive the display of meaningful antigens to patients' CD8 and CD4 T cells.

10
15

TABLE OF SEQUENCES AND SEQ ID NUMBERS (SEQUENCE LISTING):

SEQ ID NO	Sequence	Remarks
SEQ ID NO: 1	VAAARPVLPAPAI SDWERARRRH	KRAS-uORF1 (short)
SEQ ID NO: 2	VAAAKVAAARPVLPAPAI SDWERARRRH	KRAS-uORF1 (long)
SEQ ID NO: 3	VGRKEGAARARV SLLPEFGTAEVHLPAPSAVRAA RPGF	TPX2-uORF1 (long)
SEQ ID NO: 4	TAEVHLPAPSAVRAARPGF	TPX2-uORF1 (short)
SEQ ID NO: 5	GGDKGRLVGVAERQVPCRFLRP	AURKA-uORF2
SEQ ID NO: 6	NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPV TLDVL	CEACAM5 fragment
SEQ ID NO: 7	PDSSYLSGANLNLSCHSAS	CEACAM5 fragment
SEQ ID NO: 8	PQYSWRINGIPQQHTQVLFIKITPNNNGTYAC FVSNLATGRNNSIVKSITVSASGTSPGLSA	CEACAM5 fragment
SEQ ID NO: 9	NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPV TLDVLPDSSYLSGANLNLSCHSASPQYSWRINGI PQQHTQVLFIKITPNNNGTYACFVSNLATGRN NSIVKSITVSASGTSPGLSA	CEACAM5 fragment fusion construct
SEQ ID NO: 10	TGVLSLFLGGAVVSLQYVRPSALRTLDDQSAKD	DUOXA2 fragment
SEQ ID NO: 11	TEYKLVVVGADGVGKSALTIQLIQ	KRAS-G12D fragment
SEQ ID NO: 12	TEYKLVVVGAVGVGKSALTIQLIQ	KRAS-G12V fragment
SEQ ID NO: 13	NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPV TLDVLPDSSYLSGANLNLSCHSASPQYSWRINGI PQQHTQVLFIKITPNNNGTYACFVSNLATGRN NSIVKSITVSASGTSPGLSATEYKLVVVGADGVG KSALTIQLIQVAAARPVLPAPAI SDWERARRRHV GRKEGAARARV SLLPEFGTAEVHLPAPSAVRAAR PGFGGDKGRLVGVAERQVPCRFLRPTGVLSLFL GGAVVSLQYVRPSALRTLDDQSAKD	multi-antigenic domain ATP150
SEQ ID NO: 14	NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPV TLDVLPDSSYLSGANLNLSCHSASPQYSWRINGI PQQHTQVLFIKITPNNNGTYACFVSNLATGRN NSIVKSITVSASGTSPGLSATEYKLVVVGAVGVG KSALTIQLIQVAAARPVLPAPAI SDWERARRRHV GRKEGAARARV SLLPEFGTAEVHLPAPSAVRAAR PGFGGDKGRLVGVAERQVPCRFLRPTGVLSLFL GGAVVSLQYVRPSALRTLDDQSAKD	multi-antigenic domain ATP152

SEQ ID NO: 15	KRYKNRVASRKSRKFKQLLQHYREVAAAKSEN DRLRLLK	cell penetrating peptide (Z13)
SEQ ID NO: 16	STVHEILSKLSLEGDHSTPPSAYGSVKPYTNFDAE	TLR agonist (Anaxa)
SEQ ID NO: 17	MKRYKNRVASRKSRKFKQLLQHYREVAAAKSS ENDRLRLLKLNRTLTLFNVTRNDARAYVSGIQNS VSANRSDPVTLDVLPDSSYLSGANLNLSCHSASP QYSWRINGIPQQHTQVLFIKITPNNNGTYACF VSNLATGRNNSIVKSITVSASGTSPGLSATEYKLV VVGADGVGKSALTILQVAAARPVLPAPASD WERARRRHVGRKEGAARARVSLPEFGTAEVHL PAPSAVRAARPGFGGDKGRLVGVAERQVPCRF LRPTGVLSLFLGGAVVSLQYVRPSALRTLLDQSA KDSTVHEILSKLSLEGDHSTPPSAYGSVKPYTNF AE	ATP150
SEQ ID NO: 18	MKRYKNRVASRKSRKFKQLLQHYREVAAAKSS ENDRLRLLKLNRTLTLFNVTRNDARAYVSGIQNS VSANRSDPVTLDVLPDSSYLSGANLNLSCHSASP QYSWRINGIPQQHTQVLFIKITPNNNGTYACF VSNLATGRNNSIVKSITVSASGTSPGLSATEYKLV VVGAVGVGKSALTILQVAAARPVLPAPASD WERARRRHVGRKEGAARARVSLPEFGTAEVHL PAPSAVRAARPGFGGDKGRLVGVAERQVPCRF LRPTGVLSLFLGGAVVSLQYVRPSALRTLLDQSA KDSTVHEILSKLSLEGDHSTPPSAYGSVKPYTNF AE	ATP152
SEQ ID NO: 19	MNRTLTLFNVTRNDARAYVSGIQNSVSANRSDP VTLDVLPDSSYLSGANLNLSCHSASPQYSWRIN GIPQQHTQVLFIKITPNNNGTYACFVSNLATG RNNSIVKSITVSASGTSPGLSATEYKLVVVGADG VGKSALTILQTEYKLVVVGAVGVGKSALTILQ QVAAARPVLPAPASDWERARRRHVGRKEGA RARVSLPEFGTAEVHLPAPSAVRAARPGFGGDK GRLVGVAERQVPCRF LRPTGVLSLFLGGAVVSL QYVRPSALRTLLDQSAKD	multi-antigenic domain VSV-GP154
SEQ ID NO: 20	QAEPDRAHYNIVTFSSKS	multi-antigenic domain Mad25
SEQ ID NO: 21	KRYKNRVASRKSRKFKQLLQHYREVAAAK	cell penetrating peptide (Z14)
SEQ ID NO: 22	KRYKNRVASRKSRKFK	cell penetrating peptide (Z15)
SEQ ID NO: 23	REVAAAKSSENDRLRLLK	cell penetrating peptide (Z18)

SEQ ID NO: 24	STVHEILCKLSLEGDHSTPPSAYGSVKPYTNFDAE	TLR agonist (Anaxa variant)
SEQ ID NO: 25	MGQIVTMFEALPHIIDVINIVIIIVLIITSIKAVYNF ATCGILALVSFLFLAGRSCGMYGLNGPDIYKGVY QFKSVEFDMSHLNLTMPNACSANNSSHYYISMG SSGLELTFTNDSILNHNFCNLTSAFNKKTFDHTL MSIVSSLHLSIRGNSNHKAVSCDFNNGITIQYNL SFSDPQSAISQCRTFRGRVLDMFRTAFGGKYMR SGWGWAGSDGKTTWCSTSYQYLIQNRWE NHCRYAGPFGMSRILFAQEKTFLTRRLAGTFT WTLSDSSGVENPGGYCLTKWMILAAELKCFGNT AVAKCNVNHDEEFCMDMLRLIDYNKAALSKFKQ DVESALHVFKTTVNSLISDQLLMRNHLRDLMGV PYCNYSKFWYLEHAKTGETSVPKCWLVTNGSYL NETHFSDQIEQEADNMITEMLRKDYIKRQGSTPL ALMDLLMFSTSAYLISIFLHLVKIPTHRHKGGSCP KPHRLTNKGICSCGAFKVPGVKTIWKRR	GP of LCMV
SEQ ID NO: 26	MSVTVKRIIDNTVVVPKLPANEDPVEYPADYFRK SKEIPLYINTTKSLSDLRGVYQGLKSGNVSIHV NSYLYGALKDIRGKLDKDWSSFGINIGKAGDTI GIFDLVSLKALDGVLPDGVSDASRTSADDKWLP LYLLGLYRVGRTQMPEYRKKLMDGLTNQCKMI NEQFEPLVPEGRDIFDVWGNDNSNYTKIVAARD MFFHMFKKHECASFRYGTIVSRFKDCAALATFG HLCKITGMSTEDVTTWILNREVADEMVMMLP GQEIDKADSYMPYLIDFGLSSKSPYSSVKNPAFH FWGQLTALLRSTRARNARQPDDIEYTSLTAGL LYAYAVGSSADLAQQFCVGDNKYTPDDSTGGL TTNAPPQGRDVVEWLGWFEQNRKPTPDMM QYAKRAVMSLQGLREKTIGKYAKSEFDK	vesicular stomatitis virus nucleoprotein (N)
SEQ ID NO: 27	MDNLTKVREYLKSYSRLDQAVGEIDEIAQRAEK SNYELFQEDGVEEHTKPSYFQAADDSDTESEPEIE DNQGLYAPDPEAEQVEGFIQGPLDDYADEEVD VVFTSDWKQPELESDEHGKTLRLTSPEGLSGEQK SQWLSTIKAVVQSAKYWNLAECTFEASGEGVIM KERQITPDVYKVTVMNTHPSQSEAVSDVWSLS KTSMTFQPKKASLQPLTISLDELFSRGEFISVGGD GRMSHKEAILLGLRYKKLYNQARVKYSL	vesicular stomatitis virus phosphoprotein (P)

SEQ ID NO: 28	MEVHDFETDEFNDFNEDDYATREFLNPDERMTY LNHADYNLNSPLISDDIDNLIRKFNSLPIPSMWD SKNWDGVLEMLTSCQANPIPTSQMHKWMGS WLMSDNHDASQGYSFLHEVDKEAEITFDVVETF IRGWGNKPIEYIKKERWTDSEKILAYLCQKFLDLH KLTILNAVSEVELLNLARTFKGKVRSSHGTNIC RIRVPSLGPTFISEGWAYFKKLDILMDRNFLLMVK DVIIGRMQTVLSMVCRIDNLFSEQDIFSLLNIYRI GDKIVERQGNFSYDLIKMVEPICNLKMKLARES RPLVPQFPHFENHIKTSVDEGAKIDRGIRFLHDQI MSVKTVDLTLVIYGSFRHWGHPFIDYYTGLEKLH SQVTMKKDIDVSYAKALASDLARIVLFQQFNDH KKWFEVNGDLLPHDHPFKSHVKENTWPTAAQV QDFGDKWHELPLIKCFEIPDLLDPSIISDKSHSM NRSEVLKHVRMNPNTPIPSKKVLQTMMLDTKATN WKEFLKEIDEKGLDDDDLIIGLKGERELKLAGRF FSLMSWKLREYFVITEYLIKTHFVPMFKGLTMAD DLTAVIKKMLDSSSGQGLKSYEAICIANHIDYEK WNNHQRKLSNGPVFRVMGQFLGYPSLIERTHEF FEKSLIYYNGRPDLMRVHNNTLINSTSQRVCWQ GQEGGLEGLRQKGWSILNLLVIQREAKIRNTAV KVLAQGDNQVICTQYKTKKSRNVVELQGALNQ MVSNNEKIMTAIKIGTGKLGLLINDDETMQSAD YLNYGKIPIFRGVIRGLETKRWSRVTCVTNDQIPT CANIMSSVSTNALTVAHFAENPINAMIQYNYFG TFARLLLMMHDPALRQSLYEVQDKIPGLHSSTFK YAMLYLDPSIGGVSGMSLSRFLIRAFDPVTELSF WRFIHVHARSEHLKEMSAVFGNPEIAKFRITHIDK LVEDPTSLNIAMGMSPANLLKTEVKKCLIESRQTI RNQVIKDATIYLYHEEDRLRSFLWSINPLFPRFLSE FKSGTFLGVADGLISLFQNSRTIRNSFKKKYHREL DDLIVRSEVSSLTHLGKLHLRRGCKMWTCSAT HADTLRYKSWGRTVIGTTVPHPLEMLGPQHRKE TPCAPCNTSGFNYSVHCPDGIHDFSSRGPLP AYLGSKTSESTSILQPWERESKVPLIKRATRLRDAIS WFEVPDSKLAMTILSNIHSLTGEEWTKRQHGFKR TGSALHREFSTRMSHGGFASQSTAALTRLMATT DTMRDLGDQNFDFLFQATLLYAQITTTVARDG WITSCTDHYHIACKSCLRPiEEITLDSSMDYTPPD VSHVLKTRWRNGEGSWGQEIKQIYPLEGNWKNL APAEQSYQVGRCIGFLYGDLAYRKSTHAEDSSLF	vesicular stomatitis virus large protein (L)
---------------	--	---

	PLSIQGRIRGRGFLKGLLDGLMRASCCQVIHRRS LAHLKR PANAVYGGLIYLIDKLSVSPFLSLTRSGP IRDELETIPHKIPTSYP TSNRDMGVIVRNYFKYQC RLIEKGKYRSHYSQLWLFSDVLSIDFIGPFSISTLL QILYKPFLSGKDKNELRELANLSSLLRSGEGWEDI HVKFFTKDILLCP EIRHACKFGIAKDNNKDMSY PPWGRESRG TITTIPVYYTTTPYPKMLEMPPRIQN PLLSGIRLGQLPTGAHYKIRSILHGMGIHYRDFLS CGDGSGG MTAALLRENVHSRGIFNSLLELSGSV MRGASPEPPSALET LGGDKSRCVNGETCWEYPS DLCDPRTWDYFLRLKAGLGLQIDLIVMDMEVR DSSTSLKIETNVRNYVHRILDEQGVLIYKTYGTI CESEKNAV TILGPMFKTVDLVQTEFSSSQ TSEVY MVCKGLKKLIDEPNP DWSSINESWKNLYAFQSSE QEFARAKKVSTYFTLTGIPSQFIPDPFVNIETMLQI FGVPTGVSHAAALKSSDRPADLLTISLFYMAISYY NINHIRVGPIPPNPPSDGIAQNVGIAITGISFWLS LMEKDIPLYQQCLAVIQQSFP IRWEAVSVKGGY KQKWSTRGDGLPKDTRISDSLAPIGNWIRSLELV RNQVRLNPFNEILFNQLCRTVDNHLKWSNLRR NTGMIEWINRRISKEDRSILMLKSDLHEENSWRD	
SEQ ID NO: 29	MSSLKILGLKGKGKSKKLGIAPPPYEEDTSMEY APSAPIDKSYFGVDEMDTYDPNQLRYEKFFFTVK MTVRSNRPFRTYSDVAAVSHWDHMYIGMAG KRPFYKILAF LGSSNLKATPAVLADQGQPEYHAH CEGRAYLPHRMGKTPPMLNVPEHFRRPFNIGLY KGTIELTMTIYDDESLEAAPMIWDHFNSSKFSDFR EKALMFG LIVEKKASGAWVLDSIGHFK	vesicular stomatitis virus matrix protein (M)

SEQ ID NO: 30	UGCUUCUGUUUGUUUGGUAUAAUAGUA AUUUUCCGAGUCCUCUUUGAAAUUGUCA UUAGUUUUACAGACAAUGUCAGUUCUCU UAGUAAACUGUUGUGUCAGCAUCAAGGUU UUGAAGGACGUUUACUCCUAGGUCACCU UAUGGGCCGUCUAAUGAAGUCUUUUAGU UUCCUCUAAGGAGAAAUGUAGUUAUGAU GUUUUUCAAACAGUCUAGAUUCUCCUAUA CAGAUGGUUCCGGAGUUUAGGCCUUUAC AUAGUUAGUAUGUACAGUUGUCGAUGAA CAUACCUCGUAAUUUCCUGUAGGCCCCAU UCAACCUAUUUCUAACCAGUUCAAAGCCU UAUUUUGUAGCCCUUUCGUCCCCUAUGUU AGCCUUAUAAACUGGAACAUAGGAACUUU CGGGACCUGCCGCAUGAAGGUCUACCUCU UAGCCUACGAAGGUCUUGGUCGCGUCUA CUGUUUACCAACGGAAACAUAGAUGAACC GAAUAUGUCUCACCCGUCUUGUGUUUAC GGACUUAUGUCUUUUUUCGAGUACCUAC CCGACUGUUUAGUUACGUUUUACUAGUU ACUUGUCAAAACUUGGAGAACACGGUCUUC CAGCACUGUAAAAACUACACACCCCUUUAC UGUCAUUAUGUGUUUUUAAACAGCGACG UCACCUGUACAAGAAGGUGUACAAGUUUU UUGUACUUACACGGAGCAAGUCUAUGCCU UGAUAACAAAGGUCUAAGUUUCUAAACAG ACGUAACCGUUGUAAACCUGUGGAGACGU UUUAUUGGCCUUACAGAUGUCUUCUACA UUGCUGGACCUAGAACUUGGCUCUUCAAC GUCUACUUUACCAGGUUUACUACGAAGG UCCGGUUCUUUAAACUGUCCGGCUAAGU AUGUACGGAAUAAACUAGCUGAAACCUAA CAGAAGAUUCAGAGGUUAUAAGAAGGCAGU UUUUGGGACGGAAGGUGAAGACCCCGU UAACUGUCGAGAAGACGAGUCUAGGUGG UCUCGUUCCUACGGGCUGUCGGACUAC UGUAAACUUAUGUAGAGAAUGAUGUCG UCCAAACAUAUGCGAAUACGUCAUCCUA GGAGACGGCUGAACCGUGUUGUCAAAACA CAACCUCUAUUGUUUAUGUGAGGUCUAC UAUCAUGGCCUCCUAAACUGCUGAUUACGU GGCGGUGUCCCGUCUCUACACCAGCUUAC CGAGCCUACCAAACUUCUAGUUUUUGUCUU UUGGCUGAGGACUUAACUACGUCAUACGC UUUUCUCGUCAGUACAGUGACGUUCCGG AUUCUCUCUUCUGUUAACCGUUCUACGA UUCAGUCUUAACUGUUUACUGGGUAU	RNA sequence VSV- GP154 (complement without reverse)
---------------	--	---

UAAGAGUCUAGUGGAUAAUUAUAAUACG AUGUAUACUUUUUUUGAUUGUCUUAAGU ACCUAUUAGAGUGUUUUCAAGCACUCAUA GAGUUCAGGAUAAGAGCAGACCUAGUCCG CCAUCCUCUCUAUCUACUCUAGCUUCGUG UUGCUCGACUUUUCAGGUUAAUACUCAAC AAGGUUCUCCUACCUCACCUUCUCGUAUG AUUCGGGAGAAUAAAAGUCCGUCGUCUAC UAAGACUGUGUCUUAAGACUUGGUCUUUA ACUUCUGUUAGUUCCGAACAUACGUGGU CUAGGUCUUCGACUCGUUCAACUCCGAA AU AUGUCCCCGGAAUUCUACUGAUACGUC UACUCCUUCACCUACAACAUAUAAUGAAGCC UGACCUUUGUCGGACUCGAACUUAAGACU GCUCGUACCUUUCUGGAAUGCCAACUGUA GCGGUCUCCCAAUUCACCUCUCGUCUUU AGGGUCACCGAAAGCUGCUAAUUUCGUCA GCACGUUUCACGGUUUAUGACCUUAGACC GUCUCACGUGUAAACUUCGUAGCCCUUU CCCCAGUAAUACUUCUCCGCGGUCUAUUG AGGCCUACAUAUAUUCAGUGAGGUCACU ACUUGUGUGUAGGCAGGGUUAGUCUUCG UCAUAGUCUACAAACCAGAGAGAGUUUCU GUAGGUACUGAAAGGUUGGGUUCUUUCG UUCAGAAGUCGGAGAGUGGUUAUAGGAAC CUACUUAACAAGAGUAGAUCUCCUCUCAA GUAGAGACAGCCUCCACUGCCUGCUUACA GAGUAUUUCUCCGGUAGGACGAGCCGGA CUCUAUGUUUUUCAACAUGUUAGUCCGC UCUCAGUUUAUAAGAGACAUCUGAUACUU UUUUUCAUUGUCUAUAGUGCUAGAUUCA CAAUAGGGUUAGGUAAAGUAGUACUCAAG GAAUUUCUUCUAAGAGCCAGACUUCCCCU UCCAUUCUUUAGAUUCUUUAAUCCCUA GCGUGGUGGGGGAAUACUUCUCCUGUGA UCGUACCUCAUACGAGGCUCGCGAGGUUA ACUGUUUAGGAUAAAACCUCAACUGCUCU ACCUGUGGAUACUAGGCUUAGUUAUUC UAUACUCUUUAAGAAGAAAUGUCACUUUU ACUGCCAAUCUAGAUUAGCAGGCAAGUCU UGUAUGAGUCUACACCGUCGGCGACAUAG GGUAACCCUAGUGUACAUGUAGCCUUAAC GUCCCUUUGCAGGGAAGAUGUUUUAGAA CCGAAAAAACCAAGAAGAUUAGAUUUC GGUGAGGUCGCCAUAAACCGUCUAGUCCA GUUGGUCUCAUAGUGCGAGUGACGCUUC CGUCCCGAAUAAACGGUGUAUCCUACCCC	
---	--

UUCUGGGGAGGGUACGAGUUACAUGGUC UCGUGAAGUCUUCUGGUAAGUUUAUACC AGAAAUGUUCCCUUGCUAACUCGAGUGU UACUGGUAGAUGCUACUACUCAGUGACCU UCGUCGAGGAUACUAGACCCUAGUAAAGU UAAGAAGGUUUAAAAGACUAAAGUCUCUC UUCCGGAAUUACAAACCGGACUACAGCU CUUUUUCGGUAGACCUCGCACCCAGGACC UGAGAUAGCCGGUGAAGUUUACUCGAUC AGAUUGAAGAUCCAAGACUUGUUAGGGG CCAAAUGAGUCAGAGGGGAUUAAGGUCG GAGAGCUUGUUGAUUAUAGGACAGAAAA GAUAGGGGAUACUUUUUUUGAUUGUCUCU AGCUAGACAAAUGCGCAGUGCCUAGGGGG CCCGACGUCCUUAAGCGGUGGUACCCGGU CUAGCACUGGUACAAGCUCCGGGACGGGG UGUAGUAGCUGCUCCACUAGUUGUAGCA CUAGUAGCACGAGUAGUAGUAGUGGUCG UAGUUCCGGCACAUGUUGAAGCGGUGGA CGCCGUAGGACCGGGACCACUCGAAGGAC AAGGACCGGCCGUCUUCGACGCCGUACAU GCCGGACUUACCGGGGCUAUAGAUGUUC CCGCACAUGGUCAAGUUCUCGCACCUCAA GCUGUACUCGGUGGACUUGGAGUGGUAC GGGUUGCGGACGUCGCGGUUGUUUACGG UGGUGAUGUAGUCGUACCCGUCGUCGCC GGACCUCAACUGGAAGUGGUUGCUGUCG UAGGACUUGGUGUUGAAGACGUUGGAGU GGUCGCGGAAGUUGUUCUUUUGGAAGCU GGUGUGGGAGUACUCGUAGCACUCGUCG GACGUGGACUCGUAGUCUCCGUUGUCGU UGGUGUUCCGGCACUCGACGCUGAAGUU GUUGCCGUAGUGGUAGGUCAUGUUGGAC UCGAAGUCGCUAGGAGUCUCGCGGUAGU CGGUCACGUCUUGGAAGUCUCCGUCUCAC GACCUGUACAAGUCUUGGCGGAAGCCGCC GUUCAUGUACUCUUCGCCGACCCCGACCC GGCCGUCGCUGCCGUUCUGGUGGACCAC GUCGGUCUGGUCGAUGGUCAUGGAGUAG UAGGUCUUGUCUUGGACCCUCUUGGUGA CGUCUAUGCGGCCUGGAAAGCCGUACUCG UCUUAAGGACAAGCGGGUCCUCUUUUGGU UCAAGGAGUGGUCCUCUGACCGGCCGUG GAAGUGGACCUGGGACUCGCUGUCGUCG CCGCACCUCUUGGGACCGCCGAUGACGGA GUGGUUCACCUACUAGGACCGGCGGCUC GACUUCACGAAGCCGUUGUGGCGGCACCG	
--	--

GUUCACGUUGCACUUGGUGCUGCUCUC AAGACGCUGUACGACUCUGAGUAGCUGAU GUUGUUCCGGCGGGACUCGUUCAAGUUC GUCCUGCACCUCUCGCGGGACGUGCACAA GUUCUGGUGGCACUUGUCGGAGUAGUCG CUGGUCGACGAGUACUCUUUGGUGGACU CUCUGGAGUACCCGCACGGGAUGACGUU GAUGUCGUUCAAGACCAUAGACCUCGUGC GGUUCUGGCCGCUCUGGUCGCACGGGUU CACGACCGACCACUGGUUACCGUCGAUGG ACUUGCUCUGGGUGAAGUCGCUGGUCUA GCUCGUCCUUCGGCUGUUGUACUAGUGG CUCUACGACUCCUCCUGAUGUAGUUCUC UGUCCCGUCGUGGGGGGACCGGGAGUAC CUAGACGAGUACAAGUCGUGGUCGCGGA UGGAGUAGUCGUAGAAGGACGUGGACCA CUUCUAGGGGUGGGUGUCUGUGUAGUU CCCGCCGUCGACGGGGUUCGGGGUGUCU GAGUGGUUGUUCCCGUAGACGUCGACGC CGCGGAAGUUCCACGGGCCGCACUUUUG GUAGACCUUCUCCUCUAUUCGCCGGCGAU GCUGGAGCUGAUACUUUUUUUGAUUGUC UAUAGGAGCUGCGGUGGUACUUGGCCUG UGACUGGGACAAGUUGCACUGGGCCUUA CUGCGGUCUCGGAUGCACAGGCCGUAGG UCUUAUCGCACAGGCGGUUAUCUUCGCU GGGGCACUGUGACCUGCACGACGGACUUA CGUCGAUAGACUCGCCGCGGUUGGACUU GGACUCGACGGUAAGACGGAGAGGUGUC AUGUCGACCGCCUAGUUGCCUUAAGGGAG UCGUCGUGUGGGUCCACGACAAUAGCG GUUCUAGUGGGGAUUGUUGUUGCCGUG GAUGCGGACGAAGCACAGGUUAGACCGGU GGCCGUCUUUGUUGUCGUAGCACUUCAG GUAGUGGCACAGACGGUCGCCGUGUAGA GGACCUGACUCGCGGUGUCUCAUGUUCG ACCACCAACACCCUCGACUGCCGCACCCUU UUUCGCGGGACUGUUAGGUCGACUAGGU CUGGCUCAUAUUUGAGCACCAGCAGCCGC GACAACCGCAACCGUUUAGACGAGACUGG UAAGUCGAGUAAGUCCACCGGCGACGGUC UGGACAAGAAGGACGAGGACGGUAGUCGC UGACCCUUUCUGAUCUUCUUCGUGCA GCCGGCCUUUCUUCGCGGGCGAUCUCGA UCUCAAUCCGACGACGGUCUCAAACCGUG GCGGCUUCAAGUAGACGGACGAGGAUCGC GACAGUCUCGGCGGUCUGGUCCUAAACCG	
---	--

CCUCUAUCCCCGUCUGAGCACCCGCAACG UCUUUCUGUCCACGGGACGUCUAAGGAC UCCGGUUGUCCGCACGACUCGGACAAAGA ACCGCCUCGACACCACAGGGACGUCAUGC AGUCUGGAUCGCGAGACUCUUGGGACGA CCUAGUCUCGCGGUUCCUAACUAGUUGA UCGGUCUAAGAAGUACAAACCUGGUUUAG UUGAACACUAUGGUACGAGUUUCUCCGG AGUUAUUUAUAAACUCAAAAAUUAAAAUAC UUUUUUUUGAUUGUCGUUAGUACCUUCAG GUGCUAAAACUCUGGCUGCUCAAGUUACU AAAGUUACUUCUACUGAUACGGUGUUCU CUUAAGGACUUAGGGCUACUCGCGUACU GCAUGAACUUAGUACGACUAAUGUUGGAC UUAAGAGGAGAUUAAUCACUACUUAACU GUUAAAUUAGUCCUUUAAGUUAAGAGAA GGUUAAGGGAGCUACACCCUAUCAUUCU GACCCUACCUCAAGAACUCUACAAUUGCA GUACAGUUCGGUUAGGGUAGGGUUGUAG AGUCUACGUUUUUACCUACCCUUAACCA AUUACAGACUAAUAGUACUACGGUCAGUU CCCAUAUCAAAAAAUGUACUUCACCUGUU UCUCCGUCUUUAUUGUAAACUGCACCACC UCUGGAAGUAGGCGCCGACCCCGUUGUU UGGUUAAACUUUAGUAGUUUUUCCUUUCU ACCUGACUGAGUAAGUUUUAAGAGCGAAU AAACACAGUUUUUCAAACCUGAAUGUGU UCAACUGUAAUUAGAAUUUACGACAGAGA CUCCACCUUAACGAGUUGAACCGCUCCUG AAAGUUUCCGUUUCAGUCUUCUUAAGA GUACCUUGCUUGUAUACGUCCUAAUCCCA AGGGUCGAACCCAGGAUGAAAUAAGUC UCCUACCCGAAUGAAGUUCUUUGAACUA UAAGAUUACCUGGCUUUGAAAGACAAUUA CCAGUUUCUACACUAAUAUCCCUCCUACG UUUGCCACGAUAGGUACCAUACAUCUUUAU CUGUUGGACAAGAGUCUCGUUCUGUAGA AGAGGGAAGAUUUUAUAGAUGUCUUAACC UCUAUUUUUAACACCUCUCCGUCUUUUAA AAAGAAUACUGAACUAAUUUUUACCACCUU GGCUAUACGUUGAACUUCGACUACUUUAA UCGUUCUCUUAGUUCGGGAAAUACAGGGU GUUAAGGGAGUAAAACUUUUAGUAUAGU UCUGAAGACAACUACUUCUCCCGUUUUUAA CUGGCUCCAUAUUCUAAGGAGGUACUAG UCUAUUACUCACACUUUUUGUCACCUAGAG UGUGACCACUAAAUAACCUAGCAAGUCUGU	
--	--

AACCCAGUAGGAAAAUAUCUAAUAAUGU GACCUGAUCUUUUUAAUGUAAGGGUUCA UUGGUACUUCUUUCUAUAACUACACAGUA UACGUUUUCGUGAACGUUCACUAAAUCGA GCCUAACAAGAUAAAGUUGUCAAGUUACU AGUAUUUUUCACCAAGCACUUACCUCUGA ACGAGGGAGUACUAGUAGGGAAAUUUUC AGUACAAUUUCUUUUUAUGUACCGGGUGU CGACGAGUUCAAGUUCUAAAACCUCUAUU UACCGUACUUGAAGGCGACUAAUUUACAA AACUUUAUGGGCUGAAUGAUCUGGGUAG CUAUUAUAUGAGACUGUUUUCAGUAAGU UACUUAUCCAGUCUCCACAACUUUGUACA GGCUUACUUAGGCUUGUGAGGAUAGGGA UCAUUUUUCCACAACGUCUGAUACAACCU GUGUUUCCGAUGGUUAACCUUUCUAAA GAAUUUCUCUAACUACUCUCCCCGAAUCU ACUACUACUAGAUUAAUAACCAGAAUUUC CUUUCUCUCCCUUGACUUAACCGUCCA UCUAAAAAGAGGGAUUACAGAACCUUUA CGCUCUUAUGAAACAUUAAUGGCUUAUA ACUAAUUCUGAGUAAAGCAGGGAUACAAA UUUCCGGACUGUUACCGCCUGCUAGAUU GACGUCAGUAAUUUUUCUACAAUCUAAGG AGUAGGCCGGUCCUAACUUCAGUAUACU CCGUUAAACGUUACGGUUAGUGUAACUAA UGCUUUUUACCUUAUUGGUGGUUCCCU CAAUAGUUUGCCGGGUCACAAGGCUCAAU ACCCGGUCAAGAAUCCAUAAGGUAGGAAU UAGCUCUCUUGAGUACUAAAAAACUCUU UUCAGAAUAUAUGAUGUUACCUUCUGGU CUGAACUACGCACAAGUGUUGUUGUGUG ACUAGUUAAGUUGGAGGGUUGCUCAAAC AACCGUUCCUGUUCUCCACCUGACCUUC CAGAUGCCGUUUUCCUACCUCAUAGGAG UUAGAUGACCAUAAGUUUCUCUCCGAUU UUAGUCUUUGUGACGACAGUUUCAGAAC CGUGUCCACUAUUAGUUCUAAUAAACGUG UGUCAUAUUUUGCUUCUUUAGCUCUUUG CAACAUCUUAUUGUCCCACGAGAGUUAGU UUACCAAAGAUUAUUACUCUUUUAAUACU GACGUUAGUUUUUACCCUGUCCCUUCA UCCUGAAAACUAUUUACUGCUACUCUGAU ACGUUAGACGUCUAAUGAACUUAUACCU UUUUUAUGGCUAAAAGGCACCUCACUAAUC UCCCAAUCUCUGGUUCUCUACCAGUGCUC ACUGAACACAGUGGUUACUGGUUUUUGG	
---	--

GUGAACACGAUUUAUUAUACUCGAGUCAAA GGUGUUUACGAGAGUGGCAUCGAGUAAA ACGACUCUUGGGUUAGUUACGGUACUAU GUCAUGUUAUAAAACCCUGUAAACGAUC UGAGAACAACUACUACGUACUAGGACGAG AAGCAGUUAGUAACAUACUUCAAGUUCUA UUCUAUGGCCCCGAACGUGUCAAGAUGAAA GUUUUAUGCGGUACAACAUAUAAACCUGGGAA GGUAACCUCUCACAGCCCCGUACAGAAACA GGUCCAAAAACUAAUCUCGGAAGGGGUCUA GGGCAUUGUCUUUCAGAGAGUAAGACCU CUAAGUAGGUACAUGUACGAGCUUCACUC GUAGACUCCUCUACUCACGUCAUAAACC UUUGGGGCUCUAUCGGUUCAAAGCUUAU UGAGUGUAUCUGUUCGAUCAUCUUCUAG GUUGGAGAGACUUGUAGCGAUACCCUUA CUCAGGUCGCUUGAACAAUUUCUGACUCC AAUUUUUUACGAAUUAGCUUAGUUCUGU UUGGUAGUCCUUGGUCCACUAAUCCUA CGUUGGUUAUUAUAAACAUAGUACUUCUCCU AGCCGAGUCUUCAAAGAAUACCAGUUAUU UAGGAGACAAGGGAUCUAAAAAUUCACUU AAGUUUAGUCCGUGAAAAAACCCUCAGCG UCUGCCCCGAGUAGUCAGAUAAAGUUUUAA GAGCAUGAUAAAGCCUUGAGGAAAUUCUU UUUCAUAGUAUCCCUUAACCUACUAAACU AACACUCCUCACUCCAUAAGGAGAAACUGU GUAAAUCCCUUUGAAGUAAACUCUUCCCC UAGUACAUUUUACACCGUACAAGUCGAU GAGUACGACUGUGUAAUUCUAUGUUUAG GACCCCGGCAUGUCAAUAAACCCUGUUGAC AUGGGGUAGGUAAUCUUUACAACCCAGG UGUUGUAGCUUUUCUCUGAGGAACACGU GGUACAUUGUGUAGUCCCAAGUUAUACA AAGACACGUAAACAGGUCUGCCCCUAGGUAC UGCAGAAAUCAAGUGCCCCUGGUAAACGGA CGAAUAGAUAUCCAGAUUUUGUAGACUUAG AUGUAGAUAAAACGUCGGAACCCUUUCCC UUUCGUUUCAGGGUGACUAAUUUUCUCG AUGUGCAGAAUCUCUACGAUAGAGAACCA AACAAUUGGGCUGAGAUUUUGAUCGUUA CUGAUUAGAAAGAUUGUAGGUGAGAAAU UGUCCGCUUCUUAACCGGUUUUCCGUCG UACCCAAGUUUUCUUGUCCAGACGGGAA GUAUCCAAAAGCUGUAGAGCCUACUCGGU ACCACCAAGCGUAGAGUCUCGUGACGUC GUAACUGGUCCAACUACCGUUGAUGUCU	
--	--

GUGGUACUCCCUAGACCCUCUAGUCUUA AGCUGAAAAUAAGGUUCGUUGCAACGAG AUACGAGUUUAUGGUGGUGACAACGUU CUCUGCCUACCUAGUGGUCAACAUGUCUA GUAUAGUAUAACGGACAUUCAGGACAAA CUCUGGGUAUCUUCUCUAGUGGGACCUG AGUUCAUACCUGAUGUGCGGGGGUCUAC AUAGGGUACACGACUUCUGUACCUCCUUA CCCCUCCAAGCACCCCUGUUCUCUAUUU UGUCUAGAUAGGAAAUCUUCCCUUAACCU UCUUAUUCGUGGACGACUCGUUAGGAU AGUUCAGCCGUCUACAUAUCCAAAAGAU UACCUCUGAACCGCAUAUCUUUUAGAUGA GUACGGCUCUGUCAAGAGAUAAAGGAGA UAGAUUGUUCCAGCAUAAUCUCCAGCUC CAAAGAAUUUCCCCAACGAUCUGCCUAAU UACUCUCGUUCAACGACGGUUCAUUAUG UGGCCUCUUCAGACCGAGUAAACUUCUCC GGCCGGUUGCGUCACAUGCCUCCAAACUA AAUGAACUAAUUAUUAACUCACAUAGUG GAGGUAAGGAAAGAGAAUGAUCUAGUCCU GGAUAAUCUCUGCUUAAUCUUUGCUAAG GGGUGUUCUAGGGUUGGAGGAUAGGCU GUUCGUUGGCACUAUACCCCCACUAACAG UCUUUAAUGAAGUUUAUGGUUACGGCAG AUUAACUUUCCCCUUUAUGUCUAGUGU AAUAAGUGUAAUACCAUAAGAGUCUAC AGAAUAGGUUUCUGAAGUAACCUGGUAAG AGAUAAAGGUGGUGGGAGAACGUUUAGG AUAUGUUCGGUAAAAUAGACCCUUUCUA UUCUUACUCAACUCUCUCGACCGUUUAGA AAGAAGUAACGAUUCUAGUCCUCUCCCCA CCCUUCUGUAUGUACACUUUAAGAAGUG GUUCCUGUAUAAUAAACACAGGUCUCCUUU AGUCUGUACGAACGUUCAAGCCCUAACGA UCCUAUUAAUUAUUCUGUACUCGAUAG GGGAACCCCUUCCCUUAGGUCUCCUGU UAAUGUUGUUAGGGACAAUAAUAGCU GGUGGGGAUUGGUUUCUACGAUCUCUA CGGAGGUUCUAGGUUUUAGGGGACGAC AGGCCUUAGUCCAACCCGGUUAAGGUU GACCGCGAGUAAUAAUUUAAGCCUCAU AAUGUACCUUACCCUUAGGUAAUGUCCCU GAAGAACUCAACACCUCUGCCGAGGCCUC CCUACUGACGACGUAAUGAUGCUCUUUA CACGUUUCGUCUCCUUAUAAGUUAUCAGA CAAUCUUAUAGUCCAGUCAGUACGCUC	
--	--

CGCGGAGAGGACUCGGGGGGGUCACGGGA UCUUUGAAAUCCUCCUCUAUUUAGCUCUA CACAUUUACCACUUUGUACAACCCUUAUA GGUAGACUGAAUACACUGGGUUCCUGAAC CCUGAUAAAGGAGGCUGAGUUUCGUCCG AACCCCGAAGUUUAACUAAAUAACAUA CCUAUACCUUCAAGCCCUAAGAAGAUGAU CGGACUUUUAAACUCUGCUUACAAUCUUUA AUACACGUGGCCUAAAACCUACUCGUUCC UCAAAAUUAGAUGUUCUGAAUACCUUGUA UAUAAACACUCUCGCUUUUCUUAACGUCAU UGUUAGGAACCAGGGUACAAGUUCUGCCA GCUGAAUCAAGUUUGUCUAAAUCAUCA GAGUUUGCAGACUUCAUAUAUACCAUACA UUUCCAAACUUCUUUAAUUAAGCUACUUG GGUUAGGGCUAACCAGAAGGUAGUUACU UAGGACCUUUUUGGACAUGCGUAAGGUC AGUAGUCUUGUCCUAAACGGUCUCGUU UCUUCCAAUCAUGUAUGAAAUGGAACUGU CCAUAAGGGAGGGUUAAGUAAGGACUAG GAAAACAUUUGUAACUCUGAUACGAUGUU UAUAAGCCUCAUGGGUGCCCCACACAGAGU ACGCCGACGGAAUUUUAGUAGACUAUCUG GACGUCUAAAUAACUGGUAAUCGGAAAAA AUUAACCGCUAAUAUAGCAUAAUAUUGUA GUUAGUAUAGUCUCAUCCUGGCUAUGGA GGCUUGGGGGGUAGUCUACCUUAACGUG UUUUACACCCCUAGCGAUUUUGACCAUAU UCGAAAACCGACUCAAAACUACCUCUUUCU GUAAGGUGAUUAAGUUGUCACAAAUCGU CAUAAGGUCGUUAGUAAGGGCUAAUCCAC CCUCCGACAAAGUCAUUUUCUCCUAUGU UCGUCUUCACCUCAUGAUCUCCACUACCC GAGGGUUUUCUAUGGGCUUAAAGUCUGA GGAACCGGGGUAGCCCUUGACCUAGUC UAGAGACCUUAACCAGGCUUUGGUUCAAG CAGAUUUAGGUAAGUUACUCUAGAACAAG UUAGUCGAUACAGCAUGUCACCUAUUAGU AAACUUUACCAGUUUAAACGCUUCUUUGU GUCCUUAACUAAACUACCUAGUUAUCUGCU UAAAGUUUUCUUCUGGCCAGAUUAGACU ACAACUUCUCACUGGAUGUGCUCCUUUU GAGAACCUCUCUAAUUUUUUUAGUACUCCU CUGAGGUUUGAAAUUCUAUCUUUUUUUG AAACUAGGAAUUCUGGGAGAACACCAAAA AUAAAAAUAGACCAAACACCAGAAGCA	
--	--

SEQ ID NO: 31	EVMLVESGGGLVQPGGSLRLSCTASGFTFSASA MSWVRQAPGKGLEWVAYISGGGGDTYYSSSV KGRFTISRDNANKNSLYLQMNSLRAEDTAVYYCA RHSNVNYYAMDYWGQGTLVTVSSASTKGPSVF PLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNS GALTSGVHTFPAVLQSSGLYSLSSVVPSSSLGT KTYTCNVDPKPSNTKVDKRVEVKYGPCCPCPA PEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDV SQEDPEVQFNWYVDGVEVHNAKTKPREEQFNS TYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSS IEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLT CLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD SDGSFFLYSRLTVDKSRWQEGNVFSCSV MHEAL HNHYTQKSLSLGLG	heavy chain of anti- PD1 antibody PD1-1
SEQ ID NO: 32	EIVLTQSPATLSLSPGERATMSCRASENIDTSGISF MNWYQQKPGQAPKLLIYVASNQGSIGIPARFSG SGSGTDFLTISRLEPEDFAVYYCQQSKEVPWTF GQGTELEIKRTVAAPSVFIFPPSDEQLKSGTASVV CLLNNFYPRKAKVQWKVDNALQSGNSQESVTE QDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQ GLSSPVTKSFNRGEC	light chain of anti-PD1 antibody PD1-1
SEQ ID NO: 33	EVMLVESGGGLVQPGGSLRLSCTASGFTFSASA MSWVRQAPGKGLEWVAYISGGGGDTYYSSSV KGRFTISRDNANKNSLYLQMNSLRAEDTAVYYCA RHSNPNYYAMDYWGQGTLVTVSSASTKGPSVF PLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNS GALTSGVHTFPAVLQSSGLYSLSSVVPSSSLGT KTYTCNVDPKPSNTKVDKRVEVKYGPCCPCPA PEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDV SQEDPEVQFNWYVDGVEVHNAKTKPREEQFNS TYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSS IEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLT CLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD SDGSFFLYSRLTVDKSRWQEGNVFSCSV MHEAL HNHYTQKSLSLGLG	heavy chain of anti- PD1 antibody PD1-2
SEQ ID NO: 34	EIVLTQSPATLSLSPGERATMSCRASENIDTSGISF MNWYQQKPGQAPKLLIYVASNQGSIGIPARFSG SGSGTDFLTISRLEPEDFAVYYCQQSKEVPWTF GQGTELEIKRTVAAPSVFIFPPSDEQLKSGTASVV CLLNNFYPRKAKVQWKVDNALQSGNSQESVTE QDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQ GLSSPVTKSFNRGEC	light chain of anti-PD1 antibody PD1-2

SEQ ID NO: 35	EVMLVESGGGLVQPGGSLRLSCTASGFTFSKSA MSWVRQAPGKGLEWVAYISGGGGDTYYSSSV KGRFTISRDNANKNSLYLQMNSLRAEDTAVYYCA RHSNVNYYAMDYWGQGT LVTVSSASTKGPSVF PLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNS GALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGT KTYTCNVDPKPSNTKVDKRVESKYGPPCPPCPA PEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDV SQEDPEVQFNWYVDGVEVHNAKTKPREEQFNS TYRVVSVLT VLVHQLDNLNGKEYKCKVSNKGLPSS IEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLT CLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD SDGSFFLYSRLTVDKSRWQEGNVFSCSV MHEAL HNHYTQKSLSLGLG	heavy chain of anti- PD1 antibody PD1-3
SEQ ID NO: 36	EIVLTQSPATLSLSPGERATMSCRASENIDVSGISF MNWYQQKPGQAPKLLIYVASNQGSIGIPARFSG SGSGTDFTLTISRLEPEDFAVYYCQQSKEVPWTF GQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVV CLLNNFYPREAKVQWKVDNALQSGNSQESVTE QDSKSTYLSSTLTLSKADYEKHKVYACEVTHQ GLSSPVTKSFNRGEC	light chain of anti-PD1 antibody PD1-3
SEQ ID NO: 37	GTGGCCGCTGCCAGACCTGTTCTTCCTGCTC CTGCCATCAGCGACTGGGAAAGAGCTAGAA GAAGGCAC	KRAS-uORF1
SEQ ID NO: 38	GTCGGCCGGAAAGAAGGCGCCGCTAGAGCT AGAGTTAGCCTGCTGCCAGAGTTTGGCACCG CCGAAGTTCATCTGCCTGCTCCTAGCGCTGT CAGAGCCGCCAGACCAGGATT	TPX2-uORF1
SEQ ID NO: 39	GGCGGAGATAAGGGCAGACTCGTGGGCGTT GCAGAAAGACAGGTGCCCTGCAGATTCCTGA GGCCA	AURKA-uORF2
SEQ ID NO: 40	AACCGGACACTGACCCTGTTCAACGTGACCC GGAATGACGCCAGAGCCTACGTGTCCGGCA TCCAGAATAGCGTGTCCGCCAATAGAAGCGA CCCCGTGACACTGGACGTGCTG	CEACAM5 fragment
SEQ ID NO: 41	CCTGATAGCAGCTATCTGAGCGGCGCCAACC TGAACCTGAGCTGCCATTCTGCCTCT	CEACAM5 fragment
SEQ ID NO: 42	CCACAGTACAGCTGGCGGATCAACGGAATCC CTCAGCAGCACACCCAGGTGCTGTTTATCGC CAAGATCACCCCTAACAACAACGGCACCTAC GCCTGCTTCGTGTCCAATCTGGCCACCGGCA GAAACAACAGCATCGTGAAGTCCATCACCGT GTCTGCCAGCGGCACATCTCCTGGACTGAGC GCC	CEACAM5 fragment

SEQ ID NO: 43	GGCGGAGATAAGGGCAGACTCGTGGGCGTT GCAGAAAGACAGGTGCCCTGCAGATTCCTGA GGCCA	DUOXA2 fragment
SEQ ID NO: 44	ACAGAGTACAAGCTGGTGGTGTGGGAGCT GACGGCGTGGGAAAAAGCGCCCTGACAATC CAGCTGATCCAG	KRAS-G12D fragment
SEQ ID NO: 45	ACCGAGTATAAACTCGTGGTCGTCCGGCGCTG TTGGCGTTGGCAAATCTGCTCTGACCATTGAG CTCATTGAG	KRAS-G12V fragment
SEQ ID NO: 46	KRYKNRVASRKSRAKFKQLLQHYREVAAAKSSEN DRLRLLLNRTLTLEFNVTRNDARAYVSGIQNSVS ANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQY SWRINGIPQQHTQVLFIKITPNNNGTYACFVS NLATGRNNSIVKSITVSASGTSPGLSATEYKLVVV GADGVGKSALTIQLIQVAAARPVLPAPAISDWER ARRRHVGRKEGAARARVSLPEFGTAEVHLPAPS AVRAARPGFGGDKGRLVGVAERQVPCRFLRPT GVLSLFLGGAVVSLQYVRPSALRTLTDQSAKDST VHEILSKLSLEGDHSTPPSAYGSVKPYTNFDAE	ATP150 (v2)
SEQ ID NO: 47	KRYKNRVASRKSRAKFKQLLQHYREVAAAKSSEN DRLRLLLNRTLTLEFNVTRNDARAYVSGIQNSVS ANRSDPVTLDVLPDSSYLSGANLNLSCHSASPQY SWRINGIPQQHTQVLFIKITPNNNGTYACFVS NLATGRNNSIVKSITVSASGTSPGLSATEYKLVVV GAVGVGKSALTIQLIQVAAARPVLPAPAISDWER ARRRHVGRKEGAARARVSLPEFGTAEVHLPAPS AVRAARPGFGGDKGRLVGVAERQVPCRFLRPT GVLSLFLGGAVVSLQYVRPSALRTLTDQSAKDST VHEILSKLSLEGDHSTPPSAYGSVKPYTNFDAE	ATP152 (v2)
SEQ ID NO: 48	NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPV TLDVLPDSSYLSGANLNLSCHSASPQYSWRINGI PQQHTQVLFIKITPNNNGTYACFVSNLATGRN NSIVKSITVSASGTSPGLSATEYKLVVVGADGVG KSALTIQLIQTEYKLVVVGAVGVGKSALTIQLIQV AAARPVLPAPAISDWERARRRHVGRKEGAARAR VSLPEFGTAEVHLPAPSAVRAARPGFGGDKGRL VGVAERQVPCRFLRPTGVLSLFLGGAVVSLQYV RPSALRTLTDQSAKD	multi-antigenic domain VSV-GP154 (v2)
SEQ ID NO: 49	ACGAAGACAAACAAACCATTATTATCATTAAA AGGCTCAGGAGAACTTTAACAGTAATCAAA ATGTCTGTTACAGTCAAGAGAATCATTGACAA CACAGTCGTAGTTCCAAAACCTCCTGCAAATG AGGATCCAGTGGAATACCCGGCAGATTACTT CAGAAAATCAAAGGAGATTCTCTTTACATCA ATACTACAAAAAGTTTGTCTCAGATCTAAGAGG	cDNA sequence of VSV-GP154

ATATGTCTACCAAGGCCTCAAATCCGGAAAT GTATCAATCATACATGTCAACAGCTACTTGTA TGGAGCATTAAAGGACATCCGGGGTAAGTT GGATAAAGATTGGTCAAGTTTCGGAATAAAC ATCGGGAAAGCAGGGGATACAATCGGAATA TTTGACCTTGTATCCTTGAAAGCCCTGGACG GCGTACTTCCAGATGGAGTATCGGATGCTTC CAGAACCAGCGCAGATGACAAATGGTTGCCT TTGTATCTACTTGGCTTATACAGAGTGGGCA GAACACAAATGCCTGAATACAGAAAAAAGCT CATGGATGGGCTGACAAATCAATGCAAAATG ATCAATGAACAGTTTGAACCTCTTGTGCCAGA AGGTCGTGACATTTTGTATGTGTGGGGAAAT GACAGTAATTACACAAAAATTGTCGCTGCAG TGGACATGTTCTTCCACATGTTCAAAAAACAT GAATGTGCCTCGTTCAGATACGGAACATTG TTTCCAGATTCAAAGATTGTGCTGCATTGGCA ACATTTGGACACCTCTGCAAAATAACCGGAA TGTCTACAGAAGATGTAACGACCTGGATCTT GAACCGAGAAGTTGCAGATGAAATGGTCCA AATGATGCTTCCAGGCCAAGAAATTGACAAG GCCGATTCATACATGCCTTATTTGATCGACTT TGGATTGTCTTCTAAGTCTCCATATTCTTCCGT CAAAAACCCTGCCTTCCACTTCTGGGGGCAA TTGACAGCTCTTCTGCTCAGATCCACCAGAG CAAGGAATGCCCCGACAGCCTGATGACATTGA GTATACATCTCTTACTACAGCAGGTTTGTGT ACGCTTATGCAGTAGGATCCTCTGCCGACTT GGCACAACAGTTTTGTGTTGGAGATAACAAA TACTCTCAGATGATAGTACCGGAGGATTGA CGACTAATGCACCGCCACAAGGCAGAGATG TGGTCGAATGGCTCGGATGGTTTGAAGATCA AAACAGAAAACCGACTCCTGATATGATGCAG TATGCGAAAAGAGCAGTCATGTCACTGCAAG GCCTAAGAGAGAAGACAATTGGCAAGTATG CTAAGTCAGAATTTGACAAATGACCCTATAAT TCTCAGATCACCTATTATATATTATGCTACATA TGAAAAAACTAACAGATATCATGGATAATCT CACAAAAGTTCGTGAGTATCTCAAGTCCTATT CTCGTCTGGATCAGGCGGTAGGAGAGATAG ATGAGATCGAAGCACAACGAGCTGAAAAGT CCAATTATGAGTTGTTCCAAGAGGATGGAGT GGAAGAGCATACTAAGCCCTCTTATTTTCAG GCAGCAGATGATTCTGACACAGAATCTGAAC CAGAAATTGAAGACAATCAAGGCTTGTATGC ACCAGATCCAGAAGCTGAGCAAGTTGAAGG CTTTATACAGGGGCCTTTAGATGACTATGCA	
---	--

GATGAGGAAGTGGATGTTGTATTTACTTCGG ACTGGAAACAGCCTGAGCTTGAATCTGACGA GCATGGAAAGACCTTACGGTTGACATCGCCA GAGGGTTTAAGTGGAGAGCAGAAATCCCAG TGGCTTTCGACGATTAAAGCAGTCGTGCAAA GTGCCAAATACTGGAATCTGGCAGAGTGCAC ATTTGAAGCATCGGGAGAAGGGGTCATTATG AAGGAGCGCCAGATAACTCCGGATGTATATA AGGTCACTCCAGTGATGAACACACATCCGTC CCAATCAGAAGCAGTATCAGATGTTTGGTCT CTCTCAAAGACATCCATGACTTTCCAACCCAA GAAAGCAAGTCTTCAGCCTCTCACCATATCCT TGGATGAATTGTTCTCATCTAGAGGAGAGTT CATCTCTGTCGGAGGTGACGGACGAATGTCT CATAAAGAGGCCATCCTGCTCGGCCTGAGAT ACAAAAAGTTGTACAATCAGGCGAGAGTCAA ATATTCTCTGTAGACTATGAAAAAAGTAACA GATATCACGATCTAAGTGTTATCCCAATCCAT TCATCATGAGTTCCTTAAAGAAGATTCTCGGT CTGAAGGGGAAAGGTAAGAAATCTAAGAAA TTAGGGATCGCACCACCCCTTATGAAGAGG ACACTAGCATGGAGTATGCTCCGAGCGCTCC AATTGACAAATCCTATTTTGGAGTTGACGAGA TGGACACCTATGATCCGAATCAATTAAGATAT GAGAAATTCTTCTTTACAGTGAAAATGACGGT TAGATCTAATCGTCCGTTCAGAACATACTCAG ATGTGGCAGCCGCTGTATCCCATTGGGATCA CATGTACATCGGAATGGCAGGGGAAACGTCC CTTCTACAAAATCTTGGCTTTTTTGGGTTCTC TAATCTAAAGGCCACTCCAGCGGTATTGGCA GATCAAGGTCAACCAGAGTATCACGCTCACT GCGAAGGCAGGGCTTATTTGCCACATAGGAT GGGGAAGACCCCTCCCATGCTCAATGTACCA GAGCACTTCAGAAGACCATTCAATATAGGTC TTTACAAGGGAACGATTGAGCTCACAATGAC CATCTACGATGATGAGTCACTGGAAGCAGCT CCTATGATCTGGGATCATTTCATTCTTCCAA ATTTTCTGATTCAGAGAGAAGGCCTTAATGT TTGGCCTGATTGTCGAGAAAAAGGCATCTGG AGCGTGGGTCCTGGACTCTATCGGCCACTTC AAATGAGCTAGTCTAACTTCTAGCTTCTGAAC AATCCCCGGTTTACTCAGTCTCCCCTAATTCC AGCCTCTCGAACAATAATATCCTGTCTTTTCT ATCCCTATGAAAAAACTAACAGAGATCGAT CTGTTTACGCGTCACGGATCCCCCGGGCTGC AGGAATTCGCCACCATGGGCCAGATCGTGA CCATGTTTCGAGGCCCTGCCCCACATCATCGA	
---	--

CGAGGTGATCAACATCGTGATCATCGTGCTC ATCATCATCACCAGCATCAAGGCCGTGTACA ACTTCGCCACCTGCGGCATCCTGGCCCTGGT GAGCTTCCTGTTCTGCGCCGGCAGAAAGCTGC GGCATGTACGGCCTGAATGGCCCCGATATCT ACAAGGGCGTGTACCAGTTCAAGAGCGTGG AGTTCGACATGAGCCACCTGAACCTCACCAT GCCCAACGCCTGCAGCGCCAACAATAGCCA CCACTACATCAGCATGGGCAGCAGCGGCCT GGAGTTGACCTTCACCAACGACAGCATCCTG AACCACAACCTTCTGCAACCTCACCAGCGCCTT CAACAAGAAAACCTTCGACCACACCCTCATG AGCATCGTGAGCAGCCTGCACCTGAGCATCA GAGGCAACAGCAACCACAAGGCCGTGAGCT GCGACTTCAACAACGGCATCACCATCCAGTA CAACCTGAGCTTCAGCGATCCTCAGAGCGCC ATCAGCCAGTGCAGAACCTTCAGAGGCAGA GTGCTGGACATGTTCAGAACCGCCTTCGGCG GCAAGTACATGAGAAGCGGCTGGGGCTGG GCCGGCAGCGACGGCAAGACCACCTGGTGC AGCCAGACCAGCTACCAGTACCTCATCATCC AGAACAGAACCTGGGAGAACCACTGCAGAT ACGCCGGACCTTTCGGCATGAGCAGAATCCT GTTTCGCCCAGGAGAAAACCAAGTTCCTCACC AGGAGACTGGCCGGCACCTTCACCTGGACC CTGAGCGACAGCAGCGGCGTGGAGAACCCT GGCGGCTACTGCCTCACCAAGTGGATGATCC TGGCCGCCGAGCTGAAGTGCTTCGGCAACA CCGCCGTGGCCAAGTGCAACGTGAACCACG ACGAGGAGTTCTGCGACATGCTGAGACTCAT CGACTACAACAAGGCCGCCCTGAGCAAGTTC AAGCAGGACGTGGAGAGCGCCCTGCACGTG TTCAAGACCACCGTGAACAGCCTCATCAGCG ACCAGCTGCTCATGAGAAACCACCTGAGAGA CCTCATGGGCGTGCCCTACTGCAACTACAGC AAGTTCTGGTATCTGGAGCACGCCAAGACCG GCGAGACCAGCGTGCCCAAGTGCTGGCTGG TGACCAATGGCAGCTACCTGAACGAGACCCA CTTCAGCGACCAGATCGAGCAGGAAGCCGA CAACATGATCACCGAGATGCTGAGGAAGGA CTACATCAAGAGACAGGGCAGCACCCCCCT GGCCCTCATGGATCTGCTCATGTTCAGCACC AGCGCCTACCTCATCAGCATCTTCCTGCACCT GGTGAAGATCCCCACCCACAGACACATCAAG GGCGGCAGCTGCCCCAAGCCCCACAGACTC ACCAACAAGGGCATCTGCAGCTGCGGCGCC TTCAAGGTGCCCCGGCGTGAAAACCATCTGGA	
---	--

AGAGGAGATAAGCGGCCGCTACGACCTCGA CTATGAAAAAACTAACAGATATCCTCGACG CCACCATGAACCGGACACTGACCCTGTTCAA CGTGACCCGGAATGACGCCAGAGCCTACGT GTCCGGCATCCAGAATAGCGTGTCCGCCAAT AGAAGCGACCCCGTGACACTGGACGTGCTG CCTGATAGCAGCTATCTGAGCGGCGCCAACC TGAACCTGAGCTGCCATTCTGCCTCTCCACA GTACAGCTGGCGGATCAACGGAATCCCTCA GCAGCACACCCAGGTGCTGTTTATCGCCAAG ATCACCCCTAACAACAACGGCACCTACGCCT GCTTCGTGTCCAATCTGGCCACCGGCAGAAA CAACAGCATCGTGAAGTCCATCACCGTGTCT GCCAGCGGCACATCTCCTGGACTGAGCGCC ACAGAGTACAAGCTGGTGGTTGTGGGAGCT GACGGCGTGGGAAAAAGCGCCCTGACAATC CAGCTGATCCAGACCGAGTATAAACTCGTGG TCGTCCGGCGCTGTTGGCGTTGGCAAATCTGC TCTGACCATTGAGCTCATTGAGGTGGCCGCT GCCAGACCTGTTCTTCCTGCTCCTGCCATCAG CGACTGGGAAAGAGCTAGAAGAAGGCACGT CGGCCGGAAAGAAGGCGCCGCTAGAGCTA GAGTTAGCCTGCTGCCAGAGTTTGGCACCGC CGAAGTTCATCTGCCTGCTCCTAGCGCTGTC AGAGCCGCCAGACCAGGATTTGGCGGAGAT AAGGGCAGACTCGTGGGCGTTGCAGAAAGA CAGGTGCCCTGCAGATTCCTGAGGCCAACA GGCGTGCTGAGCCTGTTTCTTGGCGGAGCT GTGGTGTCCCTGCAGTACGTCAGACCTAGCG CTCTGAGAACCCTGCTGGATCAGAGCGCCAA GGATTGATCAACTAGCCAGATTCTTCATGTTT GGACCAAATCAACTTGTGATACCATGCTCAA AGAGGCCTCAATTATATTTGAGTTTTTAATTT TATGAAAAAACTAACAGCAATCATGGAAGT CCACGATTTTGAGACCGACGAGTTCAATGAT TTCAATGAAGATGACTATGCCACAAGAGAAT TCCTGAATCCCGATGAGCGCATGACGTA GAATCATGCTGATTACAACCTGAATTCTCCTC TAATTAGTGATGATATTGACAATTTAATCAGG AAATTCAATTCTCTTCCAATTCCCTCGATGTG GGATAGTAAGAACTGGGATGGAGTTCTTGA GATGTTAACGTCATGTCAAGCCAATCCCATCC CAACATCTCAGATGCATAAATGGATGGGAAG TTGGTTAATGTCTGATAATCATGATGCCAGTC AAGGGTATAGTTTTTTACATGAAGTGGACAA AGAGGCAGAAATAACATTTGACGTGGTGGA GACCTTCATCCGCGGCTGGGGCAACAAACC	
--	--

AATTGAATACATCAAAAAGGAAAGATGGACT GACTCATTCAAAATTCTCGCTTATTTGTGTCAA AAGTTTTGGACTTACACAAGTTGACATTAAT CTTAAATGCTGTCTCTGAGGTGGAATTGCTCA ACTTGGCGAGGACTTTCAAAGGCAAAGTCAG AAGAAGTTCTCATGGAACGAACATATGCAGG ATTAGGGTTCAGCTTGGGTCCTACTTTTAT TTCAGAAGGATGGGCTTACTTCAAGAACTT GATATTCTAATGGACCGAACTTTCTGTTAAT GGTCAAAGATGTGATTATAGGGAGGATGCA AACGGTGCTATCCATGGTATGTAGAATAGAC AACCTGTTCTCAGAGCAAGACATCTTCTCCCT TCTAAATATCTACAGAATTGGAGATAAAATTG TGGAGAGGCAGGGAAATTTTTCTTATGACTT GATTAATGTTGGTGAACCGATATGCAACTTG AAGCTGATGAAATTAGCAAGAGAATCAAGGC CTTTAGTCCCACAATTCCCTCATTTTGAAAATC ATATCAAGACTTCTGTTGATGAAGGGGCAAA AATTGACCGAGGTATAAGATTCCTCCATGATC AGATAATGAGTGTGAAAACAGTGGATCTCAC ACTGGTGATTTATGGATCGTTCAGACATTGG GGTCATCCTTTTATAGATTATTACACTGGACT AGAAAAATTACATTCCCAAGTAACCATGAAG AAAGATATTGATGTGTCATATGCAAAAGCACT TGCAAGTGATTTAGCTCGGATTGTTCTATTTT AACAGTTCAATGATCATAAAAAGTGGTTCGT GAATGGAGACTTGCTCCCTCATGATCATCCCT TAAAAGTCATGTTAAAGAAAATACATGGCCC ACAGCTGCTCAAGTTCAAGATTTTGGAGATA AATGGCATGAACTCCGCTGATTAAATGTTTT GAAATACCCGACTTACTAGACCCATCGATAAT ATACTCTGACAAAAGTCATTCAATGAATAGGT CAGAGGTGTTGAAACATGTCCGAATGAATCC GAACACTCCTATCCCTAGTAAAAGGTGTTG CAGACTATGTTGGACACAAAGGCTACCAATT GGAAAGAATTTCTTAAAGAGATTGATGAGAA GGGCTTAGATGATGATGATCTAATTATTGGTC TTAAAGGAAAGGAGAGGGAAGTGAAGTTGG CAGGTAGATTTTCTCCCTAATGTCTTGAAA TTGCGAGAATACTTTGTAATTACCGAATATT GATAAAGACTCATTTCTGTCCTATGTTTAAAG GCCTGACAATGGCGGACGATCTAACTGCAGT CATTAAAAAGATGTTAGATTCCTCATCCGGCC AAGGATTGAAGTCATATGAGGCAATTTGCAT AGCCAATCACATTGATTACGAAAAATGGAAT AACCACCAAAGGAAGTTATCAAACGGCCCAG TGTTCCGAGTTATGGGCCAGTTCTTAGGTTAT

CCATCCTTAATCGAGAGAACTCATGAATTTT TGAGAAAAGTCTTATATACTACAATGGAAGA CCAGACTTGATGCGTGTTCAACAACACAC TGATCAATTCAACCTCCCAACGAGTTTGTTGG CAAGGACAAGAGGGTGGACTGGAAGGTCTA CGGCAAAAAGGATGGAGTATCCTCAATCTAC TGGTTATTCAAAGAGAGGGCTAAAATCAGAAA CACTGCTGTCAAAGTCTTGGCACAAGGTGAT AATCAAGTTATTTGCACACAGTATAAAACGAA GAAATCGAGAAACGTTGTAGAATTACAGGGT GCTCTCAATCAAATGGTTTCTAATAATGAGAA AATTATGACTGCAATCAAATAGGGACAGGG AAGTTAGGACTTTTGATAAATGACGATGAGA CTATGCAATCTGCAGATTACTTGAATTATGGA AAAATACCGATTTTCCGTGGAGTGATTAGAG GGTTAGAGACCAAGAGATGGTCACGAGTGA CTTGTGTCACCAATGACCAAATACCCACTTGT GCTAATATAATGAGCTCAGTTTCCACAAATGC TCTACCGTAGCTCATTTTGCTGAGAACCCAA TCAATGCCATGATACAGTACAATTATTTGGG ACATTTGCTAGACTCTTGTTGATGATGCATGA TCCTGCTCTTCGTCAATCATTGTATGAAGTTC AAGATAAGATACCGGGCTTGCACAGTTCTAC TTTCAAATACGCCATGTTGTATTTGGACCCTT CCATTGGAGGAGTGTGCGGGCATGTCTTTGTC CAGGTTTTTGATTAGAGCCTTCCCAGATCCCG TAACAGAAAAGTCTCTCATTCTGGAGATTCATC CATGTACATGCTCGAAGTGAGCATCTGAAGG AGATGAGTGCAGTATTTGGAAACCCCGAGAT AGCCAAGTTTCGAATAACTCACATAGACAAG CTAGTAGAAGATCCAACCTCTCTGAACATCG CTATGGGAATGAGTCCAGCGAACTTGTTAAA GACTGAGGTTAAAAAATGCTTAATCGAATCA AGACAAACCATCAGGAACCAGGTGATTAAG GATGCAACCATATATTTGTATCATGAAGAGG ATCGGCTCAGAAGTTTCTTATGGTCAATAAAT CCTCTGTTCCCTAGATTTTTAAGTGAATTCAAA TCAGGCACTTTTTTGGGAGTCCGACACGGGC TCATCAGTCTATTTCAAATCTCGTACTATTC GGAACCTCCTTAAGAAAAAGTATCATAGGGA ATTGGATGATTTGATTGTGAGGAGTGAGGTA TCCTCTTTGACACATTTAGGGAAACTTCATTT GAGAAGGGGATCATGTAAAATGTGGACATG TTCAGCTACTCATGCTGACACATTAAGATACA AATCCTGGGGCCGTACAGTTATTGGGACAAC TGTACCCCATCCATTAGAAATGTTGGGTCCAC AACATCGAAAAGAGACTCCTTGTGCACCATG
--

TAACACATCAGGGTTCAATTATGTTTCTGTGC ATTGTCCAGACGGGATCCATGACGTCTTTAG TTCACGGGGACCATTGCCTGCTTATCTAGGG TCTAAAACATCTGAATCTACATCTATTTTGCAG CCTTGGGAAAGGGAAAGCAAAGTCCCACTG ATTAAGAGCTACACGTCTTAGAGATGCTAT CTCTTGGTTTGTGAACCCGACTCTAACTAG CAATGACTATACTTTCTAACATCCACTCTTTAA CAGGCGAAGAATGGACCAAAAGGCAGCATG GGTTCAAAAGAACAGGGTCTGCCCTTCATAG GTTTTCGACATCTCGGATGAGCCATGGTGGG TTCGCATCTCAGAGCACTGCAGCATTGACCA GGTTGATGGCAACTACAGACACCATGAGGG ATCTGGGAGATCAGAATTCGACTTTTTATT CAAGCAACGTTGCTCTATGCTCAAATTACCAC CACTGTTGCAAGAGACGGATGGATCACCAGT TGTACAGATCATTATCATATTGCCTGTAAGTC CTGTTTGAGACCCATAGAAGAGATCACCTG GACTCAAGTATGGACTACACGCCCCCAGATG TATCCCATGTGCTGAAGACATGGAGGAATGG GGAAGGTTCTGTTGGGACAAGAGATAAAACA GATCTATCCTTTAGAAGGGAATTGGAAGAAT TTAGCACCTGCTGAGCAATCCTATCAAGTCG GCAGATGTATAGGTTTTCTATATGGAGACTTG GCGTATAGAAAATCTACTCATGCCGAGGACA GTTCTCTATTTCTCTATCTATAAAGGTCGTA TTAGAGGTCGAGGTTTCTTAAAAGGGTTGCT AGACGGATTAATGAGAGCAAGTTGCTGCCAA GTAATACACCGGAGAAGTCTGGCTCATTTGA AGAGGCCGGCCAACGCAGTGTACGGAGGTT TGATTTACTTGATTGATAAATTGAGTGTATCA CCTCCATTCCCTTCTCTTACTAGATCAGGACCT ATTAGAGACGAATTAGAAACGATTCCCCACA AGATCCCAACCTCCTATCCGACAAGCAACCG TGATATGGGGGTGATTGTCAGAAATTACTTC AAATACCAATGCCGTCTAATTGAAAAGGGAA AATACAGATCACATTATTCACAATTATGGTTAT TCTCAGATGTCTTATCCATAGACTTCATTGGA CCATTCTCTATTTCCACCACCCTCTTGCAAATC CTATAAAGCCATTTTTATCTGGGAAAGATAA GAATGAGTTGAGAGAGCTGGCAAATCTTTCT TCATTGCTAAGATCAGGAGAGGGGTGGGAA GACATACATGTGAAATTCTTCACCAAGGACAT ATTATTGTGTCCAGAGGAAATCAGACATGCTT GCAAGTTCTGGGATTGCTAAGGATAATAATAA AGACATGAGCTATCCCCCTTGGGGAAGGGA ATCCAGAGGGACAATTACAACAATCCCTGTTT

ATTATACGACCACCCCTTACCCAAAGATGCTA GAGATGCCTCCAAGAATCCAAAATCCCCTGC TGTCCGGAATCAGGTTGGGCCAATTACCAAC TGGCGCTCATTATAAAATTTCGGAGTATATTAC ATGGAATGGGAATCCATTACAGGGACTTCTT GAGTTGTGGAGACGGCTCCGGAGGGATGAC TGCTGCATTACTACGAGAAAATGTGCATAGC AGAGGAATATTCAATAGTCTGTTAGAATTATC AGGGTCAGTCATGCGAGGCGCCTCTCCTGA GCCCCCAGTGCCCTAGAACTTTAGGAGGA GATAAATCGAGATGTGTAAATGGTGAAACAT GTTGGGAATATCCATCTGACTTATGTGACCCA AGGACTTGGGACTATTTCTCCGACTCAAAG CAGGCTTGGGGCTTCAAATTGATTTAATTGTA ATGGATATGGAAGTTCGGGATTCTTCTACTA GCCTGAAAATTGAGACGAATGTTAGAAATTA TGTGCACCGGATTTTGGATGAGCAAGGAGTT TTAATCTACAAGACTTATGGAACATATATTTGT GAGAGCGAAAAGAATGCAGTAACAATCCTTG GTCCCATGTTCAAGACGGTCGACTTAGTTCA AACAGAATTTAGTAGTTCTCAAACGTCTGAAG TATATATGGTATGTAAAGGTTTGAAGAAATTA ATCGATGAACCCAATCCCGATTGGTCTTCCAT CAATGAATCCTGGAAAAACCTGTACGCATT CAGTCATCAGAACAGGAATTTGCCAGAGCAA AGAAGGTTAGTACATACTTTACCTTGACAGGT ATCCCTCCCAATTCATTCTGATCCTTTTGTA AACATTGAGACTATGCTACAAATATTCGGAGT ACCCACGGGTGTGTCTCATGCGGCTGCCTTA AAATCATCTGATAGACCTGCAGATTTATTGAC CATTAGCCTTTTTTATATGGCGATTATATCGTA TTATAACATCAATCATATCAGAGTAGGACCGA TACCTCCGAACCCCCCATCAGATGGAATTGC ACAAAATGTGGGGATCGCTATAACTGGTATA AGCTTTTGGCTGAGTTTGATGGAGAAAGACA TTCCACTATATCAACAGTGTTTAGCAGTTATC CAGCAATCATTCCCGATTAGGTGGGAGGCTG TTTCAGTAAAAGGAGGATACAAGCAGAAGTG GAGTACTAGAGGTGATGGGCTCCCAAAAGA TACCCGAATTTACAGACTCCTTGGCCCCAATCG GGAATGATCAGATCTCTGGAATTGGTCCG AAACCAAGTTCGTCTAAATCCATTCAATGAGA TCTTGTTCAATCAGCTATGTCGTACAGTGGAT AATCATTGAAATGGTCAAATTTGCGAAGAAA CACAGGAATGATTGAATGGATCAATAGACGA ATTTCAAAGAAGACCGGTCTATACTGATGTT GAAGAGTGACCTACACGAGGAAAACCTTG	
--	--

	GAGAGATTAAAAAATCATGAGGAGACTCCAA ACTTTAAGTATGAAAAAACTTTGATCCTTAA GACCCTCTTGTGGTTTTATTTTTATCTGGTT TTGTGGTCTTCGT	
SEQ ID NO: 50	NRTLTLFNVTRNDARAYVSGIQNSVSANRSDPV TLDVLPDSSYLSGANLNLSCHSASPQYSWRINGI PQHTQVLFIAKITPNNNGTYACFVSNLATGRN NSIVKSITVSASGTSPGLSATEYKLVVVGADGVG KSALTIQLIQ	multi-antigenic domain ATP132
SEQ ID NO: 51	MKRYKNRVASRKSRKFKQLQHYREVAAAKSS ENDRLRLLLKNRTLTLFNVTRNDARAYVSGIQNS VSANRSDPVTLDVLPDSSYLSGANLNLSCHSASP QYSWRINGIPQHTQVLFIAKITPNNNGTYACF VSNLATGRNNSIVKSITVSASGTSPGLSATEYKLV VVGADGVGKSALTIQLIQSTVHEILSKLSLEGDHS TPPSAYGSVKPYTNFDAE	ATP132

CLAIMS

1. A composition comprising
 - (i) at least one tumor antigen, or a fragment or sequence variant thereof, wherein the tumor antigen is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2);
 - (ii) a nucleic acid encoding said tumor antigen;
 - (iii) an antigen-presenting cell (APC) containing (i) or (ii); or
 - (iv) a T-cell expressing either a T-cell receptor or a CAR T-cell receptor targeting said tumor antigen.
2. The composition according to claim 1, wherein the composition further comprises at least one tumor antigen selected from the group consisting of CEACAM5, DUOXA2, and KRAS, or a fragment or sequence variant thereof.
3. The composition according to claim 2, wherein KRAS, or the fragment thereof, is KRAS-G12D, or a fragment thereof, or KRAS-G12V, or a fragment thereof.
4. The composition according to any one of the previous claims, wherein the tumor antigen selected from CEACAM5, DUOXA2, KRAS, KRAS-uORF1, TPX2-uORF1 and AURKA-uORF2, or the fragment or sequence variant thereof, comprises a CD4+ and/or a CD8+ epitope.
5. The composition according to any one of the previous claims, wherein the tumor antigen comprises at least one of amino acid, wherein the amino acid sequence is
 - an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity;

- an amino acid sequence according to SEQ ID NO: 1 or 2, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3 or 4, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity; or
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity.
6. A peptide comprising:
- a) a cell penetrating peptide;
 - b) a multi-antigenic domain comprising at least one tumor antigen, or a fragment or sequence variant thereof,
wherein the tumor antigen is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1), of the TPX2 mRNA (TPX2-uORF1) or of the AURKA mRNA (AURKA-uORF2); and
 - c) a TLR peptide agonist.
7. The peptide according to claim 6, wherein the multi-antigenic domain further comprises at least one tumor antigen selected from the group consisting of CEACAM5 or a fragment thereof, DUOXA2 or a fragment thereof, and KRAS or a fragment thereof, wherein each fragment has a minimum length of 8 amino acids.
8. The peptide according to claim 6 or 7, wherein the multi-antigenic domain comprises three or more different antigens, or fragments thereof having a minimum length of 8 amino acids, in particular 3, 4, 5, 6, 7, 8, 9, 10 or more different antigens or fragments thereof.
9. The peptide according to any one of claims 6 to 8, wherein the multi-antigenic domain comprises
- a tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1),

- a tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-uORF1) and
 - a tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2).
10. The peptide according to any one of claims 6 to 9, wherein the multi-antigenic domain comprises, preferably in N- to C-terminal direction:
- CEACAM5 or a fragment thereof having a minimum length of 8 amino acids,
 - KRAS or a fragment thereof having a minimum length of 8 amino acids,
 - KRAS-uORF1, or a fragment thereof having a minimum length of 8 amino acids,
 - TPX2-uORF1, or a fragment thereof having a minimum length of 8 amino acids,
 - AURKA-uORF2, or a fragment thereof having a minimum length of 8 amino acids, and
 - DUOXA2 or a fragment thereof having a minimum length of 8 amino acids.
11. The peptide according to any one of claims 6 to 10, wherein KRAS, or the fragment thereof, is KRAS-G12D, or a fragment thereof, or KRAS-G12V, or a fragment thereof.
12. The peptide according to any one of claims 6 to 11, wherein the fragment of CEACAM5 comprises an amino acid sequence according to any one of SEQ ID NOs 6 – 8, or a sequence variant thereof having at least 70% sequence identity, preferably wherein the multi-antigenic domain comprises an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity.
13. The peptide according to any one of claims 6 to 12, wherein the multi-antigenic domain comprises at least one amino acid sequence selected from the group consisting of:
- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity;

- an amino acid sequence according to SEQ ID NO: 1 or 2, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3 or 4, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity; or
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity.
14. The peptide according to any one of claims 6 to 13, wherein the multi-antigenic domain comprises, preferably in N- to C-terminal direction:
- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, and/or an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 1 or 2, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3 or 4, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity; and
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity.
15. The peptide according to any one of claims 6 to 14, wherein the multi-antigenic domain comprises an amino acid sequence according to SEQ ID NO: 13 or 14, or a sequence variant thereof having at least 70% sequence identity.
16. The peptide according to any one of claims 6 to 15, wherein the cell penetrating peptide has an amino acid sequence comprising or consisting of an amino acid sequence

according to SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 22, or SEQ ID NO: 23, or sequence variants thereof having at least 70% sequence identity.

17. The peptide according to any one of claims 6 to 16, wherein the cell penetrating peptide has an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 15 or sequence variants thereof having at least 70% sequence identity.
18. The peptide according to any one of claims 6 to 17, wherein the TLR peptide agonist is a TLR2 peptide agonist and/or a TLR4 peptide agonist.
19. The peptide according to any one of claims 6 to 18, wherein the TLR peptide agonist comprises or consists of an amino acid sequence according to SEQ ID NO: 16 or 24; or a sequence variant thereof having at least 70% sequence identity.
20. The peptide according to any one of claims 6 to 19, wherein
 - a) the cell penetrating peptide has an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 22 or SEQ ID NO: 23, or sequence variants thereof having at least 70% sequence identity;
 - b) the multi-antigenic domain comprises
 - a fragment of CEACAM5 having a minimum length of 8 amino acids,
 - a fragment of DUOXA2 having a minimum length of 8 amino acids,
 - a fragment of KRAS having a minimum length of 8 amino acids,
 - KRAS-uORF1, or a fragment thereof having a minimum length of 8 amino acids,
 - TPX2-uORF1, or a fragment thereof having a minimum length of 8 amino acids, and/or
 - AURKA-uORF2, or a fragment thereof having a minimum length of 8 amino acids;and
 - c) the TLR peptide agonist is a TLR2 peptide agonist and/or a TLR4 peptide agonist.

21. The peptide according to any one of claims 6 to 20, wherein the peptide comprises, preferably in N-terminal to C-terminal direction:
- a) a cell penetrating peptide having an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 15, or sequence variants thereof having at least 70% sequence identity;
 - b) a multi-antigenic domain comprising
 - an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, and/or an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity; and
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity; and
 - c) a TLR peptide agonist having an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 16, or sequence variants thereof having at least 70% sequence identity.
22. The peptide according to any one of claims 6 to 21, wherein the peptide comprises in N-terminal to C-terminal direction:
- a) a cell penetrating peptide having an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 15, or sequence variants thereof having at least 90% sequence identity;
 - b) a multi-antigenic domain comprising

- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 90% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 90% sequence identity, and/or an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 90%% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 90% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 90% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 90% sequence identity; and
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 90% sequence identity; and
- c) a TLR peptide agonist having an amino acid sequence comprising or consisting of an amino acid sequence according to SEQ ID NO: 16, or sequence variants thereof having at least 90% sequence identity.
23. The peptide according to any one of claims 6 to 22, wherein the peptide comprises or consists of an amino acid sequence according to SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 46, or SEQ ID NO: 47, preferably of an amino acid sequence according to SEQ ID NO: 17 or 18, or a sequence variant thereof having at least 70% sequence identity.
24. A recombinant vesicular stomatitis virus (VSV) encoding a multi-antigenic domain comprising at least one tumor antigen, or fragment or sequence variant thereof, wherein the tumor antigen is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF), of the TPX2 mRNA (TPX2-uORF) or of the AURKA mRNA (AURKA-uORF).
25. The vesicular stomatitis virus (VSV) according to claim 24, wherein the multi-antigenic domain further comprises at least one tumor antigen selected from the group consisting

of CEACAM5 or a fragment thereof, DUOXA2 or a fragment thereof, and KRAS or a fragment thereof, wherein each fragment has a minimum length of 8 amino acids.

26. The vesicular stomatitis virus (VSV) according to claim 24 or 25, wherein the multi-antigenic domain comprises three or more different antigens, or fragments thereof having a minimum length of 8 amino acids, in particular 3, 4, 5, 6, 7, 8, 9, 10 or more different antigens or fragments thereof.
27. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 26, wherein the multi-antigenic domain comprises
 - a tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the KRAS mRNA (KRAS-uORF1),
 - a tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the TPX2 mRNA (TPX2-uORF1) and
 - a tumor antigen, which is encoded in a 5'-upstream open reading frame (uORF) within the 5' UTR of the AURKA mRNA (AURKA-uORF2).
28. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 27, wherein the multi-antigenic domain comprises, preferably in N- to C-terminal direction:
 - CEACAM5 or a fragment thereof having a minimum length of 8 amino acids,
 - KRAS or a fragment thereof having a minimum length of 8 amino acids,
 - KRAS-uORF1, or a fragment thereof having a minimum length of 8 amino acids,
 - TPX2-uORF1, or a fragment thereof having a minimum length of 8 amino acids,
 - AURKA-uORF2, or a fragment thereof having a minimum length of 8 amino acids,
 - and
 - DUOXA2 or a fragment thereof having a minimum length of 8 amino acids.
29. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 28, wherein KRAS, or the fragment thereof, is KRAS-G12D, or a fragment thereof, or KRAS-G12V, or a fragment thereof.

30. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 29, wherein the fragment of CEACAM5 comprises an amino acid sequence according to any one of SEQ ID NOs 6 – 8, or a sequence variant thereof having at least 70% sequence identity, preferably wherein the multi-antigenic domain comprises an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity.
31. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 30, wherein the multi-antigenic domain comprises at least one of amino acid sequence selected from the group consisting of:
- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 12, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 1 or 2, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3 or 4, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity; or
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity.
32. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 31, wherein the multi-antigenic domain comprises, preferably in N- to C-terminal direction:
- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 70% sequence identity, and/or an amino acid sequence according

- to SEQ ID NO: 12, and/or a sequence variant thereof having at least 70% sequence identity;
- an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 70% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 70% sequence identity; and
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 70% sequence identity.
33. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 32, wherein the multi-antigenic domain comprises in N- to C-terminal direction:
- an amino acid sequence according to SEQ ID NO: 9, or a sequence variant thereof having at least 90% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 11, or a sequence variant thereof having at least 90% sequence identity, and/or an amino acid sequence according to SEQ ID NO: 12, and/or a sequence variant thereof having at least 90% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 1, or a sequence variant thereof having at least 90% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 3, or a sequence variant thereof having at least 90% sequence identity;
 - an amino acid sequence according to SEQ ID NO: 5, or a sequence variant thereof having at least 90% sequence identity; and
 - an amino acid sequence according to SEQ ID NO: 10, or a sequence variant thereof having at least 90% sequence identity.
34. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 33, wherein the multi-antigenic domain comprises an amino acid sequence according to SEQ ID NO: 19 or SEQ ID NO: 48, preferably an amino acid sequence according to SEQ ID NO: 19, or a sequence variant thereof having at least 70% sequence identity.

35. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 34, wherein the vesicular stomatitis virus (VSV) is an oncolytic vesicular stomatitis virus (VSV).
36. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 35, wherein the vesicular stomatitis virus (VSV) is replication-competent.
37. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 36, wherein the gene coding for the glycoprotein G is replaced by the gene coding for the glycoprotein GP of Lymphocyte choriomeningitis virus (LCMV), wherein the glycoprotein GP of LCMV preferably comprises the amino acid sequence according to SEQ ID NO: 25, or a functional sequence variant thereof which is at least 80%, 85%, 90%, 95% identical thereto.
38. The vesicular stomatitis virus (VSV) according to any one of any one of claims 24 to 37, wherein
- it encodes in its genome the multi-antigenic domain as defined in any one of claims 25 to 34,
 - it encodes in its genome a vesicular stomatitis virus nucleoprotein (N), large protein (L), phosphoprotein (P), and matrix protein (M),
 - the gene coding for the glycoprotein G of the vesicular stomatitis virus is replaced by the gene coding for the glycoprotein GP of lymphocyte choriomeningitis virus (LCMV), and/or
 - the glycoprotein G of the vesicular stomatitis virus is replaced by the glycoprotein GP of LCMV.
39. The vesicular stomatitis virus (VSV) according to any one of any one of claims 24 to 38, wherein
- it encodes in its genome a vesicular stomatitis virus nucleoprotein (N) comprising an amino acid as set forth in SEQ ID NO: 26, or a sequence variant thereof having at least 80%, 85%, 90%, 92%, 94%, 96%, 98% sequence identity;

- it encodes in its genome a vesicular stomatitis virus phosphoprotein (P) comprising an amino acid as set forth in SEQ ID NO: 27, or a sequence variant thereof having at least 80%, 85%, 90%, 92%, 94%, 96%, 98% sequence identity;
 - it encodes in its genome a vesicular stomatitis virus large protein (L) comprising an amino acid as set forth in SEQ ID NO: 28, or a sequence variant thereof having at least 80%, 85%, 90%, 92%, 94%, 96%, 98% sequence identity;
 - it encodes in its genome a vesicular stomatitis virus matrix protein (M) comprising an amino acid as set forth in SEQ ID NO: 29, or a sequence variant thereof having at least 80%, 85%, 90%, 92%, 94%, 96%, 98% sequence identity;
 - it encodes in its genome a vesicular stomatitis virus the multi-antigenic domain as defined in any one of claims 25 to 34,
 - the gene coding for the glycoprotein G of the vesicular stomatitis virus is replaced by the gene coding for the glycoprotein GP of lymphocyte choriomeningitis virus (LCMV), and/or
 - the glycoprotein G is replaced by the glycoprotein GP of LCMV.
40. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 39, and wherein it encodes in its genome
- a phosphoprotein (P) comprising the amino acid sequence according to SEQ ID NO: 27, or a sequence variant thereof having at least 70% sequence identity,
 - a nucleoprotein (N) comprising the amino acid sequence according to SEQ ID NO: 26, or a sequence variant thereof having at least 70% sequence identity,
 - a matrix protein (M) comprising the amino acid sequence according to SEQ ID NO: 29, or a sequence variant thereof having at least 70% sequence identity,
 - a large protein (L) comprising the amino acid sequence according to SEQ ID NO: 28, or a sequence variant thereof having at least 70% sequence identity,
 - a glycoprotein (GP) comprising the amino acid sequence according to SEQ ID NO: 25, or a sequence variant thereof having at least 70% sequence identity, and
 - a multi-antigenic domain which comprises the amino acid sequence according to SEQ ID NO: 19 or SEQ ID NO: 48 or preferably the amino acid sequence according to SEQ ID NO: 19, or a sequence variant thereof having at least 70% sequence identity.

41. The vesicular stomatitis virus (VSV) according to any one of claims 24 to 40, wherein its RNA genome comprises or consists of an RNA sequence according to SEQ ID NO: 30, or a sequence variant thereof having at least 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity.
42. A vaccine comprising
 - (i) the composition according to any one of claims 1 to 5;
 - (ii) the peptide according to any one of claims 6 to 23; or
 - (iii) the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41.
43. The vaccine according to claim 42, wherein the vaccine comprises
 - (i) the peptide according to any one of claims 6 to 23; and
 - (ii) the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41.
44. A kit comprising
 - (i) the peptide according to any one of claims 6 to 23; and
 - (ii) the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41.
45. A combination comprising
 - (i) the peptide according to any one of claims 6 to 23; and
 - (ii) the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41.
46. The vaccine according to claim 43, the kit according to claim 44, or the combination according to claim 45, wherein the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) comprises an antigen, or a fragment or sequence variant thereof, which is contained in the multi-antigenic domain of the peptide.
47. The vaccine, the kit, or the combination according to claim 46, wherein the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV)

comprises the amino acid sequences of each of the antigens, or fragments or sequence variant thereof, which are contained in the multi-antigenic domain of the peptide.

48. The vaccine according to claim 43, the kit according to claim 44, or the combination according to claim 45, wherein the multi-antigenic domain of the peptide comprises an antigen, or a fragment or sequence variant thereof, which is contained in the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV).
49. The vaccine, the kit, or the combination according to claim 48, wherein the multi-antigenic domain of the peptide comprises the amino acid sequences of each of the antigens, or fragments or sequence variant thereof, of the multi-antigenic domain which is encoded in the genome of the vesicular stomatitis virus (VSV).
50. The vaccine, the kit, or the combination according to any one of claims 43 to 49, wherein
 - the multi-antigenic domain of the peptide comprises an amino acid sequence according to SEQ ID NO: 13 or 14, or a sequence variant thereof having at least 70% sequence identity, and
 - the multi-antigenic domain encoded in the genome of the vesicular stomatitis virus (VSV) comprises the amino acid sequence according to SEQ ID NO: 19 or SEQ ID NO: 48, preferably the amino acid sequence according to SEQ ID NO: 19, or a sequence variant thereof having at least 70% sequence identity.
51. The vaccine, the kit, or the combination according to any one of claims 43 to 50, wherein
 - the peptide comprises an amino acid sequence according to SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 46, or SEQ ID NO: 47, preferably an amino acid sequence according to SEQ ID NO: 17 or 18, or a sequence variant thereof having at least 70% sequence identity, and
 - the vesicular stomatitis virus (VSV) comprises an RNA genome comprising an RNA sequence according to SEQ ID NO: 30, or a sequence variant thereof having at least 70% sequence identity.

52. The vaccine, the kit, or the combination according to any one of claims 43 to 51, further comprising
- (iii) an inhibitor of the PD-1/PD-L1 pathway.
53. The vaccine, the kit, or the combination according to claim 52, wherein the inhibitor of the PD-1/PD-L1 pathway is selected from the group consisting of pembrolizumab; nivolumab; pidilizumab; cemiplimab; PDR-001; atezolizumab; avelumab; durvalumab, ezabenlimab, an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 31 and a light chain comprising the amino acid sequence of SEQ ID NO: 32; an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 33 and a light chain comprising the amino acid sequence of SEQ ID NO: 34; and an antibody comprising a heavy chain comprising the amino acid sequence of SEQ ID NO: 35 and a light chain comprising the amino acid sequence of SEQ ID NO: 36.
54. The composition according to any one of claim 1 to 5, the peptide according to any one of claims 6 to 23, the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41, the vaccine according to any one of claims 42, 43 and 46 to 53, the kit according to any one of claims 44 and 46 to 53, or the combination according to any one of claims 45 to 53 for use in medicine.
55. The composition according to any one of claim 1 to 5, the peptide according to any one of claims 6 to 23, the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41, the vaccine according to any one of claims 42, 43 and 46 to 53, the kit according to any one of claims 44 and 46 to 53, or the combination according to any one of claims 45 to 53 for use in the treatment of cancer.
56. The composition, the peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to claim 55, wherein the cancer is a cancer of the gastrointestinal tract (GI).

57. The composition, the peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to claim 55 or 56, wherein the cancer is selected from the group consisting of anal cancer; appendix cancer; cholangiocarcinoma/bile duct cancer, in particular extrahepatic bile duct cancer; gastrointestinal carcinoid tumor; colorectal cancer, in particular colon cancer, rectal cancer and metastatic colorectal cancer; esophageal cancer; gallbladder cancer; gastric (stomach) cancer; gastrointestinal stromal tumor (GIST), and pancreatic cancer, such as pancreatic ductal adenocarcinoma.
58. The composition, the peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 55 to 57, wherein the cancer is selected from the group consisting of colon cancer, rectal cancer, colorectal cancer, metastatic colorectal cancer, pancreatic cancer, and pancreatic ductal adenocarcinoma.
59. The peptide for use according to any one of claims 54 to 58, wherein the peptide is administered in combination with the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41.
60. The vesicular stomatitis virus (VSV) for use according to any one of claims 54 to 58, wherein the vesicular stomatitis virus (VSV) is administered in combination with the peptide according to any one of claims 6 to 23.
61. The peptide for use according to claim 59 or the vesicular stomatitis virus (VSV) for use according to claim 61, wherein the peptide and the vesicular stomatitis virus (VSV) are as defined in any one of claims 46 to 51.
62. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 54 to 61, wherein the peptide and the vesicular stomatitis virus (VSV) are each administered at least once, preferably the peptide is administered prior to the administration of the vesicular stomatitis virus (VSV).

63. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 54 to 62, wherein the peptide is administered at least twice, preferably prior to and subsequent to the administration of the vesicular stomatitis virus (VSV).
64. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 54 to 63, wherein the peptide and the vesicular stomatitis virus (VSV) are administered in the order K-V-K, K-V-K-K, K-V-K-K-K, or K-V-K-K-K-K.
65. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 54 to 64, wherein the treatment schedule comprises a single administration of the vesicular stomatitis virus (VSV).
66. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 54 to 65, wherein the peptide and the vesicular stomatitis virus (VSV) are administered via same or different route, preferably intravenously, subcutaneously, or intramuscularly.
67. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 54 to 66, wherein the peptide is administered subcutaneously, and the vesicular stomatitis virus (VSV) are administered intravenously, or intratumorally, preferably, intravenously.
68. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to any one of claims 54 to 67, further comprising administration of an inhibitor of the PD-1/PD-L1 pathway, preferably as defined in claim 53.
69. The peptide, the vesicular stomatitis virus (VSV), the vaccine, the kit or the combination for use according to claim 68, wherein the inhibitor of the PD-1/PD-L1 pathway is administered concomitantly, sequentially or alternately with the peptide or the vesicular stomatitis virus (VSV).

70. Use of the composition according to any one of claim 1 to 5, the peptide according to any one of claims 6 to 23, the vesicular stomatitis virus (VSV) according to any one of claims 24 to 41, the vaccine according to any one of claims 42, 43 and 46 to 53, the kit according to any one of claims 44 and 46 to 53, or the combination according to any one of claims 45 to 53 for the manufacture of a medicament for the treatment of cancer.
71. A method for ameliorating, treating, or reducing (the risk of occurrence) of a cancer, or for inducing or enhancing an anti-tumor response, the method comprising administration of (an effective amount of) composition according to any one of claim 1 to 5, the peptide according to any one of claims 6 to 23, the vesicular stomatitis virus (VSV) according to any one of claims 25 to 41, the vaccine according to any one of claims 42, 43 and 46 to 53, the kit according to any one of claims 44 and 46 to 53, or the combination according to any one of claims 45 to 53 to a subject in need thereof.

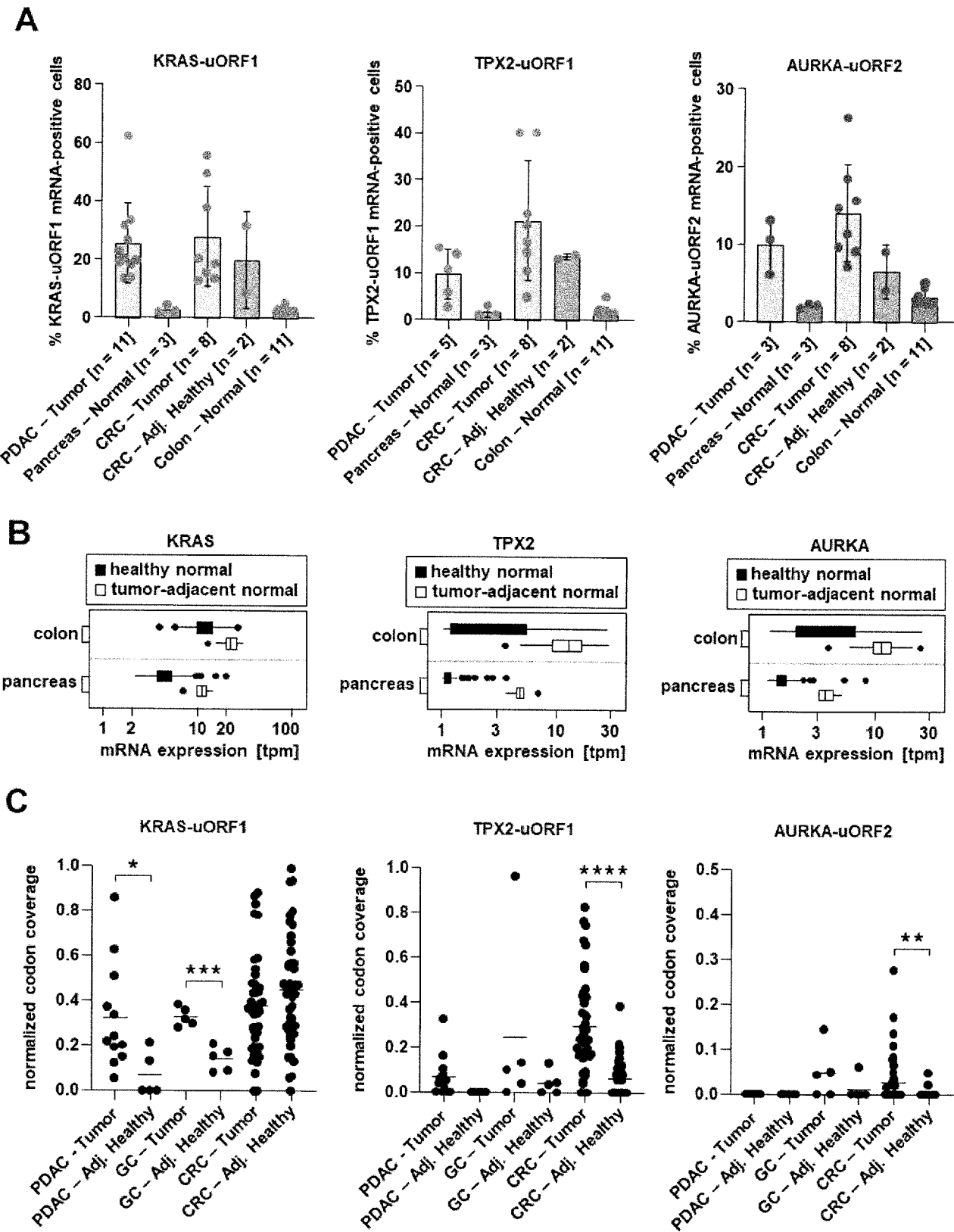


Figure 1

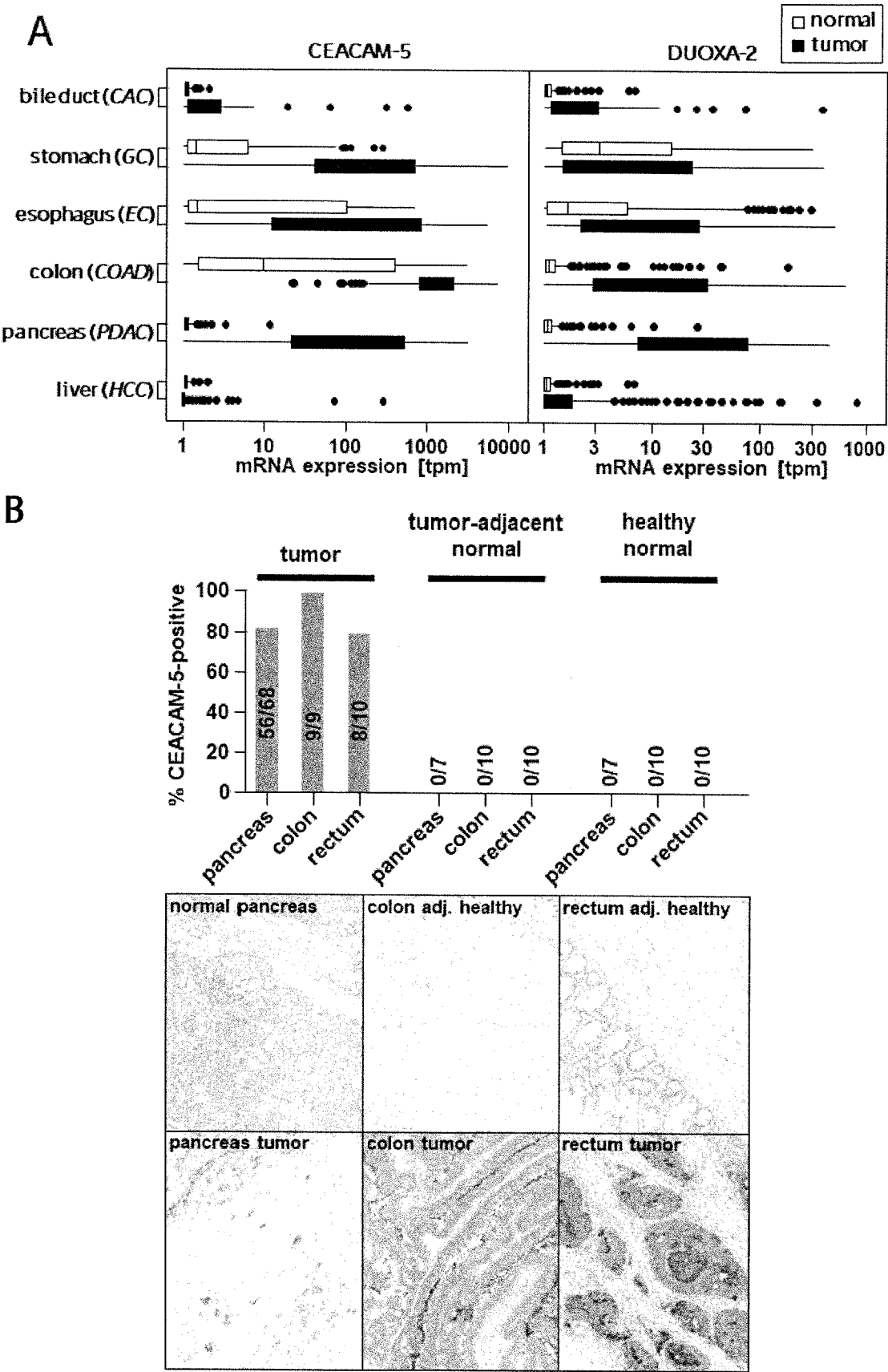


Figure 2

C

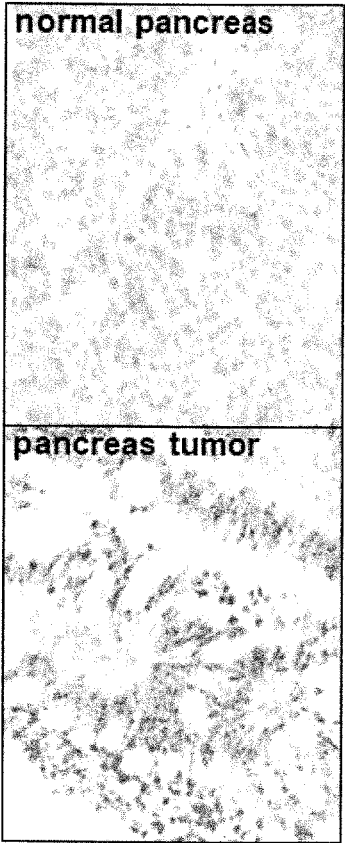
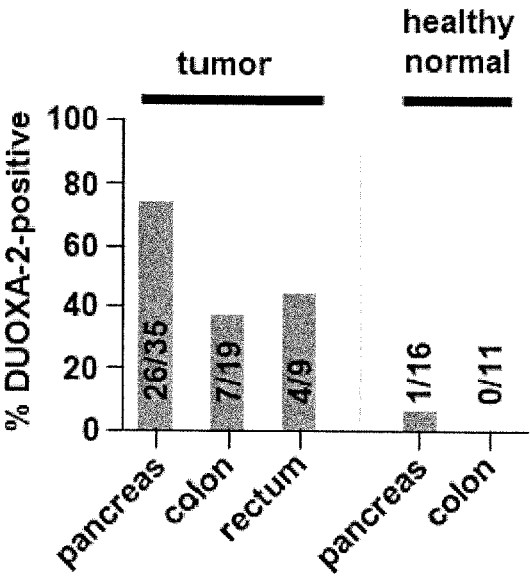


Figure 2 continued

4/10

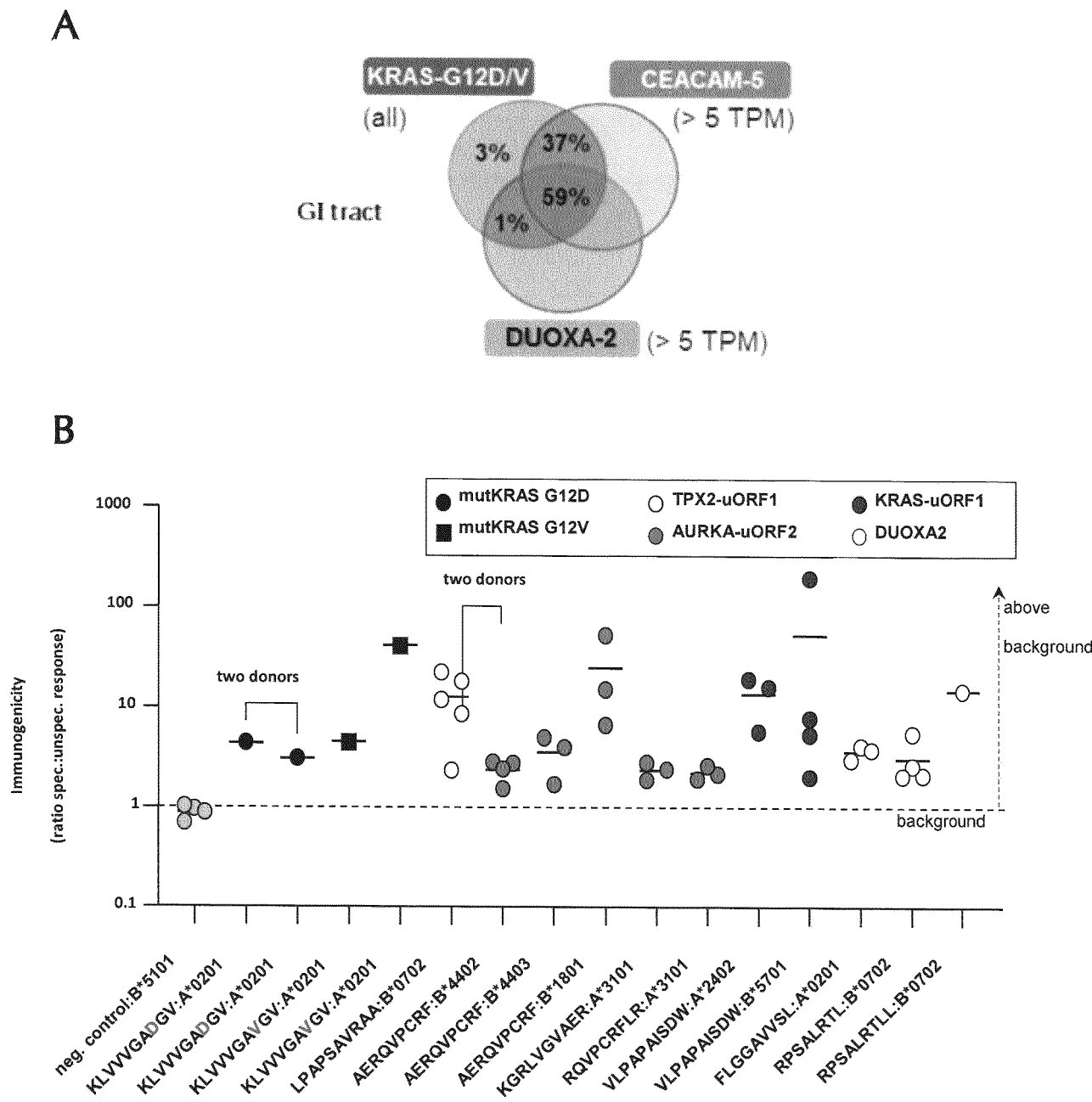


Figure 3

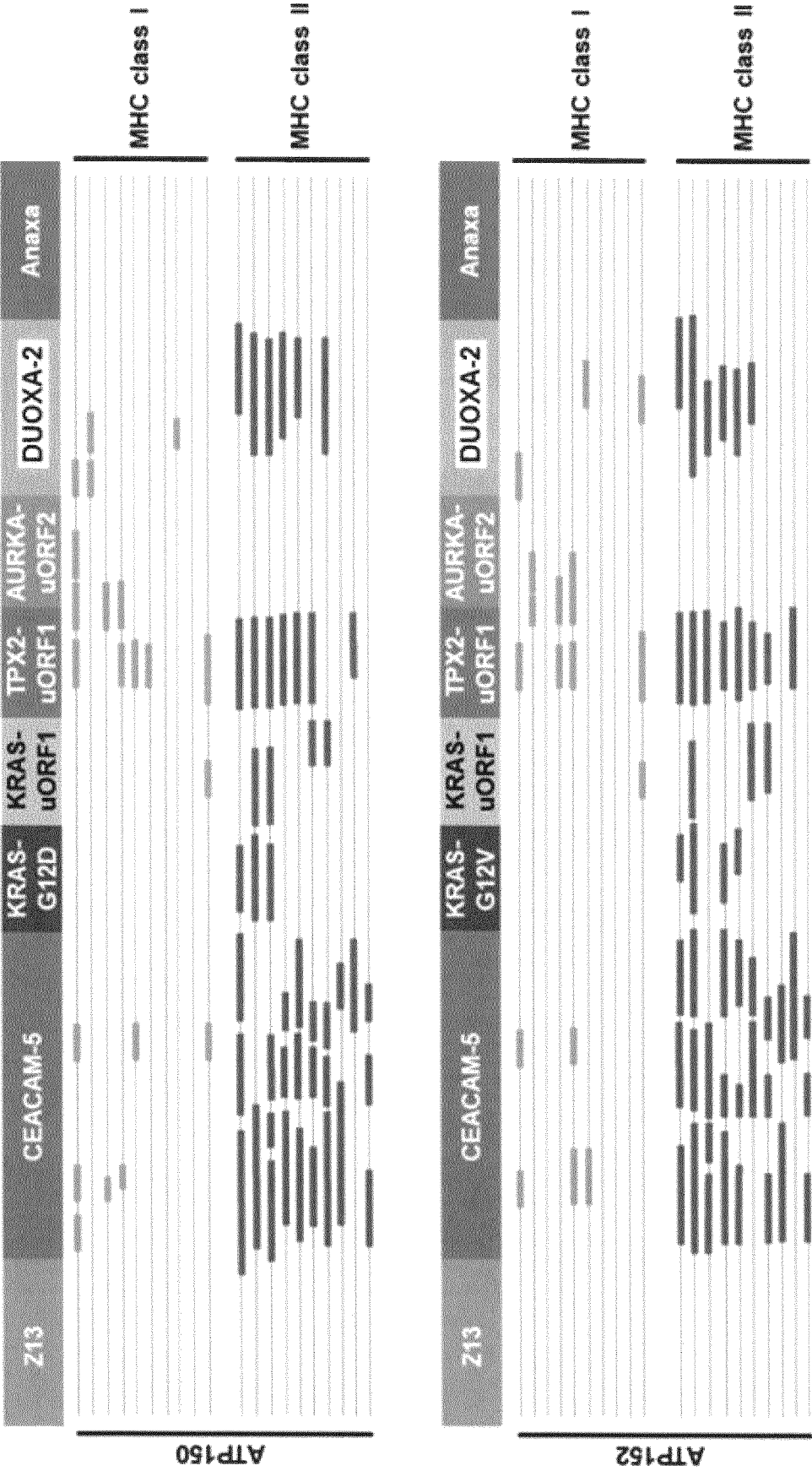
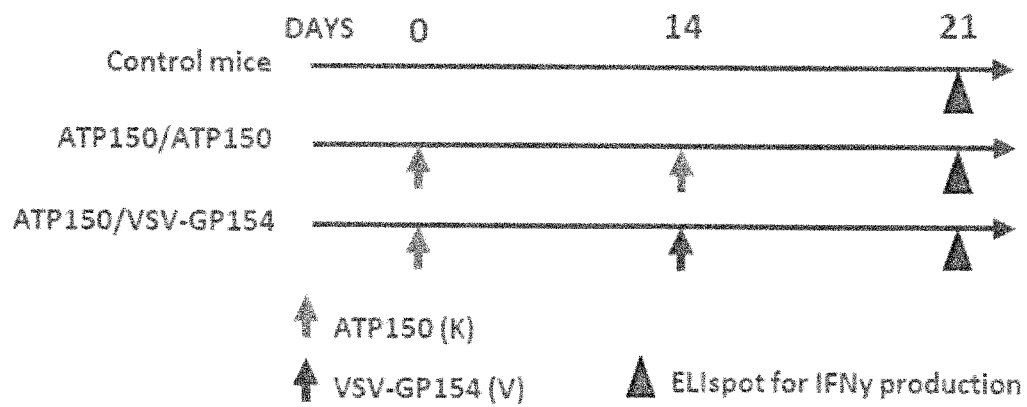


Figure 4

A



B

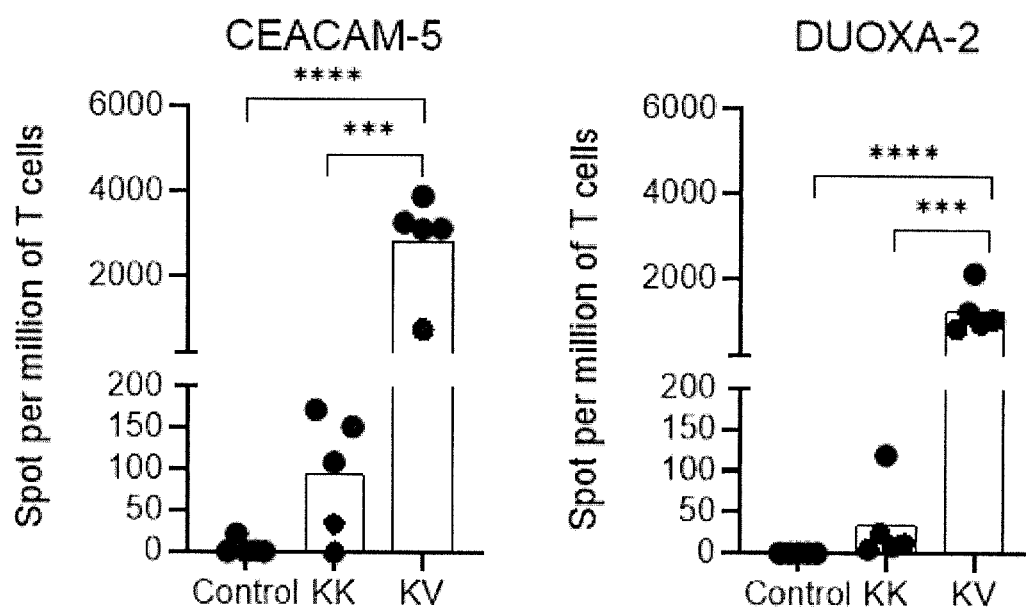


Figure 5

7/10

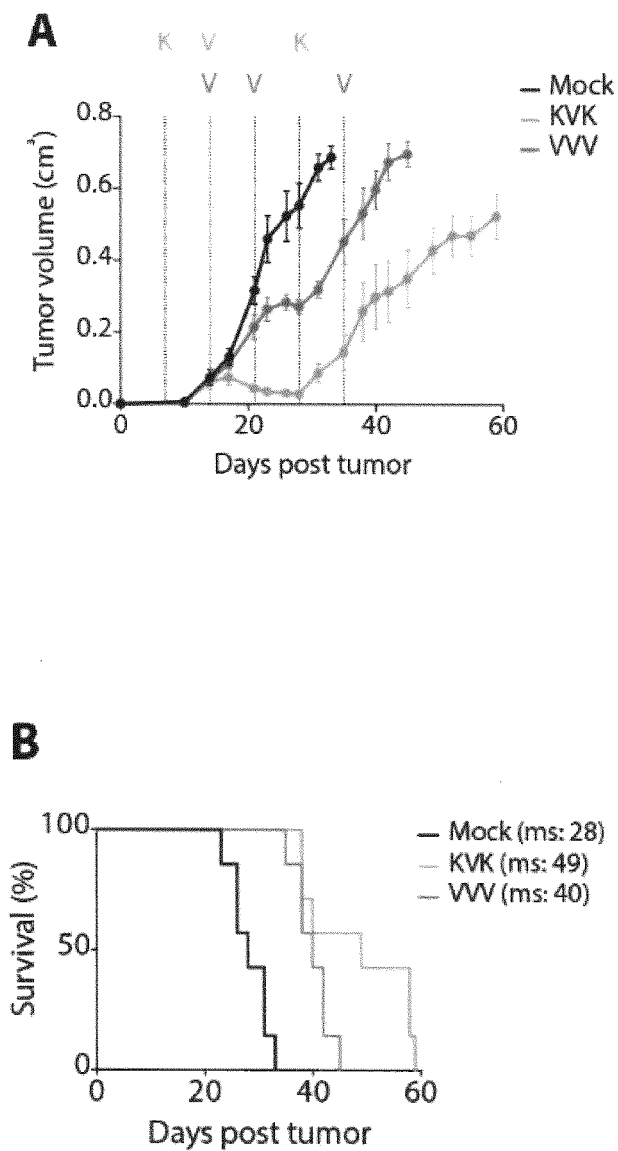
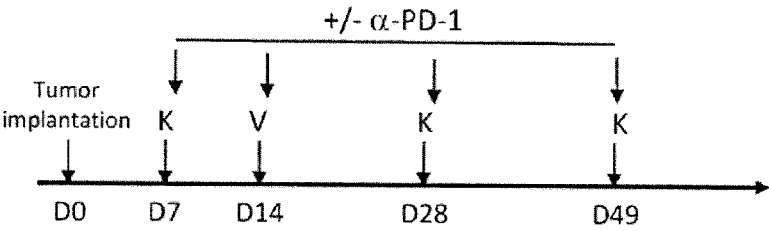


Figure 6

A



B

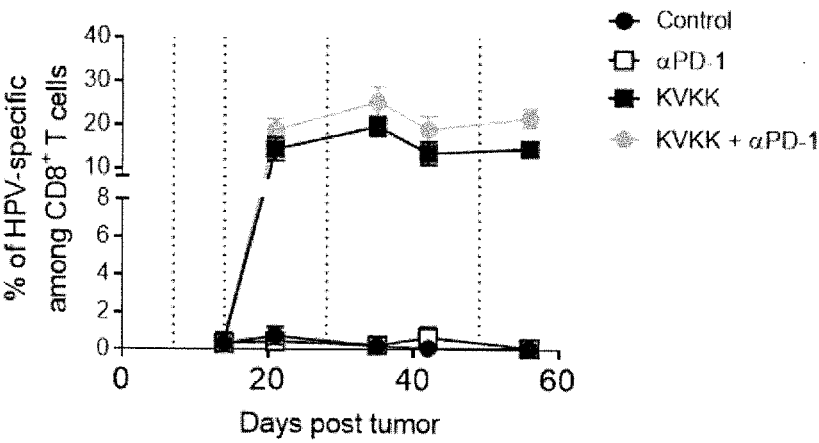
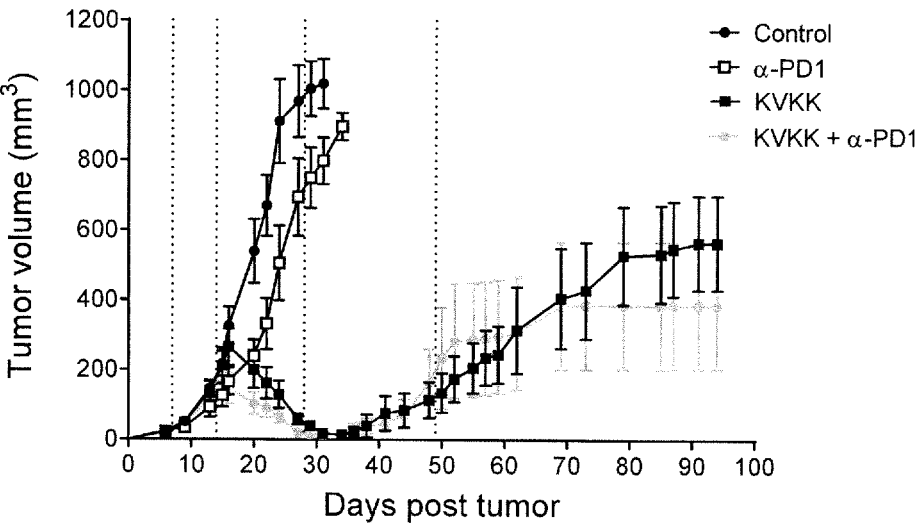


Figure 7

A



B

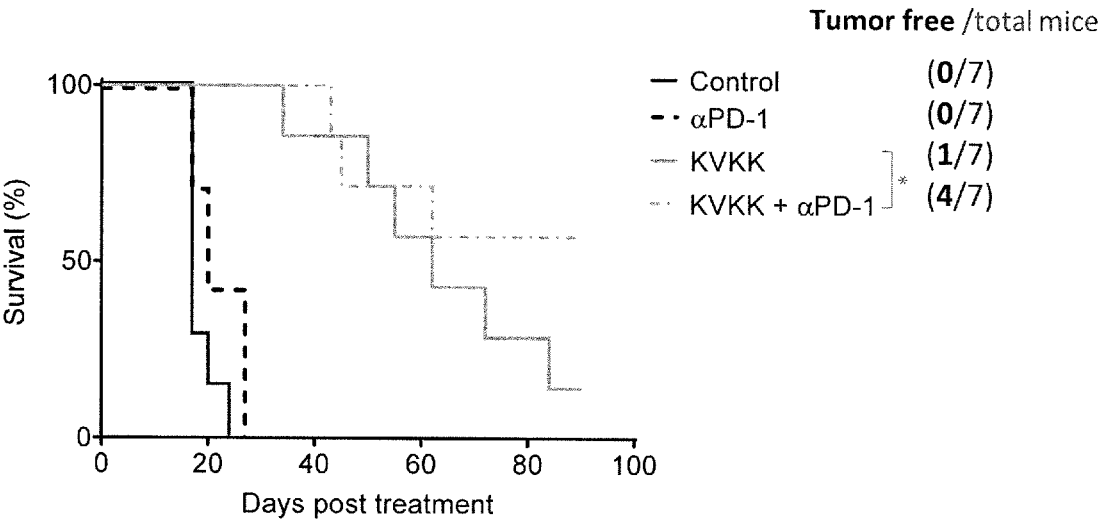


Figure 8

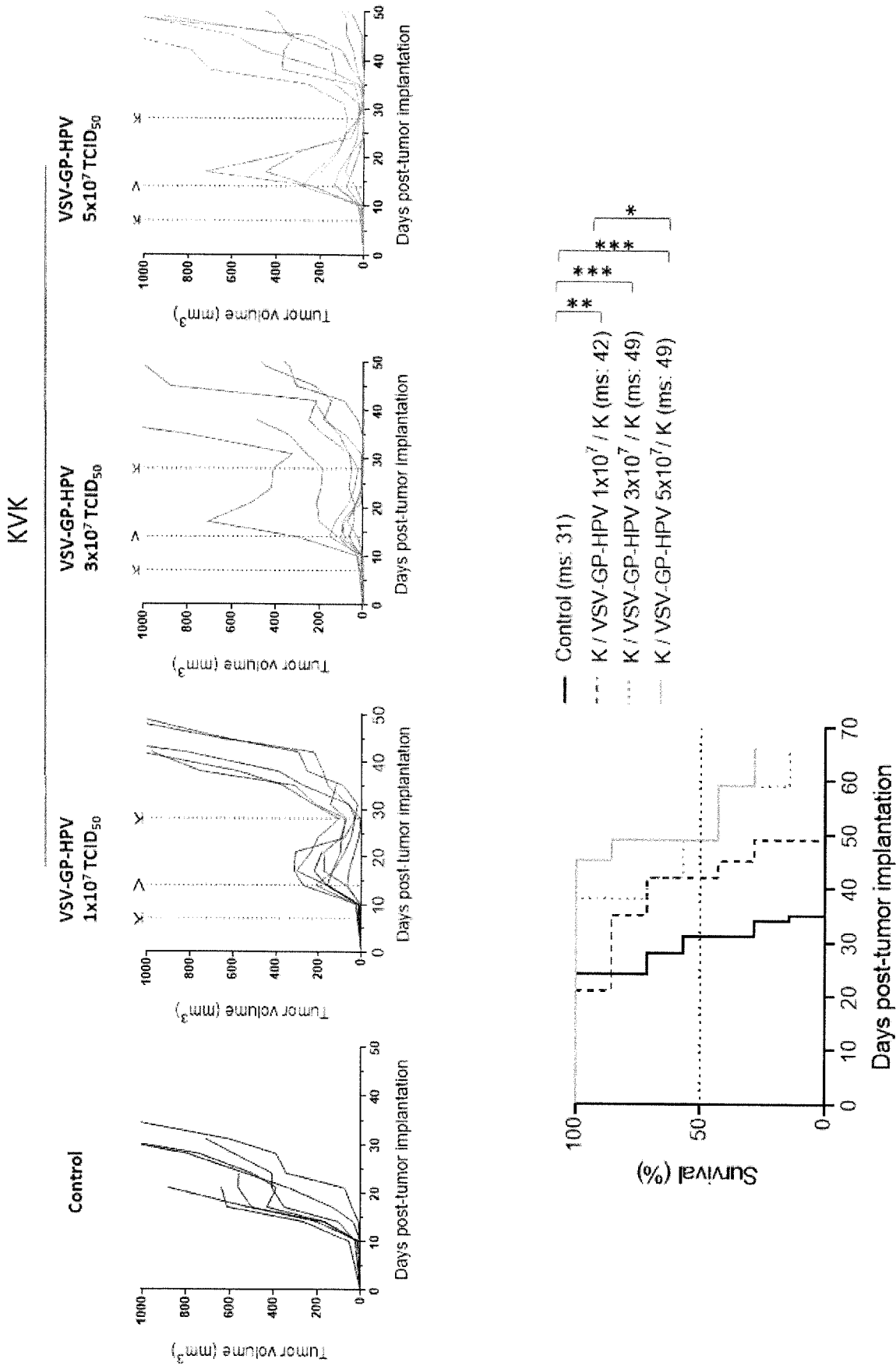


Figure 9

Sequence Listing

1	Sequence Listing Information	
1-1	File Name	AM05P021EPA sequence listing ST26.xml
1-2	DTD Version	V1_3
1-3	Software Name	WIPO Sequence
1-4	Software Version	2.1.2
1-5	Production Date	2022-10-26
1-6	Original free text language code	
1-7	Non English free text language code	
2	General Information	
2-1	Current application: IP Office	
2-2	Current application: Application number	
2-3	Current application: Filing date	
2-4	Current application: Applicant file reference	AM05P021EPA
2-5	Earliest priority application: IP Office	
2-6	Earliest priority application: Application number	
2-7	Earliest priority application: Filing date	
2-8en	Applicant name	Boehringer Ingelheim International GmbH
2-8	Applicant name: Name Latin	
2-9en	Inventor name	
2-9	Inventor name: Name Latin	
2-10en	Invention title	TUMOR ANTIGENS, COMPOUNDS COMPRISING THE TUMOR ANTIGENS AND USES THEREOF
2-11	Sequence Total Quantity	51

3-1	Sequences		
3-1-1	Sequence Number [ID]	1	
3-1-2	Molecule Type	AA	
3-1-3	Length	23	
3-1-4	Features Location/ Qualifiers	source 1..23 mol_type=protein organism=Homo sapiens	
	NonEnglishQualifier Value		
3-1-5	Residues	VAAARPVLPA PAISDWERAR RRH	23
3-2	Sequences		
3-2-1	Sequence Number [ID]	2	
3-2-2	Molecule Type	AA	
3-2-3	Length	28	
3-2-4	Features Location/ Qualifiers	source 1..28 mol_type=protein organism=Homo sapiens	
	NonEnglishQualifier Value		
3-2-5	Residues	VAAAKVAAAR PVLPAPAISD WERARRRH	28
3-3	Sequences		
3-3-1	Sequence Number [ID]	3	
3-3-2	Molecule Type	AA	
3-3-3	Length	38	
3-3-4	Features Location/ Qualifiers	source 1..38 mol_type=protein organism=Homo sapiens	
	NonEnglishQualifier Value		
3-3-5	Residues	VGRKEGAARA RVSLLEPFGT AEVHLPAPSA VRAARPGF	38
3-4	Sequences		
3-4-1	Sequence Number [ID]	4	
3-4-2	Molecule Type	AA	
3-4-3	Length	19	
3-4-4	Features Location/ Qualifiers	source 1..19 mol_type=protein organism=Homo sapiens	
	NonEnglishQualifier Value		
3-4-5	Residues	TAEVHLPAPS AVRAARPGF	19
3-5	Sequences		
3-5-1	Sequence Number [ID]	5	
3-5-2	Molecule Type	AA	
3-5-3	Length	22	
3-5-4	Features Location/ Qualifiers	source 1..22 mol_type=protein organism=Homo sapiens	
	NonEnglishQualifier Value		
3-5-5	Residues	GGDKGRLVGV AERQVPCRFL RP	22
3-6	Sequences		
3-6-1	Sequence Number [ID]	6	
3-6-2	Molecule Type	AA	
3-6-3	Length	38	
3-6-4	Features Location/ Qualifiers	REGION 1..38 note=CEACAM5 fragment source 1..38 mol_type=protein organism=synthetic construct	
	NonEnglishQualifier Value		
3-6-5	Residues	NRTLTLFNVT RNDARAYVSG IQNSVSANRS DPVTLDVL	38
3-7	Sequences		
3-7-1	Sequence Number [ID]	7	
3-7-2	Molecule Type	AA	
3-7-3	Length	19	
3-7-4	Features Location/ Qualifiers	REGION 1..19 note=CEACAM5 fragment source 1..19 mol_type=protein organism=synthetic construct	

3-7-5	NonEnglishQualifier Value Residues	PDSSYLSGAN LNLCHSAS	19
3-8	Sequences		
3-8-1	Sequence Number [ID]	8	
3-8-2	Molecule Type	AA	
3-8-3	Length	63	
3-8-4	Features Location/ Qualifiers	REGION 1..63 note=CEACAM5 fragment source 1..63 mol_type=protein organism=synthetic construct	
3-8-5	NonEnglishQualifier Value Residues	PQYSWRINGI PQQHTQVLFI AKITPNNNGT YACFVSNLAT GRNNSIVKSI TVSASGTSPG 60 LSA	63
3-9	Sequences		
3-9-1	Sequence Number [ID]	9	
3-9-2	Molecule Type	AA	
3-9-3	Length	120	
3-9-4	Features Location/ Qualifiers	REGION 1..120 note=CEACAM5 fragment fusion construct source 1..120 mol_type=protein organism=synthetic construct	
3-9-5	NonEnglishQualifier Value Residues	NRTLTLFNVT RNDARAYVSG IQNSVSNRS DPVTLDVLPD SSYLSGANLN LSCHSASPQY 60 SWRINGIPQQ HTQVLFIKI TPNNGTYAC FVSNLATGRN NSIVKSITVS ASGTSPGLSA 120	
3-10	Sequences		
3-10-1	Sequence Number [ID]	10	
3-10-2	Molecule Type	AA	
3-10-3	Length	33	
3-10-4	Features Location/ Qualifiers	REGION 1..33 note=DUOXA2 fragment source 1..33 mol_type=protein organism=synthetic construct	
3-10-5	NonEnglishQualifier Value Residues	TGVLSLFLGG AVVSLQYVRP SALRTLDDQS AKD	33
3-11	Sequences		
3-11-1	Sequence Number [ID]	11	
3-11-2	Molecule Type	AA	
3-11-3	Length	24	
3-11-4	Features Location/ Qualifiers	REGION 1..24 note=KRAS-G12D fragment source 1..24 mol_type=protein organism=synthetic construct	
3-11-5	NonEnglishQualifier Value Residues	TEYKLVVGA DVGKSAIT QLIQ	24
3-12	Sequences		
3-12-1	Sequence Number [ID]	12	
3-12-2	Molecule Type	AA	
3-12-3	Length	24	
3-12-4	Features Location/ Qualifiers	REGION 1..24 note=KRAS-G12V fragment source 1..24 mol_type=protein organism=synthetic construct	
3-12-5	NonEnglishQualifier Value Residues	TEYKLVVGA VVGKSAIT QLIQ	24
3-13	Sequences		
3-13-1	Sequence Number [ID]	13	
3-13-2	Molecule Type	AA	
3-13-3	Length	260	
3-13-4	Features Location/ Qualifiers	REGION 1..260 note=multi-antigenic domain ATP150	

3-13-5	NonEnglishQualifier Value Residues	source 1..260 mol_type=protein organism=synthetic construct NRTLTLFNVT RNDARAYVSG IQNSVSNRS DPVTLDVLPD SSYLSGANLN LSCHSASPQY 60 SWRINGIPQQ HTQVLFIKI TPNNGTYAC FVSNLATGRN NSIVKSITVS ASGTSPGLSA 120 TEYKLVVVG DVGKGSALTI QLIQVAAARP VLPAPASDW ERARRRHVGR KEGAAARVS 180 LLPEFGTAEV HLPAPSAVRA ARPGFGGDKG RLVGVAERQV PCRFLRPTGV LSLFLGGAVV 240 SLQYVRPSAL RTLLDQSAKD 260
3-14 3-14-1 3-14-2 3-14-3 3-14-4 3-14-5	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers NonEnglishQualifier Value Residues	14 AA 260 REGION 1..260 note=multi-antigenic domain ATP152 source 1..260 mol_type=protein organism=synthetic construct NRTLTLFNVT RNDARAYVSG IQNSVSNRS DPVTLDVLPD SSYLSGANLN LSCHSASPQY 60 SWRINGIPQQ HTQVLFIKI TPNNGTYAC FVSNLATGRN NSIVKSITVS ASGTSPGLSA 120 TEYKLVVVG VVGKGSALTI QLIQVAAARP VLPAPASDW ERARRRHVGR KEGAAARVS 180 LLPEFGTAEV HLPAPSAVRA ARPGFGGDKG RLVGVAERQV PCRFLRPTGV LSLFLGGAVV 240 SLQYVRPSAL RTLLDQSAKD 260
3-15 3-15-1 3-15-2 3-15-3 3-15-4 3-15-5	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers NonEnglishQualifier Value Residues	15 AA 42 REGION 1..42 note=cell penetrating peptide (Z13) source 1..42 mol_type=protein organism=synthetic construct KRYKNRVASR KSRKFKQLL QHYREVAAK SSENDRLRL LK 42
3-16 3-16-1 3-16-2 3-16-3 3-16-4 3-16-5	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers NonEnglishQualifier Value Residues	16 AA 35 REGION 1..35 note=TLR agonist (Anaxa) source 1..35 mol_type=protein organism=synthetic construct STVHEILSKL SLEGDHSTPP SAYGSVKPYT NFDAE 35
3-17 3-17-1 3-17-2 3-17-3 3-17-4 3-17-5	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers NonEnglishQualifier Value Residues	17 AA 338 REGION 1..338 note=ATP150 source 1..338 mol_type=protein organism=synthetic construct MKRYKNRVAS RKSRAKFKQL LQHYREVAAA KSENDRLRL LLKNRTLTLF NVTRNDARAY 60 VSGIQNSVSA NRSDPVTLDV LPDSSYLSGA NLNLSCHSAS PQYSWRINGI PQQHTQVLFI 120 AKITPNNGT YACFVSNLAT GRNNSIVKSI TVSASGTSPG LSATEYKLVV VGADGVGKSA 180 LTIQLIQVAA ARPVLPAI SDWERARRR VGRKEGAARA RVSLLEPFGT AEVHLPAPSA 240 VRAARPGFGG DKGRLVGVAE RQVPCRFLRP TGVLSLFLGG AVVSLQYVRP SALRTLLDQS 300 AKDSTVHEIL SKLSLEGDHS TPPSAYGSVK PYTNFDAE 338
3-18 3-18-1 3-18-2	Sequences Sequence Number [ID] Molecule Type	18 AA

3-18-3	Length	338
3-18-4	Features Location/ Qualifiers	REGION 1..338 note=ATP152 source 1..338 mol_type=protein organism=synthetic construct
3-18-5	NonEnglishQualifier Value Residues	MKRYKNRVAS RKSRAKFKQL LQHYREVAAA KSENDRLRL LLKNRTLTLF NVTRNDARAY 60 VSGIQNSVSA NRSDPVTLDV LPDSSYLSGA NLNLSCHSAS PQYSWRINGI PQQHTQVLFI 120 AKITPNNNGT YACFVSNLAT GRNNSIVKSI TVSASGTSPG LSATEYKLVV VGAVGVGKSA 180 LTIQLIQVAA ARPVLPAI SDWERARRR VGRKEGAARA RVSLLEPEFGT AEVHLPAPSA 240 VRAARPGFGG DKGRLVGVAE RQVPCRFLRP TGVLSLFLGG AVVSLQYVRP SALRTLDDQS 300 AKDSTVHEIL SKLSLEGDHS TPPSAYGSVK PYTNFDAE 338
3-19	Sequences	
3-19-1	Sequence Number [ID]	19
3-19-2	Molecule Type	AA
3-19-3	Length	285
3-19-4	Features Location/ Qualifiers	REGION 1..285 note=multi-antigenic domain VSV-GP154 source 1..285 mol_type=protein organism=synthetic construct
3-19-5	NonEnglishQualifier Value Residues	MNRTLTLFNV TRNDARAYVS GIQNSVSA NRSDPVTLDVLP DSSYLSGANL NLSCHSASPQ 60 YSWRINGIPQ QHTQVLFIK ITPNNNGTYA CFVSNLATGR NNSIVKSITV SASGTSPGLS 120 ATEYKLVVVG ADGVGKSALT IQLIQTEYKL VVGAVGVGK SALTILQIQV AAARPVLPA 180 AISDWERARR RHVGRKEGAA RARVSLLEPEF GTAEVHLPAP SAVRAARPGF GGDKGRLVG 240 AERQVPCRFL RPTGVLSLFL GGAVVSLQYV RPSALRTLLD QSAKD 285
3-20	Sequences	
3-20-1	Sequence Number [ID]	20
3-20-2	Molecule Type	AA
3-20-3	Length	18
3-20-4	Features Location/ Qualifiers	REGION 1..18 note=multi-antigenic domain Mad25 source 1..18 mol_type=protein organism=synthetic construct
3-20-5	NonEnglishQualifier Value Residues	QAEPDRAHYN IVTFSSKS 18
3-21	Sequences	
3-21-1	Sequence Number [ID]	21
3-21-2	Molecule Type	AA
3-21-3	Length	30
3-21-4	Features Location/ Qualifiers	REGION 1..30 note=cell penetrating peptide (Z14) source 1..30 mol_type=protein organism=synthetic construct
3-21-5	NonEnglishQualifier Value Residues	KRYKNRVASR KSRKFKQLL QHYREVAAAK 30
3-22	Sequences	
3-22-1	Sequence Number [ID]	22
3-22-2	Molecule Type	AA
3-22-3	Length	17
3-22-4	Features Location/ Qualifiers	REGION 1..17 note=cell penetrating peptide (Z15) source 1..17 mol_type=protein organism=synthetic construct
3-22-5	NonEnglishQualifier Value Residues	KRYKNRVASR KSRKFK 17
3-23	Sequences	
3-23-1	Sequence Number [ID]	23
3-23-2	Molecule Type	AA
3-23-3	Length	19

3-23-4	Features Location/ Qualifiers	REGION 1..19 note=cell penetrating peptide (Z18) source 1..19 mol_type=protein organism=synthetic construct
3-23-5	NonEnglishQualifier Value Residues	REVAAAKSSE NDRLRLLLK 19
3-24	Sequences	
3-24-1	Sequence Number [ID]	24
3-24-2	Molecule Type	AA
3-24-3	Length	35
3-24-4	Features Location/ Qualifiers	REGION 1..35 note=TLR agonist (Anaxa variant) source 1..35 mol_type=protein organism=synthetic construct
	NonEnglishQualifier Value Residues	STVHEILCKL SLEGDHSTPP SAYGSVKPYT NFDAE 35
3-25	Sequences	
3-25-1	Sequence Number [ID]	25
3-25-2	Molecule Type	AA
3-25-3	Length	498
3-25-4	Features Location/ Qualifiers	REGION 1..498 note=GP of LCMV source 1..498 mol_type=protein organism=synthetic construct
	NonEnglishQualifier Value Residues	MGQIVTMFEA LPHIIDEVIN IIVIIVLIIT SIKAVYNFAT CGILALVSFL FLAGRSCGMY 60 GLNGPDIYKG VYQFKSVEFD MSHLNLTMPN ACSANNSHHY ISMGSSGLEL TFTNDSILNH 120 NFCNLTSAFN KKTFDHTLMS IVSSLHLSIR GNSNHKAVSC DFNNGITIYQ NLSFSDPQSA 180 ISQCRTRFRGR VLDMFRTAFG GKYMRSGWGW AGSDGKTTWC SQTSYQYLII QNRTWENHCR 240 YAGPFGMSRI LFAQEKTFL TRRLAGTFTW TLDSSGVEN PGGYCLTKWM ILAAELKCFG 300 NTAVAKCNVN HDEEFCMDLR LIDYNKAALS KFKQDVESAL HVFKTTVNSL ISDQLLMRNH 360 LRDLMGVPYC NYSKFWYLEH AKTGETSVPK CWLVTNGSYL NETHFSDQIE QEADNMITEM 420 LRKDYIKRQG STPLALMDLL MFSTSAYLIS IFLHLVKIPT HRHIKGGSCP KPHRLTNKGI 480 CSCGAFKVPV VKTIWKRR 498
3-26	Sequences	
3-26-1	Sequence Number [ID]	26
3-26-2	Molecule Type	AA
3-26-3	Length	422
3-26-4	Features Location/ Qualifiers	source 1..422 mol_type=protein organism=vesicular stomatitis virus
	NonEnglishQualifier Value Residues	MSVTVKRIID NTVVVPKLPA NEDPVEYPAD YFRKSKEIPL YINTTKSLSD LRGYVYQGLK 60 SGNVSIHVN SYLYGALKDI RGKLDKDWSS FGINIGKAGD TIGIFDLVSL KALDGVLPDG 120 VSDASRTSAD DKWLPLYLLG LYRVGRTQMP EYRKKLMDGL TNQCKMINEQ FEPLVPEGRD 180 IFDVWGNDSN YTKIVAADVDM FFHMFKKHEC ASFRYGTIVS RFKDCAALAT FGHLCKITGM 240 STEDVTTWIL NREVADEMVG MMLPGQEIDK ADSYMPYLID FGLSSKSPYS SVKNPAFHFV 300 GQLTALLRS TRARNARQPD DIEYTSLT TA GLLYAYAVGS SADLAQQFCV GDNKYTPDDS 360 TGGLTTNAPP QGRDVVEWLG WFEDQNRKPT PDMMQYAKRA VMSLQGLREK TIGKYAKSEF 420 DK 422
3-27	Sequences	
3-27-1	Sequence Number [ID]	27
3-27-2	Molecule Type	AA
3-27-3	Length	265
3-27-4	Features Location/ Qualifiers	source 1..265 mol_type=protein organism=vesicular stomatitis virus
	NonEnglishQualifier Value Residues	MDNLTKVREY LKSYSRLDQA VGEIDEIEAQ RAEKSNYELF QEDGVEEHTK PSYFQAADD 60 DTESEPEIED NQGLYAPDPE AEQVEGFIQG PLDDYADEEV DVVFTSDWKQ PELESDEHGK 120 TLRLTSPEGL SGEQKSQWLS TIKAVVQSAK YWNLAECTFE ASGEGVIMKE RQITPDVYKV 180

		TPVMNTHPSQ SEAVSDVWSL SKTSMTFQPK KASLQPLTIS LDELFSSRGE FISVGGDGRM 240 SHKEAILLGL RYKKLYNQAR VKYSL 265
3-28	Sequences	
3-28-1	Sequence Number [ID]	28
3-28-2	Molecule Type	AA
3-28-3	Length	2109
3-28-4	Features Location/ Qualifiers	source 1..2109 mol_type=protein organism=vesicular stomatitis virus
3-28-5	NonEnglishQualifier Value Residues	MEVHDFETDE FDNFNEDDYA TREFLNPDER MTYLNHADYN LNSPLISDDI DNLIRKFNSL 60 PIPSMWD SKN WDGWLEMLTS CQANPIPTSQ MHKWMGSWLM SDNHIDASQGY SFLHEVDKEA 120 EITFDVVETF IRGWGNKPIE YIKKERWTDS FKILAYLCQK FLDLHKLTLI LNAVSEVELL 180 NLARTFKGKV RRSSHGNTIC RIRVPSLGPT FISEGWAYFK KLDILMDRNF LLMVKDVIIG 240 RMQTVLSMVC RIDNLFSEQD IFSLLNIYRI GDKIVERQGN FSYDLIKMVE PICNLKLMKL 300 ARESRLVPQ FPHFENHIKT SVDEGAKIDR GIRFLHDQIM SVKTVDLTLV IYGSFRHWGH 360 PFIDYYTGLE KLHSQVTMCK DIDVSYAKAL ASDLARIVLF QQFNDHKKWF VNGDLLPHDH 420 PFKSHVKENT WPTAAQVQDF GDKWHELPLI KCFEIPDLLD PSIIYSDKSH SMNRSEVLKH 480 VRMNPNTPIP SKKVLQTMLD TKATNWKEFL KEIDEKGLDD DDLIIGLKGK ERELKLAGRF 540 FSLMSWKLRE YFVITEYLIK THFVPMFKGL TMADDLTAVI KKMLDSSSGQ GLKSYEAIKI 600 ANHIDYEKWN NHQRKLSNGP VFRVMGQFLG YPSLIERTHE FFEKSLIYYN GRPDLMRVHN 660 NTLINSTSQR VCWQGGEGGL EGLRQKGWSI LNLLVIQREA KIRNTAVKVL AQGDNQVICT 720 QYKTKKSRNV VELQGALNQM VSNNEKIMTA IKIGTGKGL LINDDETMS ADYLNYGKIP 780 IFRGVIRGLE TKRWSRVTCV TNDQIPTCAN IMSSVSTNAL TVAHFAENPI NAMIQYNYFG 840 TFARLLLMMH DPALRQSLYE VQDKIPGLHS STFKYAMLYL DPSIGGVSGM SLSRFLIRAF 900 PDPVTESLSF WRFIHVHARS EHLKEMSAVF GNPEIAKFRI THIDKLVEDP TSLNIAMGMS 960 PANLLKTEVK KCLIESRQTI RNQVIKDATI YLYHEEDRLR SFLWSINPLF PRFLSEFKSG 1020 TFLGVADGLI SLFQNSRTIR NSFKKKYHRE LDDLIVRSEV SSLTHLGKLH LRRGCKMWT 1080 CSATHADTLR YKSWGRTVIG TTVPHPLEML GPQHRKETPC APCNTSGFNY VSVHCPDGIH 1140 DVFSSRGPLP AYLGSKTSES TSILQPWERE SKVPLIKRAT RLRDAISWFV EPDSKLAMTI 1200 LSNIHSLTGE EWTQRQHGFK RTGSALHRFS TSRMSHGGFA SQSTAALTRL MATTDTMRLD 1260 GDQNFDFLFQ ATLLYAQITT TVARDGWITS CTDHYHIACK SCLRPIEEIT LDSSMDYTPP 1320 DVSHVLKTRW NGEQSWGQEI KQIYPLEGNW KNLAPAEQSY QVGRICIGFLY GDLAYRKSTH 1380 AEDSSLFPLS IQGRIRGRGF LKGLLDGLMR ASCCQVIHRR SLAHLKRPAV AVYGGGLIYLI 1440 DKLSVSPFPL SLTRSGPIRD ELETIPHKIP TSYPTSNRDM GVIVRNYFKY QCRLIEKGKY 1500 RSHYSQLWLF SDVLSIDFIG PFSISTTLLQ ILYKPFPSGK DKNELRELAN LSSLLRSGEG 1560 WEDIHVKFFT KDILLCPPEI RHACKFGIAK DNNKDMSYPP WGRESRGTIT TIPVYTTTTP 1620 YPKMLEMPPR IQNPLLSGIR LGQLPTGAHY KIRSILHGMG IHYRDFLSCG DSGSGMTAAL 1680 LRENVHSGRI FNSLLELSGS VMRGASPEPP SALETGGDK SRCVNGETCW EYPSDLCDPR 1740 TWDYFLRLKA GLGLQIDLIV MDMEVRDSSS SLKIETNVRN YVHRILDEQG VLIYKTYGTY 1800 ICESEKNAV TILGPMFKTVD LVQTEFSSSQ TSEVYMVCKG LKKLIDEPNP DWSSINESWK 1860 NLAYFQSSEQ EFARAKKVST YFTLTGIPSQ FIPDPFVNIE TMLQIFGVPT GVSHAAALKS 1920 SDRPADLLTI SLFYMAIISY YNINHVRVGP IPPNPPSDGI AQNVGIAITG ISFWLSLMEK 1980 DIPLYQQCLA VIQQSFPIRW EAVSVKGGYK QKWSTRGDGL PKDTRISDSL APIGNWIRSL 2040 ELVRNQVRLN PFNEILFNQL CRTVDNHLKW SNLRRNTGMI EWINRRISKE DRSILMLKSD 2100 LHEENSWRD 2109
3-29	Sequences	
3-29-1	Sequence Number [ID]	29
3-29-2	Molecule Type	AA
3-29-3	Length	229
3-29-4	Features Location/ Qualifiers	source 1..229 mol_type=protein organism=vesicular stomatitis virus
3-29-5	NonEnglishQualifier Value Residues	MSSLKKILGL KGKGGKSKKL GIAPPPYEED TSMEYAPSAP IDKSYFGVDE MDYDPNQRL 60 YEKFFFTVKM TVRSNRPFRT YSDVAAAVSH WDHYIGMAG KRPFYKILAF LGSSNLKATP 120 AVLADQGQPE YHAHCEGRAY LPHRMGKTPP MLNVPEHFRR PFNIGLYKGT IELTMTIYDD 180 ESLEAAPMIW DHFNSSKFSD FREKALMFGL IVEKKASGAW VLDSIGHFK 229
3-30	Sequences	
3-30-1	Sequence Number [ID]	30
3-30-2	Molecule Type	RNA
3-30-3	Length	12057
3-30-4	Features Location/ Qualifiers	misc_feature 1..12057 note=VSV-GP154 (complement without reverse) source 1..12057

3-30-5	NonEnglishQualifier Value Residues	mol_type=other RNA organism=synthetic construct									
		tgcttctgtt	tgtttggtaa	taatagtaat	tttccgagtc	ctcttttgaaa	ttgtcattag	60			
		ttttacagac	aatgtcagtt	ctcttagtaa	ctgttggtgc	agcatcaagg	ttttgaagga	120			
		cgtttactcc	taggtcacct	tatgggccgt	ctaataaggt	cttttagttt	cctctaagga	180			
		gaaatgtagt	tatgatgttt	ttcaaacagt	ctagattctc	ctatacagat	ggttccggag	240			
		tttaggcctt	tacatagtta	gtatgtacag	ttgtcgatga	acataacctc	taatttcctg	300			
		taggccccat	tcaacctatt	tctaaccagt	tcaaagcctt	attttagacc	ctttcgtccc	360			
		ctatgttagc	cttataaact	ggaacatagg	aactttcggg	acctgccgca	tgaaggtcta	420			
		cctcatagcc	tacgaaggtc	ttggtcgcgt	ctactgttta	ccaacggaaa	catagatgaa	480			
		ccgaatatgt	ctcacccgtc	ttgtgtttac	ggacttatgt	cttttttcga	gtacctaccc	540			
		gactgtttag	ttacgtttta	ctagttactt	gtcaaacttg	gagaacacgg	tcttccagca	600			
		ctgtaaaaac	tacacacccc	tttactgtca	ttaatgtgtt	tttaacagcg	acgtcacctg	660			
		tacaagaagg	tgtacaagtt	ttttgtactt	acacggagca	agtctatgcc	ttgataacaa	720			
		aggtctaagt	ttctaacacg	acgtaaccgt	tgtaaacctg	tggagacggt	ttattggcct	780			
		tacagatgtc	ttctacattg	ctggacctag	aacttggctc	ttcaacgtct	actttaccag	840			
		gtttactacg	aagggtccgt	tctttaactg	ttccggctaa	gtatgtacgg	aataaactag	900			
		ctgaaaccta	acagaagatt	cagaggtata	agaaggcagt	ttttgggacg	gaagggtgaag	960			
		acccccgtta	actgtcgaga	agacgagtct	aggtgggtct	gttccttacg	ggctgtcgga	1020			
		ctactgtaac	tcatatgtag	agaatgatgt	cgtccaaaca	acatgcgaat	acgtcatcct	1080			
		aggagacggc	tgaaccgtgt	tgtcaaaaaca	caacctctat	tgtttatgtg	aggtctacta	1140			
		tcatggcctc	ctaactgctg	attacgtggc	ggtgttccgt	ctctacacca	gcttaccgag	1200			
		cctaccaaac	ttctagtttt	gtcttttggc	tgaggactat	actacgtcat	acgcttttct	1260			
		cgtcagtaca	gtgacgttcc	ggattctctc	ttctgttaac	cgttcatacg	attcagtctt	1320			
		aaactgttta	ctgggataatt	aagagtctag	tggataatat	ataatacgat	gtatactttt	1380			
		tttgattgtc	tatagtacct	attagagtgt	tttcaagcac	tcatagagtt	caggataaga	1440			
		gcagacctag	tccgccatcc	tctctatcta	ctctagcttc	gtgttgctcg	acttttcagg	1500			
		tttaactca	acaaggttct	cctacctcac	cttctcgat	gattcgggag	aataaaagtc	1560			
		cgtcgtctac	taagactgtg	tcttagactt	ggtctttaac	ttctgttagt	tccgaacata	1620			
		cggtgtctag	gtcttcgact	cgttcaactt	ccgaaatatg	tccccgaaa	tctactgata	1680			
		cgctactacc	ttcacctaca	acataaatga	agcctgacct	ttgtcggact	cgaacttaga	1740			
		ctgctctgac	ctttctggaa	tgccaactgt	agcgtctctc	caaattcacc	tctcgtcttt	1800			
		aggttcaccg	aaagctgcta	atttcgtcag	cacgtttcac	ggtttatgac	cttagaccgt	1860			
		ctcacgtgta	aacttcgtag	ccctcttccc	cagtaatact	tcctcgcggt	ctattgaggc	1920			
		ctacatatat	tccagtgagg	tcactacttg	tgtgtaggca	gggttagtct	tcgtcatagt	1980			
		ctacaaacca	gagagagttt	ctgtaggtag	tgaaggttg	ggttctttcg	ttcagaagtc	2040			
		ggagagtgg	ataggaacct	acttaacaag	agtagatctc	ctctcaagta	gagacagcct	2100			
		ccactgcctg	cttacagagt	atttctccgg	taggacgagc	cggactctat	gtttttcaac	2160			
		atgttagtcc	gctctcagtt	tataagagac	atctgatact	ttttttcatt	gtctatagt	2220			
		ctagattcac	aatagggtta	ggtaagtagt	actcaaggaa	tttcttctaa	gagccagact	2280			
		tcccccttcc	attctttaga	ttctttaatc	cctagcgtgg	tgggggaata	cttctcctgt	2340			
		gatcgtacct	catacgaggc	tcgcgagggt	aactgtttag	gataaaacct	caactgctct	2400			
		acctgtggat	actaggctta	gttaattcta	tactctttaa	gaagaaatgt	cacttttact	2460			
		gccaatctag	attagcagcg	aagtcttgta	tgagtctaca	ccgtcggcga	catagggtaa	2520			
		ccctagtgtg	catgtagcct	taccgtccct	ttgcagggaa	gatgttttag	aaccgaaaaa	2580			
		acccaagaag	attagatttc	cgttgaggtc	gccataaccg	tctagtcca	gttggtctca	2640			
		tagtgcgagt	gacgcttccg	tcccgaataa	acggtgtatc	ctacccttcc	tggggagggt	2700			
		acgagttaca	tggctctctg	aagtcttctg	gtaagttata	tccagaaatg	ttcccttgct	2760			
		aactcgagt	ttactggtag	atgctactac	tcagtgaacct	tcgtcgagga	tactagacct	2820			
		tagtaaaagt	aagaagggtt	aaaagactaa	agtctctctt	ccggaattac	aaaccggact	2880			
		aacagctctt	tttccgtaga	cctcgcaccc	aggacctgag	atagccggtg	aagtttactc	2940			
		gatcagattg	aagatcgaag	acttgttagg	ggccaaatga	gtcagagggg	attaagggtc	3000			
		gagagcttgt	tgattatagg	acagaaaaga	tagggatact	ttttttgatt	gtctctagct	3060			
		agacaaatgc	gcagtgccta	ggggggccga	cgctcttaag	cgggtgtacc	cggcttagca	3120			
		ctggtacaag	ctccgggacg	gggtgtagta	gctgctccac	tagttgtagc	actagtagca	3180			
		cgagtagtag	tagtggtcgt	agttccggca	catgttgaag	cgggtggacgc	cgtaggacct	3240			
		ggaccactcg	aaggacaagg	accggccgtc	ttcgacgccg	tacatgccgg	acttaccggg	3300			
		gctatagatg	ttcccgca	tgggtcaagtt	ctcgcacctc	aagctgtact	cgggtggactt	3360			
		ggagtgttac	gggttgcgga	cgctcgcggt	gttatcgggt	gtgatgtagt	cgtaccctgc	3420			
		gtcgcgggac	ctcaactgga	agtgggtgct	gtcgtaggac	ttgggtgtga	agacgttgga	3480			
		gtggtcgcgg	aagttgttct	tttggaagct	ggtgtgggag	tactcgtagc	actcgtcgga	3540			
		cgtggactcg	tagtctccgt	gtcgttgggt	gttccggcac	tcgacgtgta	agttgttgcc	3600			
		gtagtggtag	gtcatgttgg	actcgaagtc	gctaggagtc	tcgcggtagt	cggtcacgtc	3660			
		ttggaagtct	ccgtctcacg	acctgtacaa	gtcttggcgg	aagccgccgt	tcatgtactc	3720			

ttcgccgacc	ccgaccgcgc	cgctcgctgcc	gttctggtgg	accacgtcgg	tctggtcgat	3780
ggatcatggag	tagtaggtct	tgtcttggac	cctcttgggtg	acgtctatgc	ggcctggaaa	3840
gccgtactcg	tcttaggaca	agcgggtcct	cttttgggttc	aaggagtggg	cctctgaccg	3900
gccgtggaag	tggacctggg	actcgctgtc	gtcgccgcac	ctcttgggac	cgccgatgac	3960
ggagtgggtc	acctactagg	accggcggtc	cgacttcacg	aagccgttgt	ggcggcaccg	4020
gttcacgttg	cacttgggtc	tgtctctcaa	gacgctgtac	gactctgagt	agctgatgtt	4080
gttccggcgg	gactcgttca	agttcgtcct	gcacctctcg	cgggacgtgc	acaagtctctg	4140
gtggcacttg	tcggagtagt	cgctgggtcga	cgagtactct	ttgggtggact	ctctggagta	4200
cccgcacggg	atgacgttga	tgtcgttcaa	gaccatagac	ctcgtgcggg	tctggccgct	4260
ctggctgcac	gggttcacga	ccgaccactg	gttaccgtcg	atggacttgc	tctgggtgaa	4320
gtcgctgggtc	tagctcgtcc	ttcgggtgtt	gtactagtgg	ctctacgact	ccttcctgat	4380
gtagttctct	gtcccgtcgt	gggggggaccg	ggagtacctta	gacgagtaca	agtcgtgggtc	4440
gcggatggag	tagtcttaga	aggacgtgga	ccacttctag	gggtgggtgt	ctgtgtagtt	4500
cccgccgtcg	acgggggttcg	gggtgtctga	gtgggtgttc	ccgtagacgt	cgacgccgcg	4560
gaagtccac	ggcccgact	tttggtagac	cttctctct	attcgccggc	gatgtggag	4620
ctgatacttt	ttttgattgt	ctataggagc	tgcggtggta	cttggcctgt	gactgggaca	4680
agttgcactg	ggccttactg	cggtctcgga	tgcacaggcc	gtaggtctta	tcgcacaggc	4740
ggttatcttc	gctggggcac	tgtgacctgc	acgacggact	atcgtcgata	gactcgccgc	4800
ggttggactt	ggactcgacg	gtaagacgga	gaggtgtcat	gtcgaccgcc	tagttgcctt	4860
agggagtctg	cgtgtgggtc	cacgacaaat	agcgggttcta	gtggggattg	ttgttgccgt	4920
ggatgcggac	gaagcacagg	ttagaccggt	ggccgtcttt	gttgtcgtag	cacttcagggt	4980
agtggcacag	acggtcgccg	tgtagaggac	ctgactcgcg	gtgtctcatg	ttcgaccacc	5040
aacaccctcg	actgccgcac	cctttttcgc	gggactgtta	ggtcgactag	gtctgggtca	5100
tatttgagca	ccagcagccg	cgacaaccgc	aaccgtttag	acgagactgg	taagtgcagt	5160
aagtccaccg	gcgacggtct	ggacaagaag	gacgaggacg	gtagtcgctg	accctttctc	5220
gatcttcttc	cgtgcagccg	gcctttcttc	cgcgccgacg	tcgatctcaa	tcggacgacg	5280
gtctcaaacc	gtggcggctt	caagtagacg	gacgaggatc	gcgacagtct	cggcgggtctg	5340
gtcctaaacc	gcctctattc	ccgtctgagc	acccgcaacg	tctttctgtc	cacgggacgt	5400
ctaaggactc	cggttgtccg	cacgactcgg	acaaagaacc	gcctcgacac	cacagggacg	5460
tcatgcagtc	tggatcgcg	gactcttggg	acgacctagt	ctcgcggttc	ctaactagtt	5520
gatcgggtcta	agaagtacaa	acctgggtta	gttgaacact	atgggtacgag	tttctccgga	5580
gttaataata	actcaaaaat	taaaaatact	ttttttgatt	gtcgttagta	ccttcagggtg	5640
ctaaaactct	ggctgctcaa	gttactaaag	ttacttctac	tgatacgggtg	ttctcttaag	5700
gacttagggc	tactcgcgta	ctgcatgaac	ttagtacgac	taatgttgg	cttaagaggga	5760
gattaatcac	tactataact	gttaaatttag	tcctttaagt	taagagaagg	ttaagggagc	5820
tacaccctat	cattcttgac	cctacctcaa	gaactctaca	attgcagtag	agttcgggtta	5880
gggtaggggt	gtagagtcta	cgtatttacc	tacccttcaa	ccaattacag	actattagta	5940
ctacgggtcag	ttcccatatc	aaaaaatgta	cttcacctgt	ttctccgtct	ttattgtaaa	6000
ctgcaccacc	tctggaagta	ggcgccgacc	ccgttgtttg	gttaacttat	gtagtttttc	6060
ctttctacct	gactgagtaa	gttttaagag	cgaataaaca	cagttttcaa	aaacctgaat	6120
gtgttcaact	gtaattagaa	tttacgacag	agactccacc	ttaacgagtt	gaaccgctcc	6180
tgaaggtttc	cgtttcagtc	ttcttcaaga	gtaccttgct	tgtatacgtc	ctaataccaa	6240
gggtcgaacc	caggatgaaa	ataaagtctt	cctaccgcaa	tgaagtctct	tgaactataa	6300
gattacctgg	ctttgaaaga	caattaccag	tttctacact	aatatccctc	ctacgttttg	6360
cacgataggt	accatacatc	ttatctgttg	gacaagagtc	tcgttctgta	gaagagggaa	6420
gatttataga	tgtcttaacc	tctattttaa	cacctctccg	tccctttaaa	aagaatactg	6480
aactaatttt	accaccttgg	ctatacgttg	aacttcgact	actttaatcg	ttctcttagt	6540
tccggaatc	aggggtgttaa	gggagtaaaa	cttttagtat	agttctgaag	acaactactt	6600
ccccgttttt	aactggctcc	atatttctaag	gaggtactag	tctattactc	acacttttgt	6660
cacctagagt	gtgaccacta	aatacctagc	aagtctgtaa	ccccagtagg	aaaatatcta	6720
ataatgtgac	ctgatctttt	taatgtaagg	gttcattgggt	acttctttct	ataactacac	6780
agtatacgtt	ttcgtgaacg	ttcactaaat	cgagcctaac	aagataaagt	tgtcaagtta	6840
ctagtatttt	tcaccaagca	cttacctctg	aacgagggag	tactagtagg	gaaattttca	6900
gtacaatttc	ttttatgtac	cgggtgtcga	cgagtccaag	ttctaaaacc	tctatttacc	6960
gtacttgaag	gcgactaatt	tacaaaactt	tatgggtcga	atgatctggg	tagctattat	7020
atgagactgt	tttcagtaag	ttactttatc	agtctccaca	actttgtaca	ggcttactta	7080
ggcttgtgag	gatagggatc	atttttccac	aacgtctgat	acaacctgtg	tttccgatgg	7140
ttaacctttc	ttaaagaatt	tctctaacta	ctcttcccg	atctactact	actagattaa	7200
taaccagaat	ttcctttcct	ctcccttgac	ttcaaccgtc	catctaaaaa	gagggattac	7260
agaaccttta	acgtctttat	gaaacattaa	tggcttataa	actatttctg	agtaaagcag	7320
ggatacaaat	ttccggactg	ttaccgcctg	ctagattgac	gtcagtaatt	tttctacaat	7380
ctaaggagta	ggccggttcc	taacttcagt	atactccggt	aaacgtatcg	gttagtgtaa	7440
ctaatacttt	ttaccttatt	gggtgtttcc	ttcaatagtt	tgcgggttca	caaggctcaa	7500
taccgggtca	agaatccaat	aggtaggaat	tagctctctt	gagtacttaa	aaaactcttt	7560
tcagaatata	tgatgttacc	ttctgggtcg	aactacgcac	aagtgttgtt	gtgtgactag	7620

ttaagttgga	gggttgctca	aacaaccgtt	cctgttctcc	cacctgacct	tccagatgcc	7680
gtttttccta	cctcatagga	gttagatgac	caataagttt	ctctccgatt	ttagtctttg	7740
tgacgacagt	ttcagaaccg	tgttccacta	ttagttcaat	aaacgtgtgt	catattttgc	7800
ttcttttagct	ctttgcaaca	tcttaatgtc	ccacgagagt	tagtttacca	aagattatta	7860
ctcttttaat	actgacgtta	gttttatccc	tgtcccttca	atcctgaaaa	ctatttactg	7920
ctactctgat	acgttagacg	tctaataaac	ttaatacctt	tttatggcta	aaaggcacct	7980
cactaatctc	ccaatctctg	gttctctacc	agtgtctact	gaacacagtg	gttactgggt	8040
tatgggtgaa	cacgattata	ttactcgagt	caaagggtgt	tacgagagtg	gcatcgagta	8100
aaacgactct	tgggttagtt	acggtactat	gtcatgttaa	taaaaccctg	taaacgatct	8160
gagaacaact	actacgtact	aggacgagaa	gcagttagta	acatacttca	agttctattc	8220
tatggcccga	acgtgtcaag	atgaaagttt	atgcggtaca	acataaacct	gggaaggtaa	8280
cctcctcaca	gccgtacag	aaacaggtcc	aaaaactaat	ctcgggaagg	tctagggcat	8340
tgtctttcag	agagtaagac	ctctaagtag	gtacatgtac	gagcttcaact	cgtagacttc	8400
ctctactcac	gtcataaacc	tttggggctc	tatcggttca	aagcttattg	agtgtatctg	8460
ttcgatcatc	ttctaggttg	gagagacttg	tagcgatacc	cttactcagg	tcgcttgaac	8520
aatttctgac	tccaattttt	tacgaattag	cttagttctg	tttggttagtc	cttggtccac	8580
taattcctac	gttgggtatat	aaacatagta	cttctcctag	ccgagtcctc	aaagaatacc	8640
agttatttag	gagacaagg	atctaaaaat	tcacttaagt	ttagtccgtg	aaaaaacctt	8700
cagcgtctgc	ccgagtagtc	agataaagtt	ttaagagcat	gataagcctt	gaggaaattc	8760
tttttcatag	tatcccttaa	cctactaaac	taacactcct	cactccatag	gagaacttgt	8820
gtaaatccct	ttgaagtaaa	ctcttcccct	agtacatttt	acacctgtac	aagtcgatga	8880
gtacgactgt	gtaattctat	gtttaggacc	ccggcatgtc	aataaccctg	ttgacatggg	8940
gtaggtaatc	tttacaaccc	agggtgtgta	gcttttctct	gaggaacacg	tggtacattg	9000
tgtagtccca	agttaatata	aagacacgta	acaggctctg	cctaggtagt	gcagaaatca	9060
agtggccctg	gtaacggacg	aatagatccc	agattttgta	gacttagatg	tagataaaac	9120
gtcggaaacc	tttccctttc	gtttcagggt	gactaatttt	ctcgtatgtc	agaatctcta	9180
cgatagagaa	ccaaacaact	tgggctgaga	tttgatcggt	actgatatga	aagattgtag	9240
gtgagaaatt	gtccgcttct	tacctgggtt	tccgtcgtac	ccaagttttc	ttgtcccaga	9300
cggaagtat	ccaaaagctg	tagagcctac	tcggtagcac	ccaagcgtag	agtctcgtga	9360
cgtcgtaact	ggtccaacta	ccgttgatgt	ctgtggtact	ccctagaccc	tctagtctta	9420
aagctgaaaa	ataagggtcg	ttgcaacgag	atagcagttt	aatgggtggg	acaacgttct	9480
ctgcctacct	agtgggtcaac	atgtctagta	atagtataac	ggacattcag	gacaaactct	9540
gggtatcttc	tctagtggga	cctgagtcca	tacctgatgt	gcgggggtct	acatagggta	9600
cacgacttct	gtacctcctt	acccttcca	agcaccctg	ttctctattt	tgtctagata	9660
ggaaatcttc	ccttaacctt	cttaaatcgt	ggacgactcg	ttaggatagt	tcagccgtct	9720
acatatccaa	aagatatacc	tctgaaccgc	atatctttta	gatgagtacg	gctcctgtca	9780
agagataaag	gagatagata	tgttccagca	taatctccag	ctccaaagaa	ttttcccaac	9840
gatctgccta	attactctcg	ttcaacgacg	gttcattatg	tggcctcttc	agaccgagta	9900
aacttctccg	gccggttgcg	tcacatgcct	ccaaactaaa	tgaactaact	atttaactca	9960
catagtggag	gtaaggaaag	agaatgatct	agtcctggat	aatctctgct	taatctttgc	10020
taaggggtgt	tctaggggtg	gaggataggc	tgttcgttgg	cactataccc	ccactaacag	10080
tctttaatga	agtttatggt	tacggcagat	taacttttcc	cttttatgtc	tagtgtaata	10140
agtgttaata	ccaataagag	tctacagaat	aggatatctg	agtaacctgg	taagagataa	10200
agggtgggtg	agaacgttta	ggatatgttc	ggtaaaaata	gaccctttct	attcttactc	10260
aactctctcg	accgtttaga	aagaagtaac	gattctagtc	ctctccccac	ccttctgtat	10320
gtacacttta	agaagtgggt	cctgtataat	aacacaggtc	tccttttagtc	tgtacgaacg	10380
ttcaagccct	aacgattcct	attattattt	ctgtactcga	tagggggaac	cccttccctt	10440
aggctctcct	gttaatgttg	ttagggacaa	ataatatgct	ggtggggaat	gggtttctac	10500
gatctctacg	gaggttctta	ggttttaggg	gacgacaggc	cttagtccaa	cccgggtaat	10560
ggttgaccgc	gagtaatatt	ttaagcctca	tataatgtac	cttaccctta	ggtaatgtcc	10620
ctgaagaact	caacacctct	gccgaggcct	ccctactgac	gacgtaatga	tgctctttta	10680
cacgtatcgt	ctccttataa	gttatcagac	aatcttaata	gtcccagtc	gtacgctccg	10740
cggagaggac	tcgggggggtc	acgggatcct	tgaatccctc	ctctatttag	ctctacacat	10800
ttaccacttt	gtacaacctt	tataggtaga	ctgaatacac	tgggttccctg	aaccttgata	10860
aaggaggctg	agtttcgtcc	gaaccccgaa	gtttaactaa	attaacatta	cctataacctt	10920
caagccctaa	gaagatgac	ggacttttaa	ctctgcttac	aatctttaat	acacgtggcc	10980
taaaacctac	tcgttcctca	aaattagatg	ttctgaatac	cttgatatata	aacactctcg	11040
cttttcttac	gtcattgtta	ggaaccagg	tacaagtctt	gccagctgaa	tcaagtttgt	11100
cttaaatcat	caagagtttg	cagacttcat	atataccata	catttccaaa	cttctttaat	11160
tagctacttg	ggttagggct	aaccagaagg	tagttactta	ggaccttttt	ggacatgcgt	11220
aaggtcagta	gtcttgctct	taaacggtct	cgtttcttcc	aatcatgtat	gaaatggaac	11280
tgtccataag	ggagggttaa	gtaaggacta	ggaaaacatt	tgtaaactctg	atacgtgtgt	11340
tataagcctc	atgggtgccc	acacagagta	cgccgacgga	atttttagtag	actatctgga	11400
cgtctaaata	actggtaatc	ggaaaaaata	taccgctaata	atagcataat	attgtagtta	11460
gtatagtctc	atcctggcta	tggaggcttg	gggggtagtc	taccttaacg	tgttttacac	11520

		ccctagcgat attgaccata ttcgaaaacc gactcaaact acctctttct gtaaggtgat 11580 atagtgtgca caaatcgta ataggtcgtt agtaagggt aatccaccct ccgacaaagt 11640 cattttcctc ctatgttcgt cttcacctca tgatctccac taccggaggg ttttctatgg 11700 gcttaaagtc tgaggaaccg gggtagccc ttgacctagt ctagagacct taaccaggct 11760 ttggttcaag cagatttagg taagtactc tagaacaagt tagtcgatac agcatgtcac 11820 ctattagtaa actttaccag tttaaagcgt tctttgtgct cttactaact tacctagtta 11880 tctgcttaaa gttttcttct ggccagatat gactacaact tctcactgga tgtgctcctt 11940 ttgagaacct ctctaatttt ttagtactcc tctgaggttt gaaattcata ctttttttga 12000 aactaggaat tctgggagaa caccaaaaat aaaaaataga ccaaacacc agaagca 12057
3-31	Sequences	
3-31-1	Sequence Number [ID]	31
3-31-2	Molecule Type	AA
3-31-3	Length	446
3-31-4	Features Location/ Qualifiers	REGION 1..446 note=heavy chain of anti-PD1 antibody PD1-1 source 1..446 mol_type=protein organism=synthetic construct
3-31-5	NonEnglishQualifier Value Residues	EVMLVESGGG LVQPGGSLRL SCTASGFTFS ASAMSWVRQA PGKGLEWVAY ISGGGGDTYY 60 SSSVKGRFTI SRDANKSLY LQMNSLRAED TAVYYCARHS NVNYYAMDYW GQGLTLTVSS 120 ASTKGPSVFP LAPCSRSTSE STAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS 180 GLYSLSSVVT VPSSSLGTKT YTCNVDHKPS NTKVDKRVES KYGPPCPPCP APEFLGGPSV 240 FLFPPKPKDT LMISRTPEVT CVVVDVSQED PEVQFNWYVD GVEVHNAKTK PREEQFNSTY 300 RVVSVLTVLH QDWLNGKEYK CKVSNKGLPS SIEKISKAK GQPREPQVYT LPPSQEEMTK 360 NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTPPVLDs DGSFFLYSRL TVDKSRWQEG 420 NVFSCSVME ALHNHYTQKS LSLSLG 446
3-32	Sequences	
3-32-1	Sequence Number [ID]	32
3-32-2	Molecule Type	AA
3-32-3	Length	218
3-32-4	Features Location/ Qualifiers	REGION 1..218 note=light chain of anti-PD1 antibody PD1-1 source 1..218 mol_type=protein organism=synthetic construct
3-32-5	NonEnglishQualifier Value Residues	EIVLTQSPAT LSLSPGERAT MSCRASENID TSGISFMNWX QQKPGQAPKL LIYVASNQGS 60 GIPARFSGSG SGTDFTLTIS RLEPEDFAVY YCQSQKEVPW TFGQGTKLEI KRTVAAPSVF 120 IFPPSDEQLK SGTASVCLL NNFYPREAKV QWKVDNALQS GNSQESVTEQ DSKDSTYSL 180 STLTLSKADY EKHKVYACEV THQGLSPVT KSFNRGEC 218
3-33	Sequences	
3-33-1	Sequence Number [ID]	33
3-33-2	Molecule Type	AA
3-33-3	Length	446
3-33-4	Features Location/ Qualifiers	REGION 1..446 note=heavy chain of anti-PD1 antibody PD1-2 source 1..446 mol_type=protein organism=synthetic construct
3-33-5	NonEnglishQualifier Value Residues	EVMLVESGGG LVQPGGSLRL SCTASGFTFS ASAMSWVRQA PGKGLEWVAY ISGGGGDTYY 60 SSSVKGRFTI SRDANKSLY LQMNSLRAED TAVYYCARHS NPNYYAMDYW GQGLTLTVSS 120 ASTKGPSVFP LAPCSRSTSE STAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS 180 GLYSLSSVVT VPSSSLGTKT YTCNVDHKPS NTKVDKRVES KYGPPCPPCP APEFLGGPSV 240 FLFPPKPKDT LMISRTPEVT CVVVDVSQED PEVQFNWYVD GVEVHNAKTK PREEQFNSTY 300 RVVSVLTVLH QDWLNGKEYK CKVSNKGLPS SIEKISKAK GQPREPQVYT LPPSQEEMTK 360 NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTPPVLDs DGSFFLYSRL TVDKSRWQEG 420 NVFSCSVME ALHNHYTQKS LSLSLG 446
3-34	Sequences	
3-34-1	Sequence Number [ID]	34
3-34-2	Molecule Type	AA
3-34-3	Length	218
3-34-4	Features Location/ Qualifiers	REGION 1..218 note=light chain of anti-PD1 antibody PD1-2

3-34-5	NonEnglishQualifier Value Residues	source 1..218 mol_type=protein organism=synthetic construct EIVLTQSPAT LSLSPGERAT MSCRASENID TSGISFMNWy QQKPGQAPKL LIYVASNQGS 60 GIPARFSGSG SGTDFTLTIS RLEPEDFAVY YCQQSKEVPW TFGQGTKLEI KRTVAAPSVF 120 IFPPSDEQLK SGTASVVCLL NNFYPREAKV QWKVDNALQS GNSQESVTEQ DSKDSTYSLs 180 STLTLSKADY EKHKVYACEV THQGLSSPVT KSFNRGEC 218
3-35 3-35-1 3-35-2 3-35-3 3-35-4	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers	35 AA 446 REGION 1..446 note=heavy chain of anti-PD1 antibody PD1-3 source 1..446 mol_type=protein organism=synthetic construct
3-35-5	NonEnglishQualifier Value Residues	EVMLVESGGG LVQPGGSLRL SCTASGFTFS KSAMSWVRQA PGKGLEWVAY ISGGGGDTYY 60 SSSVKGRFTI SRDnaknSLY LQMNSLRAED TAVYYCARHS NVNYYAMDyW GQGTlVTVSS 120 ASTKGPSVFP LAPCSRSTSE STAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS 180 GLYSLSSVVT VPSSSLGTKT YTCNVdHKPS NTKVDKRVES KYGPPCPPCP APEFLGGPSV 240 FLFPPKPKDT LMISRTPEVT CVVVDVSQED PEVQFNWYVD GVEVHNAKTK PREEQFNSTY 300 RVVSVLTVLH QDWLNGKEYK CKVSNKGLPS SIEKTIKAK GQPREPQVYT LPPSQEEMTK 360 NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTTPVLDs DGSFFLYSRL TVDKSRWQEG 420 NVFSCSVME ALHnHYTQKS LSLSLG 446
3-36 3-36-1 3-36-2 3-36-3 3-36-4	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers	36 AA 218 REGION 1..218 note=light chain of anti-PD1 antibody PD1-3 source 1..218 mol_type=protein organism=synthetic construct
3-36-5	NonEnglishQualifier Value Residues	EIVLTQSPAT LSLSPGERAT MSCRASENID VSGISFMNWy QQKPGQAPKL LIYVASNQGS 60 GIPARFSGSG SGTDFTLTIS RLEPEDFAVY YCQQSKEVPW TFGQGTKLEI KRTVAAPSVF 120 IFPPSDEQLK SGTASVVCLL NNFYPREAKV QWKVDNALQS GNSQESVTEQ DSKDSTYSLs 180 STLTLSKADY EKHKVYACEV THQGLSSPVT KSFNRGEC 218
3-37 3-37-1 3-37-2 3-37-3 3-37-4	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers	37 DNA 69 source 1..69 mol_type=other DNA organism=Homo sapiens
3-37-5	NonEnglishQualifier Value Residues	gtggccgctg ccagacctgt tcttcctgct cctgccatca gcgactggga aagagctaga 60 agaaggcac 69
3-38 3-38-1 3-38-2 3-38-3 3-38-4	Sequences Sequence Number [ID] Molecule Type Length Features Location/ Qualifiers	38 DNA 114 source 1..114 mol_type=other DNA organism=Homo sapiens
3-38-5	NonEnglishQualifier Value Residues	gtcggccgga aagaaggcgc cgctagagct agagttagcc tgctgccaga gtttggcacc 60 gccgaagtgc atctgcctgc tcttagcgct gtcagagccg ccagaccagg attt 114
3-39 3-39-1 3-39-2 3-39-3	Sequences Sequence Number [ID] Molecule Type Length	39 DNA 66

3-39-4	Features Location/ Qualifiers	source 1..66 mol_type=other DNA organism=Homo sapiens
3-39-5	NonEnglishQualifier Value Residues	ggcggagata agggcagact cgtgggcgtt gcagaaagac aggtgccctg cagattcctg 60 aggcca 66
3-40	Sequences	
3-40-1	Sequence Number [ID]	40
3-40-2	Molecule Type	DNA
3-40-3	Length	114
3-40-4	Features Location/ Qualifiers	misc_feature 1..114 note=CEACAM5 fragment source 1..114 mol_type=other DNA organism=synthetic construct
3-40-5	NonEnglishQualifier Value Residues	aaccggacac tgaccctgtt caacgtgacc cggaatgacg ccagagccta cgtgtccggc 60 atccagaata gcgtgtccgc caatagaagc gaccccgatga cactggacgt gctg 114
3-41	Sequences	
3-41-1	Sequence Number [ID]	41
3-41-2	Molecule Type	DNA
3-41-3	Length	57
3-41-4	Features Location/ Qualifiers	misc_feature 1..57 note=CEACAM5 fragment source 1..57 mol_type=other DNA organism=synthetic construct
3-41-5	NonEnglishQualifier Value Residues	cctgatagca gctatctgag cggcgccaac ctgaacctga gctgccattc tgcctct 57
3-42	Sequences	
3-42-1	Sequence Number [ID]	42
3-42-2	Molecule Type	DNA
3-42-3	Length	189
3-42-4	Features Location/ Qualifiers	misc_feature 1..189 note=CEACAM5 fragment source 1..189 mol_type=other DNA organism=synthetic construct
3-42-5	NonEnglishQualifier Value Residues	ccacagtaca gctggcggat caacggaatc cctcagcagc acaccaggt gctgtttatc 60 gccaagatca ccctaacaa caacggcacc tacgcctgct tcgtgtccaa tctggccacc 120 ggcagaaaca acagcatcgt gaagtcacatc accgtgtctg ccagcggcac atctcctgga 180 ctgagcgcc 189
3-43	Sequences	
3-43-1	Sequence Number [ID]	43
3-43-2	Molecule Type	DNA
3-43-3	Length	66
3-43-4	Features Location/ Qualifiers	misc_feature 1..66 note=DUOXA2 fragment source 1..66 mol_type=other DNA organism=synthetic construct
3-43-5	NonEnglishQualifier Value Residues	ggcggagata agggcagact cgtgggcgtt gcagaaagac aggtgccctg cagattcctg 60 aggcca 66
3-44	Sequences	
3-44-1	Sequence Number [ID]	44
3-44-2	Molecule Type	DNA
3-44-3	Length	72
3-44-4	Features Location/ Qualifiers	misc_feature 1..72 note=KRAS-G12D fragment source 1..72 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	

3-44-5	Residues	acagagtaca agctggtggt tgtgggagct gacggcgtgg gaaaaagcgc cctgacaatc 60 cagctgatcc ag 72
3-45	Sequences	
3-45-1	Sequence Number [ID]	45
3-45-2	Molecule Type	DNA
3-45-3	Length	72
3-45-4	Features Location/ Qualifiers	misc_feature 1..72 note=KRAS-G12V fragment source 1..72 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-45-5	Residues	accgagtata aactcgtggt cgtcggcgct gttggcgctg gcaaactctgc tctgaccatt 60 cagctcattc ag 72
3-46	Sequences	
3-46-1	Sequence Number [ID]	46
3-46-2	Molecule Type	AA
3-46-3	Length	337
3-46-4	Features Location/ Qualifiers	REGION 1..337 note=ATP150 (v2) source 1..337 mol_type=protein organism=synthetic construct
	NonEnglishQualifier Value	
3-46-5	Residues	KRYKNRVASR KSRAKFKQLL QHYREVAALK SSENDRLRL LKNRTLTLFN VTRNDARAYV 60 SGIQNSVSAN RSDPVTLDVL PDSSYLSGAN LNLSCHSASP QYSWRINGIP QQHTQVLFIA 120 KITPNNGTY ACFVSNLATG RNNSIVKSIT VSASGTSPGL SATEYKLVV GADGVGKSAL 180 TIQLIQVAAA RPVLPAPAI DWERARRRHV GRKEGAARAR VSLLEPFGTA EVHLPAPSAV 240 RAARPGFGGD KGRLVGVAER QVPCRFLRPT GVLSLFLGGA VVSLQYVRPS ALRTLDDQSA 300 KDSTVHEILS KLSLEGDHST PPSAYGSVKP YTNFDAE 337
3-47	Sequences	
3-47-1	Sequence Number [ID]	47
3-47-2	Molecule Type	AA
3-47-3	Length	337
3-47-4	Features Location/ Qualifiers	REGION 1..337 note=ATP152 (v2) source 1..337 mol_type=protein organism=synthetic construct
	NonEnglishQualifier Value	
3-47-5	Residues	KRYKNRVASR KSRAKFKQLL QHYREVAALK SSENDRLRL LKNRTLTLFN VTRNDARAYV 60 SGIQNSVSAN RSDPVTLDVL PDSSYLSGAN LNLSCHSASP QYSWRINGIP QQHTQVLFIA 120 KITPNNGTY ACFVSNLATG RNNSIVKSIT VSASGTSPGL SATEYKLVV GAVGVGKSAL 180 TIQLIQVAAA RPVLPAPAI DWERARRRHV GRKEGAARAR VSLLEPFGTA EVHLPAPSAV 240 RAARPGFGGD KGRLVGVAER QVPCRFLRPT GVLSLFLGGA VVSLQYVRPS ALRTLDDQSA 300 KDSTVHEILS KLSLEGDHST PPSAYGSVKP YTNFDAE 337
3-48	Sequences	
3-48-1	Sequence Number [ID]	48
3-48-2	Molecule Type	AA
3-48-3	Length	284
3-48-4	Features Location/ Qualifiers	REGION 1..284 note=multi-antigenic domain VSV-GP154 (v2) source 1..284 mol_type=protein organism=synthetic construct
	NonEnglishQualifier Value	
3-48-5	Residues	NRTLTLFNVT RNDARAYVSG IQNSVSANRS DPVTLDVLPD SSYLSGANLN LSCHSASPQY 60 SWRINGIPQQ HTQVLFIAKI TPNNGTYAC FVSNLATGRN NSIVKSITVS ASGTSPGLSA 120 TEYKLVVGA DVGKSAITL QLIQTEYKLV VVGAVGVGKS ALTIQLIQVA AARPVLPAPA 180 ISDWERARRR HVGRKEGAAR ARVSLLEPFG TAEVHLPAPS AVRAARPGFG GDKGRLVGVA 240 ERQVPCRFLR PTGVLSLFLG GAVVSLQYVR PSALRTLDDQ SAKD 284
3-49	Sequences	
3-49-1	Sequence Number [ID]	49
3-49-2	Molecule Type	DNA
3-49-3	Length	12057

3-49-4	Features Location/ Qualifiers	misc_feature 1..12057 note=cDNA sequence of VSV-GP154 source 1..12057 mol_type=other DNA organism=synthetic construct
3-49-5	NonEnglishQualifier Value Residues	acgaagacaa acaaaccatt attatcatta aaaggctcag gagaaacttt aacagtaatc 60 aaaatgtctg ttacagtcaa gagaatcatt gacaacacag tcgtagttcc aaaacttcct 120 gcaaatgagg atccagtgga ataccggca gattacttca gaaaatcaaa ggagattcct 180 ctttacatca atactacaaa aagtttgtca gatctaagag gatatgtcta ccaaggcctc 240 aaatccggaa atgtatcaat catacatgtc aacagctact tgtatggagc attaaaggac 300 atccggggta agttggataa agatttgtca agtttcggaa taaacatcgg gaaagcaggg 360 gatacaatcg gaatatttga ccttgtatcc ttgaaagccc tggacggcgt acttcagat 420 ggagtatcgg atgcttcag aaccagcgca gatgacaaat ggttgctttt gtatctactt 480 ggcttataca gagtggcgag aacacaaatg cctgaataca gaaaaaagct catggatggg 540 ctgacaaatc aatgcaaaat gatcaatgaa cagtttgaac ctcttgtgcc agaaggtcgt 600 gacatttttg atgtgtgggg aaatgacagt aattacacaa aaattgtcgc tgcagtggac 660 atgttcttcc acatgttcaa aaaacatgaa tgtgcctcgt tcagatacgg aactattgtt 720 tccagattca aagattgtgc tgcattggca acatttggac acctctgcaa aataaccgga 780 atgtctacag aagatgtaac gacctggatc ttgaaccgag aagttgcaga tgaatgggtc 840 caaatgatgc ttccaggcca agaaattgac aaggccgatt catacatgcc ttatttgatc 900 gactttggat tgtcttctaa gtctccatat tcttccgtca aaaaccctgc cttccacttc 960 tgggggcaat tgacagctct tctgtcaga tccaccagag caaggaaatgc ccgacagcct 1020 gatgacattg agtatacatc tcttactaca gcaggtttgt tgtacgctta tgcagtagga 1080 tcctctgccg acttggcaca acagttttgt gttggagata acaaatcac tccagatgat 1140 agtaccggag gattgacgac taatgcaccg ccacaaggca gagatgtggt cgaatggctc 1200 ggatggtttg aagatcaaaa cagaaaaccg actcctgata tgatgcagta tgcgaaaaga 1260 gcagtcatgt cactgcaagg cctaagagag aagacaattg gcaagtatgc taagtcagaa 1320 tttgacaaat gaccctataa ttctcagatc acctattata tattatgcta catatgaaaa 1380 aaactaacag atatcatgga taatctcaca aaagtctcgt agtatctcaa gtcctattct 1440 cgctctggatc aggcggtagg agagatagat gagatcgaag cacaacgagc tgaagggtcc 1500 aattatgagt tgttccaaga ggatggagtg gaagagcata ctaagccctc ttattttcag 1560 gcagcagatg attctgacac agaattctgaa ccagaaattg aagacaatca aggcttgtat 1620 gcaccagatc cagaagctga gcaagttgaa ggctttatac agggggccttt agatgactat 1680 gcagatgagg aagtggatgt tgtatttact tcggactgga aacagcctga gcttgaatct 1740 gacgagcatg gaaagacctt acggttgaca tcgccagagg gtttaagtgg agagcagaaa 1800 tccagtggtc tttcgacgat taaagcagtc gtgcaaagtg ccaaatactg gaatctggca 1860 gagtgcacat ttgaagcatc gggagaaggg gtcattatga aggagcgcca gataactccg 1920 gatgtatata aggtcactcc agtgatgaac acacatccgt cccaatcaga agcagtatca 1980 gatgtttggt ctctctcaaa gacatccatg actttccaac ccaagaaagc aagtcttcag 2040 cctctcacca tatccttggg tgaattgttc tcatctagag gagagtcat ctctgtcgga 2100 ggtgacggac gaatgtctca taaagaggcc atcctgctcg gcctgagata caaaaagttg 2160 tacaatcagg cgagagtcaa atattctctg tagactatga aaaaaagtaa cagatatcac 2220 gatctaagtg ttatcccaat ccattcatca tgagttcctt aaagaagatt ctcggtctga 2280 aggggaaagg taagaaatct aagaaattag ggatcgacc acccccttat gaagaggaca 2340 ctgacatgga gtatgctccg agcgctccaa ttgacaaatc ctattttgga gttgacgaga 2400 tggacaccta tgatccgaat caattaagat atgagaaatt ctcttttaca gtgaaaatga 2460 cggttagatc taatcgctcc ttcagaacat actcagatgt ggcagccgct gtatcccatt 2520 gggatcacat gtacatcgga atggcaggga aacgtccctt ctacaaaatc ttggcttttt 2580 tgggttcttc taatctaaag gccactccag cggtattggc agatcaagggt caaccagagt 2640 atcacgctca ctgcaaggc agggcttatt tgccacatag gatggggaag acccctccca 2700 tgcctaagt accagagcac ttcagaagac cattcaatat aggtctttac aagggaacga 2760 ttgagctcac aatgaccatc tacgatgatg agtactgga agcagctcct atgatctggg 2820 atcattttcaa ttcttccaaa ttttctgatt tcagagagaa ggccttaatg tttggcctga 2880 ttgtcgagaa aaaggcatct ggagcgtggg tcctggactc tatcgccac ttcaaatgag 2940 ctagtctaac ttctagcttc tgaacaatcc ccggtttact cagtctcccc taattccagc 3000 ctctcgaaca actaatatcc tgtcttttct atccctatga aaaaaactaa cagagatcga 3060 tctgtttacg cgtcacggat cccccgggt gcaggaaatc gccaccatgg gccagatcgt 3120 gaccatgttc gaggccctgc ccacatcat cgacgagggt atcaacatcg tgatcatcgt 3180 gctcatcatc atcaccagca tcaaggccgt gtacaacttc gccacctgcg gcatcctggc 3240 cctggtgagc ttcctgttcc tggccggcag aagctgcggc atgtacggcc tgaatggccc 3300 cgatatctac aaggcggtgt accagttcaa gagcgtggag ttcgacatga gccacctgaa 3360 cctcaccatg cccaacgcct gcagcgccaa caatagccac cactacatca gcatgggagc 3420 cagcggcctg gatttgacct tcaccaacga cagcatcctg aaccacaact tctgcaacct 3480 caccagcgcc ttcaacaaga aaaccttcga ccacaccctc atgagcatcg tgagcagcct 3540

gcacctgagc	atcagaggca	acagcaacca	caaggccgtg	agctgcgact	tcaacaacgg	3600
catcaccatc	cagtacaacc	tgagcttcag	cgatcctcag	agcgccatca	gccagtgcag	3660
aaccttcaga	ggcagagtgc	tgacatgtt	cagaaccgcc	ttcggcgcca	agtacatgag	3720
aagcggctgg	ggctgggccc	gcagcgacgg	caagaccacc	tggtgcagcc	agaccagcta	3780
ccagtacctc	atcatccaga	acagaacctg	ggagaaccac	tgagatacgc	ccggaccttt	3840
cggcatgagc	agaatcctgt	tcgcccagga	gaaaaccaag	ttcctcacca	ggagactggc	3900
cggcaccttc	acctggaccc	tgagcgacag	cagcggcgtg	gagaaccttg	cgggctactg	3960
cctcaccaag	tgatgatcc	tgccgcgcca	gctgaagtgc	ttcggcaaca	ccgccgtggc	4020
caagtgaac	gtgaaccacg	acgaggagtt	ctgcgacatg	ctgagactca	tcgactacaa	4080
caaggccgcc	ctgagcaagt	tcaagcagga	cgtggagagc	gccctgcacg	tgttcaagac	4140
caccgtgaac	agcctcatca	gcgaccagct	gctcatgaga	aaccacctga	gagacctcat	4200
gggcgtgccc	tactgcaact	acagcaagtt	ctggtatctg	gagcacgcca	agaccggcga	4260
gaccagcgtg	ccaagtgtct	ggctgggtgac	caatggcagc	tacctgaacg	agaccacttt	4320
cagcgaccag	atcgagcagg	aagccgacaa	catgatcacc	gagatgctga	ggaaggacta	4380
catcaagaga	cagggcagca	ccccctggc	cctcatggat	ctgctcatgt	tcagaccag	4440
cgctacctc	atcagcatct	tcctgcacct	ggtgaagatc	cccaccacaa	gacacatcaa	4500
gggcggcagc	tgcccaagc	cccacagact	caccaaccaag	ggcatctgca	gctgcggcgc	4560
cttcaaggtg	cccggcgtga	aaacctctg	gaagaggaga	taagcggccc	ctacgacctc	4620
gactatgaaa	aaaactaaca	gatatcctcg	acgccaccat	gaaccggaca	ctgacctgt	4680
tcaacgtgac	ccggaatgac	gccagagcct	acgtgtccgg	catccagaat	agcgtgtccg	4740
ccaatagaag	cgaccccggt	acactggacg	tgctgcctga	tagcagctat	ctgagcggcg	4800
ccaacctgaa	cctgagctgc	cattctgcct	ctccacagta	cagctggcgg	atcaacggaa	4860
tcctcagca	gcacaccag	gtgctgttta	tcgccaagat	caccctaacc	aacaacggca	4920
cctacgcctg	cttcgtgtcc	aatctggcca	ccggcagaaa	caacagcatc	gtgaagtcca	4980
tcaccgtgtc	tgccagcggc	acatctcctg	gactgagcgc	cacagagtac	aagctggtgg	5040
ttgtgggagc	tgacggcgtg	ggaaaaagcg	ccctgacaat	ccagctgacg	cagaccgagt	5100
ataaactcgt	ggtcgtcggc	gctgttggcg	ttggcaaatc	tgctctgacc	attcagctca	5160
ttcaggtggc	cgctgccaga	cctgttcttc	ctgctcctgc	catcagcgac	tgggaaagag	5220
ctagaagaag	gcacgtcggc	cggaaagaag	gcgccgctag	agctagagtt	agcctgctgc	5280
cagagtttgg	caccgccgaa	gttcatctgc	ctgctcctag	cgctgtcaga	gccgcagac	5340
caggatttgg	cggagataag	ggcagactcg	tgggcgttgc	agaaagacag	gtgccttgca	5400
gattcctgag	gccaacaggc	gtgctgagcc	tgtttcttgg	cggagctgtg	gtgtccctgc	5460
agtacgtcag	acctagcgct	ctgagaaccc	tgctggatca	gagcgccaag	gattgatcaa	5520
ctagccagat	tcttcatgtt	tggaccaaat	caacttgtga	taccatgctc	aaagaggcct	5580
caattatatt	tgagttttta	atttttatga	aaaaaactaa	cagcaatcat	ggaagtccac	5640
gattttgaga	ccgacgagtt	caatgatttc	aatgaagatg	actatgccac	aagagaattc	5700
ctgaatcccg	atgagcgcat	gacgtacttg	aatcatgctg	attacaacct	gaattctcct	5760
ctaattagtg	atgatattga	caattttaatc	aggaaattca	attctcttcc	aattccctcg	5820
atgtgggata	gtaagaactg	ggatggagtt	cttgagatgt	taacgtcatg	tcaagccaat	5880
cccatcccaa	catctcagat	gcataaatgg	atgggaagtt	ggttaatgtc	tgataatcat	5940
gatgccagtc	aagggtatag	ttttttacat	gaagtggaca	aagaggcaga	aataacattt	6000
gacgtggtgg	agaccttcat	ccgcggctgg	ggcaacaaac	caattgaata	catcaaaaag	6060
gaaagatgga	ctgactcatt	caaaattctc	gcttatttgt	gtcaaaagtt	tttgacttta	6120
cacaagtgtg	cattaatctt	aaatgctgtc	tctgaggtgg	aattgtctaa	cttgcgagg	6180
actttcaag	gcaaagtcag	aagaagttct	catggaacga	acatatgcag	gattaggggt	6240
cccagcttgg	gtcctacttt	tatttcagaa	ggatgggctt	acttcaagaa	acttgatatt	6300
ctaattggacc	gaaactttct	gttaatggtc	aaagatgtga	ttatagggag	gatgcaaacg	6360
gtgctatcca	tggtatgtag	aatagacaac	ctgttctcag	agcaagacat	cttctccctt	6420
ctaaatatct	acagaattgg	agataaaaatt	gtggagaggg	agggaaattt	ttcttatgac	6480
ttgattaaaa	tggtggaacc	gatatgcaac	ttgaagctga	tgaattagc	aagagaatca	6540
aggcctttag	tcccacaatt	ccctcatttt	gaaaatcata	tcaagacttc	tgttgatgaa	6600
ggggcaaaaa	ttgaccgagg	tataagattc	ctccatgatc	agataatgag	tgtgaaaaca	6660
gtggatctca	cactggtgat	ttatggatcg	ttcagacatt	ggggtcatcc	ttttatagat	6720
tattacactg	gactagaaaa	attacattcc	caagtaacca	tgaagaaaga	tattgatgtg	6780
tcatatgcaa	aagcacttgc	aagtgattta	gctcggattg	ttctatttca	acagttcaat	6840
gatcataaaa	agtggttcgt	gaatggagac	ttgctccctc	atgatcatcc	ctttaaaagt	6900
catgttaaag	aaaatacatg	gcccacagct	gctcaagttc	aagatttttg	agataaatgg	6960
catgaacttc	cgctgattaa	atgttttgaa	ataccgact	tactagaccc	atcgataata	7020
tactctgaca	aaagtcattc	aatgaatagg	tcagaggtgt	tgaacatgt	ccgaatgaat	7080
ccgaacactc	ctatccctag	taaaaagggt	ttgcagacta	tgttgacac	aaaggctacc	7140
aattggaaag	aatttcttaa	agagattgat	gagaagggtc	tagatgatga	tgatctaatt	7200
attggtctta	aaggaaagga	gagggaactg	aagttggcag	gtagattttt	ctccctaatt	7260
tcttggaat	tcgagaata	ctttgttaatt	accgaatatt	tgataaagac	tcatttcgtc	7320
cctatgttta	aaggcctgac	aatggcggac	gatctaactg	cagtcattaa	aaagatgtta	7380
gattcctcat	ccggccaagg	attgaagtca	tatgaggcaa	tttgcatagc	caatcacatt	7440

gattacgaaa	aatggaataa	ccaccaaaagg	aagttatcaa	acggcccagtt	gttcccgagtt	7500
atggggccagtt	tcttaggtta	tccatcctta	atcgagagaa	ctcatgaatt	ttttgagaaa	7560
agtcttatat	actacaatgg	aagaccagac	ttgatgcgtg	ttcacaca	cacactgac	7620
aattcaacct	ccaacgagt	ttgttgcaa	ggacaagagg	gtggactgga	aggtctacgg	7680
caaaaaggat	ggagtatcct	caatctactg	gttattcaaa	gagaggctaa	aatcagaaac	7740
actgctgtca	aagtcttggc	acaagggtgat	aatcaagtta	tttgcacaca	gtataaaacg	7800
aagaaatcga	gaaacgttgt	agaattacag	ggtgctctca	atcaaatggt	ttctaataat	7860
gagaaaatta	tgactgcaat	caaaataggg	acaggggaagt	taggactttt	gataaatgac	7920
gatgagacta	tgcaatctgc	agattacttg	aattatggaa	aaataccgat	tttccgtgga	7980
gtgattagag	ggttagagac	caagagatgg	tcacgagtga	cttgtgtcac	caatgaccaa	8040
ataccactt	gtgctaatat	aatgagctca	gtttccacaa	atgctctcac	cgtagctcat	8100
tttgctgaga	accaatcaa	tgccatgata	cagtacaatt	attttgggac	atttgctaga	8160
ctcttggtga	tgatgcatga	tctgctctt	cgtcaatcat	tgtatgaagt	tcaagataag	8220
ataccgggct	tgacagttc	tactttcaaa	tacgccatgt	tgtatttggga	cccttccatt	8280
ggaggagtg	cgggcatgtc	tttgtccagg	tttttgatta	gagccttccc	agatcccgta	8340
acagaaagtc	tctcattctg	gagattcatc	catgtacatg	ctcgaagtga	gcatctgaag	8400
gagatgagtg	cagtatttgg	aaaccccgag	atagccaagt	ttcgaataac	tcacatagac	8460
aagctagtag	aagatccaac	ctctctgaac	atcgctatgg	gaatgagtcc	agcgaacttg	8520
ttaaagactg	aggttaaaaa	atgcttaatc	gaatcaagac	aaaccatcag	gaaccagggtg	8580
attaaggatg	caaccatata	tttgtatcat	gaagaggatc	ggctcagaag	tttcttatgg	8640
tcaataaatc	ctctgttccc	tagattttta	agtgaattca	aatcaggcac	ttttttggga	8700
gtcgcagacg	ggctcatcag	tctatttcaa	aattctcgta	ctattcgga	ctcctttaag	8760
aaaaagtatc	atagggaatt	ggatgatttg	attgtgagga	gtgagggtatc	ctctttgaca	8820
catttaggga	aacttcattt	gagaagggga	tcattgtaaaa	tgtggacatg	ttcagctact	8880
catgctgaca	cattaagata	caaactcctgg	ggccgtacag	ttattgggac	aactgtaccc	8940
catccattag	aaatgttggg	tccacaacat	cgaaaagaga	ctccttgtgc	accatgtaac	9000
acatcagggt	tcaattatgt	ttctgtgcat	tgtccagacg	ggatccatga	cgtctttagt	9060
tcacggggac	cattgcctgc	ttatctaggg	tctaaaacat	ctgaatctac	atctattttg	9120
cagccttggg	aaaggggaaag	caaagtcca	ctgattaaaa	gagctacacg	tcttagagat	9180
gctatctctt	ggtttgttga	accgactct	aaactagcaa	tgactatact	ttctaacatc	9240
cactcttta	caggcgaaaga	atggaccaaa	aggcagcatg	ggttcaaaag	aacagggtct	9300
gcccttcata	ggttttcgac	atctcggatg	agccatgggtg	ggttcgcac	tcagagcact	9360
gcagcattga	ccagggtgat	ggcaactaca	gacaccatga	gggatctggg	agatcagaat	9420
ttcgactttt	tattccaagc	aacggttctc	tatgctcaaa	ttaccaccac	tgttgcaaga	9480
gacggatgga	tcaccagttg	tacagatcat	tatcatattg	cctgtaagtc	ctgtttgaga	9540
cccatagaag	agatcacctt	ggactcaagt	atggactaca	cgccccaga	tgtatcccat	9600
gtgctgaaga	catggaggaa	tggggaagg	tcgtggggac	aagagataaa	acagatctat	9660
cctttagaag	ggaattggaa	gaatttagca	cctgctgagc	aatcctatca	agtcggcaga	9720
tgtataggtt	ttctatatgg	agacttggcg	tatagaaaat	ctactcatgc	cgaggacagt	9780
tctctatttc	ctctatctat	acaaggctgt	attagagggtc	gagggtttctt	aaaagggttg	9840
ctagacggat	taatgagagc	aagttgtgtc	caagtaatac	accggagaag	tctggctcat	9900
ttgaagaggc	cgccaacgc	agtgtacgga	ggtttgattt	acttgattga	taaattgagt	9960
gtatcacctc	cattcctttc	tcttactaga	tcaggaccta	ttagagacga	attagaaacg	10020
attccccaca	agatcccaac	ctcctatccg	acaagcaacc	gtgatatggg	ggtgattgtc	10080
agaaattact	tcaaatacca	atgccgtcta	attgaaaagg	gaaaatacag	atcacattat	10140
tcacaattat	ggttattctc	agatgtctta	tccatagact	tcattggacc	attctctatt	10200
tccaccaccc	tcttgcaaat	cctatacaag	ccatttttat	ctgggaaaga	taagaatgag	10260
ttgagagagc	tggaatctt	ttcttcattg	ctaagatcag	gagagggtgtg	ggaagacata	10320
catgtgaaat	tcttcaccaa	ggacatatta	ttgtgtccag	aggaaatcag	acatgcttgc	10380
aagttcggga	ttgctaagga	taataataaa	gacatgagct	atcccccttg	gggaagggaa	10440
tccagaggga	caattacaac	aatccctgtt	tattatacga	ccacccttta	cccaaagatg	10500
ctagagatgc	ctccaagaat	ccaaaatccc	ctgctgtccg	gaatcagggt	gggccaatta	10560
ccaactggcg	ctcattataa	aattcggagt	atattacatg	gaatgggaat	ccattacagg	10620
gacttcttga	gttgtggaga	cggctccgga	gggatgactg	ctgcattact	acgagaaaat	10680
gtgcatagca	gaggaatatt	caatagtctg	ttagaattat	cagggtcagt	catgcgaggc	10740
gcctctcctg	agccccccag	tgccctagaa	actttaggag	gagataaatc	gagatgtgta	10800
aatggtgaaa	catgttggga	atatccatct	gacttatgtg	acccaaggac	ttgggactat	10860
ttcctccgac	tcaaagcagg	cttggggctt	caaattgatt	taattgtaat	ggatatggaa	10920
gttcgggatt	cttctactag	cctgaaaatt	gagacgaatg	ttagaaatta	tgtgcaccgg	10980
attttggatg	agcaaggagt	tttaatctac	aagacttatg	gaacatatat	ttgtgagagc	11040
gaaaagaatg	cagtaacaat	ccttggctcc	atgttcaaga	cggctgactt	agttcaaaca	11100
gaatttagta	gttctcaaac	gtctgaagta	tatatggtat	gtaaagggtt	gaagaaatta	11160
atcgatgaac	ccaatcccga	ttggtcttcc	atcaatgaat	cctggaaaaa	cctgtacgca	11220
ttccagtcac	cagaacagga	atttgccaga	gcaaagaagg	ttagtacata	ctttaccttg	11280
acaggttatc	cctcccaatt	cattcctgat	ccttttgtaa	acattgagac	tatgctacaa	11340

		atattcggag taccacgagg tgtgtctcat gcggctgcct taaaatcatc tgatagacct 11400 gcagatttat tgaccattag ccttttttat atggcgatta tatcgatta taacatcaat 11460 catatcagag taggaccgat acctccgaac ccccatcag atggaattgc aaaaatgtg 11520 gggatcgcta taactgggat aagcttttgg ctgagtttga tggagaaaga cattccacta 11580 tatcaacagt gtttagcagt tatccagcaa tcattcccga ttaggtggga ggctgtttca 11640 gtaaaaggag gatacaagca gaagtggagt actagagggtg atgggctccc aaaagatacc 11700 cgaatttcag actccttggc cccaatcggg aactggatca gatctctgga attggtccga 11760 aaccaagttc gtctaaatcc attcaatgag atcttgttca atcagctatg tcgtacagt 11820 gataatcatt tgaaatggtc aaatttgcga agaaacacag gaatgattga atggatcaat 11880 agacgaattt caaaagaaga ccggtctata ctgatgttga agagtgcct acacgaggaa 11940 aactcttggg gagattaaaa aatcatgagg agactccaaa ctttaagtat gaaaaaact 12000 ttgatcctta agaccctctt gtgggttttta ttttttatct gggtttgtgg tcttcgt 12057
3-50	Sequences	
3-50-1	Sequence Number [ID]	50
3-50-2	Molecule Type	AA
3-50-3	Length	144
3-50-4	Features Location/ Qualifiers	source 1..144 mol_type=protein organism=synthetic construct
3-50-5	NonEnglishQualifier Value Residues	NRTLTLFNVT RNDARAYVSG IQNSVSANRS DPVTLDLVD SSYLSGANLN LSCHSASPQY 60 SWRINGIPQQ HTQVLFIKI TPNNGTYAC FVSNLATGRN NSIVKSITVS ASGTSPGLSA 120 TEYKLVVGA DVGKSLTI QLIQ 144
3-51	Sequences	
3-51-1	Sequence Number [ID]	51
3-51-2	Molecule Type	AA
3-51-3	Length	221
3-51-4	Features Location/ Qualifiers	source 1..221 mol_type=protein organism=synthetic construct
3-51-5	NonEnglishQualifier Value Residues	KRYKNRVASR KSRKFKQLL QHYREVAALK SSENDRLRL LKNRTLTLFN VTRNDARAYV 60 SGIQNSVSAN RSDPVTLDVL PDSSYLSGAN LNLSCHSASP QYSWRINGIP QQHTQVLFI 120 KITPNNGTY ACFVSNLATG RNNSIVKSIT VSASGTSPGL SATEYKLVV GADGVGKSAL 180 TIQLIQSTVH EILSKLSLEG DHSTPPSAYG SVKPYTNFDA E 221