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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
07 September 2018 (07.09.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/158455 A1

(51) International Patent Classification:
G01N 33/569 (2006.01) *A61K 38/00* (2006.01)

(21) International Application Number:
PCT/EP2018/055230

(22) International Filing Date:
02 March 2018 (02.03.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

17159243.9	03 March 2017 (03.03.2017)	EP
17159242.1	03 March 2017 (03.03.2017)	EP
1703809.2	09 March 2017 (09.03.2017)	GB

(71) Applicant: **TREOS BIO ZRT.** [HU/HU]; Viola utca 2, 8200 Veszprem (HU).

(72) Inventors: **LISZIEWICZ, Julianna**; Szorosi ut 13, 8220 Balatonalmadi (HU). **MOLNÁR, Levente**; Kossuth Lajos u. 20., 2363 Felsőpakony (HU). **TŐKE, Enikő**; Kossuth Lajos u. 20., 2363 Felsőpakony (HU). **TOTH, József**; Újkut u. 72, 9012 Gyor (HU). **LORINCZ, Orsolya**; Attila u. 27, fszt/13, 1047 Budapest (HU). **CSISZOVSZKI, Zsolt**; Bikszádi u. 12/A, 4/9, 1119 Budapest (HU). **SOMOGYI, Eszter**; Borgazda u. 3, 8220 Balatonalmádi (HU). **PÁNTYA, Katalin**; Üllői út 16/b 2.em. 9/b, 1085 Budapest (HU). **MEGYESI, Mónika**; Táncsics Mihály u. 63, 3441 Mezokeresztes (HU).

(74) Agent: **KIMBLIN, Nicola**; 14 South Square, Gray's Inn, London Greater London WC1R 5JJ (GB).

(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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: PEPTIDE VACCINES

(57) Abstract: The disclosure relates to polypeptides and pharmaceutical compositions comprising polypeptides that find use in the prevention or treatment of cancer, in particular breast cancer, ovarian cancer and colorectal cancer. The disclosure also relates to methods of inducing a cytotoxic T cell response in a subject or treating cancer by administering pharmaceutical compositions comprising the peptides, and companion diagnostic methods of identifying subjects for treatment. The peptides comprise T cell epitopes that are immunogenic in a high percentage of patients.

WO 2018/158455 A1

PEPTIDE VACCINES

Field

The disclosure relates to polypeptides and vaccines that find use in the prevention or treatment of cancer, in particular most breast cancers, ovarian cancers and colorectal cancers.

Background

Cancer is killing millions of people worldwide, because existing drugs do not enable effective prevention or treatment. Current checkpoint inhibitor immunotherapies that re-activate existing immune responses can provide clinical benefit for a fraction of cancer patients. Current cancer vaccines that induce new immune responses are poorly immunogenic and fail to benefit most patients.

Recent analyses of 63,220 unique tumors revealed that cancer vaccines need to be generated specifically for each patient because extensive inter-individual tumor genomic heterogeneity (Hartmaier et al. *Genome Medicine* 2017 **9**:16). Using state of art technologies it is currently not feasible to scale HLA-specific cancer vaccines to large populations.

Summary

In antigen presenting cells (APC) protein antigens are processed into peptides. These peptides bind to human leukocyte antigen molecules (HLAs) and are presented on the cell surface as peptide-HLA complexes to T cells. Different individuals express different HLA molecules and different HLA molecules present different peptides. Therefore, according to the state of the art, a peptide, or a fragment of a larger polypeptide, is identified as immunogenic for a specific human subject if it is presented by a HLA molecule that is expressed by the subject. In other words, the state of the art describes immunogenic peptides as HLA-restricted epitopes. However, HLA restricted epitopes induce T cell responses in only a fraction of individuals who express the HLA molecule. Peptides that activate a T cell response in one individual are inactive in others despite HLA allele matching. Therefore, it was previously unknown how an individual's HLA molecules present the antigen-derived epitopes that positively activate T cell responses.

As provided herein multiple HLAs expressed by an individual need to present the same peptide in order to trigger a T cell response. The fragments of a polypeptide antigen that are immunogenic for a specific individual are those that can bind to multiple class I (activate cytotoxic T cells) or class II (activate helper T cells) HLAs expressed by that individual. For example, the inventors have discovered that the presence of a T cell epitope that binds to at least three HLA type I of a subject predicts an immune response in the subject to a polypeptide.

Based on this discovery the inventors have identified the T cell epitopes from certain breast, ovarian and/or colorectal cancer associated-polypeptide antigens (cancer testis antigens (CTA)) that are capable of binding to at least three class I HLA in a high proportion of individuals. These T cell epitopes, or fragments of the antigens comprising the T cell epitopes, are useful for inducing specific immune responses against tumor cells expressing these antigens and for treating or preventing cancer.

In a first aspect the disclosure provides a polypeptide that comprises a fragment of up to 50 consecutive amino acids of

(a) a colorectal cancer-associated antigen selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, LEMD1, MAGE-A8, MAGE-A6 and MAGE-A3, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 21 to 40 and 234 to 250;

(b) an ovarian cancer-associated antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN, and AKAP-3 wherein the fragment comprises the amino acid sequence of any one of SEQ ID NOs: 272 to 301; and/or

(c) a breast cancer associated antigen selected from PIWIL-2, AKAP-4, EpCAM, BORIS, HIWI, SPAG9, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, PRAME, NY-SAR-35, MAGE-A9, NY-BR-1, SURVIVIN, MAGE-A11, HOM-TES-85 and NY-ESO-1 wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194.

In some specific cases the disclosure provides a polypeptide that

(a) is a fragment of a colorectal cancer-associated antigen selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, MAGE-A6, MAGE-A3 and LEMD1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 21 to 40 and 234 to 250; or

(b) comprises or consists of two or more fragments of one or more colorectal cancer associated antigens selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, MAGE-A6, MAGE-A3 and LEMD1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 21 to 40 and 234 to 250, optionally wherein the fragments overlap or are arranged end to end in the polypeptide; or

(c) is a fragment of a ovarian cancer-associated antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN and AKAP-3, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 272 to 301; or

(d) comprises or consists of two or more fragments of one or more ovarian cancer associated antigens selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN and AKAP-3, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 272 to 301, optionally wherein the fragments overlap or are arranged end to end in the polypeptide; or

(e) is a fragment of a breast cancer associated antigen selected from SPAG9, AKAP-4, BORIS, NY-SAR-35, NY-BR-1, SURVIVIN, MAGE-A11, PRAME, MAGE-A9, HOM-TES-85, PIWIL-2, EpCAM, HIWI, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, wherein the fragment comprises the amino acid sequence from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194; or

(f) comprises or consists of two or more fragments of one or more breast cancer associated antigens selected from SPAG9, AKAP-4, BORIS, NY-SAR-35, NY-BR-1, SURVIVIN, MAGE-A11, PRAME, MAGE-A9, HOM-TES-8, PIWIL-2, EpCAM,

HIWI, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194; optionally wherein the fragments overlap or are arranged end to end in the polypeptide and.

In some specific cases the polypeptide comprises or consists of fragments of

(a) TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, MAGE-A6, MAGE-A3 and LEMD1;

(b) PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN and AKAP-3; and/or

(c) SPAG9, AKAP-4, BORIS, NY-SAR-35, NY-BR-1, SURVIVIN, MAGE-A11, PRAME, MAGE-A9, HOM-TES-8, PIWIL-2, EpCAM, HIWI, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2;

wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 21 to 40 and 234 to 250; SEQ ID NOs: 272 to 301; and/or SEQ ID NOs: 1 to 20, 24 and 172 to 194.

In some cases the polypeptide comprises or consists of one or more amino acid sequences selected from SEQ ID NOs: 41-80, 251 to 271, 302 to 331 and 196 to 233.

In some cases the polypeptide comprises or consists of one or more amino acid sequences selected from SEQ ID NOs: 41-80, 195-233, 251-271 and 302-331 or selected from SEQ ID NOs: 81-142, 332-346, and 435-449.

In a further aspect the disclosure provides a panel of two or more polypeptides as described above, wherein each peptide comprises or consists of a different amino acid sequence selected from SEQ ID NOs: 21 to 40 and 234 to 250; or selected from SEQ ID NOs: 272 to 301; or selected from SEQ ID NOs: 1 to 20, 24 and 172 to 194; or selected from SEQ ID NOs: 1 to 40, 234 to 250, 272 to 301 and 172 to 194. In some cases the panel of polypeptides comprises or consists of one or more peptides comprising or consisting of the amino acid sequences of SEQ ID NOs: 130, 121, 131, 124, 134, 126 and/or SEQ ID NOs: 435-449.

In a further aspect the disclosure provides a pharmaceutical composition or kit having one or more polypeptides or panels of peptides as described above as active ingredients, or having a

polypeptide comprising at least two amino acid sequences selected from SEQ ID NOs: 21 to 40 and 234 to 250; SEQ ID NOs: 272 to 301; and/or SEQ ID NOs: 1 to 20, 24 and 172 to 194 as an active ingredient; or selected from SEQ ID NOs: 130, 121, 131, 124, 134, 126 and/or 435-449 as an active ingredient.

In a further aspect the disclosure provides a method of inducing immune responses, (e.g. vaccination, providing immunotherapy or inducing a cytotoxic T cell response in a subject), the method comprising administering to the subject a pharmaceutical composition, kit or the panel of polypeptides as described above. The method may be a method of treating cancer, such as breast cancer, ovarian cancer or colorectal cancer.

In further aspects, the disclosure provides

- the pharmaceutical composition, kit or panel of polypeptides described above for use in a method of inducing immune responses or for use in a method of treating cancer, optionally breast cancer, ovarian cancer or colorectal cancer; and
- use of a peptide or a panel of peptides as described above in the manufacture of a medicament for inducing immune responses or for treating cancer, optionally breast cancer, ovarian cancer or colorectal cancer.

In a further aspect the disclosure provides a method of identifying a human subject who will likely have a cytotoxic T cell response to administration of a pharmaceutical composition as described above, the method comprising

(i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition comprise a sequence that is a T cell epitope capable of binding to at least three HLA class I of the subject; and

(ii) identifying the subject as likely to have a cytotoxic T cell response to administration of the pharmaceutical composition.

In a further aspect the disclosure provides a method of identifying a subject who will likely have a clinical response to a method of treatment as described above, the method comprising

(i) determining that the active ingredient polypeptide(s) comprise two or more different amino acid sequences each of which is

a. a T cell epitope capable of binding to at least three HLA class I of the subject; and

b. a fragment of a cancer-associated antigen expressed by cancer cells of the subject; and

(ii) identifying the subject as likely to have a clinical response to the method of treatment.

In a further aspect the disclosure provides a method of determining the likelihood that a specific human subject will have a clinical response to a method of treatment according to claim 10, wherein one or more of the following factors corresponds to a higher likelihood of a clinical response:

(a) presence in the active ingredient polypeptide(s) of a higher number of amino acid sequences and/or different amino acid sequences that are each a T cell epitope capable of binding to at least three HLA class I of the subject;

(b) a higher number of target polypeptide antigens, comprising at least one amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject;

optionally wherein the target polypeptide antigens are expressed in the subject, further optionally wherein the target polypeptides antigens are in one or more samples obtained from the subject;

(c) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens and/or optionally target polypeptide antigens that have been determined to comprise at least one amino acid sequence that is both

A. comprised in in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject;

and/or

(d) a higher number of target polypeptide antigens that the subject is predicted to express, optionally a higher number of target polypeptide antigens that the subject expresses with a threshold probability, and/or optionally the target polypeptide antigens that have been determined to comprise at least one amino acid sequence that is both

A. comprised in in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject.

In some cases the cancer-associated antigens may be TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, LEMD1, MAGE-A8, MAGE-A6, MAGE-A3, PIWIL-4, WT1, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, PRAME, HIWI, PLU-1, TSGA10, ODF-4, RHOXF-2, NY-SAR-35, MAGE-A9, NY-BR-1, MAGE-A11, HOM-TES-85, NY-ESO-1 and AKAP-3. In some cases the methods above comprise the step of determining that one or more cancer-associated antigens is expressed by cancer cells of the subject. The cancer-associated antigen(s) may be present in one or more samples obtained from the subject

In some cases administration of the pharmaceutical composition or the active ingredient polypeptides of the kit may then be selected as a method of treatment for the subject. The subject may further be treated by administration of the pharmaceutical composition or the active ingredient polypeptides.

In a further aspect the disclosure provides a method of treatment as described above, wherein the subject has been identified as likely to have a clinical response or as having above a threshold minimum likelihood of having a clinical response to the treatment by the method described above.

In a further aspect the disclosure provides a method of identifying a human subject who will likely not have a clinical response to a method of treatment as described above, the method comprising

(i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition do not comprise two or more different amino acid sequences each of

which is a T cell epitope capable of binding to at least three HLA class I of the subject; and

(ii) identifying the subject as likely not to have a clinical response to the method of treatment.

The methods described above may comprise the step of determining the HLA class I genotype of the subject.

disclosureThe disclosure will now be described in more detail, by way of example and not limitation, and by reference to the accompanying drawings. Many equivalent modifications and variations will be apparent, to those skilled in the art when given this disclosure. Accordingly, the exemplary embodiments of the disclosure set forth are considered to be illustrative and not limiting. Various changes to the described embodiments may be made without departing from the scope of the disclosure. All documents cited herein, whether supra or infra, are expressly incorporated by reference in their entirety.

The present disclosure includes the combination of the aspects and preferred features described except where such a combination is clearly impermissible or is stated to be expressly avoided. As used in this specification and the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to “a peptide” includes two or more such peptides.

Section headings are used herein for convenience only and are not to be construed as limiting in any way.

Description of the Figures

Fig. 1

ROC curve of HLA restricted PEPI biomarkers.

Fig. 2

ROC curve of ≥1 PEPI3+ Test for the determination of the diagnostic accuracy.

Fig. 3

Distribution of HLA class I PEPI3+ compared to CD8+ T cell responses measured by a state of art assay among peptide pools used in the CD8+ T cell response assays. A: HLA class I restricted PEPI3+s. The 90% Overall Percent of Agreement (OPA) among the T cell responses and PEPI3+ peptides demonstrate the utility of the invented peptides for prediction of vaccine induced T cell response set of individuals. B: Class I HLA restricted epitopes (PEPI1+). The OPA between predicted epitopes and CD8+ T cell responses was 28% (not statistically significant). Darkest grey: True positive (TP), both peptide and T cell responses were detected; Light grey: False negative (FN), only T cell responses were detected; Lightest grey: False positive (FP), only peptide were detected; Dark grey: True negative (TN): neither peptides nor T cell responses were detected.

Fig. 4

Distribution of HLA class II PEPIs compared to CD4+ T cell responses measured by a state of art assay among peptide pools used in the assays. A: HLA class II restricted PEPI4+s. 67% OPA between PEPI4+ and CD4+ T-cell responses (p=0.002). B: The class II HLA restricted epitopes. OPA between class II HLA restricted epitopes and CD4+ T cell responses was 66% (not statistically significant). Darkest grey: True positive (TP), both peptide and T cell responses were detected; Light grey: False negative (FN), only T cell responses were detected; Lightest grey: False positive (FP), only peptide were detected; Dark grey: True negative (TN): neither peptides nor T cell responses were detected.

Fig. 5

Multiple HLA binding peptides that define the HPV-16 LPV vaccine specific T cell response set of 18 VIN-3 and 5 cervical cancer patients. HLA class I restricted PEPI3 counts (A and B) and HLA class II restricted PEPI3 counts (C and D) derived from LPV antigens of each patient. Light grey: immune responders measured after vaccination in the clinical trial; Dark grey: Immune non-responders measured after vaccination in the clinical trial. Results show that $\geqq 3$ HLA class I binding peptides predict the CD8+ T cell reactivity and $\geqq 4$ HLA class II binding peptides predict the CD4+ T cell reactivity.

Fig. 6

The multiple HLA class I binding peptides that define the HPV vaccine specific T cell response set of 2 patients. A: Four HPV antigens in the HPV vaccine. Boxes represent the length of the amino acid sequences from the N terminus to the C terminus. B: Process to identify the multiple HLA binding peptides of two patients: HLA sequences of the patients labelled as 4-digit HLA genotype right from the patient's ID. The location of the 1st amino acid of the 54 and 91 epitopes that can bind to the patient 12-11 and patient 14-5 HLAs (PEPI1+) respectively are depicted with lines. PEPI2 represents the peptides selected from PEPI1+s that can bind to multiple HLAs of a patient (PEPI2+). PEPI3 represent peptides that can bind to ≧3 HLAs of a patient (PEPI3+). PEPI4 represent peptides that can bind to ≧4 HLAs of a patient (PEPI4+). PEPI5 represent peptides that can bind to ≧5 HLAs of a patient (PEPI5+). PEPI6 represent peptides that can bind to 6 HLAs of a patient (PEPI6). C: The DNA vaccine specific PEPI3+ set of two patients characterizes their vaccine specific T cell responses.

Fig. 7

Correlation between the ≧1 PEPI3+ Score and CTL response rates of peptide targets determined in clinical trials.

Fig. 8

Correlation between the ≧1 PEPI3+ Score and the clinical Immune Response Rate (IRR) of immunotherapy vaccines. Dashed lines: 95% confidence band.

Fig 9

Correlation between the ≧2 PEPI3+ Score and Disease Control Rate (DCR) of immunotherapy vaccines. Dashed lines: 95% confidence band.

Fig. 10

Peptide hotspot analysis example: PRAME antigen hotspot on 433 patients of the Model Population. On the y axis are the 433 patients of the Model Population, on the x axis is the amino acid sequence of the PRAME antigen (CTA). Each data point represents a PEPI presented by ≧3 HLA class I of one patient starting at the specified amino acid position. The two most frequent PEPIs (called bestEPIs) of the PRAME antigen are highlighted in dark gray (peptide hotspots = PEPI Hotspots).

Fig. 11

CTA Expression Curve calculated by analyzing expression frequency data of tumor specific antigens (CTAs) in human breast cancer tissues. (No cell line data were included.)

Fig. 12

Antigen expression distribution for breast cancer based on the calculation of multi-antigen responses from expression frequencies of the selected 10 different CTAs. A: non-cumulative distribution to calculate the expected value for the number of expressed antigens (AG50). This value shows that probably 6.14 vaccine antigens will be expressed by breast tumor cells. B: cumulative distribution curve of the minimum number of expressed antigens (CTA expression curve). This shows that minimum 4 vaccine antigens will be expressed with 95% probability in breast cancer cell (AG95).

Fig. 13

PEPI representing antigens: breast cancer vaccine-specific CTA antigens with ≥1 PEPI, called as "AP") distribution within the Model Population (n=433) for breast cancer vaccine. A: non-cumulative distribution of AP where the average number of APs is: AP50=5.30, meaning that in average almost 6 CTAs will have PEPIs in the Model Population. B: cumulative distribution curve of the minimum number of APs in the Model Population (n=433). This shows that at least one vaccine antigen will have PEPIs in 95% of the Model Population (n=433) (AP95=1).

Fig. 14

PEPI represented expressed antigen (breast cancer vaccine-specific CTA antigens expressed by the tumor, for which ≥1 PEPI is predicted, called as "AGP") distribution within the model population (n=433) calculated with CTA expression rates for breast cancer. A: non-cumulative distribution of AGP where the expected value for number expressed CTAs represented by PEPI is AGP50=3.37. AGP50 is a measure of the effectiveness of the disclosed breast cancer vaccine in attacking breast tumor in an unselected patient population. AGP50 = 3.37 means that at least 3 CTAs from the vaccine will probably be expressed by the breast tumor cells and present PEPIs in the Model Population. B: cumulative distribution curve of the minimum number of AGPs in the

Model Population (n=433) shows that at least 1 of the vaccine CTAs will present PEPIs in 92% of the population and the remaining 8% of the population will likely have no AGP at all (AGP95=0, AGP92=1).

Fig. 15

CTA Expression Curve calculated by analyzing expression frequency data of tumor specific antigens (CTAs) in human colorectal cancer tissues. (No cell line data were included.)

Fig. 16

Antigen expression distribution for colorectal cancer based on the calculation of multi-antigen responses from expression frequencies of the selected 7 different CTAs. A: non-cumulative distribution to calculate the expected vale for the number of expressed vaccine antigens in colorectal cancers (AG50). This value shows that probably 4.96 vaccine antigens will be expressed by colorectal tumor cells. B: cumulative distribution curve of the minimum number of expressed antigens (CTA expression curve). This shows that minimum 3 antigens will be expressed with 95% probability in the colorectal cancer cell (AG95).

Fig. 17

PEPI represented antigen (colorectal cancer vaccine-specific CTA antigens for which ≥1 PEPI is predicted. Called as "AP") distribution within the model population (n=433) for colorectal cancer. A: non-cumulative distribution of AP where the average number of APs is: AP50=4.73, meaning that in average 5 CTAs will be represented by PEPIs in the model population B: cumulative distribution curve of the minimum number of APs in the model population (n=433). This shows that 2 or more antigens will be represented by PEPIs in 95% of the model population (n=433) (AP95=2).

Fig. 18

PEPI represented expressed antigen (colorectal cancer vaccine-specific CTA antigens expressed by the tumor, for which ≥1 PEPI is predicted. Called as "AGP") distribution within the model population (n=433) calculated with CTA expression rates for colorectal cancer. A: non-cumulative distribution of AGP where the expected value for number expressed CTAs represented by PEPI is AGP50=2.54. AGP50 is a measure of the effectiveness of the disclosed colorectal cancer vaccine in attacking colorectal tumors in an unselected patient population. AGP50 = 2.54 means that at

least 2-3 CTAs from the vaccine will probably be expressed by the colorectal tumor cells and present PEPIs in the Model Population. B: cumulative distribution curve of the minimum number of AGPs in the Model Population (n=433) shows that at least 1 of the vaccine CTAs will be expressed and also present PEPIs in 93% of the population (AGP93=1).

Fig 19

Schematic showing exemplary positions of amino acids in overlapping HLA class I- and HLA class-II binding epitopes in a 30-mer peptide.

Fig. 20

Antigenicity of PolyPEPI1018 CRC Vaccine in a general population. The antigenicity of PolyPEPI1018 in a subject is determined by the AP count, which indicates the number of vaccine antigens that induce T cell responses in a subject. The AP count of PolyPEPI1018 was determined in each of the 433 subjects in the Model Population using the PEPI Test, and the AP50 count was then calculated for the Model Population. The AP50 of PolyPEPI1018 in the Model Population is 4.73. The mean number of immunogenic antigens (i.e., antigens with ≥1 PEPI) in PolyPEPI1018 in a general population is 4.73. Abbreviations: AP = antigens with ≥1 PEPI. Left Panel: Cumulative distribution curve. Right Panel: Distinct distribution curve.

Fig. 21

Effectiveness of PolyPEPI1018 CRC Vaccine in a general population. Vaccine induced T cells can recognize and kill tumor cells if a PEPI in the vaccine is presented by the tumor cell. The number of AGPs (expressed antigens with PEPI) is an indicator of vaccine effectiveness in an individual, and is dependent on both the potency and antigenicity of PolyPEPI1018. The mean number of immunogenic CTAs (i.e., AP [expressed antigens with ≥1 PEPI]) in PolyPEPI1018 is 2.54 in the Model Population. The likelihood that PolyPEPI1018 induces T cell responses against multiple antigens in a subject (i.e., mAGP) in the Model Population is 77%.

Fig. 22

Probability of vaccine antigen expression in the XYZ patient's tumor cells. There is over 95% probability that 5 out of the 12 target antigens in the vaccine regimen is expressed in the patient's tumor. Consequently, the 12 peptide vaccines together can induce immune responses against at

least 5 ovarian cancer antigens with 95% probability (AGP95). It has 84% probability that each peptide will induce immune responses in the XYZ patient. AGP50 is the mean (expected value) =7.9 (it is a measure of the effectiveness of the vaccine in attacking the tumor of XYZ patient).

Fig. 23

MRI findings of patient XYZ treated with personalised (PIT) vaccine. This late stage, heavily pretreated ovarian cancer patient had an unexpected objective response after the PIT vaccine treatment. These MRI findings suggest that PIT vaccine in combination with chemotherapy significantly reduced her tumor burden. The patient now continues the PIT vaccine treatment.

Fig. 24

Probability of vaccine antigen expression in the ABC patient's tumor cells. There is over 95% probability that 4 out of the 13 target antigens in the vaccine is expressed in the patient's tumor. Consequently, the 12 peptide vaccines together can induce immune responses against at least 4 breast cancer antigens with 95% probability (AGP95). It has 84% probability that each peptide will induce immune responses in the ABC patient. AGP50 is the mean (expected value) of the discrete probability distribution = 6.45 (it is a measure of the effectiveness of the vaccine in attacking the tumor of ABC patient).

Description of the SequencesSEQ ID NOs: 1 to 20 set forth 9 mer T cell epitopes described in Table 17.

SEQ ID NOs: 21 to 40 set forth 9 mer T cell epitopes described in Table 20.

SEQ ID NOs 41 to 60 set forth 15 mer T cell epitopes described in Table 17.

SEQ ID NOs 61 to 80 set forth 15 mer T cell epitopes described in Table 20.

SEQ ID NOs: 81 to 111 set forth breast cancer vaccine peptides described in Table 18a.

SEQ ID NOs 112 to 142 set forth the colorectal cancer vaccine peptides described in Table 21a.

SEQ ID NOs 143 to 158 set forth breast cancer, colorectal cancer and/or ovarian cancer associated antigens.

SEQ ID NOs 159 to 171 set forth the additional peptide sequences described in Table 10.

SEQ ID NOs 172 to 194 set forth further 9 mer T cell epitopes described in Table 17.

SEQ ID NOs 195 to 233 set forth further 15 mer T cell epitopes described in Table 17.
SEQ ID NOs 234 to 250 set forth further 9 mer T cell epitopes described in Table 20.
SEQ ID NOs 251 to 271 set forth further 15 mer T cell epitopes described in Table 20.
SEQ ID NOs: 272 to 301 set forth the 9 mer T cell epitopes described in Table 23.
SEQ ID NOs: 302 to 331 set forth the 15 mer T cell epitopes described in Table 23.
SEQ ID NOs: 332 to 346 set forth the ovarian cancer vaccine peptides set forth in Table 24.
SEQ ID NOs: 347 to 361 set forth further breast cancer, colorectal cancer and/or ovarian cancer associated antigens.
SEQ ID NOs: 362 to 374 set forth personalised vaccine peptides designed for patient XYZ described in Table 38.
SEQ ID NOs: 375 to 386 set forth personalised vaccine peptides designed for patient ABC described in Table 41.
SEQ ID NOs 387 to 434 set forth further 9 mer T cell epitopes described in Table 32
SEQ ID NOs: 435 to 449 set forth further breast cancer vaccine peptides described in Table 18a.

Detailed Description

HLA Genotypes

HLAs are encoded by the most polymorphic genes of the human genome. Each person has a maternal and a paternal allele for the three HLA class I molecules (HLA-A*, HLA-B*, HLA-C*) and four HLA class II molecules (HLA-DP*, HLA-DQ*, HLA-DRB1*, HLA-DRB3*/4*/5*). Practically, each person expresses a different combination of 6 HLA class I and 8 HLA class II molecules that present different epitopes from the same protein antigen. The function of HLA molecules is to regulate T cell responses. However up to date it was unknown how the HLAs of a person regulate T cell activation.

The nomenclature used to designate the amino acid sequence of the HLA molecule is as follows: gene name*allele:protein number, which, for instance, can look like: HLA-A*02:25. In this example, “02” refers to the allele. In most instances, alleles are defined by serotypes – meaning that the proteins of a given allele will not react with each other in serological assays.

Protein numbers ("25" in the example above) are assigned consecutively as the protein is discovered. A new protein number is assigned for any protein with a different amino acid sequence (e.g. even a one amino acid change in sequence is considered a different protein number). Further information on the nucleic acid sequence of a given locus may be appended to the HLA nomenclature, but such information is not required for the methods described herein.

The HLA class I genotype or HLA class II genotype of an individual may refer to the actual amino acid sequence of each class I or class II HLA of an individual, or may refer to the nomenclature, as described above, that designates, minimally, the allele and protein number of each HLA gene. An HLA genotype may be determined using any suitable method. For example, the sequence may be determined via sequencing the HLA gene loci using methods and protocols known in the art. Alternatively, the HLA set of an individual may be stored in a database and accessed using methods known in the art.

Some subjects may have two HLA alleles that encode the same HLA molecule (for example, two copies for HLA-A*02:25 in case of homozygosity). The HLA molecules encoded by these alleles bind all of the same T cell epitopes. For the purposes of this disclosure "binding to at least two HLA molecules of the subject" as used herein includes binding to the HLA molecules encoded by two identical HLA alleles in a single subject. In other words, "binding to at least two HLA molecules of the subject" and the like could otherwise be expressed as "binding to the HLA molecules encoded by at least two HLA alleles of the subject".

Polyeptides

The disclosure relates to polypeptides that are derived from CTAs and that are immunogenic for a high proportion of the human population.

As used herein, the term "polypeptide" refers to a full-length protein, a portion of a protein, or a peptide characterized as a string of amino acids. As used herein, the term "peptide" refers to a short polypeptide comprising between 2, or 3, or 4, or 5, or 6, or 7, or 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15 and 10, or 11, or 12, or 13, or 14, or 15, or 20, or 25, or 30, or 35, or 40, or 45, or 50 or 55 or 60 amino acids.

The terms "fragment" or "fragment of a polypeptide" as used herein refer to a string of amino acids or an amino acid sequence typically of reduced length relative to the or a reference polypeptide and comprising, over the common portion, an amino acid sequence identical to the reference polypeptide. Such a fragment according to the disclosure may be, where appropriate, included in a larger polypeptide of which it is a constituent. In some cases the fragment may comprise the full length of the polypeptide, for example where the whole polypeptide, such as a 9 amino acid peptide, is a single T cell epitope. In some cases the fragments referred to herein may be between 2, or 3, or 4, or 5 or 6 or 7 or 8 or 9 and 20, or 25, or 30, or 35, or 40, or 45, or 50 amino acids.

As used herein, the term "epitope" or "T cell epitope" refers to a sequence of contiguous amino acids contained within a protein antigen that possess a binding affinity for (is capable of binding to) one or more HLAs. An epitope is HLA- and antigen-specific (HLA-epitope pairs, predicted with known methods), but not subject specific. An epitope, a T cell epitope, a polypeptide, a fragment of a polypeptide or a composition comprising a polypeptide or a fragment thereof is "immunogenic" for a specific human subject if it is capable of inducing a T cell response (a cytotoxic T cell response or a helper T cell response) in that subject. In some cases the helper T cell response is a Th1-type helper T cell response. In some cases an epitope, a T cell epitope, a polypeptide, a fragment of a polypeptide or a composition comprising a polypeptide or a fragment thereof is "immunogenic" for a specific human subject if it is more likely to induce a T cell response or immune response in the subject than a different T cell epitope (or in some cases two different T cell epitopes each) capable of binding to just one HLA molecule of the subject.

The terms "T cell response" and "immune response" are used herein interchangeably, and refer to the activation of T cells and/or the induction of one or more effector functions following recognition of one or more HLA-epitope binding pairs. In some cases an "immune response" includes an antibody response, because HLA class II molecules stimulate helper responses that are involved in inducing both long lasting CTL responses and antibody responses. Effector functions include cytotoxicity, cytokine production and proliferation. According to the present

disclosure, an epitope, a T cell epitope, or a fragment of a polypeptide is immunogenic for a specific subject if it is capable of binding to at least two, or in some cases at least three, class I or at least two, or in some cases at least three or at least four class II HLAs of the subject.

For the purposes of this disclosure we have coined the term "personal epitope", or "PEPI" to distinguish subject specific epitopes from HLA specific epitopes. A "PEPI" is a fragment of a polypeptide consisting of a sequence of contiguous amino acids of the polypeptide that is a T cell epitope capable of binding to one or more HLA class I molecules of a specific human subject. In other cases a "PEPI" is a fragment of a polypeptide consisting of a sequence of contiguous amino acids of the polypeptide that is a T cell epitope capable of binding to one or more HLA class II molecules of a specific human subject. In other words a "PEPI" is a T cell epitope that is recognised by the HLA set of a specific individual, and is consequently specific to the subject in addition to the HLA and the antigen. In contrast to an "epitope", which is specific only to HLA and the antigen, PEPIs are specific to an individual because different individuals have different HLA molecules which each bind to different T cell epitopes. This subject specificity of the PEPIs allows to make personalized cancer vaccines.

"PEPI1" as used herein refers to a peptide, or a fragment of a polypeptide, that can bind to one HLA class I molecule (or, in specific contexts, HLA class II molecule) of an individual. "PEPI1+" refers to a peptide, or a fragment of a polypeptide, that can bind to one or more HLA class I molecule of an individual.

"PEPI2" refers to a peptide, or a fragment of a polypeptide, that can bind to two HLA class I (or II) molecules of an individual. "PEPI2+" refers to a peptide, or a fragment of a polypeptide, that can bind to two or more HLA class I (or II) molecules of an individual, i.e. a fragment identified according to a method of the disclosure.

"PEPI3" refers to a peptide, or a fragment of a polypeptide, that can bind to three HLA class I (or II) molecules of an individual. "PEPI3+" refers to a peptide, or a fragment of a polypeptide, that can bind to three or more HLA class I (or II) molecules of an individual.

"PEPI4" refers to a peptide, or a fragment of a polypeptide, that can bind to four HLA class I (or II) molecules of an individual. "PEPI4+" refers to a peptide, or a fragment of a polypeptide, that can bind to four or more HLA class I (or II) molecules of an individual.

"PEPI5" refers to a peptide, or a fragment of a polypeptide, that can bind to five HLA class I (or II) molecules of an individual. "PEPI5+" refers to a peptide, or a fragment of a polypeptide, that can bind to five or more HLA class I (or II) molecules of an individual.

"PEPI6" refers to a peptide, or a fragment of a polypeptide, that can bind to all six HLA class I (or six HLA class II) molecules of an individual.

Generally speaking, epitopes presented by HLA class I molecules are about nine amino acids long and epitopes presented by HLA class II molecules are about fifteen amino acids long. For the purposes of this disclosure, however, an epitope may be more or less than nine (for HLA Class I) or fifteen (for HLA Class II) amino acids long, as long as the epitope is capable of binding HLA. For example, an epitope that is capable of binding to class I HLA may be between 7, or 8 or 9 and 9 or 10 or 11 amino acids long. An epitope that is capable of binding to a class II HLA may be between 13, or 14 or 15 and 15 or 16 or 17 amino acids long.

A given HLA of a subject will only present to T cells a limited number of different peptides produced by the processing of protein antigens in an APC. As used herein, "display" or "present", when used in relation to HLA, references the binding between a peptide (epitope) and an HLA. In this regard, to "display" or "present" a peptide is synonymous with "binding" a peptide.

Using techniques known in the art, it is possible to determine the epitopes that will bind to a known HLA. Any suitable method may be used, provided that the same method is used to determine multiple HLA-epitope binding pairs that are directly compared. For example, biochemical analysis may be used. It is also possible to use lists of epitopes known to be bound by a given HLA. It is also possible to use predictive or modelling software to determine which epitopes may be bound by a given HLA. Examples are provided in Table 1. In some cases a T cell epitope is capable of binding to a given HLA if it has an IC50 or predicted IC50 of less than 5000 nM, less than 2000 nM, less than 1000 nM, or less than 500 nM.

Table 1 - Example software for determining epitope-HLA binding

EPITOPE PREDICTION TOOLS	WEB ADDRESS
BIMAS, NIH	www-bimas.cit.nih.gov/molbio/hla_bind/
PPAPROC, Tubingen Univ.	
MHCPred, Edward Jenner Inst. of Vaccine Res.	
EpiJen, Edward Jenner Inst. of Vaccine Res.	http://www.ddg-pharmfac.net/epijen/EpiJen/EpiJen.htm
NetMHC, Center for Biological Sequence Analysis	http://www.cbs.dtu.dk/services/NetMHC/
SVMHC, Tubingen Univ.	http://abi.inf.uni-tuebingen.de/Services/SVMHC/
SYFPEITHI, Biomedical Informatics, Heidelberg	http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePredictio n.htm
ETK EPITOOLKIT, Tubingen Univ.	http://etk.informatik.uni-tuebingen.de/epipred/
PREDEP, Hebrew Univ. Jerusalem	http://margalit.huji.ac.il/Teppred/mhc-bind/index.html
RANKPEP, MIF Bioinformatics	http://bio.dfci.harvard.edu/RANKPEP/
IEDB, Immune Epitope Database	http://tools.immuneepitope.org/main/html/tcell_tools.html
EPITOPE DATABASES	WEB ADDRESS
MHCBN, Institute of Microbial Technology, Chandigarh, INDIA	http://www.imtech.res.in/raghava/mhcbn/
SYFPEITHI, Biomedical Informatics, Heidelberg	http://www.syfpeithi.de/
AntiJen, Edward Jenner Inst. of Vaccine Res.	http://www.ddg-pharmfac.net/antijen/AntiJen/antijenhomepage.htm
EPIMHC database of MHC ligands, MIF Bioinformatics	http://immunax.dfci.harvard.edu/epimhc/
IEDB, Immune Epitope Database	http://www.iedb.org/

In some embodiments the peptides of the disclosure may comprise or consist of one or more fragments of one or more CTAs. CTAs are not typically expressed beyond embryonic development in healthy cells. In healthy adults, CTA expression is limited to male germ cells that do not express HLAs and cannot present antigens to T cells. Therefore, CTAs are considered expressional neoantigens when expressed in cancer cells.

CTAs are a good choice for cancer vaccine targets because their expression is (i) specific for tumor cells, (ii) more frequent in metastases than in primary tumors and (iii) conserved among metastases of the same patient (Gajewski ed. Targeted Therapeutics in Melanoma. Springer New York. 2012).

The peptides of the disclosure may comprise or consist of one or more fragments of one or more breast cancer associated antigens selected from SPAG9 (SEQ ID NO: 143), AKAP-4 (SEQ ID NO: 144), BORIS (SEQ ID NO: 145), NY-SAR-35 (SEQ ID NO: 146), NY-BR-1 (SEQ ID NO: 147), SURVIVIN (SEQ ID NO: 148), MAGE-A11 (SEQ ID NO: 149), PRAME (SEQ ID NO: 150), MAGE-A9 (SEQ ID NO: 151), HOM-TES-85 (SEQ ID NO: 152), PIWIL-2 (SEQ ID NO: 349), EpCAM (SEQ ID NO: 154), HIWI (SEQ ID NO: 350), PLU-1 (SEQ ID NO: 351), TSGA10 (SEQ ID NO: 351), ODF-4 (SEQ ID NO: 352), SP17 (SEQ ID NO:354), RHOXF-2 (SEQ ID NO: 355), and NY-ESO-1 (SEQ ID NO: 356); one or more ovarian cancer-associated antigens selected from PIWIL-4 (SEQ ID NO: 357), WT1 (SEQ ID NO: 358), EpCAM (SEQ ID NO: 154), BORIS (SEQ ID NO: 145), AKAP-4 (SEQ ID NO: 144), OY-TES-1 (SEQ ID NO: 359), SP17 (SEQ ID NO: 354), PIWIL-2 (SEQ ID NO: 349), PIWIL-3 (SEQ ID NO: 360), SPAG9 (SEQ ID NO: 143), PRAME (SEQ ID NO: 150), HIWI (SEQ ID NO: 350), SURVIVIN (SEQ ID NO: 148), and AKAP-3 (SEQ ID NO: 361); and/or one or more colorectal cancer-associated antigens selected from TSP50 (SEQ ID NO: 153), EpCAM (SEQ ID NO: 154), SPAG9 (SEQ ID NO: 143), CAGE1 (SEQ ID NO: 155), FBXO39 (SEQ ID NO: 156), SURVIVIN (SEQ ID NO: 148), MAGE-A8 (SEQ ID NO 157), MAGE-A6 (SEQ ID NO: 158), LEMD1 (SEQ ID NO:348) and MAGE-A3 (SEQ ID NO: 347). In some cases the peptide comprises or consists of one or more amino acid sequences selected from SEQ ID NOs: 41-80, or from SEQ ID NOs: 41-80, 195-233, 251-271 and 302-331 that are optimised for T cell activation / binding to all HLA types across the population.

In some cases the amino acid sequence is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the target polypeptide antigen, in other words that are not the same sequence of consecutive amino acids found adjacent to the selected fragments in the target polypeptide antigen. In some cases the sequence is flanked by up to 41 or 35 or 30 or

25 or 20 or 15 or 10, or 9 or 8 or 7 or 6 or 5 or 4 or 3 or 2 or 1 additional amino acid at the N and/or C terminus or between target polypeptide fragments. In other cases each polypeptide either consists of a fragment of a target polypeptide antigen, or consists of two or more such fragments arranged end to end (arranged sequentially in the peptide end to end) or overlapping in a single peptide (where two or more of the fragments comprise partially overlapping sequences, for example where two PEPIs in the same polypeptide are within 50 amino acids of each other).

When fragments of different polypeptides or from different regions of the same polypeptide are joined together in an engineered peptide there is the potential for neoepitopes to be generated around the join or junction. Such neoepitopes encompass at least one amino acid from each fragment on either side of the join or junction, and may be referred to herein as junctional amino acid sequences. The neoepitopes may induce undesired T cell responses against healthy cells (autoimmunity). The polypeptides may be designed, or the polypeptides may be screened, to avoid, eliminate or minimise neoepitopes that correspond to a fragment of a protein expressed in normal healthy human cells and/or neoepitopes that are capable of binding to at least two, or in some cases at least three, or at least four HLA class I molecules of the subject, or in some cases at least two, or at least three or four or five HLA class II molecules of the subject. In some cases the peptide is designed, or the polypeptide screened, to eliminate polypeptides having a junctional neoepitope that is capable of binding in more than a threshold percentage of human subjects in an intent-to-treat population, to at least two HLA class I molecules expressed by individual subjects of the population. In some cases the threshold is 20%, or 15%, or 10%, or 5%, or 2%, or 1%, or 0.5% of said population. Alignment may be determined using known methods such as BLAST algorithms. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/).

The presence in a vaccine or immunotherapy composition of at least two polypeptide fragments (epitopes) that can bind to at least three HLA class I of an individual (≥2 PEPI3+) is predictive for a clinical response. In other words, if ≥2 PEPI3+ can be identified within the active ingredient polypeptide(s) of a vaccine or immunotherapy composition, then an individual

is a likely clinical responder. The at least two multiple HLA-binding PEPIs of the composition polypeptides may both target a single antigen (e.g a polypeptide vaccine comprising two multiple HLA-binding PEPIs derived from a single tumor associated antigen targeted by the vaccine) or may target different antigens (e.g. a polypeptide vaccine comprising one multiple HLA-binding PEPI derived from one tumor associated antigen and a second multiple HLA-binding PEPI derived from a different tumor associated antigen).

Without wishing to be bound by theory, the inventors believe that one reason for the increased likelihood of deriving clinical benefit from a vaccine/immunotherapy comprising at least two multiple-HLA binding PEPIs, is that diseased cell populations, such as cancer or tumor cells or cells infected by viruses or pathogens such as HIV, are often heterogenous both within and between effected subjects. A specific cancer patient, for example, may or may not express or overexpress a particular cancer associated target polypeptide antigen of a vaccine, or their cancer may comprise heterogeneous cell populations, some of which (over-)express the antigen and some of which do not. In addition, the likelihood of developing resistance is decreased when more multiple HLA-binding PEPIs are included or targeted by a vaccine/immunotherapy because a patient is less likely to develop resistance to the composition through mutation of the target PEPI(s).

Currently most vaccines and immunotherapy compositions target only a single polypeptide antigen. However according to the present disclosure it is in some cases beneficial to provide a pharmaceutical composition that targets two or more different polypeptide antigens. For example, most cancers or tumors are heterogeneous, meaning that different cancer or tumor cells of a subject (over-)express different antigens. The tumour cells of different cancer patients also express different combinations of tumour-associated antigens. The anti-cancer immunogenic compositions that are most likely to be effective are those that target multiple antigens expressed by the tumor, and therefore more cancer or tumor cells, in an individual human subject or in a population.

The beneficial effect of combining multiple bestEPIs in a single treatment (administration of one or more pharmaceutical compositions that together comprise multiple PEPIs), can be

illustrated by the personalised vaccine polypeptides described in Examples 15 and 16 below. Exemplary CTA expression probabilities in ovarian cancer are as follows: BAGE: 30%; MAGE A9: 37%; MAGE A4: 34%; MAGE A10: 52%. If patient XYZ were treated with a vaccine comprising PEPIs in only BAGE and MAGE A9, then the probability of having a mAGP (multiple expressed antigens with PEPI) would be 11%. If patent XYZ were treated with a vaccine comprising only PEPIs for the MAGE A4 and MAGE A10 CTAs, then the probability of having a multiAGP would be 19%. However if a vaccine contained all 4 of these CTAs (BAGE, MAGE A9, MAGE A4 and MAGE A10), then the probability of having a mAGP would be 50%. In other words the effect would be greater than the combined probabilities of mAGP for both two-PEPI treatments (probability mAGP for BAGE/MAGE + probability mAGP for MAGE A4 and MAGE A10). Patient XYZ's PIT vaccine described in Example 21 contains a further 9 PEPIs, and thus, the probability of having a mAGP is over 99.95%.

Likewise exemplary CTA expression probabilities in breast cancer are as follows: MAGE C2: 21%; MAGE A1: 37%; SPC1: 38%; MAGE A9: 44%. Treatment of patient ABC with a vaccine comprising PEPIs in only MAGE C2: 21% and MAGE A1 has a mAGP probability of 7%. Treatment of patient ABC with a vaccine comprising PEPIs in only SPC1: 38%; MAGE A9 has a mAGP probability of 11%. Treatment of patient ABC with a vaccine comprising PEPIs in MAGE C2: 21%; MAGE A1: 37%; SPC1: 38%; MAGE A9 has a mAGP probability of 44% (44 > 7 + 11). Patient ABC's PIT vaccine described in Example 22 contains a further 8 PEPIs, and thus, the probability of having a mAGP is over 99.93%.

Accordingly in some cases, the polypeptide or panel of polypeptides of the disclosure or an active ingredient polypeptide of a pharmaceutical composition or kit of the disclosure may comprise or consist of any combination of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25 fragments of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25 one or more of the cancer associated antigens, or CTAs, such as the CTA discussed above. Each fragment comprises or consists of a different target epitope having an amino acid sequence selected from SEQ ID NOs: 1-40; or selected from SEQ ID NOs: 1 to 20; or selected from SEQ ID NOs: 21 to 40; or selected from SEQ ID NOs: 1-

20, 24 and 172-194; or selected from SEQ ID NOs: 21-40 and 234-250; or selected from SEQ ID NOs: 272-301; or selected from SEQ ID NOs: 1-40, 172-194 and 234-250; or selected from SEQ ID NOs: 21-40, 234-250 and 272-301; or selected from SEQ ID NOs: 1-20, 24, 172-194 and 272-301; or selected from SEQ ID NOs: 1-40, 172-194, 234-250 and 272-301; or selected from SEQ ID NOs: 41-60, 64 and 195-233; or selected from SEQ ID NOs: 61-80 and 251-271; or selected from SEQ ID NOs: 302-331; or selected from SEQ ID NOs: 41-80, 195-233 and 251-271; or selected from SEQ ID NOs: 61-80, 251-271 and 302 to 331; or selected from SEQ ID NOs: 41-60, 64, 191-233 and 302 to 331; or selected from SEQ ID NOs: 41-80, 195-233, 251-271 and 332-346; or selected from SEQ ID NOs: 1-20, 24, 41-60, 64, 172-194 and 195-233; or selected from SEQ ID NOs: 21-40, 61-80, 234-250 and 251-271; or selected from SEQ ID NOs: 271-331; or selected from SEQ ID NOs: 1-80, 172-194, 195-233, 234-250 and 251-271; or selected from SEQ ID NOs: 21-40, 61-80, 234-250, 251-271, 272-301 and 302-331; or selected from SEQ ID NOs: 1-80, 172-233, 234-271 and 272-331; or selected from SEQ ID NOs: 81-111 and 435-449; or selected from SEQ ID NOs: 112-142; or selected from SEQ ID NOs: 332-346; or selected from SEQ ID NOs: 81-142; or selected from SEQ ID NOs: 112-142 and 332-346; or selected from SEQ ID NOs: 81-111, 435-449 and 332-346; or selected from SEQ ID NOs: 81-142 and 332-346; or selected from SEQ ID NOs: 41-60, 64, 81-111, 435-449 and 195-233; or selected from SEQ ID NOs: 61-80, 112-142 and 251-271; or selected from SEQ ID NOs: 302-346; or selected from SEQ ID NOs: 41-142, 195-233 and 251-271; or selected from SEQ ID NOs: 61-80, 112-142, 251-271 and 302-346; or selected from SEQ ID NOs: 41-60, 64, 81-111, 435-449, 195-233 and 302-346; or selected from SEQ ID NOs: 41-142, 195-233, 251-271 and 302-346; or selected from SEQ ID NOs: 1-20, 24, 41-60, 64, 81-111, 435-449 and 172-233; or selected from SEQ ID NOs: 21-40, 61-80, 112-142, or 234-271; or selected from SEQ ID NOs: 272-346; or selected from SEQ ID NOs: 1-142 and 172-271; or selected from SEQ ID NOs: 21-40, 61-80, 112-142 and 234-346; or selected from SEQ ID NOs: 1-20, 24, 41-60, 64, 81-111, 435-449, 172-233 and 272 to 346; or selected from SEQ ID NOs: 1-142 and 172-346; or selected from SEQ ID NOs: 1 to 2, or to 3, or 4, or 5, or 6, or 7, or 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16, or 17, or 18, or 19, or SEQ ID NOs: 20 to 21, or to 22, or 23, or 24, or 25, or 26, or 27, or 28, or

29, or 30, or 31, or 32, or 33, or 34, or 35, or 36, or 37, or 38, or 39; or a different amino acid sequences selected from SEQ ID NOs: 41 to 80, or SEQ ID NOs: 41 to 60, or SEQ ID NOs: 61-80; or SEQ ID NOs: 41 to 42, or to 43, or to 44, or to 45, or to 46, or to 47, or to 48, or to 49, or 50, or 51, or 52, or 53, or 54, or 55, or 56, or 57, or 58, or 59, SEQ ID NOs: 60 to 61, or to 62, or to 63, or to 64, or to 65, or to 66, or to 67, or to 68, or to 69, or to 70, or to 71, or to 72, or to 73, or to 74, or to 75, or to 76, or to 77, or to 78, or to 79; a different amino acid sequences selected from SEQ ID NOs: 81 to 142; or selected from SEQ ID NOs: 81 to 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 105, 106, 107, 108, 109, 110, or 111; or selected from SEQ ID NOs: 81 to 105; or selected from SEQ ID NOs: 99, 100, 92, 93, 101, 103, 104, 105 and 98; or selected from SEQ ID NOs: 112 to 142; or selected from SEQ ID NOs: 112 to 113, 114, 115, 116, 117, 118, 119, 120, 121, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141 or 142; or selected from SEQ ID NOs: 112 to 134; or selected from SEQ ID NOs: 121, 124, 126, 127, 130, 131, 132, 133 and 134; or selected from SEQ ID NOs: 1 to 2, or to 3, or 4, or 5, or 6, or 7, or 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16, or 17, or 18, or 19, or SEQ ID NOs: 20 to 21, or to 22, or 23, or 24, or 25, or 26, or 27, or 28, or 29, or 30, or 31, or 32, or 33, or 34, or 35, or 36, or 37, or 38, or 39; or a different amino acid sequences selected from SEQ ID NOs: 41 to 80, or SEQ ID NOs: 41 to 60, or SEQ ID NOs: 61-80; or SEQ ID NOs: 41 to 42, or to 43, or to 44, or to 45, or to 46, or to 47, or to 48, or to 49, or 50, or 51, or 52, or 53, or 54, or 55, or 56, or 57, or 58, or 59, SEQ ID NOs: 60 to 61, or to 62, or to 63, or to 64, or to 65, or to 66, or to 67, or to 68, or to 69, or to 70, or to 71, or to 72, or to 73, or to 74, or to 75, or to 76, or to 77, or to 78, or to 79; a different amino acid sequences selected from SEQ ID NOs: 81 to 142; or selected from SEQ ID NOs: 81 to 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 105, 106, 107, 108, 109, 110, or 111; or selected from SEQ ID NOs: 81 to 105; or selected from SEQ ID NOs: 99, 100, 92, 93, 101, 103, 104, 105 and 98; or selected from SEQ ID NOs: 112 to 142; or selected from SEQ ID NOs: 112 to 113, 114, 115, 116, 117, 118, 119, 120, 121, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141 or 142; or selected from SEQ ID NOs: 112 to 134; or selected from SEQ ID NOs: 121, 124, 126, 127, 130, 131, 132, 133 and 134;

or selected from SEQ ID Nos: 130, 121, 131, 124, 134, 126; or selected from SEQ ID NO: 435-449; or selected from any of these groups of sequences excluding SEQ ID NOs: 12, 32, 19 and/or 39, and/or SEQ ID NOs: 21, 41, 23 and/or 43 and/or SEQ ID NOs: 172, 177, 195 and/or 203, and/or SEQ ID NOs: 1, 41 and/or 197, and/or SEQ ID NOs: 4, 44 and/or 201, and/or SEQ ID NOs: 1, 4, 44, 197 and/or 201, and/or SEQ ID NOs: 1, 41, 197, 184 and/or 212, and/or SEQ ID NOs: 3, 43 and/or 200, and/or SEQ ID NOs: 3, 43, 200, 7 and/or 47, and/or SEQ ID NOs: 10, 50 and/or 220, and/or SEQ ID NOs: 24, 64 and/or 202, and/or SEQ ID NOs: 6, 46 and/or 209, and/or SEQ ID NOs: 182, 210, 185 and/or 213, and/or SEQ ID NOs: 14, 54, 225 and 226, and/or SEQ ID NOs: 190, 218, 11, 51 and/or 219, and/or SEQ ID NOs: 12, 224 and/or 52, and/or SEQ ID NOs:192, 227 and/or 228, and/or SEQ ID NOs:17, 229, 230 and/or 57, and/or SEQ ID NOs: 21, 252, 61 and/or 253, and/or SEQ ID NOs: 23, 63 and/or 256, and/or SEQ ID NOs: 21, 252, 61, 253, 23, 63 and/or 256, and/or SEQ ID NOs: 237 and/or 238, and/or SEQ ID NOs: 26 and/or 240, and/or SEQ ID NOs: 242, 244, 263 and/or 265, and/or SEQ ID NOs: 29, 69 and/or 259, and/or SEQ ID NOs: 24, 64 and/or 255, and/or SEQ ID NOs: 236, 257 and/or 258, and/or SEQ ID NOs: 27, 67, 241 and/or 262, and/or SEQ ID NOs: 252, 249 and/or 264, and/or SEQ ID NOs: 35, 250 and/or 75, and/or SEQ ID NOs: 252, 249, 264, 35, 250 and/or 75, and/or SEQ ID NOs: 36, 266 and/or 76, and/or SEQ ID NOs: 36, 266, 76, 39 and/or 79, and/or SEQ ID NOs: 38, 268 and/or 78, and/or SEQ ID NOs: 38, 268, 78, 246 and/or 270, and/or SEQ ID NOs: 245, 269, and/or 248, and/or SEQ ID NOs: 245, 269, 248, 40 and/or 80, and/or SEQ ID NOs: 272, 302, 281 and/or 311, and/or SEQ ID NOs: 276, 306, 300 and/or 330, and/or SEQ ID NOs: 276, 306, 289 and/or 319, and/or SEQ ID NOs: 277, 307, 283 and/or 313, and/or SEQ ID NOs: 277, 307, 290 and/or 320, and/or SEQ ID NOs: 282, 312, 297 and/or 327, or any other combinations of the sequences disclosed herein that are within 50-60 amino acids of each other in any one or more of the antigens of SEQ ID NOs: 143-158 and 347 to 351; and/or SEQ ID NOs: 18, 19 and/or 20 and/or SEQ ID NOs: 34-40; and/or SEQ ID NOs corresponding to peptides shown in Table 17, 20 and/or 23 having a N%*B% value of less than 12% or 13% or 14% or 17.6% or 17.8% or 18% or 20% or 21% or 22% or 22.2% or 24% or 25% or 27% or 28% or 30% or 31% or 31.5% or 32% or 32.5% or 35%. In some cases the panel of peptides comprises or consists of one or more

polypeptides comprising or consisting of the amino acid sequences of SEQ ID NOs: 130, 121, 131, 124, 134, 126 and/or SEQ ID NOs: 435-449.

In some cases the disclosure provides a panel of any two or more of the peptides or groups of peptides described above. For example the panel may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more such peptides. In some cases the panel comprises or consists of peptides comprising or consisting of all or any combination of the amino acid sequences of SEQ ID NOs: 99, 100, 92, 93, 101, 103, 104, 105 and 98; or the amino acid sequences of SEQ ID NOs: 121, 124, 126, 127, 130, 131, 132, 133 and 134. In some cases the panel comprises or consists of peptides comprising or consisting of all or any combination of the amino acid sequences of SEQ ID NOs: SEQ ID NOs: 130, 121, 131, 124, 134, 126 and/or SEQ ID NOs: 435-449.

Pharmaceutical Compositions, Methods of Treatment and Modes of Administration

In some aspects the disclosure relates to a pharmaceutical composition, kit, or panels of polypeptides as described above having one or more polypeptides as active ingredient(s). These may be for use in a method of inducing an immune response, treating, vaccinating or providing immunotherapy to a subject, and the pharmaceutical composition may be a vaccine or immunotherapy composition. Such a treatment comprises administering one or more polypeptides or pharmaceutical compositions that together comprise all of the active ingredient polypeptides of the treatment to the subject. Multiple polypeptides or pharmaceutical compositions may be administered together or sequentially, for example all of the pharmaceutical compositions or polypeptides may be administered to the subject within a period of 1 year, or 6 months, or 3 months, or 60 or 50 or 40 or 30 days.

The term “active ingredient” as used herein refers to a polypeptide that is intended to induce an immune response and may include a polypeptide product of a vaccine or immunotherapy composition that is produced *in vivo* after administration to a subject. For a DNA or RNA immunotherapy composition, the polypeptide may be produced *in vivo* by the cells of a subject to whom the composition is administered. For a cell-based composition, the polypeptide

may be processed and/or presented by cells of the composition, for example autologous dendritic cells or antigen presenting cells pulsed with the polypeptide or comprising an expression construct encoding the polypeptide. The pharmaceutical composition may comprise a polynucleotide or cell encoding one or more active ingredient polypeptides.

The composition/kit may optionally further comprise at least one pharmaceutically acceptable diluent, carrier, or preservative and/or additional polypeptides that do not comprise any PEPIs. The polypeptides may be engineered or non-naturally occurring. The kit may comprise one or more separate containers each containing one or more of the active ingredient peptides. The composition/kit may be a personalised medicine to prevent, diagnose, alleviate, treat, or cure a disease of an individual, such as a cancer.

The immunogenic or pharmaceutical compositions or kits described herein may comprise, in addition to one or more immunogenic peptides, a pharmaceutically acceptable excipient, carrier, diluent, buffer, stabiliser, preservative, adjuvant or other materials well known to those skilled in the art. Such materials are preferably non-toxic and preferably do not interfere with the pharmaceutical activity of the active ingredient(s). The pharmaceutical carrier or diluent may be, for example, water containing solutions. The precise nature of the carrier or other material may depend on the route of administration, e.g. oral, intravenous, cutaneous or subcutaneous, nasal, intramuscular, intradermal, and intraperitoneal routes.

The pharmaceutical compositions of the disclosure may comprise one or more "pharmaceutically acceptable carriers". These are typically large, slowly metabolized macromolecules such as proteins, saccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, sucrose (Paoletti et al., 2001, Vaccine, 19:2118), trehalose (WO 00/56365), lactose and lipid aggregates (such as oil droplets or liposomes). Such carriers are well known to those of ordinary skill in the art. The pharmaceutical compositions may also contain diluents, such as water, saline, glycerol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present. Sterile pyrogen-free, phosphate buffered physiologic saline is a typical carrier (Gennaro, 2000, Remington: The Science and Practice of Pharmacy, 20th edition, ISBN:0683306472).

The pharmaceutical compositions of the disclosure may be lyophilized or in aqueous form, i.e. solutions or suspensions. Liquid formulations of this type allow the compositions to be administered direct from their packaged form, without the need for reconstitution in an aqueous medium, and are thus ideal for injection. The pharmaceutical compositions may be presented in vials, or they may be presented in ready filled syringes. The syringes may be supplied with or without needles. A syringe will include a single dose, whereas a vial may include a single dose or multiple doses.

Liquid formulations of the disclosure are also suitable for reconstituting other medicaments from a lyophilized form. Where a pharmaceutical composition is to be used for such extemporaneous reconstitution, the disclosure provides a kit, which may comprise two vials, or may comprise one ready-filled syringe and one vial, with the contents of the syringe being used to reconstitute the contents of the vial prior to injection.

The pharmaceutical compositions of the disclosure may include an antimicrobial, particularly when packaged in a multiple dose format. Antimicrobials may be used, such as 2-phenoxyethanol or parabens (methyl, ethyl, propyl parabens). Any preservative is preferably present at low levels. Preservative may be added exogenously and/or may be a component of the bulk antigens which are mixed to form the composition (e.g. present as a preservative in pertussis antigens).

The pharmaceutical compositions of the disclosure may comprise detergent e.g. Tween (polysorbate), DMSO (dimethyl sulfoxide), DMF (dimethylformamide). Detergents are generally present at low levels, e.g. <0.01%, but may also be used at higher levels, e.g. 0.01 – 50%.

The pharmaceutical compositions of the disclosure may include sodium salts (e.g. sodium chloride) and free phosphate ions in solution (e.g. by the use of a phosphate buffer).

In certain embodiments, the pharmaceutical composition may be encapsulated in a suitable vehicle either to deliver the peptides into antigen presenting cells or to increase the stability. As will be appreciated by a skilled artisan, a variety of vehicles are suitable for delivering a pharmaceutical composition of the disclosure. Non-limiting examples of suitable structured fluid delivery systems may include nanoparticles, liposomes, microemulsions,

micelles, dendrimers and other phospholipid-containing systems. Methods of incorporating pharmaceutical compositions into delivery vehicles are known in the art.

In order to increase the immunogenicity of the composition, the pharmacological compositions may comprise one or more adjuvants and/or cytokines.

Suitable adjuvants include an aluminum salt such as aluminum hydroxide or aluminum phosphate, but may also be a salt of calcium, iron or zinc, or may be an insoluble suspension of acylated tyrosine, or acylated sugars, or may be cationically or anionically derivatised saccharides, polyphosphazenes, biodegradable microspheres, monophosphoryl lipid A (MPL), lipid A derivatives (e.g. of reduced toxicity), 3-O-deacylated MPL [3D-MPL], quil A, Saponin, QS21, Freund's Incomplete Adjuvant (Difco Laboratories, Detroit, Mich.), Merck Adjuvant 65 (Merck and Company, Inc., Rahway, N.J.), AS-2 (Smith-Kline Beecham, Philadelphia, Pa.), CpG oligonucleotides, bioadhesives and mucoadhesives, microparticles, liposomes, polyoxyethylene ether formulations, polyoxyethylene ester formulations, muramyl peptides or imidazoquinolone compounds (e.g. imiquamod and its homologues). Human immunomodulators suitable for use as adjuvants in the disclosure include cytokines such as interleukins (e.g. IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc), macrophage colony stimulating factor (M-CSF), tumour necrosis factor (TNF), granulocyte, macrophage colony stimulating factor (GM-CSF) may also be used as adjuvants.

In some embodiments, the compositions comprise an adjuvant selected from the group consisting of Montanide ISA-51 (Seppic, Inc., Fairfield, N.J., United States of America), QS-21 (Aquila Biopharmaceuticals, Inc., Lexington, Mass., United States of America), GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete and incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT).

By way of example, the cytokine may be selected from the group consisting of a transforming growth factor (TGF) such as but not limited to TGF-α and TGF-β; insulin-like growth factor-I and/or insulin-like growth factor-II; erythropoietin (EPO); an osteoinductive

factor; an interferon such as but not limited to interferon-.α, -β, and -γ; a colony stimulating factor (CSF) such as but not limited to macrophage-CSF (M-CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF). In some embodiments, the cytokine is selected from the group consisting of nerve growth factors such as NGF-β; platelet-growth factor; a transforming growth factor (TGF) such as but not limited to TGF-α. and TGF-β; insulin-like growth factor-I and insulin-like growth factor-II; erythropoietin (EPO); an osteoinductive factor; an interferon (IFN) such as but not limited to IFN-α, IFN-β, and IFN-γ; a colony stimulating factor (CSF) such as macrophage-CSF (M-CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF); an interleukin (Il) such as but not limited to IL-1, IL-1.alpha., IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12; IL-13, IL-14, IL-15, IL-16, IL-17, IL-18; LIF; kit-ligand or FLT-3; angiostatin; thrombospondin; endostatin; a tumor necrosis factor (TNF); and LT.

It is expected that an adjuvant or cytokine can be added in an amount of about 0.01 mg to about 10 mg per dose, preferably in an amount of about 0.2 mg to about 5 mg per dose. Alternatively, the adjuvant or cytokine may be at a concentration of about 0.01 to 50%, preferably at a concentration of about 2% to 30%.

In certain aspects, the pharmaceutical compositions of the disclosure are prepared by physically mixing the adjuvant and/or cytokine with the peptides of the disclosure under appropriate sterile conditions in accordance with known techniques to produce the final product.

Examples of suitable compositions of the invented polypeptide fragments and methods of administration are provided in Esseku and Adeyeye (2011) and Van den Mooter G. (2006). Vaccine and immunotherapy composition preparation is generally described in Vaccine Design ("The subunit and adjuvant approach" (eds Powell M. F. & Newman M. J. (1995) Plenum Press New York). Encapsulation within liposomes, which is also envisaged, is described by Fullerton, US Patent 4,235,877.

In some embodiments, the compositions disclosed herein are prepared as a nucleic acid vaccine. In some embodiments, the nucleic acid vaccine is a DNA vaccine. In some embodiments, DNA vaccines, or gene vaccines, comprise a plasmid with a promoter and

appropriate transcription and translation control elements and a nucleic acid sequence encoding one or more polypeptides of the disclosure. In some embodiments, the plasmids also include sequences to enhance, for example, expression levels, intracellular targeting, or proteasomal processing. In some embodiments, DNA vaccines comprise a viral vector containing a nucleic acid sequence encoding one or more polypeptides of the disclosure. In additional aspects, the compositions disclosed herein comprise one or more nucleic acids encoding peptides determined to have immunoreactivity with a biological sample. For example, in some embodiments, the compositions comprise one or more nucleotide sequences encoding 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more peptides comprising a fragment that is a T cell epitope capable of binding to at least three HLA class I molecules and/or at least three HLA class II molecules of a patient. In some embodiments, the peptides are derived from an antigen that is expressed in cancer. In some embodiments the DNA or gene vaccine also encodes immunomodulatory molecules to manipulate the resulting immune responses, such as enhancing the potency of the vaccine, stimulating the immune system or reducing immunosuppression. Strategies for enhancing the immunogenicity of of DNA or gene vaccines include encoding of xenogeneic versions of antigens, fusion of antigens to molecules that activate T cells or trigger associative recognition, priming with DNA vectors followed by boosting with viral vector, and utilization of immunomodulatory molecules. In some embodiments, the DNA vaccine is introduced by a needle, a gene gun, an aerosol injector, with patches, via microneedles, by abrasion, among other forms. In some forms the DNA vaccine is incorporated into liposomes or other forms of nanobodies. In some embodiments, the DNA vaccine includes a delivery system selected from the group consisting of a transfection agent; protamine; a protamine liposome; a polysaccharide particle; a cationic nanoemulsion; a cationic polymer; a cationic polymer liposome; a cationic nanoparticle; a cationic lipid and cholesterol nanoparticle; a cationic lipid, cholesterol, and PEG nanoparticle; a dendrimer nanoparticle. In some embodiments, the DNA vaccines is administered by inhalation or ingestion. In some embodiments, the DNA vaccine is introduced into the blood, the thymus, the pancreas, the skin, the muscle, a tumor, or other sites.

In some embodiments, the compositions disclosed herein are prepared as an RNA vaccine. In some embodiments, the RNA is non-replicating mRNA or virally derived, self-amplifying RNA. In some embodiments, the non-replicating mRNA encodes the peptides disclosed herein and contains 5' and 3' untranslated regions (UTRs). In some embodiments, the virally derived, self-amplifying RNA encodes not only the peptides disclosed herein but also the viral replication machinery that enables intracellular RNA amplification and abundant protein expression. In some embodiments, the RNA is directly introduced into the individual. In some embodiments, the RNA is chemically synthesized or transcribed *in vitro*. In some embodiments, the mRNA is produced from a linear DNA template using a T7, a T3, or an Sp6 phage RNA polymerase, and the resulting product contains an open reading frame that encodes the peptides disclosed herein, flanking UTRs, a 5' cap, and a poly(A) tail. In some embodiments, various versions of 5' caps are added during or after the transcription reaction using a vaccinia virus capping enzyme or by incorporating synthetic cap or anti-reverse cap analogues. In some embodiments, an optimal length of the poly(A) tail is added to mRNA either directly from the encoding DNA template or by using poly(A) polymerase. The RNA encodes one or more peptides comprising a fragment that is a T cell epitope capable of binding to at least three HLA class I and/or at least three HLA class II molecules of a patient. In some embodiments, the fragments are derived from an antigen that is expressed in cancer. In some embodiments, the RNA includes signals to enhance stability and translation. In some embodiments, the RNA also includes unnatural nucleotides to increase the half-life or modified nucleosides to change the immunostimulatory profile. In some embodiments, the RNAs is introduced by a needle, a gene gun, an aerosol injector, with patches, via microneedles, by abrasion, among other forms. In some forms the RNA vaccine is incorporated into liposomes or other forms of nanobodies that facilitate cellular uptake of RNA and protect it from degradation. In some embodiments, the RNA vaccine includes a delivery system selected from the group consisting of a transfection agent; protamine; a protamine liposome; a polysaccharide particle; a cationic nanoemulsion; a cationic polymer; a cationic polymer liposome; a cationic nanoparticle; a cationic lipid and cholesterol nanoparticle; a cationic lipid, cholesterol, and PEG nanoparticle; a dendrimer

nanoparticle; and/or naked mRNA; naked mRNA with in vivo electroporation; protamine-complexed mRNA; mRNA associated with a positively charged oil-in-water cationic nanoemulsion; mRNA associated with a chemically modified dendrimer and complexed with polyethylene glycol (PEG)-lipid; protamine-complexed mRNA in a PEG-lipid nanoparticle; mRNA associated with a cationic polymer such as polyethylenimine (PEI); mRNA associated with a cationic polymer such as PEI and a lipid component; mRNA associated with a polysaccharide (for example, chitosan) particle or gel; mRNA in a cationic lipid nanoparticle (for example, 1,2-dioleoyloxy-3-trimethylammoniumpropane (DOTAP) or dioleoylphosphatidylethanolamine (DOPE) lipids); mRNA complexed with cationic lipids and cholesterol; or mRNA complexed with cationic lipids, cholesterol and PEG-lipid. In some embodiments, the RNA vaccine is administered by inhalation or ingestion. In some embodiments, the RNA is introduced into the blood, the thymus, the pancreas, the skin, the muscle, a tumor, or other sites, and/or by an intradermal, intramuscular, subcutaneous, intranasal, intranodal, intravenous, intrasplenic, intratumoral or other delivery route.

Polynucleotide or oligonucleotide components may be naked nucleotide sequences or be in combination with cationic lipids, polymers or targeting systems. They may be delivered by any available technique. For example, the polynucleotide or oligonucleotide may be introduced by needle injection, preferably intradermally, subcutaneously or intramuscularly. Alternatively, the polynucleotide or oligonucleotide may be delivered directly across the skin using a delivery device such as particle-mediated gene delivery. The polynucleotide or oligonucleotide may be administered topically to the skin, or to mucosal surfaces for example by intranasal, oral, or intrarectal administration.

Uptake of polynucleotide or oligonucleotide constructs may be enhanced by several known transfection techniques, for example those including the use of transfection agents. Examples of these agents include cationic agents, for example, calcium phosphate and DEAE-Dextran and lipofectants, for example, lipofectam and transfectam. The dosage of the polynucleotide or oligonucleotide to be administered can be altered.

Administration is typically in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to result in a clinical response or to show clinical benefit to the individual, e.g. an effective amount to prevent or delay onset of the disease or condition, to ameliorate one or more symptoms, to induce or prolong remission, or to delay relapse or recurrence.

The dose may be determined according to various parameters, especially according to the substance used; the age, weight and condition of the individual to be treated; the route of administration; and the required regimen. The amount of antigen in each dose is selected as an amount which induces an immune response. A physician will be able to determine the required route of administration and dosage for any particular individual. The dose may be provided as a single dose or may be provided as multiple doses, for example taken at regular intervals, for example 2, 3 or 4 doses administered hourly. Typically peptides, polynucleotides or oligonucleotides are typically administered in the range of 1 pg to 1 mg, more typically 1 pg to 10 μg for particle mediated delivery and 1 μg to 1 mg, more typically 1-100 μg, more typically 5-50 μg for other routes. Generally, it is expected that each dose will comprise 0.01-3 mg of antigen. An optimal amount for a particular vaccine can be ascertained by studies involving observation of immune responses in subjects.

Examples of the techniques and protocols mentioned above can be found in Remington's Pharmaceutical Sciences, 20th Edition, 2000, pub. Lippincott, Williams & Wilkins.

In some cases in accordance with the disclosure, more than one peptide or composition of peptides is administered. Two or more pharmaceutical compositions may be administered together/simultaneously and/or at different times or sequentially. Thus, the disclosure includes sets of pharmaceutical compositions and uses thereof. The use of combination of different peptides, optionally targeting different antigens, is important to overcome the challenges of genetic heterogeneity of tumors and HLA heterogeneity of individuals. The use of peptides of the disclosure in combination expands the group of individuals who can experience clinical benefit from vaccination. Multiple pharmaceutical compositions of peptides of the disclosure, manufactured for use in one regimen, may define a drug product.

Routes of administration include but are not limited to intranasal, oral, subcutaneous, intradermal, and intramuscular. The subcutaneous administration is particularly preferred. Subcutaneous administration may for example be by injection into the abdomen, lateral and anterior aspects of upper arm or thigh, scapular area of back, or upper ventrodorsal gluteal area.

The compositions of the disclosure may also be administered in one, or more doses, as well as, by other routes of administration. For example, such other routes include, intracutaneously, intravenously, intravascularly, intraarterially, intraperitnoeally, intrathecally, intratracheally, intracardially, intralobally, intramedullarly, intrapulmonarily, and intravaginally. Depending on the desired duration of the treatment, the compositions according to the disclosure may be administered once or several times, also intermittently, for instance on a monthly basis for several months or years and in different dosages.

Solid dosage forms for oral administration include capsules, tablets, caplets, pills, powders, pellets, and granules. In such solid dosage forms, the active ingredient is ordinarily combined with one or more pharmaceutically acceptable excipients, examples of which are detailed above. Oral preparations may also be administered as aqueous suspensions, elixirs, or syrups. For these, the active ingredient may be combined with various sweetening or flavoring agents, coloring agents, and, if so desired, emulsifying and/or suspending agents, as well as diluents such as water, ethanol, glycerin, and combinations thereof.

One or more compositions of the disclosure may be administered, or the methods and uses for treatment according to the disclosure may be performed, alone or in combination with other pharmacological compositions or treatments, for example chemotherapy and/or immunotherapy and/or vaccine. The other therapeutic compositions or treatments may for example be one or more of those discussed herein, and may be administered either simultaneously or sequentially with (before or after) the composition or treatment of the disclosure.

In some cases the treatment may be administered in combination with checkpoint blockade therapy, co-stimulatory antibodies, chemotherapy and/or radiotherapy, targeted therapy or monoclonal antibody therapy. It has been demonstrated that chemotherapy sensitizes tumors

to be killed by tumor specific cytotoxic T cells induced by vaccination (Ramakrishnan *et al. J Clin Invest.* 2010; 120(4):1111-1124). Examples for checkpoint inhibitors are CTLA-4 inhibitor, Ipilimumab and programmed cell death-1/programmed cell death ligand-1 (PD-1/PD-L1) signaling inhibitors, Nibolumab, Pembrolizumab, Atezolizumab and Durvalumab. Examples of chemotherapy agents include alkylating agents including nitrogen mustards such as mechlorethamine (HN2), cyclophosphamide, ifosfamide, melphalan (L-sarcolysin) and chlorambucil; anthracyclines; epothilones; nitrosoureas such as carmustine (BCNU), lomustine (CCNU), semustine (methyl-CCNU) and streptozocin (streptozotocin); triazenes such as decarbazine (DTIC; dimethyltriazenoimidazole-carboxamide; ethylenimines/methylmelamines such as hexamethylmelamine, thiotepa; alkyl sulfonates such as busulfan; Antimetabolites including folic acid analogues such as methotrexate (amethopterin); alkylating agents, antimetabolites, pyrimidine analogs such as fluorouracil (5-fluorouracil; 5-FU), floxuridine (fluorodeoxyuridine; FUdR) and cytarabine (cytosine arabinoside); purine analogues and related inhibitors such as mercaptopurine (6-mercaptopurine; 6-MP), thioguanine (6-thioguanine; TG) and pentostatin (2'-deoxycoformycin); epipodophylotoxins; enzymes such as L-asparaginase; biological response modifiers such as IFNα, IL-2, G-CSF and GM-CSF; platinum coordination complexes such as cisplatin (cis-DDP), oxaliplatin and carboplatin; anthracenediones such as mitoxantrone and anthracycline; substituted urea such as hydroxyurea; methylhydrazine derivatives including procarbazine (N-methylhydrazine, MIH) and procarbazine; adrenocortical suppressants such as mitotane (o,p'-DDD) and aminoglutethimide; taxol and analogues/derivatives; hormones and agonists/antagonists including adrenocorticosteroid antagonists such as prednisone and equivalents, dexamethasone and aminoglutethimide, progestin such as hydroxyprogesterone caproate, medroxyprogesterone acetate and megestrol acetate, estrogen such as diethylstilbestrol and ethinyl estradiol equivalents, antiestrogen such as tamoxifen, androgens including testosterone propionate and fluoxymesterone/equivalents, antiandrogens such as flutamide, gonadotropin-releasing hormone analogs and leuprolide and non-steroidal antiandrogens such as flutamide; natural products including vinca alkaloids such as vinblastine (VLB) and vincristine, epipodophyllotoxins such as etoposide and teniposide,

antibiotics such as dactinomycin (actinomycin D), daunorubicin (daunomycin; rubidomycin), doxorubicin, bleomycin, plicamycin (mithramycin) and mitomycin (mitomycin C), enzymes such as L-asparaginase, and biological response modifiers such as interferon alphenomes.

In some cases the method of treatment is a method of vaccination or a method of providing immunotherapy. As used herein, "immunotherapy" is the prevention or treatment of a disease or condition by inducing or enhancing an immune response in an individual. In certain embodiments, immunotherapy refers to a therapy that comprises the administration of one or more drugs to an individual to elicit T cell responses. In a specific embodiment, immunotherapy refers to a therapy that comprises the administration or expression of polypeptides that contain one or more PEPIs to an individual to elicit a T cell response to recognize and kill cells that display the one or more PEPIs on their cell surface in conjunction with a class I HLAs. In another specific embodiment, immunotherapy comprises the administration of one or more PEPIs to an individual to elicit a cytotoxic T cell response against cells that display tumor associated antigens (TAAs) or cancer testis antigens (CTAs) comprising the one or more PEPIs on their cell surface. In another embodiment, immunotherapy refers to a therapy that comprises the administration or expression of polypeptides that contain one or more PEPIs presented by class II HLAs to an individual to elicit a T helper response to provide co-stimulation to cytotoxic T cells that recognize and kill diseased cells that display the one or more PEPIs on their cell surface in conjunction with a class I HLAs. In still another specific embodiment, immunotherapy refers to a therapy that comprises administration of one or more drugs to an individual that re-activate existing T cells to kill target cells. The theory is that the cytotoxic T cell response will eliminate the cells displaying the one or more PEPIs, thereby improving the clinical condition of the individual. In some instances, immunotherapy may be used to treat tumors. In other instances, immunotherapy may be used to treat intracellular pathogen-based diseases or disorders.

In some cases the disclosure relates to the treatment of cancer or the treatment of solid tumors. In some cases the treatment is of breast cancer, ovarian cancer or colorectal cancer. In other cases the treatment may be of any other cancer or solid tumor that expresses a target tumor associated antigen of the present peptides as described herein, or any cancer in which such target

polypeptide antigens are expressed in some or a high percentage of subjects. The treatment may be of cancers or malignant or benign tumors of any cell, tissue, or organ type. The cancer may or may not be metastatic. Exemplary cancers include carcinomas, sarcomas, lymphomas, leukemias, germ cell tumors, or blastomas. The cancer may or may not be a hormone related or dependent cancer (*e.g.*, an estrogen or androgen related cancer).

Selection of polypeptides and patients

Specific polypeptide antigens, and particularly short peptides derived from such antigens that are commonly used in vaccination and immunotherapy, induce immune responses in only a fraction of human subjects. The polypeptides of the present disclosure are specifically selected to induce immune responses in a high proportion of the general population, but they may not be effective in all individuals due to HLA genotype heterogeneity. HLA genotype population heterogeneity means that the immune or clinical response rate to the vaccines described herein will differ between different human subpopulations. In some cases the vaccines described herein are for use to treat a specific or target subpopulation, for example an Asian population, or a Vietnamese, Chinese, and/or Japanese population.

The disclosure also provides a method of identifying a human subject who will likely have a cytotoxic T cell response to administration of a pharmaceutical composition comprising a peptide of the disclosure (likely responders), or of predicting the likelihood that a subject will have a cytotoxic T cell response.

As provided herein T cell epitope presentation by multiple HLAs of an individual is generally needed to trigger a T cell response. The best predictor of a cytotoxic T cell response to a given polypeptide, as determined by the inventors, is the presence of at least one T cell epitope that is presented by three or more HLA class I of an individual (≥1 PEPI3+). Accordingly the presence within the active ingredient peptides of a pharmaceutical composition of one or more T cell epitopes that is capable of binding to at least three HLA of a subject is predictive for the subject having a cytotoxic T cell response to administration of the pharmaceutical composition. The subject is a likely immune responder.

In some cases the T cell epitope that is capable of binding to at least three HLA class I of the subject has the amino acid sequence of any one of SEQ ID NOs: 1 to 40, or SEQ ID NOs: 1 to 40, 172-194, 234-250 and 272-301. In other cases the T cell epitope may have a different amino acid sequence within the one or more peptides of the pharmaceutical composition.

The inventors have further discovered that the presence in a vaccine or immunotherapy composition of at least two epitopes that can bind to at least three HLA of an individual is predictive for a clinical response. In other words, if an individual has a total of ≥2 PEPI3+ within the active ingredient polypeptide(s) of a vaccine or immunotherapy composition, and these PEPI3+s are derived from antigen sequences that are in fact expressed in the individual (for example, target tumor cells of the individual express the target tumor-associated antigens), then the individual is a likely clinical responder (i.e. a clinically relevant immune responder).

Accordingly some aspects of the disclosure relate to a method of identifying a subject who will likely have a clinical response to a method of treatment according to the disclosure, or of predicting the likelihood that a subject will have a clinical response. A "clinical response" or "clinical benefit" as used herein may be the prevention or a delay in the onset of a disease or condition, the amelioration of one or more symptoms, the induction or prolonging of remission, or the delay of a relapse or recurrence or deterioration, or any other improvement or stabilisation in the disease status of a subject. Where appropriate, a "clinical response" may correlate to "disease control" or an "objective response" as defined by the Response Evaluation Criteria In Solid Tumors (RECIST) guidelines.

In some embodiments the method comprises determining that one or more cancer-associated antigens selected from SPAG9, AKAP-4, BORIS, NY-SAR-35, NY-BR-1, SURVIVIN, MAGE-A11, PRAME, MAGE-A9, HOM-TES-85, TSP50, EpCAM, CAGE1, FBXO39, MAGE-A8 and MAGE-A6 is expressed by a cancer. For example expression of the cancer associated antigen may be detected in a sample obtained from the subject, for example a tumor biopsy, using methods that are known in the art.

The inventors have discovered that it is not sufficient that a vaccine or immunotherapy composition targets an antigen that is expressed by cancer or tumor cells of a patient, nor that the

target sequences of that antigen can bind to HLA class I of the patient (HLA restricted epitopes). The composition is likely effective only in patients that both express the target antigen and have three or more HLA class I that bind to a single T cell epitope of the target antigen. Moreover, as described above, at least two epitopes that binds to at least 3 HLAs of the patient are generally needed to induce a clinically relevant immune response.

Therefore the method further comprises determining that the active ingredient peptide(s) of the pharmaceutical composition comprise two or more different amino acid sequences each of which is a) a fragment of a cancer-associated antigen expressed by cancer cells of the subject, determined as described above; and b) a T cell epitope capable of binding to at least three HLA class I of the subject.

In some cases the T cell epitope that is capable of binding to at least three HLA class I of the subject has the amino acid sequence of any one of SEQ ID NOs: 1 to 40, or SEQ ID NOs: 1 to 40, 172-194, 234-250 and 272-301. In other cases the T cell epitope may have a different amino acid sequence within the one or more peptides of the pharmaceutical composition.

In some cases the likelihood that a subject will have a clinical response to a peptide vaccine or immunotherapy composition, such as those described herein, can be determined without knowing whether the target antigens are expressed in cancer or tumor cells of the subject and/or without determining the HLA class I genotype of the subject. Known antigen expression frequencies in the disease (e.g. MAGE-A3 in a tumor type like breast or colorectal cancer) and/or known frequencies for HLA class I and class II genotype of subjects in the target population (e.g ethnic population, general population, diseased population) may be used instead. Moreover by combining peptides that target the most frequently presented PEPIs across the population (BestEPIs) in multiple frequently expressed target antigens in the disease, as identified and described herein, it is possible to design a cancer vaccine regime that is effective for a high proportion of patients. However, using the companion diagnostic methods described herein to pre-select patients who are most likely to have a clinical response will increase clinical response rates amongst treated patients.

The likelihood that a subject will respond to treatment is increased by (i) the presence of more multiple HLA-binding PEPIs in the active ingredient polypeptides; (ii) the presence of PEPIs in more target polypeptide antigens; and (iii) expression of the target polypeptide antigens in the subject or in diseased cells of the subject. In some cases expression of the target polypeptide antigens in the subject may be known, for example if target polypeptide antigens are in a sample obtained from the subject. In other cases, the probability that a specific subject, or diseased cells of a specific subject, (over-)express a specific or any combination of target polypeptide antigens may be determined using population expression frequency data, e.g. probability of expression of an antigen in breast cancer, colorectal cancer or ovarian cancer. The population expression frequency data may relate to a subject- and/or disease-matched population or the intent-to-treat population. For example, the frequency or probability of expression of a particular cancer-associated antigen in a particular cancer or subject having a particular cancer, for example breast cancer, can be determined by detecting the antigen in tumor, e.g. breast cancer tumor samples. In some cases such expression frequencies may be determined from published figures and scientific publications. In some cases a method of the disclosure comprises a step of determining the expression frequency of a relevant target polypeptide antigen in a relevant population.

Disclosed is a range of pharmacodynamic biomarkers to predict the activity/effect of vaccines in individual human subjects as well as in populations of human subjects. These biomarkers expedite more effective vaccine development and also decrease the development cost and may be used to assess and compare different compositions. Exemplary biomarkers are as follows.

- **AG95 – potency of a vaccine**: The number of antigens in a cancer vaccine that a specific tumor type expresses with 95% probability. AG95 is an indicator of the vaccine's potency, and is independent of the immunogenicity of the vaccine antigens. AG95 is calculated from the tumor antigen expression rate data. Such data may be obtained from experiments published in peer reviewed scientific journals. Technically, AG95 is determined from the binomial distribution of antigens in the vaccine, and takes into account all possible variations and expression rates.

- **PEPI3+ count – immunogenicity of a vaccine in a subject**: Vaccine-derived PEPI3+ are personal epitopes that bind to et least 3 HLAs of a subject and induce T cell responses. PEPI3+ can be determined using the PEPI3+ Test in subjects who's complete 4-digit HLA genotype is known.
- **AP count – antigenicity of a vaccine in a subject**: Number of vaccine antigens with PEPI3+. Vaccines contain sequences from target polypeptide antigens expressed by diseased cells. AP count is the number of antigens in the vaccine that contain PEPI3+, and the AP count represents the number of antigens in the vaccine that can induce T cell responses in a subject. AP count characterizes the vaccine-antigen specific T cell responses of the subject since it depends only on the HLA genotype of the subject and is independent of the subject's disease, age, and medication. The correct value is between 0 (no PEPI presented by the antigen) and maximum number of antigens (all antigens present PEPIs).
- **AP50 – antigenicity of a vaccine in a population**: The mean number of vaccine antigens with a PEPI in a population. The AP50 is suitable for the characterization of vaccine-antigen specific T cell responses in a given population since it depends on the HLA genotype of subjects in a population.
- **AGP count – effectiveness of a vaccine in a subject**: Number of vaccine antigens expressed in the tumor with PEPI. The AGP count indicates the number of tumor antigens that vaccine recognizes and induces a T cell response against (hit the target). The AGP count depends on the vaccine-antigen expression rate in the subject's tumor and the HLA genotype of the subject. The correct value is between 0 (no PEPI presented by expressed antigen) and maximum number of antigens (all antigens are expressed and present a PEPI).
- **AGP50 – effectiveness of a cancer vaccine in a population**: The mean number of vaccine antigens expressed in the indicated tumor with PEPI (i.e., AGP) in a population. The AGP50 indicates the mean number of tumor antigens that the T cell responses induced by the vaccine can recognize. AGP50 is dependent on the expression rate of the antigens in the indicated tumor type and the immunogenicity of the antigens in the target population. AGP50 can estimate a vaccine's effectiveness in different populations and can be used to compare different vaccines in the same population. The computation of AGP50 is similar to that used for AG50, except the expression is weighted by the occurrence of the PEPI3+ in the subject on the expressed vaccine antigens. In a theoretical population, where each subject has a PEPI

from each vaccine antigen, the AGP50 will be equal to AG50. In another theoretical population, where no subject has a PEPI from any vaccine antigen, the AGP50 will be 0. In general, the following statement is valid: $0 \leq AGP50 \leq AG50$.

- **mAGP – a candidate biomarker for the selection of likely responders:** Likelihood that a cancer vaccine induces T cell responses against multiple antigens expressed in the indicated tumor. mAGP is calculated from the expression rates of vaccine-antigens in the tumor and the presence of vaccine derived PEPIs in the subject. Technically, based on the AGP distribution, the mAGP is the sum of probabilities of the multiple AGP (≥2 AGPs).

The results of a prediction as set out above may be used to inform a physician's decisions concerning treatment of the subject. Accordingly, in some cases the method of the disclosure predicts that a subject will have or is likely to have a T cell response and/or a clinical response to a treatment as described herein, and the method further comprises selecting the treatment for the human subject. In some cases a subject is selected for treatment if their likelihood of a response targeted at a predefined number of target polypeptide antigens, optionally wherein the target polypeptide antigens are (predicted to be) expressed, is above a predetermined threshold. In some cases the number of target polypeptide antigens or epitopes is two. In some cases the number of target polypeptide antigens or epitopes is three, or four, or five, or six, or seven, or eight, or nine, or ten. The method may further comprise administering the treatment to the human subject. Alternatively, the method may predict that the subject will not have an immune response and/or a clinical response and further comprise selecting a different treatment for the subject.

Further embodiments of the disclosure – (1)

1. A pharmaceutical composition comprising one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 112 to 142.
2. The pharmaceutical composition of item 1, comprising 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, or 6 or more peptides.
3. The pharmaceutical composition of item 1, comprising two peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 121 and 124.

4. The pharmaceutical composition of item 1, comprising four peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 126, 130, 131, and 134.
5. The pharmaceutical composition of item 1, comprising six peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 121, 124, 126, 130, 131, and 134.
6. The pharmaceutical composition of item 5, further comprising at least one additional peptide comprising a fragment of an antigen selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, and MAGE-A6.
7. The pharmaceutical composition of item 5, further comprising one or more additional peptides, each of the one or more additional peptides comprising a different one of the amino acid sequence of any one of SEQ ID NOs: 112-120, 122, 123, 125, 127-129, 132, 133, and 135-142.
8. The pharmaceutical composition of item 1, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
9. The pharmaceutical composition of item 8, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
10. A pharmaceutical composition comprising one or more nucleic acid molecules encoding one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 112 to 142.
11. A method of identifying and treating a human subject having cancer who will likely have a clinical response to administration of a pharmaceutical composition according to item 1, the method comprising

(i) assaying a biological sample of the subject to determine HLA genotype of the subject;

(ii) determining that the pharmaceutical composition comprises two or more sequences that are a T cell epitope capable of binding to at least three HLA class I molecules of the subject;

(iii) determining the probability that a tumor of the subject expresses one or more antigen corresponding to the T cell epitopes identified in step (ii) using population expression data for each antigen, to identify the likelihood of the subject to have a clinical response to administration of the pharmaceutical composition; and

(iv) administering the composition of item 1 to the identified subject.

12. The method of item 11, wherein the subject has colorectal cancer.

13. The method of item 11, wherein the pharmaceutical composition comprises 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, or 6 or more peptides.

14. The method of item 11, wherein the pharmaceutical composition comprises two peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 121 and 124.

15. The method of item 11, wherein the pharmaceutical composition comprises four peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 126, 130, 131, and 134.

16. The method of item 11, wherein the pharmaceutical composition comprises six peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 121, 124, 126, 130, 131, and 134.

17. The method of item 11, wherein the pharmaceutical composition further comprises at least one additional peptide comprising a fragment of an antigen selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, and MAGE-A6.

18. The method of item 11, wherein the pharmaceutical composition further comprises one or more additional peptides, each of the one or more additional peptides comprising a different one of the amino acid sequence of any one of SEQ ID NOs: 112-120, 122, 123, 125, 127-129, 132, 133, and 135-142.

19. The method of item 11, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.

20. The method of item 19, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

21. The method of item 11, further comprising administering a chemotherapeutic agent, a checkpoint inhibitor, a targeted therapy, radiation therapy, another immunotherapy, or combination thereof to the identified subject.

22. The method of item 13, further comprising prior to the administering step,

(i) assaying a tumor sample from the subject to determine that the three or more peptides of the pharmaceutical composition comprise two or more different amino acid sequences each of which is

a. a fragment of a cancer-associated antigen expressed by cancer cells of the subject as determined in step (i); and

b. a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(ii) confirming the subject as likely to have a clinical response to the method of treatment.

Further embodiments of the disclosure – (2)

Breast Cancer

1. A pharmaceutical composition comprising one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 81 to 111 and 435 to 449.

2. The pharmaceutical composition of item 1, comprising 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, 6 or more peptides, 7 or more peptides, 8 or more peptides, 9 or more peptides, 10 or more peptides, 11 or more peptides, or 12 or more peptides.

3. The pharmaceutical composition of item 1, comprising 9 peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 92, 93, 98, 99-101, and 103-105.

4. The pharmaceutical composition of item 1, further comprising at least one additional peptide comprising a fragment of an antigen selected from PIWIL-2, AKAP-4, EpCAM, BORIS, HIWI, SPAG9, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, PRAME, NY-SAR-35, MAGE-A9, NY-BR-1, SURVIVIN, MAGE-A11, HOM-TES-85 and NY-ESO-1.

5. The pharmaceutical composition of item 4, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194.

6. The pharmaceutical composition of item 4, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs:41-60 and 195-233.

7. The pharmaceutical composition of item 1, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.

8. The pharmaceutical composition of item 7, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

9. A pharmaceutical composition comprising one or more nucleic acid molecules encoding one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 81 to 111 and 435 to 449.

10. The pharmaceutical composition of item 9, wherein the one or more nucleic acid molecules encode 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, 6 or more peptides, 7 or more peptides, 8 or more peptides, 9 or more peptides, 10 or more peptides, 11 or more peptides, or 12 or more peptides.

11. The pharmaceutical composition of item 9, wherein the one or more nucleic acid molecules encode 9 peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 92, 93, 98, 99-101, and 103-105.

12. The pharmaceutical composition of item 9, wherein the one or more nucleic acid molecules encode at least one additional peptide comprising a fragment of an antigen selected from PIWIL-2, AKAP-4, EpCAM, BORIS, HIWI, SPAG9, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, PRAME, NY-SAR-35, MAGE-A9, NY-BR-1, SURVIVIN, MAGE-A11, HOM-TES-85 and NY-ESO-1.

13. The pharmaceutical composition of item 12, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194.

14. The pharmaceutical composition of item 12, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs:41-60 and 195-233.

15. The pharmaceutical composition of item 9, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.

16. The pharmaceutical composition of item 15, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

17. A method of identifying and treating a human subject having cancer who will likely have a clinical response to administration of a pharmaceutical composition according to item 1, the method comprising

(i) assaying a biological sample of the subject to determine HLA genotype of the subject;
(ii) determining that the pharmaceutical composition comprises two or more sequences that are a T cell epitope capable of binding to at least three HLA class I molecules of the subject;
(iii) determining the probability that a tumor of the subject expresses one or more antigen corresponding to the T cell epitopes identified in step (ii) using population expression data for each antigen, to identify the likelihood of the subject to have a clinical response to administration of the pharmaceutical composition; and
(iv) administering the composition of item 1 to the identified subject.

18. The method of item 17, wherein the subject has breast cancer.

19. The method of item 17, wherein the pharmaceutical composition comprises 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, 6 or more peptides, 7 or more peptides, 8 or more peptides, 9 or more peptides, 10 or more peptides, 11 or more peptides, or 12 or more peptides.

20. The method of item 17, wherein the pharmaceutical composition comprises 9 peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 92, 93, 98, 99-101, and 103-105.
21. The method of item 17, wherein the pharmaceutical composition further comprises comprising at least one additional peptide comprising a fragment of an antigen selected from PIWIL-2, AKAP-4, EpCAM, BORIS, HIWI, SPAG9, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, PRAME, NY-SAR-35, MAGE-A9, NY-BR-1, SURVIVIN, MAGE-A11, HOM-TES-85 and NY-ESO-1.
22. The method of item 21, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194.
23. The method of item 21, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs:41-60 and 195-233.
24. The method of item 17, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
25. The method of item 24, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
26. The method of item 17, further comprising administering a chemotherapeutic agent, a checkpoint inhibitor, a targeted therapy, radiation therapy, another immunotherapy, or combination thereof to the identified subject.
27. The method of item 17, further comprising prior to the administering step,

(iii) assaying a tumor sample from the subject to determine that the three or more peptides of the pharmaceutical composition comprise two or more different amino acid sequences each of which is

a. a fragment of a cancer-associated antigen expressed by cancer cells of the subject as determined in step (i); and

b. a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(iv) confirming the subject as likely to have a clinical response to the method of treatment.

28. A method of identifying and treating a human subject having cancer who will likely have an immune response to administration of a pharmaceutical composition according to item 1, the method comprising

(i) assaying a biological sample of the subject to determine HLA genotype of the subject;

(ii) determining that the pharmaceutical composition comprises one or more sequences that are a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(iii) administering the composition of item 1 to the identified subject.

29. A kit comprising:

a. a first pharmaceutical composition comprising one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 81-111 and 435 to 449; and

b. a second different pharmaceutical composition comprising one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 81-111 and 435 to 449.

30. A pharmaceutical composition comprising: a nucleic acid molecule expressing two or more polypeptides, each polypeptide comprising a fragment of up to 50 consecutive

amino acids of an antigen selected from PIWIL-2, AKAP-4, EpCAM, BORIS, HIWI, SPAG9, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, PRAME, NY-SAR-35, MAGE-A9, NY-BR-1, SURVIVIN, MAGE-A11, HOM-TES-85 and NY-ESO-1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24, and 172 to 194.

31. A pharmaceutical composition comprising one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 332-346.

32. The pharmaceutical composition of item 31, comprising 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, 6 or more peptides, 7 or more peptides, 8 or more peptides, 9 or more peptides, 10 or more peptides, 11 or more peptides, 12 or more peptides, 13 or more peptides, 14 or more peptides, or 15 or more peptides.

33. The pharmaceutical composition of item 31, comprising 15 peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 332-346.

34. The pharmaceutical composition of item 31, further comprising at least one additional peptide comprising a fragment of an antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN, and AKAP-3.

35. The pharmaceutical composition of item 34, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 272-301.

36. The pharmaceutical composition of item 34, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs:302-331.

37. The pharmaceutical composition of item 31, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.

38. The pharmaceutical composition of item 37, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene

(DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

39. A pharmaceutical composition comprising one or more nucleic acid molecules encoding one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 332-346.

40. The pharmaceutical composition of item 39, wherein the one or more nucleic acid molecules encode 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, 6 or more peptides, 7 or more peptides, 8 or more peptides, 9 or more peptides, 10 or more peptides, 11 or more peptides, 12 or more peptides, 13 or more peptides, 14 or more peptides, or 15 or more peptides.

41. The pharmaceutical composition of item 39, wherein the one or more nucleic acid molecules encode 15 peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 332-346.

42. The pharmaceutical composition of item 39, wherein the one or more nucleic acid molecules encode at least one additional peptide comprising a fragment of an antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN, and AKAP-3.

43. The pharmaceutical composition of item 42, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 272-301.

44. The pharmaceutical composition of item 42, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs:302-331.

45. The pharmaceutical composition of item 39, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.

46. The pharmaceutical composition of item 45, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

47. A method of identifying and treating a human subject having cancer who will likely have a clinical response to administration of a pharmaceutical composition according to item 28, the method comprising

(i) assaying a biological sample of the subject to determine HLA genotype of the subject;

(ii) determining that the pharmaceutical composition comprises two or more sequences that are a T cell epitope capable of binding to at least three HLA class I molecules of the subject;

(iii) determining the probability that a tumor of the subject expresses one or more antigen corresponding to the T cell epitopes identified in step (ii) using population expression data for each antigen, to identify the likelihood of the subject to have a clinical response to administration of the pharmaceutical composition; and

(iv) administering the composition of item 28 to the identified subject.

48. The method of item 47, wherein the subject has ovarian cancer.

49. The method of item 47, wherein the pharmaceutical composition comprises 2 or more peptides, 3 or more peptides, 4 or more peptides, 5 or more peptides, 6 or more peptides, 7 or more peptides, 8 or more peptides, 9 or more peptides, 10 or more peptides, 11 or more peptides, 12 or more peptides, 13 or more peptides, 14 or more peptides, or 15 or more peptides.

50. The method of item 47, wherein the pharmaceutical composition comprises 15 peptides, wherein each peptide comprises a different one of the amino acid sequences of SEQ ID NOs: 332-346.

51. The method of item 47, wherein the pharmaceutical composition further comprises comprising at least one additional peptide comprising a fragment of an antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN, and AKAP-3.

52. The method of item 51, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 272-301.

53. The method of item 51, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs:302-331.

54. The method of item 47, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.

55. The method of item 54, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosamide, bacillus Calmette-Guerin (BCG), corynbacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenezene (DNCB), keyhole limpet hemocyanins (KLH), Freunds adjuvant (complete), Freunds adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

56. The method of item 47, further comprising administering a chemotherapeutic agent, a checkpoint inhibitor, a targeted therapy, radiation therapy, another immunotherapy, or combination thereof to the identified subject.

57. The method of item 47, further comprising prior to the administering step,

assaying a tumor sample from the subject to determine that the three or more peptides of the pharmaceutical composition comprise two or more different amino acid sequences each of which is

a. a fragment of a cancer-associated antigen expressed by cancer cells of the subject as determined in step (i); and

b. a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

confirming the subject as likely to have a clinical response to the method of treatment.

58. A method of identifying and treating a human subject having cancer who will likely have an immune response to administration of a pharmaceutical composition according to item 31, the method comprising

(i) assaying a biological sample of the subject to determine HLA genotype of the subject;

(ii) determining that the pharmaceutical composition comprises one or more sequences that are a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(iii) administering the composition of item 31 to the identified subject.

59. A kit comprising:

a. a first pharmaceutical composition comprising one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 332-346; and

b. a second different pharmaceutical composition comprising one or more peptides, wherein each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 332-346.

60. A pharmaceutical composition comprising: a nucleic acid molecule expressing two or more polypeptides, each polypeptide comprising a fragment of up to 50 consecutive amino acids of an antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN, and AKAP-3,

wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 272-301.

Examples

Example 1 – HLA-epitope binding prediction process and validation

Predicted binding between particular HLA and epitopes (9 mer peptides) was based on the Immune Epitope Database tool for epitope prediction (www.iedb.org).

The HLA I-epitope binding prediction process was validated by comparison with HLA I-epitope pairs determined by laboratory experiments. A dataset was compiled of HLA I-epitope pairs reported in peer reviewed publications or public immunological databases.

The rate of agreement with the experimentally determined dataset was determined (Table 2). The binding HLA I-epitope pairs of the dataset were correctly predicted with a 93% probability. Coincidentally the non-binding HLA I-epitope pairs were also correctly predicted with a 93% probability.

Table 2. Analytical specificity and sensitivity of the HLA-epitope binding prediction process.

HLA-epitope pairs	*True epitopes (n=327)* *(Binder match)*	*False epitopes (n=100)* *(Non-binder match)*
HIV	91% (32)	82% (14)
Viral	100% (35)	100% (11)
Tumor	90% (172)	94% (32)
Other (fungi, bacteria, etc.)	100% (65)	95% (36)
All	*93% (304)*	*93% (93)*

The accuracy of the prediction of multiple HLA binding epitopes was determined. Based on the analytical specificity and sensitivity using the 93% probability for both true positive and true negative prediction and 7% (=100% - 93%) probability for false positive and false negative prediction, the probability of the existence of a multiple HLA binding epitope in a person can be calculated. The probability of multiple HLA binding to an epitope shows the relationship between the number of HLAs binding an epitope and the expected minimum number of real binding. Per PEPI definition three is the expected minimum number of HLA to bind an epitope (bold).

Table 3. Accuracy of multiple HLA binding epitopes predictions.

Expected minimum number of real HLA binding	Predicted number of HLAs binding to an epitope						
	0	1	2	3	4	5	6
1	35%	95%	100%	100%	100%	100%	100%
2	6%	29%	90%	99%	100%	100%	100%
3	1%	4%	22%	84%	98%	100%	100%
4	0%	0%	2%	16%	78%	96%	99%
5	0%	0%	0%	1%	10%	71%	94%
6	0%	0%	0%	0%	0%	5%	65%

The validated HLA-epitope binding prediction process was used to determine all HLA-epitope binding pairs described in the Examples below.

Example 2 – Epitope presentation by multiple HLA predicts cytotoxic T lymphocyte (CTL) response

The presentation of one or more epitopes of a polypeptide antigen by one or more HLA I of an individual is predictive for a CTL response was determined.

The study was carried out by retrospective analysis of six clinical trials, conducted on 71 cancer and 9 HIV-infected patients (Table 4)[1-7]. Patients from these studies were treated with an HPV vaccine, three different NY-ESO-1 specific cancer vaccines, one HIV-1 vaccine and a CTLA-4 specific monoclonal antibody (Ipilimumab) that was shown to reactivate CTLs against

NY-ESO-1 antigen in melanoma patients. All of these clinical trials measured antigen specific CD8+ CTL responses (immunogenicity) in the study subjects after vaccination. In some cases, correlation between CTL responses and clinical responses were reported.

No patient was excluded from the retroactive study for any reason other than data availability. The 157 patient datasets (Table 4) were randomized with a standard random number generator to create two independent cohorts for training and evaluation studies. In some cases the cohorts contained multiple datasets from the same patient, resulting in a training cohort of 76 datasets from 48 patients and a test/validation cohort of 81 datasets from 51 patients.

Table 4. Summary of patient datasets

Clinical trial	Immunotherapy	Target Antigen	Disease	# Patients*	# Data sets (#antigen x #patient)	Immunoassay performed in the clinical trials**	HLA genotyping method	Ref
1	VGX-3100	HPV16-E6 HPV16-E7 HPV18-E6 HPV18-E7 HPV16/18	Cervical cancer	17/18	5 x 17	IFN-γ ELISPOT	High Resolution SBT	1
2	HIVIS vaccine	HIV-1 Gag HIV-1 RT	AIDS	9/12	2 x 9	IFN-γ ELISPOT	Low-Medium Resolution SSO	2
3	rNY-ESO-1	NY-ESO-1	Breast-and ovarian cancers, melanoma and sarcoma	18/18	1 x 18	In vitro and Ex vivo IFN-γ ELISPOT	High Resolution SBT	3 4
4	Ipilimumab	NY-ESO-1	Metastatic melanoma	19/20	1 x 19	ICS after T-cell stimulation	Low to medium resolution typing, SSP of genomic DNA, high resolution sequencing	5

5	NY-ESO-1f	NY-ESO-1 (91-110)	Esophageal-, non-small-cell lung- and gastric cancer	10/10	1 x 10	ICS after T-cell stimulation	SSO probing and SSP of genomic DNA	6
6	NY-ESO-1 overlapping peptides	NY-ESO-1 (79-173)	Esophageal- and lung cancer, malignant melanoma	7/9	1 x 7	ICS after T-cell stimulation	SSO probing and SSP of genomic DNA	7
Total	6	7		80	157	N/A		

*Number of patients used in the retrospective analysis from the original number of patient of the clinical trials.
**Immunoassays are based on T cell stimulation with antigen-specific peptide pools and quantify the released cytokines by different techniques.
CT: Clinical trial; SBT: Sequence Based Typing; SSO: Sequence-Specific Oligonucleotide; ICS: Intracellular cytokine staining; SSP: Sequence-specific priming

The reported CTL responses of the training dataset were compared with the HLA I restriction profile of epitopes (9 mers) of the vaccine antigens. The antigen sequences and the HLA I genotype of each patient were obtained from publicly available protein sequence databases or peer reviewed publications and the HLA I-epitope binding prediction process was blinded to patients' clinical CTL response data. The number of epitopes from each antigen predicted to bind to at least 1 (PEPI1+), or at least 2 (PEPI2+), or at least 3 (PEPI3+), or at least 4 (PEPI4+), or at least 5 (PEPI5+), or all 6 (PEPI6) HLA class I molecules of each patient was determined and the number of HLA bound were used as classifiers for the reported CTL responses. The true positive rate (sensitivity) and true negative rate (specificity) were determined from the training dataset for each classifier (number of HLA bound) separately.

ROC analysis was performed for each classifier. In a ROC curve, the true positive rate (Sensitivity) was plotted in function of the false positive rate (1-Specificity) for different cut-off points (FIG. 1). Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold (epitope (PEPI) count). The area under the ROC curve (AUC) is a measure of how well the classifier can distinguish between two diagnostic groups (CTL responder or non-responder).

The analysis unexpectedly revealed that predicted epitope presentation by multiple class I HLAs of a subject (PEPI2+, PEPI3+, PEPI4+, PEPI5+, or PEPI6), was in every case a better predictor of CTL response than epitope presentation by merely one or more HLA class I (PEPI1+, AUC = 0.48, Table 5).

Table 5. Determination of diagnostic value of the PEPI biomarker by ROC analysis

Classifiers	AUC
PEPI1+	0.48
PEPI2+	0.51
PEPI3+	0.65
PEPI4+	0.52
PEPI5+	0.5
PEPI6+	0.5

The CTL response of an individual was best predicted by considering the epitopes of an antigen that could be presented by at least 3 HLA class I of an individual (PEPI3+, AUC = 0.65, Table 5). The threshold count of PEPI3+ (number of antigen-specific epitopes presented by 3 or more HLA of an individual) that best predicted a positive CTL response was 1 (Table 6). In other words, at least one antigen-derived epitope is presented by at least 3 HLA class I of a subject (≥1 PEPI3+), then the antigen can trigger at least one CTL clone, and the subject is a likely CTL responder. Using the ≥1 PEPI3+ threshold to predict likely CTL responders ("≥1 PEPI3+ Test") provided 76% diagnostic sensitivity (Table 12).

Table 6. Determination of the ≥1 PEPI3+ threshold to predict likely CTL responders in the training dataset.

	PEPI3+ Count											
	1	2	3	4	5	6	7	8	9	10	11	12
Sensitivity:	**0.76**	0.60	0.31	0.26	0.14	0.02	0	0	0	0	0	0

1-Specificity:	**0.59**	0.24	0.21	0.15	0.09	0.06	0.06	0.03	0.03	0.03	0.03	0.03

Example 3 – Validation of the ≥1 PEPI3+ Test

The test cohort of 81 datasets from 51 patients was used to validate the ≥1 PEPI3+ threshold to predict an antigen-specific CTL response. For each dataset in the test cohort it was determined whether the ≥1 PEPI3+ threshold was met (at least one antigen-derived epitope presented by at least three class I HLA of the individual). This was compared with the experimentally determined CTL responses reported from the clinical trials (Table 7).

The clinical validation demonstrated that a PEPI3+ peptide induce CTL response in an individual with 84% probability. 84% is the same value that was determined in the analytical validation of the PEPI3+ prediction, epitopes that binds to at least 3 HLAs of an individual (Table 3). These data provide strong evidences that immune responses are induced by PEPIs in individuals.

Table 7. Diagnostic performance characteristics of the ≥1 PEPI3+ Test (n=81).

Performance characteristic		Description	Result
Positive predictive value (PPV)	100%[A/(A + B)]	The likelihood that an individual that meets the ≥1 PEPI3+ threshold has antigen-specific CTL responses after treatment with immunotherapy.	*84%*
Sensitivity	100%[A / (A+C)]	The proportion of subjects with antigen-specific CTL responses after treatment with immunotherapy who meet the ≥1 PEPI3+ threshold.	*75%*
Specificity	100%[D / (B + D)]	The proportion of subjects without antigen-specific CTL responses after treatment with immunotherapy who do not meet the ≥1 PEPI3+ threshold.	*55%*
Negative predictive value (NPV)	100%[D/(C +D)]	The likelihood that an individual who does not meet the ≥1 PEPI3+ threshold does not have antigen-specific CTL responses after treatment with immunotherapy.	*42%*

Overall percent agreement (OPA)	100%[(A + D)/ N]	The percentage of predictions based on the ≥1 PEPI3+ threshold that match the experimentally determined result, whether positive or negative.	*70%*
Fisher's exact (p)			*0.01*

ROC analysis determined the diagnostic accuracy, using the PEPI3+ count as cut-off values (Fig. 2). The AUC value = 0.73. For ROC analysis an AUC of 0.7 to 0.8 is generally considered as fair diagnostic.

A PEPI3+ count of at least 1 (≥1 PEPI3+) best predicted a CTL response in the test dataset (Table 8). This result confirmed the threshold determined during the training (Table 5).

Table 8. Confirmation of the ≥1 PEPI3+ threshold to predict likely CTL responders in the test/validation dataset.

	PEPI3+ Count											
	1	2	3	4	5	6	7	8	9	10	11	12
Sensitivity:	**0.75**	0.52	0.26	0.23	0.15	0.13	0.08	0.05	0	0	0	0
1-Specificity:	**0.45**	0.15	0.05	0	0	0	0	0	0	0	0	0

Example 4 – The ≥1 PEPI3+ Test predicts CD8+ CTL reactivities

The ≥1 PEPI3+ Test was compared with a previously reported method for predicting a specific human subject's CTL response to peptide antigens.

The HLA genotypes of 28 cervical cancer and VIN-3 patients that received the HPV-16 synthetic long peptide vaccine (LPV) in two different clinical trials were determined from DNA samples[8 8 9 10]. The LPV consists of long peptides covering the HPV-16 viral oncoproteins E6 and E7. The amino acid sequence of the LPV was obtained from these publications. The publications also report the T cell responses of each vaccinated patient to pools of overlapping peptides of the vaccine.

For each patient epitopes (9 mers) of the LPV that are presented by at least three patient class I HLA (PEPI3+s) were identified and their distribution among the peptide pools was

determined. Peptides that comprised at least one PEPI3+ (≥1 PEPI3+) were predicted to induce a CTL response. Peptides that comprised no PEPI3+ were predicted not to induce a CTL response.

The ≥1 PEPI3+ Test correctly predicted 489 out of 512 negative CTL responses and 8 out of 40 positive CTL responses measured after vaccination (Fig. 3A). Overall, the agreement between the ≥1 PEPI3+ Test and experimentally determined CD8+ T cell reactivity was 90% ($p<0.001$).

For each patient the distribution among the peptide pools of epitopes that are presented by at least one patient class I HLA (≥1 PEPI1+, HLA restricted epitope prediction, prior art method) was also determined. ≥1 PEPI1+ correctly predicted 116 out of 512 negative CTL responses and 37 out of 40 positive CTL responses measured after vaccination (FIG. 3B). Overall, the agreement between the HLA restricted epitope prediction (≥1 PEPI1+) and CD8+ T cell reactivity was 28% (not significant).

Example 5 - Prediction of HLA class II restricted CD4+ helper T cell epitopes

The 28 cervical cancer and VIN-3 patients that received the HPV-16 synthetic long peptide vaccine (LPV) in two different clinical trials (as detailed in Example 4) were investigated for CD4+ T helper responses following LPV vaccination (FIG. 4). The sensitivity of the prediction of HLA class II restricted epitopes was 78%, since the State of Art tool predicted 84 positive responses (positive CD4+ T cell reactivity to a peptide pool for a person's DP alleles) out of 107 (sensitivity = 78%). The specificity was 22% since it could rule out 7 negative responses out of 31. Overall, the agreement between HLA-restricted class II epitope prediction and CD4+ T cell reactivity was 66%, which was statistically not significant.

Example 6 - The ≥1 PEPI3+ Test predicts T cell responses to full length LPV polypeptides

Using the same reported studies as Examples 4 and 5, the ≥1 PEPI3+ Test was used to predict patient CD8+ and CD4+ T cell responses to the full length E6 and E7 polypeptide antigens of the LPV vaccine. Results were compared to the experimentally determined responses

were reported. The Test correctly predicted the CD8+ T cell reactivity (PEPI3+) of 11 out of 15 VIN-3 patients with positive CD8+ T cell reactivity test results (sensitivity 73%, PPV 85%) and of 2 out of 5 cervical cancer patients (sensitivity 40%, PPV 100%). The CD4+ T cell reactivities (PEPI4+) were correctly predicted 100% both of VIN-3 and cervical cancer patients (Fig 5).

Class I and class II HLA restricted PEPI3+ count was also observed to correlate with the reported clinical benefit to LPV vaccinated patients. Patients with higher PEPI3+ counts had either complete or partial response already after 3 months.

Example 7 – Case Study

pGX3001 is an HPV16 based DNA vaccine containing full length E6 and E7 antigens with a linker in between. pGX3002 is an HPV18 based DNA vaccine containing full length E6 and E7 antigens with a linker in between. A Phase II clinical trial investigated the T cell responses of 17 HPV-infected patients with cervical cancer who were vaccinated with both pGX3001 and pGX3002 (VGX-3100 vaccination)[1].

Fig. 5-6 shows for two illustrative patients (patient 12-11 and patient 14-5) the position of each epitope (9 mer) presented by at least 1 (PEPI1+), at least 2 (PEPI2+), at least 3 (PEPI3+), at least 4 (PEPI4+), at least 5 (PEPI5+), or all 6 (PEPI6) class I HLA of these patients within the full length sequence of the two HPV-16 and two HPV-18 antigens.

Patient 12-11 had an overall PEPI1+ count of 54 for the combined vaccines (54 epitopes presented by one or more class I HLA). Patient 14-5 had a PEPI1+ count of 91. Therefore patient 14-5 has a higher PEPI1+ count than patient 12-11 with respect to the four HPV antigens. The PEPI1+s represent the distinct vaccine antigen specific HLA restricted epitope sets of patients 12-11 and 14-5. Only 27 PEPI1+s were common between these two patients.

For the PEPI3+ counts (number of epitopes presented by three or more patient class I HLA), the results for patients 12-11 and 14-5 were reversed. Patient 12-11 had a PEPI3+ count of 8, including at least one PEPI3+ in each of the four HPV16/18 antigens. Patient 14-5 had a PEPI3+ count of 0.

The reported immune responses of these two patients matched the PEPI3+ counts, not the PEPI1+ counts. Patient 12-11 developed immune responses to each of the four antigens post-vaccination as measured by ELISpot, whilst patient 14-5 did not develop immune responses to any of the four antigens of the vaccines. A similar pattern was observed when the PEPI1+ and PEPI3+ sets of all 17 patients in the trial were compared. There was no correlation between the PEPI1+ count and the experimentally determined T cell responses reported from the clinical trial. However, correlation between the T cell immunity predicted by the ≥1 PEPI3+ Test and the reported T cell immunity was observed. The ≥1 PEPI3+ Test predicted the immune responders to HPV DNA vaccine.

Moreover, the diversity of the patient's PEPI3+ set resembled the diversity of T cell responses generally found in cancer vaccine trials. Patients 12-3 and 12-6, similar to patient 14-5, did not have PEPI3+s predicting that the HPV vaccine could not trigger T cell immunity. All other patients had at least one PEPI3 predicting the likelihood that the HPV vaccine can trigger T cell immunity. 11 patients had multiple PEPI3+ predicting that the HPV vaccine likely triggers polyclonal T cell responses. Patients 15-2 and 15-3 could mount high magnitude T cell immunity to E6 of both HPV, but poor immunity to E7. Other patients 15-1 and 12-11 had the same magnitude response to E7 of HPV18 and HPV16, respectively.

Example 8 – Design of a Model Population for conducting *in silico* trials and identifying candidate precision vaccine targets for large population

An *in silico* human trial cohort of 433 subjects with complete 4-digit HLA class I genotype (2 x HLA-A*xx:xx; 2 x HLA-B*xx:xx; 2 x HLA-C*xx:xx) and demographic information was compiled. This Model Population has subjects with mixed ethnicity having a total of 152 different HLA alleles that are representative for >85% of presently known allele G-groups.

A database of a "Big Population" containing 7,189 subjects characterized with 4-digit HLA genotype and demographic information was also established. The Big Population has 328 different HLA class I alleles. The HLA allele distribution of the Model Population significantly

correlated with the Big Population (Table 9) (Pearson p<.001). Therefore the 433 patient Model Population is representative for a 16 times larger population.

The Model Population is representative for 85% of the human race as given by HLA diversity as well as HLA frequency.

Table 9. Statistical analysis of HLA distributions in "Model Population" vs. "Big Population".

Group name 1	Group name 2	Pearson R value	Correlation	P Value
433 Model Population	7,189 Big Population	0.89	Strong	P<0.001

Example 9 –*In silico* trials based on the identification of multiple HLA binding epitopes predict the reported T cell response rates of clinical trials

The objective of this study was to determine whether a model population, such as the one described in Example 8, may be used to predict CTL reactivity rates of vaccines, i.e. used in an *in silico* efficacy trials.

Twelve peptide vaccines derived from cancer antigens that induced T cell responses in a subpopulation of subjects were identified from peer reviewed publications. These peptides have been investigated in clinical trials enrolling a total of 172 patients (4 ethnicities). T cell responses induced by the vaccine peptides have been determined from blood specimens and reported. The immune response rate as the percentage of study subjects with positive T cell responses measured in the clinical trials was determined (FIG. 7).

Table 10. Clinical trials conducted with peptide vaccines.

Peptide vaccines	*Source antigen*	*Peptide length*	*T cell assay*	*Pop. (n)*	*Ethnicity*	*Ref.*

MMNLMQPKTQQTYTYD	JUP	16mer	Multimer staining	18	Canadian	[12]
GRGSTTTNYLLDRDDYRNTSD	ADA17	21mer	Multimer staining	18	Canadian	[12]
LKKGAADGGKLDGNAKLNRSLK	BAP31	22mer	Multimer staining	18	Canadian	[12]
FPPKDDHTLKFLYDDNQRPYPP	TOP2A	22mer	Multimer staining	18	Canadian	[12]
RYRKPDYTLDDGHGLLRFKST	Abl-2	21mer	Multimer staining	18	Canadian	[12]
QRPPFSQLHRFLADALNT	DDR1	18mer	Multimer staining	18	Canadian	[12]
ALDQCKTSCALMQQHYDQTSCFSSP	ITGB8	25mer	Multimer staining	18	Canadian	[12]
STAPPAHGVTSAPDTRPAPGSTAPP	MUC-1	25mer	Proliferation	80	Canadian	[13]
YLEPGPVTA	gp100	9mer	Tetramer	18	US	[14]
MTPGTQSPFFLLLLLTVLTVV	MUC-1	21mer	Cytotoxicity	10	Israeli	[15]
SSKALQRPV	Bcr-Abl	9mer	ELISPOT	4	US	[16]
RMFPNAPYL	WT-1	9mer	Multimer staining	24	US	[17]
RMFPNAPYL (HLA-A*0201)	WT-1	9mer	Cytokine staining	18	CEU	[18]

The 12 peptides were investigated with the ≥1 PEPI3+ Test in each of the 433 subjects of the Model Population described in Example 8. The "≥1 PEPI3+ Score" for each peptide was calculated as the proportion of subjects in the Model Population having at least one vaccine derived epitope that could bind to at least three subject-specific HLA class I (≥1 PEPI3+). If the

corresponding clinical trial stratified patients for HLA allele selected population, the Model Population was also filtered for subjects with the respective allele(s) (Example: WT1, HLA-A*0201).

The experimentally determined response rates reported from the trials were compared with the ≥1 PEPI3+ Scores. The Overall Percentage of Agreements (OPA) were calculated on the paired data (Table 11). A linear correlation between ≥1 PEPI3+ Score and response rate (R^2 = 0.77) was observed (FIG. 7). This result shows that the identification of peptides predicted to bind to multiple HLAs of an individual is useful to predict *in silico* the outcome of clinical trials.

Table 11. Comparison of ≥1 PEPI3+ Scores and CTL response rates of 12 peptide vaccines.

Peptide vaccine	*Source antigen*	*Response rate (Clinical Trials)*	*≥1 PEPI3+ Score* (Model Population)*	*OPA*
MMNLMQPKTQQTYTYD	JUP	0%	22%	NA
GRGSTTTNYLLDRDDYRNTSD	ADA17	11%	18%	61%
LKKGAADGGKLDGNAKLNRSLK	BAP31	11%	7%	64%
FPPKDDHTLKFLYDDNQRPYPP	TOP2A	11%	39%	28%
RYRKPDYTLDDGHGLLRFKST	Abl-2	17%	12%	71%
QRPPFSQLHRFLADALNT	DDR1	17%	5%	29%
ALDQCKTSCALMQQHYDQTSCFSSP	ITGB8	28%	31%	90%
STAPPAHGVTSAPDTRPAPGSTAPP	MUC-1	20%	2%	10%
YLEPGPVTA	gp100	28%	4%	14%
MTPGTQSPFFLLLLLTVLTVV	MUC-1	90%	95%	95%

SSKALQRPV	Bcr-Abl	0%	0%	100%
RMFPNAPYL	WT-1	100%	78%	78%
RMFPNAPYL (HLA-A*0201)	WT-1	81%	61%	75%

** % subjects in the Model Population with ≧1 vaccine derived PEPI3+*

Example 10. *In silico* trials based on the identification of multiple HLA binding epitopes predict the reported T cell response rates of clinical trials II

Nineteen clinical trials with published immune response rates (IRR) conducted with peptide or DNA based vaccines were identified (Table 19). These trials involved 604 patients (9 ethnicities) and covered 38 vaccines derived from tumor and viral antigens. Vaccine antigen specific CTL responses were measured in each study patient and the response rate in the clinical study populations was calculated and reported.

Each vaccine peptide of the 19 clinical trials was investigated with the ≥1 PEPI3+ Test in each subject of the Model Population. The ≥1 PEPI3+ Score for each peptide was calculated as the proportion of subjects in the Model Population having at least one vaccine derived PEPI3+. The experimentally determined response rates reported from the trials were compared with the PEPI Scores, as in Example 9 (Table 20). A linear correlation between the response rate and ≥1 PEPI3+ Score ($R^2 = 0.70$) was observed (FIG. 8). This result confirms that the identification of peptides predicted to bind to multiple HLAs of an individual can predict T cell responses of subjects, and *in silico* trials can predict the outcome of clinical trials.

Table 12. Response rates published in clinical trials.

Immunotherapy	Type	CTL assay	Pop. (n)	Race/ Ethnicity	Ref.
StimuVax	peptide	Proliferation	80	Canadian	13
gp100 vaccine	DNA	Tetramer	18	US	14
IMA901 phase I	peptide	ELISPOT	64	CEU	19
IMA901 phase II	peptide	Multimer staining	27	CEU	
ICT107	peptide	ICC	15	US	20
ProstVac	DNA	ELISPOT	32	CEU87%, Afr. Am.12%, Hisp.1%	[2]1
Synchrotope TA2M	DNA	Tetramer	26	US	22
MELITAC 12.1	peptide	ELISPOT	167	US	23
WT1 vaccine	peptide	Tetramer	22	Japanese	24
Ipilimumab (NY-ESO-1)	checkpoint inhibitor **	ICC	19	US	5
VGX-3100	DNA	ELISPOT	17	US	1
HIVIS-1	DNA	ELISPOT	12	CEU98%, Asian1%, Hisp.1%	2
ImMucin	peptide	Cytotoxicity	10	Israeli	15
NY-ESO-1 OLP	peptide	IFN-gamma	7	Japanese	7
GVX301	peptide	Proliferation	14	CEU	25
WT1 vaccine	peptide	ELISPOT	12	US	26
WT1 vaccine	peptide	ICC	18	CEU	18
DPX-0907*	peptide	Multimer staining	18	Canadian	12
Melanoma peptide vaccine	peptide	ELISPOT	26	White	27

Table 13. Linear correlation between PEPI Score and response rate ($R^2 = 0.7$).

Immunotherapy	Clinical Trial Response Rate	≥1 PEPI3+ Score*	OPA
StimuVax (failed to show efficacy in Phase III)	20%	2%	10%
gp100 vaccine	28%	4%	14%
IMA901 phase I	74%	48%	65%
IMA901 phase II	64%	48%	75%
ICT107	33%	52%	63%
ProstVac	45%	56%	80%
Synchrotope TA2M	46%	24%	52%
MELITAC 12.1	49%	47%	96%
WT1 vaccine	59%	78%	76%
Ipilimumab (NY-ESO-1*)	72%	84%	86%
VGX-3100	78%	87%	90%
HIVIS-1	80%	93%	86%
ImMucin	90%	95%	95%
NY-ESO-1 OLP	100%	84%	84%
GVX301	64%	65%	98%
WT1 vaccine	83%	80%	96%
WT1 vaccine	81%	61%	75%
DPX-0907	61%	58%	95%
Melanoma peptide vaccine	52%	42%	81%

** % subjects in the Model Population with ≧1 vaccine derived PEPI3+*

Example 11 – *In silico* trial based on the identification of multiple HLA binding epitopes in a multi-peptide vaccine predict the reported clinical trial immune response rate

IMA901 is a therapeutic vaccine for renal cell cancer (RCC) comprising 9 peptides derived from tumor-associated peptides (TUMAPs) that are naturally presented in human cancer tissue. A total of 96 HLA-A*02+ subjects with advanced RCC were treated with IMA901 in two

independent clinical studies (phase I and phase II). Each of the 9 peptides of IMA901 have been identified in the prior art as HLA-A2-restricted epitopes. Based on currently accepted standards, they are all strong candidate peptides to boost T cell responses against renal cancer in the trial subjects, because their presence has been detected in renal cancer patients, and because the trial patients were specifically selected to have at least one HLA molecule capable of presenting each of the peptides.

For each subject in the Model population how many of the nine peptides of the IMA901 vaccine were capable of binding to three or more HLA was determined. Since each peptide in the IMA901 vaccine is a 9 mer this corresponds to the PEPI3+ count. The results were compared with the immune response rates reported for the Phase I and Phase II clinical trials (Table 14).

Table 14. Immune Response Rates in the Model Population and in two clinical trials to IMA901

Immune responses to TUMAPs	Model Population (HLA-A2+) (n=180)	Phase I (n=27)*	Phase II (n=64)*
No peptide	39%	25%	36%
1 peptide	34%	44%	38%
≧ 2 peptides	27% (MultiPEPI Score)	29%	26%
≧ 3 peptides	3%	ND	3%

**No of patients evaluated for immune responses*

The phase I and phase II study results show the variability of the immune responses to the same vaccine in different trial cohorts. Overall, however, there was a good agreement between response rates predicted by the ≥2 PEPI3+ Test and the reported clinical response rates.

In a retrospective analysis, the clinical investigators of the trials discussed above found that subjects who responded to multiple peptides of the IMA901 vaccine were significantly (p = 0.019) more likely to experience disease control (stable disease, partial response) than subjects who responded only to one peptide or had no response. 6 of 8 subjects (75%) who responded to multiple peptides experienced clinical benefit in the trial, in contrast to 14% and 33% of 0 and 1

peptide responders, respectively. The randomized phas e II trial confirmed that immune responses to multiple TUMAPs were associated with a longer overall survival.

Since the presence of PEPIs accurately predicted responders to TUMAPs, clinical responders to IMA901 are likely patients who can present ≥2 PEPIs from TUMAPs. This subpopulation is only 27% of HLA-A*02 selected patients, and according to the clinical trial result, 75% of this subpopulation is expected to experience clinical benefit. The same clinical results suggest that 100% of patients would experience clinical benefit if patient selection is based on ≥3 PEPIs from TUMAPs, albeit this population would represent only 3% of the HLA-A*02 selected patient population. These results suggest that the disease control rate (stable disease or partial response) is between 3% and 27% in the patient population which was investigated in the IMA901 clinical trials. In the absence of complete response, only a portion of these patients can experience survival benefit.

These findings explain the absence of improved survival in the Phase III IMA901 clinical trial. These results also demonstrated that HLA-A*02 enrichment of the study population was not sufficient to reach the primary overall survival endpoint in the Phase III IMA901 trial. As the IMA901 trial investigators noted, there is a need for the development of a companion diagnostic (CDx) to select likely responders to peptide vaccines. These findings also suggest that selection of patients with ≥2 TUMAP specific PEPIs may provide sufficient enrichment to demonstrate significant clinical benefit of IMA901.

Example 12 - *In silico* trial based on the identification of vaccine-derived multiple HLA binding epitopes predict reported experimental clinical response rates

Acorrelation between the ≥2 PEPI3+ Score of immunotherapy vaccines determined in the Model Population described in Example 8 and the reported Disease Control Rate (DCR, proportion of patients with complete responses and partial responses and stable disease) determined in clinical trials was determined.

Seventeen clinical trials conducted with peptide- and DNA-based cancer immunotherapy vaccines that have published Disease Control Rates (DCRs) or objective response rate (ORR)

were identified from peer reviewed scientific journals (Table 15). These trials involved 594 patients (5 ethnicities) and covered 29 tumor and viral antigens. DCRs were determined according to the Response Evaluation Criteria in Solid Tumors (RECIST), which is the current standard for clinical trials, in which clinical responses are based on changes in maximum cross-sectional dimensions[42, 43, 44.] In case there was no available DCR data, objective response rate (ORR) data was used, which is also defined according to the RECIST guidelines.

Table 16 compares the ≥2 PEPI3+ Score for each vaccine in the Model Population and the published DCR or ORR. A correlation between the predicted and measured DCR was observed providing further evidence that not only the immunogenicity but also the potency of cancer vaccines depends on the multiple HLA sequences of individuals (R2 = 0.76) (FIG. 9).

Table 15.Clinical trials selected for Disease Control Rate (DCR) prediction.

Immuno-therapy	Antigen	Sponsor	Disease	Pop. (n)	Study pop./ Ethnicity	HLA restriction	Adm. form	Dose (mg)	Dosing schedule	Assessment time (weeks)	Ref.
IMA901 phase I	9 TAAs	Immatics	Renal cell cancer	28	CEU	A02	i.d.	0.4	8x in 10 wks	12	19
IMA901 phase II	9 TAAs	Immatics	Renal cell cancer	68	CEU	A02	i.d	0.4	7x in 5 wks then 10x 3 wks	24	19
Ipilimumab	NY-ESO-1	MSKCC	Melanoma	19	US	no	i.v.	0.3 3 10	4 x every 3 wks	24	5
HPV-SLP*	HPV-16 E6, E7	Leiden University	VIN	20	CEU	no	s.c.	0.3	3 x every 3 wks	12	9
HPV-SLP*		Leiden University	HPV-related cervical cancer	5	CEU	no	s.c.	0.3	3 x every 3 wks	12 (OR)	10
gp100 - 2 peptides*	gp100	BMS	Melanoma	136	US	A*0201	s.c.	1	4 x every 3 wks	12	[28]
Immucin	Muc-1	VaxilBio	Myeloma	15	Israeli	no	s.c.	0.1	6 x every 2 wks	12**	[29]
StimuVax	Muc-1	Merck	NSCLC	80	Canadian	no	s.c.	1	8x wkly then every 6 wks	12	13, 30
VGX-3100	HPV-16&18	Inovio	HPV-related cervical cancer	125	US	no	i.m.	6	0, 4, 12 wks	36	[31]
TSPP peptide vaccine	Thymidylate synthase	Siena University	CRC, NSCLC, Gallbladder carc., Breast-, Gastric cancer	21	CEU	no	s.c.	0.1 0.2 0.3	3 x 3 wks	12	32
KIF20A-66 peptide vaccine*	KIF20A	Chiba Tokushukai Hospital	Metastatic pancreatic cancer	29	Japanese	A*2402	s.c.	1 3	2 cycles 1, 8, 15, 22 days then every 2 wks	12 (OR)	33
Peptide vaccine*	3 TAAs	Kumamoto University	HNSCC	37	Japanese	A*2402	s.c.	1	8 x wkly then every 4 wks	12	[34]
7-peptide cocktail vaccine*	7 TAAs	Kinki University	Metastatic colorectal cancer	30	Japanese	A*2402	s.c.	1	Cycles: 5 x wkly then 1 wk rest	10 (OR)	[35]
GVX301*	hTERT	University Genoa	Prostate and renal cancer	14	Japanese	A02	i.d.	0.5	1, 3, 5, 7, 14, 21, 35, 63 days	12	25
MAGE-A3 Trojan*	MAGE-A3	Abramson Cancer Center	Multiple myeloma	26	US	no	s.c.	0.3	14, 42, 90, 120, 150 days	24	[36]
PepCan	HPV-16 E6	University of Arkansas	CIN2/3	23	US	no	i.m.	0.05 0.1 0.25 0.5	4 x 3 wks	24	[37]
Melanoma peptide vaccine*	Tyrosinase, gp100	University of Virginia	Melanoma	26	US	A1, A2 or A3	s.c.	0.1	6 cycles: 0, 7, 14, 28, 35, 42 days	6	27

*Montanide ISA51 VG as adjuvant

**Disease response was assessed according to the International Myeloma Working Group response criteria[45]

Table 16. The Disease Control Rates (DCRs) and MultiPEPI Scores (predicted DCR) in 17 clinical trials.

Immunotherapy	DCR	MultiPEPI Score (Predicted DCR)	Overall Percentage of Agreement
IMA901 phase I	43%	27%	*61%*
IMA901 phase II	22%	27%	*81%*
Ipilimumab	60%	65%	*92%*
HPV-SLP	60%	70%	*86%*
HPV-SLP	62%	70%	*89%*
gp100 - 2 peptides	15%	11%	*73%*
Immucin	73%	59%	*81%*
StimuVax	0%	0%	*100%*
VGX-3100	50%	56%	*89%*
TSPP peptide vaccine	48%	31%	*65%*
KIF20A-66 peptide vaccine	26%	7%	*27%*
Peptide vaccine	27%	10%	*37%*
7-peptide cocktail vaccine	10%	9%	*90%*
GVX301	29%	7%	*24%*
MAGE-A3 Trojan	35%	10%	*29%*
PepCan	52%	26%	*50%*
Melanoma peptide vaccine	12%	6%	*50%*

Example 13 – Breast cancer vaccine design for large population and composition

We used the PEPI3+ Test described above to design peptides for use in breast cancer vaccines that are effective in a large percentage of patients, taking into account the heterogeneities of both tumour antigens and patients' HLAs.

Breast cancer CTAs were identified and ranked based on the overall expression frequencies of antigens found in breast cancer tumor samples as reported in peer reviewed publications (Chen et al. Multiple Cancer/Testis Antigens Are Preferentially Expressed in Hormone-Receptor Negative and High-Grade Breast Cancers. Plos One 2011; 6(3): e17876.; Kanojia et al. Sperm-Associated Antigen 9, a Novel Biomarker for Early Detection of Breast Cancer. Cancer Epidemiol Biomarkers Prev 2009; 18(2):630 –639.; Saini et al. A Novel Cancer Testis Antigen, A-Kinase Anchor Protein 4 (AKAP4) Is a Potential Biomarker for Breast Cancer. Plos One 2013; 8(2): e57095).

Based on the ranked expression rate we have selected the most frequently expressed CTA as target antigens for breast cancer vaccine. The expression rates of the selected breast cancer specific CTAs are illustrated in Figure 11.

To select immunogenic peptides from the target CTAs we used the PEPI3+ Test and the Model Population described in Example 8 to identify the 9 mer epitopes (PEPI3+s) that are most frequently presented by at least 3HLAs of the individuals in the Model Population. We refer to these epitopes herein as "bestEPIs". An illustrative example of the "PEPI3+ hotspot" analysis and bestEPI identification is shown in FIG. 10 for the PRAME antigen.

We multiplied the reported expression frequency for each CTA by the frequency of the PEPI3+ hotspots in the Model Population to identify the T cell epitopes (9 mers) that will induce a cytotoxic T cell response against breast cancer antigens in the highest proportion of individuals (Table 17). We then selected 15 mers encompassing each of the selected 9 mers (Table 17). The 15 mers were selected to bind to most HLA class II alleles of most subjects, using the process described in Example 19 below. These 15 mers can induce both CTL and T helper responses in the highest proportion of subjects.

Table 17. BestEPI list (9-mers underlined) for selecting breast cancer peptides for vaccine composition. N%: Antigen expression frequency in colorectal cancers; B%: bestEPI frequency, ie. the percentage of individuals with epitopes binding to at least 3 HLA class I of subjects in the model population (433 subjects); HLAII**: Percentage of individuals having CD4+ T cell specific PEPI4+ within normal donors (n=400); N%*B%: N% multiplied by B%.

SEQ ID NO. 9mer	SEQ ID NO. 15mer	Antigen		BestEPIs and Optimized 15 mer				
		Antigen	N%	Opt. 15mer	Opt. Position	B%	HLAII** (CD4)	B%*N%
172	195	PIWIL-2	94%	FVASINLTLTKWYSR	760	67%	93%	64%
173	196	PIWIL-2	94%	RNFYDPTSAMVLQQH	341	60%	49%	57%
1	41	AKAP4	85%	DQVNIDYLMNRPQNL	161	52%	46%	44%
1	197	AKAP4	85%	VNIDYLMNRPQNLRL	163	52%	57%	44%

174	198	EpCam	84%	RTYWIIIELKHKARE	140	51%	100%	43%
2	42	AKAP4	85%	MMAYSDTTMMSDDID	1	49%	0%	41%
3	43	BORIS	71%	MFTSSRMSSFNRHMK	263	57%	66%	40%
3	199	BORIS	71%	VCMFTSSRMSSFNRH	261	57%	96%	40%
175	200	HIWI	100%	HAFDGTILFLPKRLQ	161	39%	83%	39%
4	201	AKAP4	85%	SDLQKYALGFQHALS	116	46%	81%	39%
4	44	AKAP4	85%	LQKYALGFQHALSPS	118	46%	88%	39%
24	64	SPAG9	88%	GTGKLGFSFVRITAL	1137	44%	94%	39%
24	202	SPAG9	88%	KLGFSFVRITALMVS	1140	44%	100%	39%
5	45	SPAG9	88%	AQKMSSLLPTMWLGA	962	43%	69%	38%
176	203	PIWIL-2	94%	YSRVVFQMPHQEIVD	772	40%	77%	38%
177	204	HIWI	100%	GFTTSILQYENSIML	251	37%	86%	37%
178	205	PLU-1	82%	LRYRYTLDDLYPMMN	732	45%	84%	37%
179	206	TSGA10	70%	YSSNAYHMSSTMKPN	653	48%	33%	34%
180	207	TSGA10	70%	LQKVQFEKVSALADL	494	46%	97%	32%
181	208	PLU-1	82%	NRTSYLHSPFSTGRS	1321	38%	37%	31%
6	46	SPAG9	88%	GNILDSFTVCNSHVL	779	36%	4%	31%
6	209	SPAG9	88%	LDSFTVCNSHVLCIA	782	36%	6%	31%
7	47	BORIS	71%	NMAFVTSGELVRHRR	319	44%	75%	31%
182	210	ODF-4	63%	NSPLPFQWRITHSFR	63	49%	35%	30%
183	211	SP17	47%	AFAAAYFESLLEKRE	37	65%	100%	30%
184	212	AKAP4	85%	DLSFYVNRLSSLVIQ	216	36%	100%	30%
185	213	ODF-4	63%	QDGRLLSSTLSLSSN	41	47%	75%	29%
186	214	RHOXF-2	60%	WEEAYTFEGARYYIN	62	48%	79%	29%
187	215	PLU-1	82%	EKAMARLQELLTVSE	955	34%	69%	28%
188	216	HIWI	100%	RSIAGFVASINEGMT	642	28%	57%	28%
8	48	PRAME	53%	LERLAYLHARLRELL	457	52%	100%	28%
189	217	RHOXF-2	60%	SDYAVHPMSPVGRTS	132	43%	5%	26%
190	218	NY-SAR-35	55%	MMQMFGLGAISLILV	184	46%	69%	25%
11	51	NY-SAR-35	55%	FSSSGTTSFKCFAPF	163	45%	0%	25%
11	219	NY-SAR-35	55%	LRHKCCFSSSGTTSF	157	45%	1%	25%
9	49	SPAG9	88%	SGAVMSERVSGLAGS	16	28%	9%	25%
10	220	BORIS	71%	RFTQSGTMKIHILQK	406	35%	69%	25%
10	50	BORIS	71%	HTRFTQSGTMKIHIL	404	35%	80%	25%
191	221	EpCam	84%	QTLIYYVDEKAPEFS	246	28%	34%	24%

13	222	NY-SAR-35	55%	FVLANGHILPNSENA	97	42%	6%	23%
13	53	NY-SAR-35	55%	CSGSSYFVLANGHIL	91	42%	78%	23%
13	223	NY-SAR-35	55%	SSYFVLANGHILPNS	94	42%	85%	23%
12	224	MAGE-A9	44%	FMFQEALKLKVAELV	102	49%	100%	22%
12	52	MAGE-A9	44%	QLEFMFQEALKLKVA	99	49%	100%	22%
14	54	PRAME	53%	RHSQTLKAMVQAWPF	64	37%	38%	20%
14	225	PRAME	53%	HSQTLKAMVQAWPFT	65	37%	37%	20%
14	226	PRAME	53%	QTLKAMVQAWPFTCL	67	37%	85%	20%
15	55	NY-BR-1	47%	YSCDSRSLFESSAKI	424	39%	0%	18%
16	56	Survivin	66%	TAKKVRRAIEQLAAM	127	26%	26%	17%
192	227	MAGE-A11	59%	SHSYVLVTSLNLSYD	286	26%	100%	15%
192	228	MAGE-A11	59%	TSHSYVLVTSLNLSY	285	26%	100%	15%
17	229	MAGE-A11	59%	AMDAIFGSLSDEGSG	184	23%	0%	14%
17	230	MAGE-A11	59%	ESFSPTAMDAIFGSL	178	23%	0%	14%
17	57	MAGE-A11	59%	SPTAMDAIFGSLSDE	181	23%	0%	14%
18	58	HOM-TES-85	47%	MASFRKLTLSEKVPP	1	29%	51%	13%
19	59	MAGE-A9	44%	SSISVYYTLWSQFDE	67	30%	97%	13%
20	231	NY-BR-1	47%	KPSAFEPATEMQKSV	582	27%	0%	12%
20	60	NY-BR-1	47%	PGKPSAFEPATEMQK	580	27%	0%	12%
193	232	NY-ESO-1	9%	SRLLEFYLAMPFATP	85	52%	98%	5%
194	233	NY-ESO-1	9%	FYLAMPFATPMEAEL	90	51%	96%	5%

Then we designed thirty-one 30 mer peptides (Table 18a). The 30 mers may each consist of two optimized 15 mer fragments, generally from different frequent CTAs, arranged end to end, each fragment comprising one of the 9 mers (BestEPIs) from Table 17 . Nine of these 30 mer peptides were selected for a panel of peptides, referred to as PolyPEPI915 (Table 18b). Expression frequencies for the 10 CTAs targeted by PolyPEPI915, singly and in combination, are shown in FIG. 11.

Table 18a. – 30mer breast cancer vaccine peptides

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
81	BCV900-2-1	AKAP4	LQKYALGFQHALSPSMMAYSDTTMMSDDID	69%	88%
82	BCV900-2-2	BORIS/AKAP4	VCMFTSSRMSSFNRHVNIDYLMNRPQNLRL	76%	97%

83	BCV900-2-3	BORIS	NMAFVTSGELVRHRRHTRFTQSGTMKIHIL	57%	92%
84	BCV900-2-4	SPAG9	LDSFTVCNSHVLCIAKLGFSFVRITALMVS	58%	100%
85	BCV900-2-5	SPAG9/NY-SAR-35	AQKMSSLLPTMWLGAMMQMFGLGAISLILV	66%	83%
86	BCV900-2-6	PRAME	LERLAYLHARLRELLQTLKAMVQAWPFTCL	71%	100%
87	BCV900-2-7	NY-SAR-35	SSYFVLANGHILPNSLRHKCCFSSSGTTSF	64%	85%
88	BCV900-2-8	Survivin/MAGE-A9	TAKKVRRAIEQLAAMQLEFMFQEALKLKVA	58%	100%
89	BCV900-2-9	MAGE-A11/NY-BR-1	TSHSYVLVTSLNLSYYSCDSRSLFESSAKI	65%	100%
90	BCV900-3-1	SPAG9/BORIS	LDSFTVCNSHVLCIAVCMFTSSRMSSFNRH	65%	96%
91	BCV900-3-2	NY-SAR-35/PRAME	LRHKCCFSSSGTTSFQTLKAMVQAWPFTCL	59%	85%
92	BCV900-3-3	NY-BR-1/SURVIVIN	YSCDSRSLFESSAKITAKKVRRAIEQLAAM	55%	26%
93	BCV900-3-4	AKAP-4/BORIS	MMAYSDTTMMSDDIDHTRFTQSGTMKIHIL	72%	80%
94	BCV900-3-5	SPAG9/AKAP-4	AQKMSSLLPTMWLGALQKYALGFQHALSPS	64%	92%
95	BCV900-3-6	MAGE-A11/BORIS	TSHSYVLVTSLNLSYNMAFVTSGELVRHRR	61%	100%
96	BCV900-3-7	NY-SAR-35/AKAP-4	MMQMFGLGAISLILVVNIDYLMNRPQNLRL	71%	84%
97	BCV900-3-8	NY-SAR-35/SPAG-9	SSYFVLANGHILPNSKLGFSFVRITALMVS	65%	100%
98	BCV900-3-9	PRAME/MAGE-A9	LERLAYLHARLRELLQLEFMFQEALKLKVA	73%	100%
99	BCV900-4-1	SPAG9/AKAP4	GNILDSFTVCNSHVLLQKYALGFQHALSPS	53%	88%
100	BCV900-4-2	BORIS/NY-SAR-35	NMAFVTSGELVRHRRFSSSGTTSFKCFAPF	65%	75%
101	BCV900-4-5	SPAG9/BORIS	AQKMSSLLPTMWLGAMFTSSRMSSFNRHMK	72%	87%
102	BCV900-4-6	MAGE-A11/PRAME	TSHSYVLVTSLNLSYHSQTLKAMVQAWPFT	60%	100%
103	BCV900-5-6	HomTes85/MageA11	MASFRKLTLSEKVPPSPTAMDAIFGSLSDE	45%	51%
104	BCV900-5-7	AKAP4/PRAME	DQVNIDYLMNRPQNLRHSQTLKAMVQAWPF	64%	67%
105	BCV900-5-8	NYSAR/SPAG9	CSGSSYFVLANGHILSGAVMSERVSGLAGS	46%	78%
106	BCV900-S-2	AKAP-4/MAGE-A9	DLSFYVNRLSSLVIQSSISVYYTLWSQFDE	60%	100%
107	BCV900-S-4	SPAG9/NY-ESO-1	SGAVMSERVSGLAGSSRLLEFYLAMPFATP	59%	98%
108	BCV900-S-6	HOM-TES-85/MAGE-A11	MASFRKLTLSEKVPPESFSPTAMDAIFGSL	46%	51%
109	BCV900-S-7	NY-ESO-1/NY-BR-1	FYLAMPFATPMEAELKPSAFEPATEMQKSV	60%	96%
110	BCV900-T-27	MAGE-A11/PRAME	AMDAIFGSLSDEGSGHSQTLKAMVQAWPFT	54%	37%
111	BCV900-T-28	NY-SAR-35/SPAG9	FVLANGHILPNSENAGTGKLGFSFVRITAL	61%	94%
435	BCV900-6-1	TSGA10 / PIWIL-2	YSSNAYHMSSTMKPNFVASINLTLTKWYSR	80%	95%
436	BCV900-6-2	PIWIL-2 / AKAP4	RNFYDPTSAMVLQQHMMAYSDTTMMSDDID	88%	49%
437	BCV900-6-3	PLU-1 / RHOXF-2	LRYRYTLDDLYPMMNSDYAVHPMSPVGRTS	67%	85%
438	BCV900-6-4	SPAG9 / EpCam	SGAVMSERVSGLAGSRTYWIIIELKHKARE	60%	100%
439	BCV900-6-5	AKAP4 / PLU-1	DLSFYVNRLSSLVIQNRTSYLHSPFSTGRS	66%	100%
440	BCV900-6-6	AKAP4 / HIWI	VNIDYLMNRPQNLRLHAFDGTILFLPKRLQ	70%	94%

441	BCV900-6-7	AKAP4 / PLU-1	SDLQKYALGFQHALSEKAMARLQELLTVSE	56%	92%
442	BCV900-6-8	HIWI / ODF-4	GFTTSILQYENSIMLQDGRLLSSTLSLSSN	61%	94%
443	BCV900-6-9	PIWIL-2 / BORIS	YSRVVFQMPHQEIVDNMAFVTSGELVRHRR	61%	85%
444	BCV900-6-10	SP17 / BORIS	AFAAAYFESLLEKREMFTSSRMSSFNRHMK	82%	100%
445	BCV900-6-11	ODF-4 / HIWI	NSPLPFQWRITHSFRRSIAGFVASINEGMT	60%	69%
446	BCV900-6-12	NY-SAR-35 / RHOXF-2	SSYFVLANGHILPNSWEEAYTFEGARYYIN	74%	93%
447	BCV900-6-13	TSGA10 / PRAME	LQKVQFEKVSALADLLERLAYLHARLRELL	68%	100%
448	BCV900-6-14	MAGE-A11 / MAGE-A9	SHSYVLVTSLNLSYDFMFQEALKLKVAELV	65%	100%
449	BCV900-6-15	BORIS / EpCam	RFTQSGTMKIHILQKQTLIYYVDEKAPEFS	53%	80%

Table 18b – Selected Breast Cancer Vaccine peptides for PolyPEPI915 panel/composition

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
99	BCV900-4-1	SPAG9/AKAP4	GNILDSFTVCNSHVLLQKYALGFQHALSPS	53%	75%
100	BCV900-4-2	BORIS/NY-SAR-35	NMAFVTSGELVRHRRFSSSGTTSFKCFAPF	65%	46%
92	BCV900-3-3	NY-BR-1/SURVIVIN	YSCDSRSLFESSAKITAKKVRRAIEQLAAM	55%	11%
93	BCV900-3-4	AKAP-4/BORIS	MMAYSDTTMMSDDIDHTRFTQSGTMKIHIL	72%	45%
101	BCV900-4-5	SPAG9/BORIS	AQKMSSLLPTMWLGAMFTSSRMSSFNRHMK	72%	50%
103	BCV900-5-6	HomTes85/MageA11	MASFRKLTLSEKVPPSPTAMDAIFGSLSDE	45%	16%
104	BCV900-5-7	AKAP4/PRAME	DQVNIDYLMNRPQNLRHSQTLKAMVQAWPF	64%	33%
105	BCV900-5-8	NYSAR/SPAG9	CSGSSYFVLANGHILSGAVMSERVSGLAGS	46%	48%
98	BCV900-3-9	PRAME/MAGE-A9	LERLAYLHARLRELLQLEFMFQEALKLKVA	73%	100%
			PolyPEPI915 (9 peptide together)	96%	100%

* Percentage of individuals having CD8+ T cell specific PEPI3+ within the HLA class I Model Population (n=433).

**Percentage of individuals having CD4+ T cell specific PEPI4+ within the normal donors (n=400).

Characterization of PolyPEPI915

Tumor heterogeneity can be addressed by including peptide sequences that target multiple CTAs in a vaccine or immunotherapy regime. The PolyPEPI915 composition targets 10 different CTAs. Based on the antigen expression rates for these 10 CTAs, we modelled the predicted average number of expressed antigens (AG50) and the minimum number of expressed antigens with 95% likelihood (AG95) in the cancer cells. 95% of individuals expressed minimum 4 of the 10 target antigens (AG95=4) as shown by the antigen expression curve in **FIG. 12**.

The AG values described above characterize a vaccine independently from the target patient population. They can be used to predict the likelihood that a specific cancer (e.g. breast cancer) expresses antigens targeted by a specific vaccine or immunotherapy composition. AG values are based on known tumor heterogeneity, but do not take HLA heterogeneity into account.

HLA heterogeneity of a certain population can be characterised from the viewpoint of an immunotherapy or vaccine composition by the number of antigens representing PEPI3+. These are the vaccine-specific CTA antigens for which ≥1 PEPI3+ is predicted, referred to herein as the "AP". The average number of antigens with PEPI3+ (AP50) shows how the vaccine can induce immune response against the antigens targeted by the composition (breast cancer vaccine specific immune response). The PolyPEPI915 composition can induce immune response against an average of 5.3 vaccine antigens (AP50=5.30) and 95% of the Model Population can induce immune response against at least one vaccine antigen (AP95=1)(FIG. 13).

Vaccines can be further characterized by AGP values that refers to antigens with PEPIs". This parameter is the combination of the previous two parameters: (1) AG is depending on the antigen expression frequencies in the specific tumor type but not on the HLA genotype of individuals in the population, and (2) AP is depending on the HLA genotype of individuals in a population without taking account the expression frequencies of the antigen. The AGP is depending on both, the expression frequencies of vaccine antigens in the disease and the HLA genotype of individuals in a population.

Combining the data of AG of breast cancer and AP in the Model Population we determined the AGP value of PolyPEPI915 that represents the probability distribution of vaccine antigens that are induce immune responses against antigens expressed in breast tumors. For PolyPEPI915, the AGP50 value in the Model Population is 3.37. The AGP92=1, means that 92% of the subjects in the Model Population induce immune responses against at least one expressed vaccine antigen (FIG. 14).

Example 14 – Patient selection using companion diagnostic for breast cancer vaccine

The likelihood that a specific patient will have an immune response or a clinical response to treatment with one or more cancer vaccine peptides, for example as described above, can be determined based on (i) the identification of PEPI3+ within the vaccine peptide(s) (9 mer epitopes capable of binding at least three HLA of the patient); and/or (ii) a determination of target antigen expression in cancer cells of the patient, for example as measured in a tumour biopsy. Ideally both parameters are determined and the optimal combination of vaccine peptides is selected for use in treatment of the patient. However, PEPI3+ analysis alone may be used if a determination of the expressed tumour antigens, for example by biopsy, is not possible, not advised, or unreliable due to biopsy error (i.e. biopsy tissue samples taken from a small portion of the tumor or metastasised tumors do not represent the complete repertoire of CTAs expressed in the patient).

Example 15 - Comparison of PolyPEPI915 with competing breast cancer vaccines

We used the *in silico* clinical trial model described in above to predict the immune response rates of competing breast cancer vaccines that investigated in clinical trials (Table 19). The immune response rate of these products were between 3% and 91%.

The single peptide vaccines were immunogenic in 3% - 23% of individuals. In comparison, peptides having an amino acid sequence selected from SEQ ID NOs: 81-111 were immunogenic in from 44% to 73% of individuals in the same cohorts. This result represents substantial improvement in immunogenicity of each peptide in PolyPEPI915.

Competing combination peptide products immune response rates were between 10 - 62%. The invented PolyPEPI915 combination product were 96% in the Model Population and 93% in a breast cancer patient population representing improvement in immunogenicity.

Table 19. Predicted immune response rates of competing breast cancer vaccines

Breast Cancer Vaccines	Sponsors	Target antigens	Predicted immune response rates*

			433 normal donors (Model Population)	90 patients with breast cancer
DPX0907 Multipeptide	ImmunoVaccine Tech.	7	58%	62%
Multipeptide vaccine	University of Virginia	5	22%	31%
Ad-sig-hMUC-1/ecdCD40L	Singapore CRI	1	91%	80%
NY-ESO-1 IDC-G305	Immune Design Corp.	1	84%	84%
6 HER2 peptide pulsed DC	University Pennsylvania	1	29%	36%
HER-2 B Cell peptide	Ohio State University	1	18%	23%
HER-2/neu ID protein	University Washington	1	10%	11%
NeuVax peptide	Galena Biopharma	1	6%	3%
StimuVax®(L-BLP25) peptide	EMD Serono	1	6%	8%
PolyPEPI915	Treos Bio	10	96%	93%

*Proportion of subjects with ≧1 PEPI3+

Another improvement of using the PolyPEPI915 vaccine is the lower chance of tumor escape. Each 30 mer peptide in PolyPEPI915 targets 2 tumor antigens. CTLs against more tumor antigens are more effective against heterologous tumor cells that CTLs against a single tumor antigen.

Another improvement is that PolyPEPI915 vaccine that individuals who likely respond to vaccination can be identified based on their HLA genotype (sequence) and optionally antigen expression in their tumor using the methods described here. Pharmaceutical compositions with PolyPEPI vaccines will not be administered to individuals whose HLA cannot present any PEPI3 from the vaccines. During clinical trials correlation will be made between the mAGP or number of AGP in the PolyPEPI915 regimen and the duration of individual's responses. A vaccine combination with > 1 AGP is most likely required to destroy heterologous tumor cells. Pharmaceutical compositions with PolyPEPI vaccines will not be administered to individuals whose HLA cannot present any PEPI3 from the vaccines.

Example 16 Colorectal cancer vaccine design and composition

We show another example for colorectal vaccine composition using the same design method demonstrated above. We used the PEPI3+ Test described above to design peptides for use in colorectal cancer vaccines that are effective in a large percentage of patients, taking into account the heterogeneities of both tumour antigens and patient HLAs.

Colorectal cancer CTAs were identified and ranked based on the overall expression frequencies of antigens found in breast cancer tumor samples as reported in peer reviewed publications (FIG. 15) (Choi J, Chang H. The expression of MAGE and SSX, and correlation of COX2, VEGF, and survivin in colorectal cancer. Anticancer Res 2012. 32(2):559-564.; Goossens-Beumer IJ, Zeestraten EC, Benard A, Christen T, Reimers MS, Keijzer R, Sier CF, Liefers GJ, Morreau H, Putter H, Vahrmeijer AL, van de Velde CJ, Kuppen PJ. Clinical prognostic value of combined analysis of Aldh1, Survivin, and EpCAM expression in colorectal cancer. Br J Cancer 2014. 110(12):2935-2944.; Li M, Yuan YH, Han Y, Liu YX, Yan L, Wang Y, Gu J. Expression profile of cancer-testis genes in 121 human colorectal cancer tissue and adjacent normal tissue. Clinical Cancer Res 2005. 11(5):1809–1814).

Based on the ranked expression rate we have selected the most frequently expressed CTA as target antigens for the colorectal cancer vaccine. The expression rates of the selected breast cancer specific CTAs are illustrated in Figure 15.

To select immunogenic peptides from the most frequently expressed colorectal cancer CTAs we used the PEPI3+ Test and the Model Population described in Example 8 to identify the "bestEPIs".

We multiplied the reported expression frequency for each CTA (N%) by the frequency of the PEPI3+ hotspots in the Model Population (B%) to identify the T cell epitopes (9 mers) that will induce an immune response against colorectal cancer antigens in the highest proportion of individuals (Table 20). We then selected 15 mers encompassing each of the selected 9 mers (Table 20). The 15 mers were selected to bind to most HLA class II alleles of most subjects, using the process described in Example 19 below. These 15 mers can induce both CTL and T helper responses in the highest proportion of subjects.

Table 20. BestEPI list (9-mers underlined) for selecting colorectal cancer peptides for vaccine composition. N%: Antigen expression frequency in colorectal cancers; B%: bestEPI frequency, ie. the percentage of individuals with epitopes binding to at least 3 HLA class I of subjects in the model population (433 subjects); HLAII**: Percentage of individuals having CD4+ T cell specific PEPI4+ within normal donors (n=400); N%*B%: N% multiplied by B%.

SEQ ID NO. 9mer	SEQ ID NO. 15mer	Antigen		BestEPIs and Optimized 15 mer				
		Antigen	N%	Opt. 15mer	Opt. Position	B%	HLAII** (CD4)	B%*N%
234	251	TSP50	89%	VCSMEGTWYLVGLVS	315	58%	72%	52%
21	252	TSP50	89%	GFSYEQDPTLRDPEA	105	51%	0%	45%
21	61	TSP50	89%	RSCGFSYEQDPTLRD	102	51%	0%	45%
21	253	TSP50	89%	YRSCGFSYEQDPTLR	101	51%	0%	45%
22	62	EpCAM	88%	VRTYWIIIELKHKAR	139	51%	100%	45%
235	254	EpCAM	88%	LLAAATATFAAAQEE	12	39%	28%	34%
24	255	SPAG9	74%	KLGFSFVRITALMVS	1140	44%	100%	33%
23	63	TSP50	89%	PSTTMETQFPVSEGK	83	36%	0%	32%
24	64	SPAG9	74%	GTGKLGFSFVRITAL	1137	44%	94%	32%
23	256	TSP50	89%	LPSTTMETQFPVSEG	82	36%	0%	32%

25	65	SPAG9	74%	AQKMSSLLPTMWLGA	962	43%	69%	32%
26	66	CAGE1	74%	LASKMHSLLALMVGL	613	42%	99%	31%
27	67	FBXO39	39%	KFMNPYNAVLTKKFQ	95	78%	43%	30%
28	68	CAGE1	74%	PKSMTMMPALFKENR	759	37%	87%	27%
238	257	SPAG9	74%	LDSFTVCNSHVLCIA	782	36%	6%	27%
236	258	SPAG9	74%	GNILDSFTVCNSHVL	779	36%	4%	26%
29	69	EpCAM	88%	YVDEKAPEFSMQGLK	251	28%	0%	25%
29	259	EpCAM	88%	QTLIYYVDEKAPEFS	246	28%	34%	25%
30	70	FBXO39	39%	FKKTMSTFHNLVSLN	216	58%	92%	23%
31	71	Survivin	86%	TAKKVRRAIEQLAAM	127	26%	26%	22%
237	260	TSP50	89%	SRTLLLALPLPLSLL	368	24%	100%	21%
32	72	SPAG9	74%	SGAVMSERVSGLAGS	16	28%	9%	21%
238	260	TSP50	89%	SRTLLLALPLPLSLL	368	23%	100%	20%
34	74	FBXO39	39%	KVNFFFERIMKYERL	284	46%	100%	18%
33	73	TSP50	89%	SRYRAQRFWSWVGQA	190	20%	88%	18%
239	261	LEMD1	56%	FIIVVFVYLTVENKS	164	30%	97%	17%
240	66	CAGE1	74%	LASKMHSLLALMVGL	613	22%	99%	16%
241	262	FBXO39	39%	RNSIRSSFISSLSFF	142	40%	100%	16%
242	263	CAGE1	74%	NIENYSTNALIQPVD	97	21%	14%	16%
243	264	Survivin	86%	MGAPTLPPAWQPFLK	1	17%	0%	15%
244	265	CAGE1	74%	RQFETVCKFHWVEAF	119	18%	45%	13%
35	75	Survivin	86%	KDHRISTFKNWPFLE	15	15%	83%	13%
36	266	MAGE-A8	44%	PEEAIWEALSVMGLY	220	20%	78%	9%
36	76	MAGE-A8	44%	SRAPEEAIWEALSVM	217	20%	6%	9%
37	77	MAGE-A8	44%	DEKVAELVRFLLRKY	113	18%	95%	8%
37	267	MAGE-A8	44%	EKVAELVRFLLRKYQ	114	18%	99%	8%
38	268	MAGE-A6	28%	KLLTQYFVQENYLEY	244	27%	98%	8%
38	78	MAGE-A6	28%	QYFVQENYLEYRQVP	248	27%	93%	8%
40	80	MAGE-A6	28%	IGHVYIFATCLGLSY	172	25%	82%	7%
39	79	MAGE-A8	44%	EFLWGPRALAETSYV	273	16%	44%	7%
245	269	MAGE-A3	23%	IGHLYIFATCLGLSY	172	28%	85%	6%
246	270	MAGE-A3	23%	KLLTQHFVQENYLEY	244	27%	77%	6%
247	271	MAGE-A8	44%	ASSSSTLIMGTLEEV	39	14%	19%	6%
248	269	MAGE-A3	23%	IGHLYIFATCLGLSY	172	25%	85%	6%
249	264	Survivin	86%	MGAPTLPPAWQPFLK	1	5%	0%	4%

250	75	Survivin	86%	KDHRISTFKNWPFLE	15	4%	83%	3%

Then we designed thirty-one 30 mer peptides (Table 21a). The 30 mers each consist of two optimized 15 mer fragments, generally from different frequent CTAs, each 30 mer generally containing at least one high frequency HLA class-II binding PEPI. The 15 mer fragments are arranged end to end, and each comprises one of the 9 mers (BestEPIs) from Table 20 as described above. Nine of these 30 mer peptides were selected for a panel of peptide vaccines, referred to as PolyPEPI1015 (Table 21b). Expression frequencies for the 8 CTAs targeted by PolyPEPI1015, singly and in combination, are shown in FIG. 15.

Table 21a – 30mer colorectal cancer vaccine peptides

SEQ ID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
112	CCV1000-1-1	TSP50	VCSMEGTWYLVGLVSYRSCGFSYEQDPTLR	71%	72%
113	CCV1000-1-2	EpCAM/TSP50	VRTYWIIIELKHKARLPSTTMETQFPVSEG	62%	100%
114	CCV1000-1-4	Survivin	TAKKVRRAIEQLAAMMGAPTLPPAWQPFLK	39%	26%
115	CCV1000-1-5	CAGE1	LASKMHSLLALMVGLPKSMTMMPALFKENR	68%	99%
116	CCV1000-1-6	Spag9	KLGFSFVRITALMVSLDSFTVCNSHVLCIA	58%	100%
117	CCV1000-1-7	FBXO39	KFMNPYNAVLTKKFQFKKTMSTFHNLVSLN	91%	92%
118	CCV1000-1-8	Spag9/FBXO39	AQKMSSLLPTMWLGAKVNFFFERIMKYERL	75%	100%
119	CCV1000-1-9	Survivin/Mage-A8	KDHRISTFKNWPFLEPEEAIWEALSVMGLY	39%	93%
120	CCV1000-2-1	TSP50	YRSCGFSYEQDPTLRVCSMEGTWYLVGLVS	71%	72%
121	CCV1000-2-2	EpCAM/Survivin	VRTYWIIIELKHKARTAKKVRRAIEQLAAM	57%	100%
122	CCV1000-2-4	TSP50/Spag9	LPSTTMETQFPVSEGKLGFSFVRITALMVS	61%	100%
123	CCV1000-2-5	Survivin/Mage-A8	MGAPTLPPAWQPFLKPEEAIWEALSVMGLY	40%	78%
124	CCV1000-2-6	CAGE1/Survivin	LASKMHSLLALMVGLKDHRISTFKNWPFLE	58%	99%
125	CCV1000-2-7	CAGE1/Spag9	PKSMTMMPALFKENRLDSFTVCNSHVLCIA	61%	87%
126	CCV1000-2-8	FBXO39	KFMNPYNAVLTKKFQKVNFFFERIMKYERL	90%	100%
127	CCV1000-2-9	Spag9/FBXO39	AQKMSSLLPTMWLGAFKKTMSTFHNLVSLN	67%	92%
128	CCV1000-3-1	TSP50	GFSYEQDPTLRDPEAVCSMEGTWYLVGLVS	71%	72%
129	CCV1000-3-7	CAGE1/Spag9	PKSMTMMPALFKENRGNILDSFTVCNSHVL	61%	87%

130	CCV1000-5-1	TSP50	PSTTMETQFPVSEGKSRYRAQRFWSWVGQA	53%	88%
131	CCV1000-5-3	EpCAM /Mage-A8	YVDEKAPEFSMQGLKDEKVAELVRFLLRKY	43%	95%
132	CCV1000-5-4	TSP50/Spag9	RSCGFSYEQDPTLRDGTGKLGFSFVRITAL	67%	94%
133	CCV1000-5-5	Mage-A8/Mage-A6	SRAPEEAIWEALSVMQYFVQENYLEYRQVP	45%	94%
134	CCV1000-5-7	CAGE1/Spag9	PKSMTMMPALFKENRSGAVMSERVSGLAGS	57%	87%
135	CCV1000-S-1	SPAG9/FBXO39	SGAVMSERVSGLAGSRNSIRSSFISSLSFF	64%	100%
136	CCV1000-S-2	CAGE1/MAGE-A8	NIENYSTNALIQPVDEKVAELVRFLLRKYQ	28%	99%
137	CCV1000-S-3	CAGE1/MAGE-A6	RQFETVCKFHWVEAFKLLTQYFVQENYLEY	46%	98%
138	CCV1000-S-5	MAGE-A8/MAGE-A3	EFLWGPRALAETSYVKLLTQHFVQENYLEY	39%	91%
139	CCV1000-S-6	MAGE-A8/EpCAM	ASSSSTLIMGTLEEVQTLIYYVDEKAPEFS	41%	41%
140	CCV1000-S-7	TSP50/MAGE-A3	SRTLLLALPLPLSLLIGHLYIFATCLGLSY	60%	100%
141	CCV1000-S-9	LEMD1/MAGE-A6	FIIVVFVYLTVENKSIGHVYIFATCLGLSY	51%	99%
142	CCV1000-S-17	EPCAM	LLAAATATFAAAQEEQTLIYYVDEKAPEFS	52%	54%

* Percentage of individuals having CD8+ T cell specific PEPI3+ within the Model Population (n=433).

**Percentage of individuals having CD4+ T cell specific PEPI4+ within normal donors (n=400).

Table 21b – Selected Colorectal Cancer Vaccine peptides for PolyPEPI1015 composition

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
130	CCV1000-5-1	TSP50	PSTTMETQFPVSEGKSRYRAQRFWSWVGQA	53%	53%
121	CCV1000-2-2	EpCAM/Survivin	VRTYWIIIELKHKARTAKKVRRAIEQLAAM	57%	98%
131	CCV1000-5-3	EpCAM /Mage-A8	YVDEKAPEFSMQGLKDEKVAELVRFLLRKY	43%	72%
132	CCV1000-5-4	TSP50/Spag9	RSCGFSYEQDPTLRDGTGKLGFSFVRITAL	67%	82%
133	CCV1000-5-5	Mage-A8/Mage-A6	SRAPEEAIWEALSVMQYFVQENYLEYRQVP	45%	76%
124	CCV1000-2-6	CAGE1/Survivin	LASKMHSLLALMVGLKDHRISTFKNWPFLE	58%	95%
134	CCV1000-5-7	CAGE1/Spag9	PKSMTMMPALFKENRSGAVMSERVSGLAGS	57%	57%
126	CCV1000-2-8	FBXO39	KFMNPYNAVLTKKFQKVNFFFERIMKYERL	90%	98%
127	CCV1000-2-9	Spag9/FBXO39	AQKMSSLLPTMWLGAFKKTMSTFHNLVSLN	67%	66%
			PolyPEPI1015 (9 peptide together)	100%	99%

* Percentage of individuals having CD8+ T cell specific PEPI3+ within the Model Population (n=433).

**Percentage of individuals having CD4+ T cell specific PEPI4+ within normal donors (n=400).

Characterization of PolyPEPI1015 colorectal cancer vaccine

Tumor heterogeneity: The PolyPEPI1015 composition targets 8 different CTAs (Fig 15). Based on the antigen expression rates for these 8 CTAs, AG50 = 5.22 and AG95 = 3 **FIG. 16**. Patient heterogeneity: the AP50=4.73 and AP95 = 2 (AP95=2) (FIG. 17). Both tumor and patient heterogeneity: AGP50 = 3.16 and AGP95 = 1 (Model Population) (FIG. 18).

Example 17 - Comparison of colorectal cancer vaccine peptides with competing colorectal cancer vaccines

We used the *in silico* clinical trial model described above to determine T cell responder rate of state of art and currently developed CRC peptide vaccines and compared to and compared to that of polyPEPI1015 (Table 22). Our PEPI3+ test demonstrate that competing vaccines can induce immune responses against one tumor antigen in a fraction of subjects (2% - 77%). However, the multi-antigen (multi-PEPI) response determination for the 2 competitor multi-antigen vaccines resulted in no or 2% responders. *% of responders are the ratio of subjects from the Model population with 1≥PEPI3+ for HLAI (CD8+ T cell responses) in case of 1, or for 2, 3, 4 or 5 antigens of the vaccine compositions. Since multi-PEPI responses correlate with clinical responses induced by tumor vaccines, it is unlikely that any of the competing vaccines will demonstrate clinical benefit in 98% of patients. In contrast, we predicted multi-PEPI responses in 95% of subjects suggesting the likelihood for clinical benefit in the majority of patients.

Table 22 Predicted immune response rates of polyPEPI1015 and competing colorectal cancer vaccines

Colorectal Cancer Vaccines	*Sponsor*	% of CD8+ T cell responders in 433 subjects*					
		Vaccine antigens (Ags)	% responders against multiple Ags				
			1 Ag	2 Ags	3 Ags	4Ags	5 Ags
Stimuvax®(L-BLP25) Peptide Vaccine	Johannes Gutenberg University Mainz	1	6%	-	-	-	-
WT1 Multipeptide Vaccine	Shinshu University, Japan	1	79%	-	-	-	-
Multiepitope Peptide Cocktail Vaccine	Kinki University	7	5%	2%	0%	0%	0%
p53 Synthetic Long Peptide Vaccine	Leiden University Medical Center	1	77%	-	-	-	-
HER-2 B Cell Peptide Vaccine	Ohio State University Comprehensive Cancer Center	1	18%	-	-	-	-
NY-ESO-1 peptide pulsed dendritic cell vaccine	Jonsson Comprehensive Cancer Center	1	0%	-	-	-	-
OCV-C02	Otsuka Pharmaceutical Co., Ltd.	2	2%	0%	-	-	-
PolyPEPI1015	**Treos Bio**	**8**	**100%**	**95%**	**87%**	**70%**	**54%**

Example 18 Ovarian cancer vaccine design and composition

We used the PEPI3+ Test to design peptides for use in ovarian cancer vaccines using essentially the same design method described in Examples 13 and 16 above.

We multiplied the reported expression frequency for CTAs associated with ovarian cancer (N%) by the frequency of the PEPI3+ hotspots in the Model Population (B%) to identify the T cell epitopes (9 mers) that will induce an immune response against ovarian cancer antigens in the highest proportion of individuals (Table 23). We then selected 15 mers encompassing each of the selected 9 mers (Table 23). The 15 mers were selected to bind to most HLA class II alleles of most subjects, using the process described in Example 20 below.

Table 23. BestEPI list (9-mers underlined) for selecting ovarian cancer peptides for vaccine composition. N%: Antigen expression frequency in colorectal cancers; B%: bestEPI frequency, ie.

the percentage of individuals with epitopes binding to at least 3 HLA class I of subjects in the model population (433 subjects); HLAII**: Percentage of individuals having CD4+ T cell specific PEPI4+ within normal donors (n=400); N%*B%: N% multiplied by B%.

SEQ ID NO. 9mer	SEQ ID NO. 15mer	Antigen		BestEPIs and Optimized 15 mer				
		Antigen	N%	Opt. 15mer	Opt. Position	B%	HLAII** (CD4)	B%*N%
272	302	PIWIL-4	90%	QGMMMSIATKIAMQM	585	79%	72%	71%
273	303	PIWIL-4	90%	KAKAFDGAILFLSQK	153	62%	80%	56%
274	304	WT1	63%	SSGQARMFPNAPYLP	121	78%	0%	49%
275	305	EpCam	92%	RTYWIIIELKHKARE	140	51%	100%	47%
276	306	BORIS	82%	MFTSSRMSSFNRHMK	263	57%	66%	46%
277	307	AKAP4	88%	QVNIDYLMNRPQNLR	162	52%	46%	46%
278	308	OY-TES-1	65%	STPMIMENIQELIRS	277	67%	82%	43%
279	309	AKAP4	88%	MMAYSDTTMMSDDID	1	49%	0%	43%
280	310	SP17	65%	AFAAAYFESLLEKRE	37	65%	100%	42%
281	311	PIWIL-4	90%	RAIQQYVDPDVQLVM	534	46%	5%	42%
282	312	PIWIL-2	61%	GFVASINLTLTKWYS	759	67%	93%	41%
283	313	AKAP4	88%	DLQKYALGFQHALSP	117	46%	82%	40%
284	314	PIWIL-3	88%	GYVTSVLQYENSITL	266	44%	54%	39%
285	315	SPAG9	90%	VREEAQKMSSLLPTM	958	43%	1%	39%
286	316	PIWIL-3	88%	MSLKGHLQSVTAPMG	523	42%	17%	37%
287	317	PIWIL-3	88%	QKSIAGFVASTNAEL	663	42%	37%	37%
288	318	PIWIL-2	61%	RNFYDPTSAMVLQQH	341	60%	49%	37%
289	319	BORIS	82%	NMAFVTSGELVRHRR	319	44%	75%	36%
290	320	AKAP4	88%	LSFYVNRLSSLVIQM	217	36%	100%	31%
291	321	PRAME	59%	LERLAYLHARLRELL	457	52%	100%	30%
292	322	BORIS	82%	RFTQSGTMKIHILQK	406	35%	69%	29%
293	323	HIWI	68%	HAFDGTILFLPKRLQ	161	39%	83%	27%
294	324	EpCam	92%	YVDEKAPEFSMQGLK	251	28%	0%	26%
295	325	SPAG9	90%	SGAVMSERVSGLAGS	16	28%	9%	25%
296	326	HIWI	68%	GFTTSILQYENSIML	251	37%	86%	25%
297	327	PIWIL-2	61%	YSRVVFQMPHQEIVD	772	40%	77%	24%
298	328	PRAME	59%	RHSQTLKAMVQAWPF	64	37%	38%	22%

299	329	Survivin	84%	AKKVRRAIEQLAAMD	128	26%	25%	22%
300	330	BORIS	82%	ERSDEIVLTVSNSNV	210	25%	2%	21%
301	331	WT1	63%	RTPYSSDNLYQMTSQ	218	32%	0%	20%

Then we designed 15 30 mer peptides (Table 24).

Table 24 – 30mer ovarian cancer vaccine peptides

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
332	OC1212-01	OY-TES-1/PIWIL-4	STPMIMENIQELIRSQGMMMSIATKIAMQM	94%	98%
333	OC1212-02	PIWIL-2/PIWIL-4	RNFYDPTSAMVLQQHKAKAFDGAILFLSQK	89%	90%
334	OC1212-03	BORIS/AKAP4	NMAFVTSGELVRHRRMMAYSDTTMMSDDID	68%	75%
335	OC1212-04	WT1/WT1	SSGQARMFPNAPYLPRTPYSSDNLYQMTSQ	84%	0%
336	OC1212-05	BORIS/HIWI	MFTSSRMSSFNRHMKHAFDGTILFLPKRLQ	67%	94%
337	OC1212-06	PIWIL-2/EpCam	YSRVVFQMPHQEIVDRTYWIIIELKHKARE	67%	100%
338	OC1212-07	AKAP4/PIWIL-4	LSFYVNRLSSLVIQMRAIQQYVDPDVQLVM	71%	100%
339	OC1212-08	AKAP4/ SP17	QVNIDYLMNRPQNLRAFAAAYFESLLEKRE	78%	100%
340	OC1212-09	PIWIL-3/PIWIL-3	GYVTSVLQYENSITLQKSIAGFVASTNAEL	64%	65%
341	OC1212-10	SPAG9/BORIS	VREEAQKMSSLLPTMRFTQSGTMKIHILQK	62%	69%
342	OC1212-11	PIWIL-2/EpCam	GFVASINLTLTKWYSYVDEKAPEFSMQGLK	74%	93%
343	OC1212-12	PIWIL-3/SPAG9	MSLKGHLQSVTAPMGSGAVMSERVSGLAGS	52%	19%
344	OC1212-13	AKAP4/PRAME	DLQKYALGFQHALSPLERLAYLHARLRELL	67%	100%
345	OC1212-14	HIWI/BORIS	GFTTSILQYENSIMLERSDEIVLTVSNSNV	49%	86%
346	OC1212-15	PRAME/Survivin	RHSQTLKAMVQAWPFAKKVRRAIEQLAAMD	48%	42%

* Percentage of individuals having CD8+ T cell specific PEPI3+ within the Model Population (n=433).

**Percentage of individuals having CD4+ T cell specific PEPI4+ within normal donors (n=400).

Example 19. Efficacy by design procedure exemplified for PolyPEPI1018 colorectal cancer vaccine

The PolyPEPI1018 Colorectal Cancer (CRC) Vaccine (PolyPEPI1018) composition is a peptide vaccine intended to be used as an add-on immunotherapy to standard-of-care CRC treatment options in patients identified as likely responders using a companion *in vitro* diagnostic test (CDx). Clinical trials are ongoing in the US and Italy to evaluate PolyPEPI1018 in metastatic colorectal cancer patients. The product contains 6 peptides (6 of the 30 mer peptides PolyPEPI1015 described in examples 16 and 17) mixed with the adjuvant Montanide. The 6 peptides were selected to induce T cell responses against 12 epitopes from 7 cancer testis antigens (CTAs) that are most frequently expressed in CRC. The 6 peptides were optimized to induce long lasting CRC specific T cell responses. Likely responder patients with T cell responses against multiple CTAs expressed in the tumor can be selected with a companion diagnostic (CDx). This example sets out the precision process used to design PolyPEPI1018. This process can be applied to design vaccines against other cancers and diseases.

A. Selection of Multiple Antigen Targets

The selection of tumor antigens is essential for the safety and efficacy of cancer vaccines. The feature of a good antigen is to have restricted expression in normal tissues so that autoimmunity is prevented. Several categories of antigen meet this requirement, including uniquely mutated antigens (e.g. p53), viral antigens (e.g. human papillomavirus antigens in cervical cancer), and differentiation antigens (e.g. CD20 in B-cell lymphoma).

The inventors selected multiple cancer testis antigens (CTAs) as target antigens since they are expressed in various types of tumor cells and testis cells, but not expressed in any other normal somatic tissues or cells. CTAs are desirable targets for vaccines for at least the following reasons:

- tumors of higher histological grade and later clinical stage often show higher frequency of CTA expression
- only a subpopulation of tumor cells express a certain CTA
- different cancer types are significantly different in their frequency of CTA expression

- tumors that are positive for a CTA often show simultaneous expression of more than one CTA
- None of the CTAs appear to be cell surface antigens, therefore these are unique targets for cancer vaccines (they are not suitable targets for antibody based immunotherapies)

To identify the target CTAs for PolyPEPI1018, the inventors built a CTA expression knowledgebase. This knowledgebase contains CTAs that are expressed in CRC ranked in order by expression rate. Correlation studies conducted by the inventors (see Example 11) suggest that vaccines which induce CTL responses against multiple antigens that are expressed in tumor cells can benefit patients. Therefore, seven CTAs with high expression rates in CRC were selected for inclusion in PolyPEPI1018 development. Details are set out in Table 25.

Table 25 Target CTAs in PolyPEPI1018 CRC vaccine

CTA Name	Expression Rate	Characterization
TSP50	89.47%	*Testis-Specific Protease-Like Protein 50* is an oncogene which induces cell proliferation, cell invasion, and tumor growth. It is frequently expressed in gastric-, breast-, cervical- and colorectal cancer samples; and rarely expressed in normal human tissues, except in spermatocytes of testes.
EpCAM	88.35%	*Epithelial Cell Adhesion Molecule* is a tumor associated antigen, which is expressed in colon cancers and over-expressed in various human carcinomas. The high expression of EpCAM in cancer-initiating stem cells makes it a valuable target for cancer vaccines. EpCAM is also expressed in at low or negligible levels in normal epithelial cells, with the exception of squamous epithelium, hepatocytes and keratinocytes.
Survivin	87.28%	*Survivin (Baculoviral IAP repeat-containing protein 5)* is a multi-tasking protein that promotes cell proliferation and inhibits apoptosis. Though it is strongly expressed in fetal tissues and necessary for normal development, it is not expressed in most adult tissues. Survivin is expressed in various cancers including carcinomas. Normal tissues that express low level survivin include thymus, $CD34^+$ bone-marrow-derived stem cells, and basal colonic epithelium. Dramatic over-expression of survivin compared with normal tissues iis observed in tumors in the lung, breast, colon, stomach, esophagus, pancreas, bladder, uterus, ovaries, large-cell non-Hodgkin's lymphoma, leukemias, neuroblastoma, melanoma and non-melanoma skin cancers.
CAGE1	74.47%	*Cancer-associated gene 1 protein* is a typical CTA, which might play a role in cell proliferation and tumorigenesis. CAGE1 is highly expressed in colorectal cancer tissues and weakly expressed in adjacent normal colorectal mucosa. In addition, CAGE1 is expressed in melanoma, hepatoma, and breast tumors. No CAGE1 protein expression is detected in healthy human tissues, other than testes.

SPAG9	74.36%	*Sperm-associated antigen 9* is involved in c-Jun N-terminal kinase-signaling and functions as a scaffold protein, thus playing an important role in cell survival, proliferation, apoptosis and tumor development. SPAG9 expression was detected in epithelial ovarian cancer (90%), breast cancer (88%), cervical cancer (82%), renal cell cancer (88%) and colorectal cancer (74%) patients. None of the adjacent noncancerous tissues showed antigen expression. SPAG9 expression is restricted to testis.
FBXO39	38.60%	*FBXO39 (BCP-20)* is a testis specific protein and is an important part of the E3 ubiquitin ligase complex. It participates in ubiquitination and has a role in regulating the cell cycle, immune responses, signaling, and proteasomal degradation of proteins. FBXO39 is expressed in colon and breast cancers. FBXO39 expression has also been detected in ovary, placenta, and lung. FBXO39 expression is 100-fold higher in testis and 1,000-fold higher in colorectal cancers compared with normal tissue.
MAGEA8	43.75%	*Melanoma-associated antigen 8* function is not known, though it may play a role in embryonal development and tumor transformation or aspects of tumor progression. MAGE-A8 gene is expressed in CRC and hepatocellular carcinoma. MAGE-A8 expression in normal tissues is restricted to the testis and the placenta.

B. Precise Targeting is Achieved by PEPI3+ Biomarker Based Vaccine Design

As described above the PEPI3+ biomarker predicts a subject's vaccine induced T cell responses. The inventors developed and validated a test to accurately identify the PEPIs from antigen sequences and HLA genotypes (Examples 1, 2, 3). The PEPI Test algorithm was used to identify the dominant PEPIs (besEPIs) from the 7 target CTAs to be included in PolyPEPI1018 CRC vaccine.

The dominant PEPIs identified with the process described here can induce CTL responses in the highest proportion of subjects:

i. Identification of all HLA class I binding PEPIs from the 7 CTA targets in each of the 433 subjects in the Model Population

ii. Identification of the dominant PEPIs (BestEPIs) that are PEPIs present in the largest subpopulation.

The 12 dominant PEPIs that are derived from the 7 CTAs in PolyPEPI1018 are presented in the Table 26. The PEPI % in Model Population indicates the proportion of 433 subjects with the indicated PEPI, i.e. the proportion of subjects where the indicated PEPI can induce CTL responses. There is very high variability (18% - 78%) in the dominant PEPIs to induce CTL responses despite the optimization steps used in the identification process.

Table 26 CRC specific HLA class I binding dominant PEPIs in PolyPEPI1018

Dominant PEPI3+ for each of the 7 CTAs in PolyPEPI1018 in CRC patients			
Peptides in PolyPEPI1018	**CRC Antigens**	**Dominant PEPI3+**	**PEPI3+% in Model Population**
CRC-P1	TSP50	TTMETQFPV	36%
		YRAQRFWSW	20%
CRC-P2	EpCAM	RTYWIIIEL	51%
	Survivin	RAIEQLAAM	26%
CRC-P3	EpCAM	YVDEKAPEF	28%
	MAGE-A8	KVAELVRFL	18%
CRC-P6	CAGE1	KMHSLLALM	42%
	Survivin	STFKNWPFL	15%
CRC-P7	CAGE1	KSMTMMPAL	37%
	SPAG9	VMSERVSGL	28%
CRC-P8	FBXO39	FMNPYNAVL	78%
		FFFERIMKY	46%

The inventors optimized each dominant PEPI to bind to most HLA class II alleles of most subjects. This should enhance efficacy, because it will induce $CD4^+$ T helper cells that can augment $CD8^+$ CTL responses and contribute to long lasting T cell responses. The example presented in Figure 4 demonstrates that PEPIs that bind to $\geqq 3$ HLA class II alleles most likely activate T helper cells.

The 15-mer peptides selected with the process described here contain both HLA class I and class II binding dominant PEPIs. Therefore, these peptides can induce both CTL and T helper responses in the highest proportion of subjects.

Process:

1. Identification the HLA class II genotype of 400 normal donors*
2. Extension of each 9-mer dominant PEPI (Table 20) on both sides with amino acids that match the source antigen
3. Prediction of HLA class II PEPIs of 400 normal donors using an IEDB algorithm
4. Selection the 15-mer peptide with the highest proportion of subject have HLA Class II binding PEPIs

5. Ensure the presence of one dominant HLA class II PEPI in each vaccine peptide when joining two 15-mer peptides

The 12 optimized 15-mer peptides derived from the 7 CTAs in PolyPEPI1018 are presented in the Table 27. These peptides have different HLA class II binding characteristics. There is a high variability (0% - 100%) in PEPI generation capacity (≧3 HLA binding) among these peptides despite such an optimized personalized vaccine design.

Table 27 Antigen specific HLA class II binding PEPIs in PolyPEPI1018.

Nr.	CRC antigens	Average HLA class II binding alleles	% subjects with ≧1 HLA class II binding	% subjects with ≧2 HLA class II binding	% subjects with ≧3 HLA class II binding	% subjects with ≧4 HLA class II binding
CRC-P1	TSP50 (83-97)	0	0%	0%	0%	0%
	TSP50 (190-204)	**4**	100%	99%	88%	**53%**
CRC-P2	**EPCAM(139-153)**	**5**	100%	100%	100%	**98%**
	SURVIVIN(127-141)	2	84%	58%	26%	11%
CRC-P3	EPCAM(251-265)	0	0%	0%	0%	0%
	MAGE-A8(113-127)	**4**	100%	100%	95%	**72%**
CRC-P6	**CAGE1(613-627)**	**5**	100%	100%	99%	**95%**
	SURVIVIN(15-29)	3	100%	97%	83%	45%
CRC-P7	**CAGE1(759-773)**	**3**	100%	98%	87%	**56%**
	SPAG9(16-30)	1	66%	35%	9%	2%
CRC-P8	FBXO39(95-109)	3	100%	94%	43%	13%
	FBXO39(284-298)	**5**	100%	100%	100%	**98%**

The 30-mer vaccine peptides have the following advantages compared to shorter peptides:

(i) Multiple precisely selected tumor specific immunogens: each 30 mer contains two precisely selected cancer specific immunogenic peptides that are capable to induce CTL

and T helper responses in the majority of the relevant population (similar to the model population).

(ii) Ensure natural antigen presentation. 30-mer long polypeptides can be viewed as pro-drugs: They are not biologically active by themselves, but are processed to smaller peptides (9 to 15 amino acid long) to be loaded into the HLA molecules of professional antigen presenting cells. The antigen presentation resulting from long peptide vaccination reflects physiological pathways for presentation in both HLA class I and class II molecules. In addition, long peptide processing in the cells is much more efficient than that of large intact proteins.

(iii) Exclude induction of tolerizing T cell responses. 9-mer peptides do not require processing by professional antigen-presenting cells and therefore bind exogenously to the HLA class I molecules. Thus, injected short peptides will bind in large numbers to HLA class I molecules of all nucleated cells that have surface HLA class I. In contrast, >20-mers long peptides are processed by antigen presenting cells before binding to HLA class I. Therefore, vaccination with long peptides is less likely to lead to tolerance and will promote the desired antitumor activity.

(iv) Induce long lasting T cell responses because it can stimulate T helper responses by binding to multiple HLA class II molecules

(v) Utility. GMP manufacturing, formulation, quality control and administration of a smaller number of peptides (each with all of the above characteristics) is more feasible than a larger number of peptides supplying different characteristics.

Each 30-mer peptide in PolyPEPI1018 consists of 2 HLA class I binding dominant PEPIs and at least one strong HLA class II binding PEPI. Strong binding PEPIs bind to 4 HLA class II alleles in >50% of individuals. Therefore, the vaccine peptides are tailored to both HLA class I and class II alleles of individual subjects in a general population (which is a relevant population for CRC vaccine design).

As demonstrated above the high HLA genotype variability in subjects results in high variability of T cell responses induced by PolyPEPI1018. This justifies the co-development of a

CDx that determines likely responders. The PEPI3+ and >2PEPI3+ biomarkers could predict the immune response and clinical responses, respectively, of subjects vaccinated with PolyPEPI1018 as detailed in Examples 11 and 12. These biomarkers will be used to co-develop a CDx which predicts likely responders to PolyPEPI1018 CRC vaccine.

Example 20 - Analysis of the composition and immunogenicity of PolyPEPI1018 CRC vaccine

Selected peptides for the PolyPEPI1018 composition are as shown in Table 28.

Table 28 - Selected Colorectal Cancer Vaccine peptides for PolyPEPI1018 composition

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
130	CCV1000-5-1	TSP50	PSTTMETQFPVSEGKSRYRAQRFWSWVGQA	53%	88%
121	CCV1000-2-2	EpCAM/Survivin	VRTYWIIIELKHKARTAKKVRRAIEQLAAM	57%	100%
131	CCV1000-5-3	EpCAM /Mage-A8	YVDEKAPEFSMQGLKDEKVAELVRFLLRKY	43%	95%
124	CCV1000-2-6	Cage/Survivin	LASKMHSLLALMVGLKDHRISTFKNWPFLE	58%	99%
134	CCV1000-5-7	Cage/Spag9	PKSMTMMPALFKENRSGAVMSERVSGLAGS	57%	87%
126	CCV1000-2-8	FBXO39	KFMNPYNAVLTKKFQKVNFFFERIMKYERL	90%	100%
			PolyPEPI1018 (6 peptide together)	98%	100%

* Percentage of individuals having HLA class I binding PEPI3+ within the Model Population (n=433).

**Percentage of individuals having HLA class II binding PEPI3+ within the Model Population (n=433).

The peptides of PolyPEPI1018 are formulated in two mixtures, MIX1 containing the peptides of SEQ ID: 130, 131 and MIX2 containing the peptides of SEQ ID: 121, 124, 134, 126. MIX 1 and MIX 2 may be administered sequentially.

Characterization of immunogenicity

The inventors used the PEPI3+ Test to characterized the immunogenicity of PolyPEPI1018 in a cohort of 37 CRC patients with complete HLA genotype data. T cell responses were predicted in each patient against the same 9 mer peptides that will be used in clinical trials. These peptides represent the 12 dominant PEPI3+ within the PolyPEPI1018 peptides. The 9 mers are shown in Table 26.

The specificity and sensitivity of PEPI3+ prediction depends on the actual number of HLAs predicted to bind a particular epitope. Specifically, the inventors have determined that the

probability that one HLA-restricted epitope induces a T cell response in a subject is typically 4%, which explains the poor sensitivity of the state-of-art prediction methods based on HLA restricted epitope prediction. Applying the PEPI3+ methodology, the inventors determined the probability that T cell response to each of the dominant PEPI3+-specific would be induced by PolyPEPI1018 in the 37 CRC patients. The results from this analysis are summarized in the Table 29.

Table 29 Probability of Dominant PEPI in the 6 Peptides of PolyPEPI1018 in 37 CRC Patients

CRC Patient	CRC-P1		CRC-P2		CRC-P3		CRC-P6		CRC-P7		CRC-P8		Expected Number of PEPIs
	TSP50 (83-97)	TSP50 (190-204)	EpCAM (139-153)	Survivin (127-141)	EpCAM (251-265)	MAGEA8 (113-127)	CAGE1 (613-627)	Survivin (15-29)	CAGE1 (759-773)	SPAG9 (16-30)	FBXO39 (95-109)	FBXO39 (284-298)	
CRC-01	22%	4%	22%	4%	22%	22%	**100%**	1%	**98%**	**84%**	**100%**	22%	**5.01**
CRC-02	22%	1%	22%	22%	22%	22%	**100%**	1%	**98%**	22%	**100%**	**98%**	**5.29**
CRC-03	**84%**	22%	**84%**	22%	22%	22%	**84%**	22%	22%	22%	**100%**	22%	**5.29**
CRC-04	22%	**84%**	22%	4%	22%	4%	**98%**	4%	4%	22%	**100%**	**84%**	**4.70**
CRC-05	22%	22%	4%	4%	22%	4%	**98%**	1%	4%	4%	**100%**	**84%**	**3.68**
CRC-06	**84%**	22%	4%	**84%**	**98%**	4%	22%	4%	4%	4%	**100%**	**98%**	**5.27**
CRC-07	22%	22%	22%	22%	22%	4%	**98%**	1%	22%	22%	**100%**	**84%**	**4.41**
CRC-08	22%	22%	22%	**98%**	**84%**	22%	**84%**	22%	22%	22%	**100%**	**84%**	**6.04**
CRC-09	22%	**84%**	**84%**	**84%**	**84%**	22%	**100%**	4%	22%	22%	**98%**	**84%**	**7.10**
CRC-10	4%	**98%**	22%	22%	4%	4%	4%	22%	22%	22%	**98%**	**84%**	**4.06**
CRC-11	22%	22%	4%	4%	22%	4%	**84%**	1%	4%	4%	**98%**	**84%**	**3.53**
CRC-12	**84%**	22%	4%	22%	4%	4%	**84%**	4%	**84%**	4%	**100%**	22%	**4.38**
CRC-13	**84%**	22%	4%	22%	**84%**	4%	**84%**	1%	1%	4%	**100%**	**98%**	**5.07**
CRC-14	22%	**84%**	4%	4%	22%	4%	**84%**	1%	4%	4%	**100%**	**84%**	**4.16**
CRC-15	**84%**	22%	22%	22%	22%	4%	**84%**	4%	22%	4%	**100%**	**84%**	**4.74**
CRC-16	4%	**84%**	4%	4%	22%	4%	**84%**	1%	4%	22%	**100%**	**84%**	**4.16**
CRC-17	**84%**	**84%**	4%	**84%**	**84%**	4%	4%	4%	4%	4%	**100%**	22%	**4.82**
CRC-18	**84%**	22%	22%	**84%**	**84%**	4%	22%	22%	4%	4%	**100%**	**84%**	**5.36**
CRC-19	22%	22%	22%	22%	22%	4%	**98%**	4%	22%	22%	**100%**	**84%**	**4.45**
CRC-20	**84%**	22%	4%	22%	**84%**	4%	**84%**	1%	4%	4%	**100%**	**98%**	**5.10**
CRC-21	22%	22%	22%	22%	**84%**	22%	**98%**	4%	4%	22%	**100%**	**84%**	**5.06**
CRC-22	22%	**98%**	**84%**	4%	22%	22%	**84%**	22%	**84%**	22%	**98%**	22%	**5.84**

CRC Patient	CRC-P1		CRC-P2		CRC-P3		CRC-P6		CRC-P7		CRC-P8		Expected Number of PEPIs
	TSP50 (83-97)	TSP50 (190-204)	EpCAM (139-153)	Survivin (127-141)	EpCAM (251-265)	MAGEA8 (113-127)	CAGE1 (613-627)	Survivin (15-29)	CAGE1 (759-773)	SPAG9 (16-30)	FBXO39 (95-109)	FBXO39 (284-298)	
CRC-23	**84%**	**84%**	**84%**	**84%**	**84%**	22%	**84%**	**84%**	**84%**	4%	**100%**	**84%**	**8.82**
CRC-24	22%	22%	4%	4%	22%	4%	**84%**	1%	4%	4%	**100%**	**84%**	**3.55**
CRC-25	22%	**84%**	22%	4%	22%	4%	**84%**	4%	22%	4%	**100%**	**84%**	**4.56**
CRC-26	**84%**	22%	4%	22%	**84%**	4%	**84%**	1%	4%	4%	**100%**	**84%**	**4.97**
CRC-27	22%	22%	4%	4%	22%	4%	**98%**	1%	4%	4%	**100%**	**84%**	**3.68**
CRC-28	**84%**	22%	4%	22%	**84%**	4%	**84%**	1%	4%	4%	**100%**	**98%**	**5.10**
CRC-29	**84%**	**84%**	4%	22%	22%	4%	**84%**	1%	22%	22%	**100%**	**84%**	**5.33**
CRC-30	**84%**	22%	4%	22%	**84%**	4%	**84%**	1%	4%	4%	**100%**	**98%**	**5.10**
CRC-31	22%	**84%**	22%	4%	4%	4%	22%	1%	4%	4%	**98%**	**84%**	**3.53**
CRC-32	**84%**	**84%**	4%	**84%**	22%	4%	4%	4%	4%	4%	**98%**	**84%**	**4.80**
CRC-33	**84%**	22%	4%	22%	**84%**	4%	**84%**	1%	4%	4%	**100%**	**98%**	**5.10**
CRC-34	22%	22%	22%	22%	22%	4%	**84%**	1%	22%	4%	**100%**	**84%**	**4.09**
CRC-35	22%	4%	4%	1%	22%	4%	4%	1%	4%	4%	**84%**	**84%**	**2.37**
CRC-36	22%	4%	4%	1%	22%	4%	4%	1%	4%	4%	**84%**	**84%**	**2.37**
CRC-37	22%	4%	4%	1%	22%	4%	4%	1%	4%	4%	**84%**	**84%**	**2.37**

Abbreviations: CRC = colorectal cancer; PEPI = personal epitope

Note: Percentages represent the likelihood of CD8+ T cell Responses Induced by PolyPEPI1018.

Overall, these results show that the most immunogenic peptide in PolyPEPI1018 is CRC-P8, which it is predicted to bind to >3 HLAs in most patients. The least immunogenic peptide, CRC-P3, binds to >1 HLA in many patients and has a 22% chance of inducing T cell responses. Since bioassays used to detect T cell responses are less accurate than PEPI3+, this calculation may be the most accurate characterization of the T cell responses in CRC patients. Though MAGE-A8 and SPAG9 were immunogenic in the Model Population used for vaccine design, MAGE-A8-specific PEPI3+ were absent in the 37 CRC patients, and only one patient (3%) had SPAG9 specific PEPI3+.

Further characterization of the predicted PolyPEPI1018 response rate in the model population described in Example 8 and in 295 CRC patients with known HLA class I genotypes are shown in Tables 30 and 31.

Table 30 – PolyPEPI1018 Response Rates in the Model Population (433 Normal donors)

PolyPEPI1018 Response Rates	>=1	>=2	>=3	>=4	>=5	>=6	>=7	>=8	>=9
Multi PEPI	98%	94%	83%	70%	52%	38%	27%	18%	11%
Multi Peptide	98%	91%	73%	52%	30%	12%	N/D	N/D	N/D
Multi Antigen	98%	92%	72%	49%	31%	14%	6%	N/D	N/D

Table 31 – PolyPEPI1018 Response Rates for 295 CRC patients

PolyPEPI1018 Response Rates	>=1	>=2	>=3	>=4	>=5	>=6	>=7	>=8	>=9
Multi PEPI	99%	96%	92%	85%	69%	53%	40%	32%	25%
Multi Peptide	99%	93%	86%	71%	49%	29%	N/D	N/D	N/D
Multi Antigen	99%	93%	86%	72%	49%	32%	13%	N/D	N/D

Characterization of toxicity – immunoBLAST

A method was developed that can be performed on any antigen to determine its potential to induce toxic immune reaction, like autoimmunity. The method is referred to herein as immunoBLAST. PolyPEPI1018 contains six 30-mer polypeptides. Each polypeptide consists of two 15-mer peptide fragments derived from antigens expressed in CRC. Neoepitopes might be generated in the joint region of the two 15-mer peptides and could induce undesired T cell responses against healthy cells (autoimmunity). This was assesses using the immunoBLAST methodology.

A 16-mer peptide for each of the 30-mer components of PolyPEP1018 was designed. Each 16-mer contains 8 amino acids from the end of the first 15 residues of the 30-mer and 8 amino acids from the beginning of the second 15 residues of the 30-mer – thus precisely spanning the joint region of the two 15-mers. These 16-mers are then analysed to identify cross-reactive regions of local similarity with human sequences using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi), which compares protein sequences to sequence databases and calculates the statistical significance of matches. 8-mers within the 16-mers were selected as the examination length since that length represents the minimum length needed for a peptide to form an epitope, and is the distance between the anchor points during HLA binding.

As shown in Figure 19, the positions of amino acids in a polypeptide are numbered. The start positions of potential 9-mer peptides that can bind to HLAs and form neoepitopes are the 8 amino acids in positions 8-15. The start positions of tumor antigen derived peptides harbored by the 15-mers that can form the pharmaceutically active epitopes are 7+7=14 amino acids at position 1-7 and 16-22. The ratio of possible neoepitope generating peptides is 36.4% (8/22).

The PEPI3+ Test was to identify neoepitopes and neoPEPI among the 9-mer epitopes in the joint region. The risk of PolyPEPI1018 inducing unwanted T cell responses was assessed in the 433 subjects in the Model Population by determining the proportion of subjects with PEPI3+ among the 9-mers in the joint region. The result of neoepitope/neoPEPI analysis is summarized in Table 32. In the 433 subjects of the Model Population, the average predicted epitope number that could be generated by intracellular processing was 40.12. Neoepitopes were frequently generated; 11.61 out of 40.12 (28.9%) epitopes are neoepitopes. Most of the peptides were able to be be identified as a neoepitope, but the number of subjects that present neoepitopes varied.

Epitopes harbored by PolyPEPI1018 create an average of 5.21 PEPI3+. These PEPIs can activate T cells in a subject. The amount of potential neoPEPIs was much lower than neoepitopes (3.7%). There is a marginal possibility that these neoPEPIs compete on T cell activation with PEPIs in some subjects. Importantly, the activated neoPEPI specific T cells had no targets on healthy tissue.

Table 32 - Identification of Potential Neoepitopes of PolyPEPI1018

PolyPEPI1018 Peptide ID:	Potential Neoepitope	Epitope & PEPI3+ binding in 433 Subjects of the Model Population							
		Epitope Binding (1 x HLA)				PEPI3+ binding (3 x HLA)			
		Sub#	Sub%	NeoEPI	NeoEPI count	Sub#	Sub%	NeoPEPI	NeoPEPI count
CRC-P1	QFPVSEGKS	0	0.0%		7	0	0.0%		3
	FPVSEGKSR	160	37.0%	X		1	0.2%	X	
	PVSEGKSRY	150	34.6%	X		0	0.0%		
	VSEGKSRYR	194	44.8%	X		1	0.2%	X	
	SEGKSRYRA	113	26.1%	X		0	0.0%		
	EGKSRYRAQ	77	17.8%	X		0	0.0%		
	GKSRYRAQR	37	8.5%	X		0	0.0%		
	KSRYRAQRF	337	77.8%	X		33	7.6%	X	
CRC-P2	IELKHKART	32	7.4%	X	7	0	0.0%		1
	ELKHKARTA	63	14.5%	X		0	0.0%		
	LKHKARTAK	59	13.6%	X		0	0.0%		
	KHKARTAKK	166	38.3%	X		1	0.2%	X	
	HKARTAKKV	0	0.0%			0	0.0%		
	KARTAKKVR	70	16.2%	X		0	0.0%		
	ARTAKKVRR	134	30.9%	X		0	0.0%		
	RTAKKVRRA	41	9.5%	X		0	0.0%		
CRC-P3	EFSMQGLKD	0	0.0%		5	0	0.0%		1
	FSMQGLKDE	188	43.4%	X		0	0.0%		
	SMQGLKDEK	138	31.9%	X		0	0.0%		
	MQGLKDEKV	16	3.7%	X		0	0.0%		
	QGLKDEKVA	0	0.0%			0	0.0%		
	GLKDEKVAE	0	0.0%			0	0.0%		
	LKDEKVAEL	186	43.0%	X		3	0.7%	X	
	KDEKVAELV	51	11.8%	X		0	0.0%		
CRC-P6	LLALMVGLK	252	58.2%	X	7	0	0.0%		1
	LALMVGLKD	86	19.9%	X		0	0.0%		
	ALMVGLKDH	65	15.0%	X		0	0.0%		
	LMVGLKDHR	97	22.4%	X		0	0.0%		
	MVGLKDHRI	67	15.5%	X		0	0.0%		
	VGLKDHRIS	0	0.0%			0	0.0%		
	GLKDHRIST	4	0.9%	X		0	0.0%		
	LKDHRISTF	195	45.0%	X		5	1.2%	X	
CRC-P7	PALFKENRS	0	0.0%		5	0	0.0%		1
	ALFKENRSG	0	0.0%			0	0.0%		
	LFKENRSGA	41	9.5%	X		0	0.0%		
	FKENRSGAV	114	26.3%	X		0	0.0%		

PolyPEPI1018 Peptide ID:	Potential Neoepitope	Epitope & PEPI3+ binding in 433 Subjects of the Model Population							
		Epitope Binding (1 x HLA)				PEPI3+ binding (3 x HLA)			
		Sub#	Sub%	NeoEPI	NeoEPI count	Sub#	Sub%	NeoPEPI	NeoPEPI count
	KENRSGAVM	261	60.3%	X		0	0.0%		
	ENRSGAVMS	0	0.0%			0	0.0%		
	NRSGAVMSE	227	52.4%	X		0	0.0%		
	RSGAVMSER	197	45.5%	X		2	0.5%	X	
CRC-P8	AVLTKKFQK	181	41.8%	X	7	0	0.0%		3
	VLTKKFQKV	208	48.0%	X		2	0.5%	X	
	LTKKFQKVN	0	0.0%			0	0.0%		
	TKKFQKVNF	25	5.8%	X		0	0.0%		
	KKFQKVNFF	250	57.7%	X		12	2.8%	X	
	KFQKVNFFF	273	63.0%	X		23	5.3%	X	
	FQKVNFFFE	163	37.6%	X		0	0.0%		
	QKVNFFFER	110	25.4%	X		0	0.0%		

Abbreviations: CRC = colorectal cancer; HLA = human leukocytic antigen; PEPI = personal epitope

Each of the 30-mer peptides in PolyPEPI1018 were released for clinical development since none of the 8-mers in the joint regions matched any human protein, except the target CTAs.

Characterisation of activity / efficacy

The inventors have developed pharmacodynamic biomarkers to predict the activity/effect of vaccines in individual human subjects as well as in populations of human subjects. These biomarkers expedite more effective vaccine development and also decrease the development cost. The inventors have the following tools:

Antigen expression knowledgebase: The inventors have collected data from experiments published in peer reviewed scientific journals regarding the tumor antigens expressed by tumor cells and organized by tumor type to create a database of CTA expression levels – CTA database (CTADB). As of April 2017, the CTADB contained data from 145 CTAs from 41,132 tumor specimens, and was organized by the CTA expression frequencies in different types of cancer.

***In silico* trial populations:** The inventors have also collected data on the HLA genotypes of several different model populations. Each individual in the populations has complete 4-digit HLA genotype and ethnicity data. The populations are summarized in Table 33.

Table 33 In silico trial populations

Population	Number of Subjects	Inclusion criteria
Model Population	433	Complete HLA class I genotype Diverse ethnicity
CRC patients	37	Complete HLA class I genotype CRC diagnosis, unknown ethnicity
"Big" Population	7,189	Complete HLA class I genotype Diverse ethnicity
Chinese Population	234	Complete HLA class I genotype Chinese ethnicity
Irish Population	999	Complete HLA class I genotype Irish ethnicity

Abbreviations: CRC = colorectal cancer; HLA = human leukocyte antigen

Using these tools (or potentially equivalent databases or model populations), the following markers can be assessed:

- **AG95 – potency of a vaccine**: The number of antigens in a cancer vaccine that a specific tumor type expresses with 95% probability. AG95 is an indicator of the vaccine's potency, and is independent of the immunogenicity of the vaccine antigens. AG95 is calculated from the tumor antigen expression rate data, which is collected in the CTADB. Technically, AG95 is determined from the binomial distribution of CTAs, and takes into account all possible variations and expression rates. In this study, AG95 was calculated by cumulating the probabilities of a certain number of expressed antigens, by the widest range of antigens where the sum of probabilities was less than or equal to 95%. The correct value isbetween 0 (no expression expected with 95% probability) and maximum number of antigens (all antigens expressed with 95% probability).
- **PEPI3+ count – immunogenicity of a vaccine in a subject**: Vaccine-derived PEPI3+ are personal epitopes that induce T cell responses in a subject. PEPI3+ can be determined using the PEPI3+ Test in subjects who's complete 4-digit HLA genotype is known.
- **AP count – antigenicity of a vaccine in a subject**: Number of vaccine antigens with PEPI3+. Vaccines like PolyPEPI1018 contain sequences from antigens expressed in tumor cells. AP count is the number of antigens in the vaccine that contain PEPI3+, and the AP count represents the number of antigens in the vaccine that can induce T cell responses in a subject. AP count characterizes the vaccine-antigen specific T cell responses of the subject since it depends only on the HLA genotype of the subject and is independent of the subject's disease, age, and medication. The correct value is between 0 (no PEPI presented by the antigen) and maximum number of antigens (all antigens present PEPIs).
- **AP50 – antigenicity of a vaccine in a population**: The mean number of vaccine antigens with a PEPI in a population. The AP50 is suitable for the characterization of vaccine-antigen specific T cell responses in a given population since it depends on the HLA genotype of

subjects in a population. Technically, the AP count was calculated in the Model Population and the binomial distribution of the result was used to calculate the AP50.

- **AGP count – effectiveness of a vaccine in a subject**: Number of vaccine antigens expressed in the tumor with PEPI. The AGP count indicates the number of tumor antigens that vaccine recognizes and induces a T cell response against (hit the target). The AGP count depends on the vaccine-antigen expression rate in the subject's tumor and the HLA genotype of the subject. The correct value is between 0 (no PEPI presented by expressed antigen) and maximum number of antigens (all antigens are expressed and present a PEPI).
- **AGP50 – effectiveness of a cancer vaccine in a population**: The mean number of vaccine antigens expressed in the indicated tumor with PEPI (i.e., AGP) in a population. The AGP50 indicates the mean number of tumor antigens that the T cell responses induced by the vaccine can recognize. AGP50 is dependent on the expression rate of the antigens in the indicated tumor type and the immunogenicity of the antigens in the target population. AGP50 can estimate a vaccine's effectiveness in different populations and can be used to compare different vaccines in the same population. The computation of AGP50 is similar to that used for AG50, except the expression is weighted by the occurrence of the PEPI3+ in the subject on the expressed vaccine antigens. In a theoretical population, where each subject has a PEPI from each vaccine antigen, the AGP50 will be equal to AG50. In another theoretical population, where no subject has a PEPI from any vaccine antigen, the AGP50 will be 0. In general, the following statement is valid: $0 \leq AGP50 \leq AG50$.
- **mAGP – a candidate biomarker for the selection of likely responders:** Likelihood that a cancer vaccine induces T cell responses against multiple antigens expressed in the indicated tumor. mAGP is calculated from the expression rates of vaccine-antigens in CRC and the presence of vaccine derived PEPIs in the subject. Technically, based on the AGP distribution, the mAGP is the sum of probabilities of the multiple AGP (≥2 AGPs).

<u>Application of these markers to assess antigenicity and effectiveness PolyPEPI1018 in Individual Patients with CRC</u>

Table 34 shows the antigenicity and effectiveness of PolyPEPI1018 in 37 CRC patients using AP and AGP50, respectively. As expected from the high variability of PolyPEPI1018 specific T cell responses (see Table 29), the AP and AGP50 have high variability. The most immunogenic antigen in PolyPEPI1018 was FOXO39; each patient had a PEPI3+. However, FOXO39 is expressed only 39% of CRC tumors, suggesting that 61% of patients will have FOXO39 specific T cell responses that do not recognize the tumor. The least immunogenic antigen was MAGE-A8; none of the 37 CRC patients had a PEPI3+ despite the antigen being expressed in 44% of CRC tumors. These results illustrate that both expression and immunogenicity of antigens can be taken into account when determining a cancer vaccine's effectiveness.

AGP50 indicates the mean number of expressed antigens in CRC tumor with PEPIs. Patients with higher AGP50 values are more likely to respond to PolyPEPI1018 since higher AGP50 values indicate that the vaccine can induce T cell responses against more antigens expressed in CRC cells.

The last column in the table 32 shows the probability of mAGP (multiple AGP; i.e., at least 2 AGPs) in each of the 37 CRC patients. The average mAGP in patients with CRC is 66%, suggesting that there is a 66% likelihood that a CRC patient will induce T cell responses against multiple antigens expressed in the tumor.

Table 34 - Antigenicity (AP count), Effectiveness (AGP50 count), and mAGP of PolyPEPI1018 in 37 CRC Patients

Antigens (CTAs) in PolyPEPI1018	TSP50	EpCAM	Survivin	CAGE1	SPAG9	FBXO39	MAGE-A8	Number of AP (AP count)	Number of AGP50 (AGP50 count)	mAGP
Expression rate	89%	88%	87%	74%	74%	39%	44%			
CRC Patients										
CRC-01	0	0	0	1	1	1	0	3	1.87	90%
CRC-02	0	0	0	1	0	1	0	2	1.13	85%
CRC-03	1	1	0	1	0	1	0	4	2.91	97%
CRC-04	1	0	0	1	0	1	0	3	2.03	91%
CRC-05	0	0	0	1	0	1	0	2	1.13	78%
CRC-06	1	1	1	1	0	1	0	5	3.78	99%
CRC-07	0	0	0	1	0	1	0	2	1.13	84%
CRC-08	0	1	1	1	0	1	0	4	2.89	98%
CRC-09	1	1	1	1	0	1	0	5	3.78	99%

CRC-10	1	0	0	0	0	1	0	2	1.28	86%
CRC-11	0	0	0	1	0	1	0	2	1.13	79%
CRC-12	1	0	0	1	0	1	0	3	2.03	88%
CRC-13	1	1	1	1	0	1	0	5	3.78	98%
CRC-14	1	0	0	1	0	1	0	3	2.03	87%
CRC-15	1	0	0	1	0	1	0	3	2.03	90%
CRC-16	1	0	0	1	0	1	0	3	2.03	85%
CRC-17	1	1	1	0	0	1	0	4	3.04	96%
CRC-18	1	1	1	1	0	1	0	5	3.78	98%
CRC-19	0	0	0	1	0	1	0	2	1.13	85%
CRC-20	1	1	1	1	0	1	0	5	3.78	98%
CRC-21	0	1	0	1	0	1	0	3	2.01	93%
CRC-22	1	1	0	1	0	1	0	4	2.91	97%
CRC-23	1	1	1	1	0	1	0	5	3.78	99%
CRC-24	0	0	0	1	0	1	0	2	1.13	82%
CRC-25	1	0	0	1	0	1	0	3	2.03	89%
CRC-26	1	1	0	1	0	1	0	4	2.91	95%
CRC-27	0	0	0	1	0	1	0	2	1.13	78%
CRC-28	1	1	1	1	0	1	0	5	3.78	98%
CRC-29	1	0	0	1	0	1	0	3	2.03	92%
CRC-30	1	1	1	1	0	1	0	5	3.78	98%
CRC-31	1	0	0	0	0	1	0	2	1.28	80%
CRC-32	1	0	1	0	0	1	0	3	2.15	91%
CRC-33	1	1	1	1	0	1	0	5	3.78	98%
CRC-34	0	0	0	1	0	1	0	2	1.13	82%
CRC-35	0	0	0	0	0	1	0	1	0.39	55%
CRC-36	0	0	0	0	0	1	0	1	0.39	55%
CRC-37	0	0	0	0	0	1	0	1	0.39	55%

Abbreviations: CRC = colorectal cancer; PEPI = personal epitope; CTA = cancer testis antigen; AP = expressed antigens with ≥1 PEPI

These biomarkers have immediate utility in vaccine development and in the routine clinical practice because they do not require invasive biopsies. Antigen expression data can be obtained from achieved tumor specimen and organized in databases. 4-digit HLA genotyping can be done from a saliva specimen. It is a validated test performed by certified laboratories worldwide for transplantation and paternity testing. These assessments will allow drug developers and physicians to gain deeper insights into the immunogenicity and activity of tumor response and the possible emergence of resistance.

Application of these markers to asses antigenicity and effectiveness PolyPEPI1018 in populations

Antigenicity of PolyPEPI1018 CRC Vaccine in a general population

The antigenicity of PolyPEPI1018 in a subject is determined by the AP count, which indicates the number of vaccine antigens that induce T cell responses in a subject. The AP count of PolyPEPI1018 was determined in each of the 433 subjects in the Model Population using the PEPI Test, and the AP50 count was then calculated for the Model Population.

As shown in Figure 20 the AP50 of PolyPEPI1018 in the Model Population is 3.62. Therefore, the mean number of immunogenic antigens (i.e., antigens with ≥1 PEPI) in PolyPEPI1018 in a general population is 3.62.

Effectiveness of PolyPEPI1018 CRC Vaccine in a general population

Vaccine induced T cells can recognize and kill tumor cells if a PEPI in the vaccine is presented by the tumor cell. The number of AGPs (expressed antigens with PEPI) is an indicator of vaccine effectiveness in an individual, and is dependent on both the potency and antigenicity of PolyPEPI1018. As shown in Figure 21, the mean number of immunogenic CTAs (i.e., AP [expressed antigens with ≥1 PEPI]) in PolyPEPI1018 is 2.54 in the Model Population.

The likelihood that PolyPEPI1018 induces T cell responses against multiple antigens in a subject (i.e., mAGP) in the Model Population is 77%.

Comparison of the PolyPEPI1018 CRC vaccine activities in different populations

Tables 35 to 37 show comparison of the immunogenicity, antigenicity, and effectiveness of PolyPEPI1018 in different populations.

Table 35 - Comparison of Immunogenicity, Antigenicity, and Effectiveness of PolyPEPI1018 in Different Sub-populations

Populations	**Number of subject**	**Number of PEPI3+**		**Number of AP**		**Number of AGP50**	
		Average	**SD**	**Average**	**SD**	**Average**	**SD**

CRC	**37**	5.16	1.98	3.19	1.31	2.21	1.13
Model	**433**	5.02	2.62	3.62	1.67	2.54	1.25
Big	**7,189**	5.20	2.82	3.75	1.74	2.66	1.30
Chinese	**324**	5.97	3.16	4.28	1.78	3.11	1.30
Irish	**999**	3.72	1.92	2.86	1.46	1.94	1.10

Abbreviations: CRC = colorectal cancer; PEPI = personal epitope; SD = standard deviation; AP = expressed antigens with ≥1 PEPI

The average number of PEPI3+ and AP results demonstrate that PolyPEPI1018 is highly immunogenic and antigenic in all populations; PolyPEPI1018 can induce an average of 3.7 – 6.0 CRC specific T cell clones against 2.9 – 3.7 CRC antigens. PolyPEPI1018 immunogenicity was similar in patients with CRC and the average population ($p>0.05$), this similarity may have been due to the small sample size of the CRC population. Additional analyses suggest that PolyPEPI1018 is significantly more immunogenic in a Chinese population compared to an Irish or a general population ($p<0.0001$). The differences in immunogenicity are also reflected in the effectiveness of the vaccine as characterized by AGP50; PolyPEPI1018 is most effective in a Chinese population and less effective in an Irish population. Since a CDx will be used to select likely responders to PolyPEPI1018, ethnic differences will only be reflected in the higher percentage of Chinese ind - viduals that might be eligible for treatment compared with Irish individuals.

Table 36 - PolyPEPI1018 CRC Vaccine, Predicted Immune Response Rates Against Multiple CRC Antigens

Population		No. subjects	PolyPEPI1018 MultiAG CTL Responses				
			≧3	≧4	≧5	≧6	7
CRC Patients	Vietnamese	211	91%	81%	56%	38%	17%
	US	44	57%	34%	20%	5%	0%
	Caucasian	83	75%	51%	30%	17%	4%
	US	400	61%	39%	25%	12%	3%
	Europe	1,386	55%	30%	18%	7%	1%
	Chinese	324	84%	68%	45%	26%	15%

Normal Donors	Okinawan (JP)	104	81%	57%	36%	16%	13%
	Japanese	45	77%	55%	34%	16%	13%

Table 37 - PolyPEPI1018 CRC Vaccine, Predicted Immune Response Rates Against Multiple CRC Antigens

Population		No. subjects	Number of PEPI		Number of AP		Number of AGP50	
			Average	SD	Average	SD	Average	SD
CRC Patients	Vietnamese	211	6.96	3.01	4.81	1.58	3.47	1.16
	US	44	4.05	2.05	3.00	1.46	2.05	1.12
	Caucasian	83	4.75	2.39	3.57	1.76	2.50	1.27
Normal Donors	US	400	4.30	2.50	3.19	1.74	2.17	1.30
	Europe	1,386	3.84	2.01	2.94	1.51	2.00	1.14
	Chinese	324	5.97	3.16	4.28	1.78	3.11	1.30
	Okinawan (JP)	104	5.29	2.58	4.01	1.63	2.91	1.19
	Japanese	45	5.31	3.27	3.67	1.77	2.66	1.29

Example 21 - Personalised Immunotherapy Composition for Treatment of Ovarian Cancer

This example describes the treatment of an ovarian cancer patient with a personalised immunotherapy composition, wherein the composition was specifically designed for the patient based on her HLA genotype based on the disclosure described herein. This Example and Example 22 below provide clinical data to support the principals regarding binding of epitopes by multiple HLA of a subject to induce a cytotoxic T cell response on which the present disclosure is based.

The HLA class I and class II genotype of metastatic ovarian adenocarcinoma cancer patient XYZ was determined from a saliva sample.

To make a personalized pharmaceutical composition for patient XYZ thirteen peptides were selected, each of which met the following two criteria: (i) derived from an antigen that is expressed in ovarian cancers, as reported in peer reviewed scientific publications; and (ii) comprises a fragment that is a T cell epitope capable of binding to at least three HLA class I of patient XYZ (Table 38). In addition, each peptide is optimized to bind the maximum number of HLA class II of the patient.

Table 38: XYZ ovarian cancer patient's personalized vaccine

XYZ's vaccine	Target Antigen	Antigen Expression	20mer peptides	MAX HLA classI	MAX HLA classII
POC01_P1	AKAP4	89%	NSLQKQLQAVLQWIAASQFN	3	5
POC01_P2	BORIS	82%	SGDERSDEIVLTVSNSNVEE	4	2
POC01_P3	SPAG9	76%	VQKEDGRVQAFGWSLPQKYK	3	3
POC01_P4	OY-TES-1	75%	EVESTPMIMENIQELIRSAQ	3	4
POC01_P5	SP17	69%	AYFESLLEKREKTNFDPAEW	3	1
POC01_P6	WT1	63%	PSQASSGQARMFPNAPYLPS	4	1
POC01_P7	HIWI	63%	RRSIAGFVASINEGMTRWFS	3	4
POC01_P8	PRAME	60%	MQDIKMILKMVQLDSIEDLE	3	4
POC01_P9	AKAP-3	58%	ANSVVSDMMVSIMKTLKIQV	3	4
POC01_P10	MAGE-A4	37%	REALSNKVDELAHFLLRKYR	3	2
POC01_P11	MAGE-A9	37%	ETSYEKVINYLVMLNAREPI	3	4
POC01_P12a	MAGE-A10	52%	DVKEVDPTGHSFVLVTSLGL	3	4
POC01_P12b	BAGE	30%	SAQLLQARLMKEESPVVSWR	3	2

Eleven PEPI3 peptides in this immunotherapy composition can induce T cell responses in XYZ with 84% probability and the two PEPI4 peptides (POC01-P2 and POC01-P5) with 98% probability, according to the validation of the PEPI Test shown in Table 3. T cell responses target 13 antigens expressed in ovarian cancers. Expression of these cancer antigens in patient XYZ was not tested. Instead the probability of successful killing of cancer cells was determined based on the probability of antigen expression in the patient's cancer cells and the positive predictive value of the ≥1 PEPI3+ Test (AGP count). AGP count predicts the effectiveness of a vaccine in a subject: Number of vaccine antigens expressed in the patient's tumor (ovarian adenocarcinoma) with PEPI. The AGP count indicates the number of tumor antigens that vaccine recognizes and induces a T cell response against the patient's tumor (hit the target). The AGP count depends on the vaccine-antigen expression rate in the subject's tumor and the HLA genotype of the subject. The correct value is between 0 (no PEPI presented by expressed antigen) and maximum number of antigens (all antigens are expressed and present a PEPI).

The probability that patient XYZ will express one or more of the 12 antigens is shown in

Fig. 22. AGP95 = 5, AGP50 = 7.9, mAGP = 100%, AP = 13.

A pharmaceutical composition for patient XYZ may be comprised of at least 2 from the 13 peptides (Table 38), because the presence in a vaccine or immunotherapy composition of at least two polypeptide fragments (epitopes) that can bind to at least three HLA of an individual (≥2 PEPI3+) was determined to be predictive for a clinical response. The peptides are synthetized, solved in a pharmaceutically acceptable solvent and mixed with an adjuvant prior to injection. It is desirable for the patient to receive personalized immunotherapy with at least two peptide vaccines, but preferable more to increase the probability of killing cancer cells and decrease the chance of relapse.

For treatment of patient XYZ the 12 peptides were formulated as 4 x 3/4 peptide (POC01/1, POC01/2, POC01/3, POC01/4). One treatment cycle is defined as administration of all 13 peptides within 30 days.

Patient history:

Diagnosis: Metastatic ovarian adenocarcinoma

Age: 51

Family anamnesis: colon and ovary cancer (mother) breast cancer (grandmother)

Tumor pathology:

BRCA1-185delAG, BRAF-D594Y, MAP2K1-P293S, NOTCH1-S2450N

- 2011: first diagnosis of ovarian adenocarcinoma; Wertheim operation and chemotherapy; lymph node removal
- 2015: metastasis in pericardial adipose tissue, excised
- 2016: hepatic metastases
- 2017: retroperitoneal and mesenteric lymph nodes have progressed; incipient peritoneal carcinosis with small accompanying ascites

Prior Therapy:

- 2012: Paclitaxel-carboplatin (6x)
- 2014: Caelyx-carboplatin (1x)
- 2016-2017 (9 months): Lymparza (Olaparib) 2x400 mg/day, oral

- 2017: Hycamtin inf. 5x2,5 mg (3x one seria/month)

  PIT vaccine treatment began on 21 April 2017.

Table 39 Patient XYZ peptide treatment schedule

Lot #		Vaccinations			
		1st cycle	2nd cycle	3rd cycle	4th cycle
POC01/1	N1727	21.04.2017	16.06.2017	30.08.2017	19.10.2017
POC01/2	N1728	28.04.2017	31.05.2017		
POC01/3	N1732		16.06.2017	02.08.2017	20.09.2017
POC01/4	N1736	15.05.2017	06.07.2017		

Patient' tumor MRI findings (Baseline April 15, 2016)

- Disease was confined primarily to liver and lymph nodes. The use of MRI limits detection of lung (pulmonary) metastasis
- May 2016 – Jan 2017: Olaparib treatment
- Dec/25/2016 (before PIT vaccine treatment) There was dramatic reduction in tumor burden with confirmation of response obtained at FU2
- Jan - Mar 2017 – TOPO protocol (topoisomerase)
- April/6/2017 FU3 demonstrated regrowth of existing lesions and appearance of new lesions leading to disease progression
- **April 21 2017 START PIT**
- Jul/21/17 (after the 2nd Cycle of PIT) FU4 demonstrated continued growth in lesions and general enlargement of pancreas and abnormal para pancreatic signal along with increased ascites
- Jul/26/17 – CBP+Gem+Avastin
- Sep/20/17 (after 3 Cycles of PIT) FU5 demonstrated reversal of lesion growth and improved pancreatic/parapancreatic signal. The findings suggest pseudo progression

- Nov 28/17 (after 4 Cycles of PIT) FU6 demonstrated best response with resolution of non target lesions

MRI data for patient XYZ is shown in Table 40 and Figure 23.

Table 40. Summary Table of Lesions Responses

Lesion/ Time Point	Baseline (%Δ from BL)	FU1 (%Δ from BL)	FU2 (%Δ from BL)	FU3 (%Δ from BL)	FU4 (%Δ from BL)	FU5 (%Δ from BL)	FU6 (%Δ from BL)	Best Response Cycle	PD Time Point
TL1	NA	-56.1	-44.4	-44.8	+109.3	-47.8	-67.3	**FU6**	**FU4**
TL2	NA	-100.0	-100.0	-47.1	-13.1	-100.0	-100.0	**FU1**	**FU3**
TL3	NA	-59.4	-62.3	-62.0	-30.9	-66.7	-75.9	**FU6**	**FU4**
TL4	NA	-65.8	-100.0	-100.0	-100.0	-100.0	-100.0	**FU2**	**NA**
SUM	NA	-66.3	-76.0	-68.9	-23.5	-78.2	-85.2	**FU6**	**FU4**

Example 22 Design of Personalised Immunotherapy Composition for Treatment of Breast Cancer

The HLA class I and class II genotype of metastatic breast cancer patient ABC was determined from a saliva sample. To make a personalized pharmaceutical composition for patient ABC twelve peptides were selected, each of which met the following two criteria: (i) derived from an antigen that is expressed in breast cancers, as reported in peer reviewed scientific publications; and (ii) comprises a fragment that is a T cell epitope capable of binding to at least three HLA class I of patient ABC (Table 41). In addition, each peptide is optimized to bind the maximum number of HLA class II of the patient. The twelve peptides target twelve breast cancer antigens. The probability that patient ABC will express one or more of the 12 antigens is shown in Figure 24.

Table 41. 12 peptides for ABC breast cancer patient

BRC09 vaccine peptides	Target Antigen	Antigen Expression	20mer peptide	MAXHLA Class I	MAXHLA Class II
PBRC01_cP1	FSIP1	49%	ISDTKDYFMSKTLGIGRLKR	3	6
PBRC01_cP2	SPAG9	88%	FDRNTESLFEELSSAGSGLI	3	2
PBRC01_cP3	AKAP4	85%	SQKMDMSNIVLMLIQKLLNE	3	6

PBRC01_cP4	BORIS	71%	SAVFHERYALIQHQKTHKNE	3	6
PBRC01_cP5	MAGE-A11	59%	DVKEVDPTSHSYVLVTSLNL	3	4
PBRC01_cP6	NY-SAR-35	49%	ENAHGQSLEEDSALEALLNF	3	2
PBRC01_cP7	HOM-TES-85	47%	MASFRKLTLSEKVPPNHPSR	3	5
PBRC01_cP8	NY-BR-1	47%	KRASQYSGQLKVLIAENTML	3	6
PBRC01_cP9	MAGE-A9	44%	VDPAQLEFMFQEALKLKVAE	3	8
PBRC01_cP10	SCP-1	38%	EYEREETRQVYMDLNNNIEK	3	3
PBRC01_cP11	MAGE-A1	37%	PEIFGKASESLQLVFGIDVK	3	3
PBRC01_cP12	MAGE-C2	21%	DSESSFTYTLDEKVAELVEF	4	2

Predicted efficacy: AGP95=4; 95% likelihood that the PIT Vaccine induces CTL responses against 4 CTAs expressed in the breast cancer cells of BRC09. Additional efficacy parameters: AGP50 = 6.3, mAGP = 100%, AP = 12.

Detected efficacy after the 1st vaccination with all 12 peptides: 83% reduction of tumor metabolic activity (PET CT data).

For treatment of patient ABC the 12 peptides were formulated as 4 x 3 peptide (PBR01/1, PBR01/2, PBR01/3, PBR01/4). One treatment cycle is defined as administration of all 12 different peptide vaccines within 30 days.

Patient history

Diagnosis: bilateral metastatic breast carcinoma: Right breast is ER positive, PR negative, Her2 negative; Left Breast is ER, PR and Her2 negative.

First diagnosis: 2013 (4 years before PIT vaccine treatment)

2016: extensive metastatic disease with nodal involvement both above and below the diaphragm. Multiple liver and pulmonar metastases.

2016-2017 treatment: Etrozole, Ibrance (Palbociclib) and Zoladex

Results

Mar 7, 2017: Prior PIT Vaccine treatment

Hepatic multi-metastatic disease with truly extrinsic compression of the origin of the choledochal duct and massive dilatation of the entire intrahepatic biliary tract. Celiac, hepatic hilar and retroperitoneal adenopathy

May 26 2017: After 1 cycle of PIT

Detected efficacy: 83% reduction of tumor metabolic activity (PET CT) liver, lung lymphnodes and other metastases.

Detected safety: Skin reactions

Local inflammation at the site of the injections within 48 hours following vaccine administrations

Follow up:

BRC-09 was treated with 5 cycles of PIT vaccine. She was feeling very well and she refused a PET CT examination in Sept 2017. In November she had symptoms, PET CT scan showed progressive disease, but she refused all treatments. In addition, her oncologist found out that she did not take Palbocyclib since spring/summer. Patient ABC passed away in Jan 2018.

The combination of pablocyclib and the personalised vaccine was likely to have been responsible for the remarkable early response observed following administration of the vaccine. Palbocyclib has been shown to improve the activity of immunotherapies by increases CTA presentation by HLAs and decreasing the proliferation of Tregs: (Goel et al. Nature. 2017:471-475). The PIT vaccine may be used as add-on to the state-of-art therapy to obtain maximal efficacy.

Example 23 – Personalised Immunotherapy Composition for treatment of patient with late stage metastatic breast cancerPatient BRC05 was diagnosed with inflammatory breast cancer on the right with extensive lymphangiosis carcinomatose.Inflammatory breast cancer (IBC) is a rare, but aggressive form of locally advanced breast cancer. It's called inflammatory breast cancer because its main symptoms are swelling and redness (the breast often looks inflamed). Most inflammatory breast cancers are invasive ductal carcinomas (begin in the milk ducts). This

type of breast cancer is associated with the expression of oncoproteins of high risk Human Papilloma Virus[1]. Indeed, HPV16 DNA was diagnosed in the tumor of this patient.

Patient's stage in 2011 (6 years prior to PIT vaccine treatment):

T4: Tumor of any size with direct extension to the chest wall and/or to the skin (ulceration or skin nodules)

pN3a: Metastases in ≥ 10 axillary lymph nodes (at least 1 tumor deposit > 2.0 mm); or metastases to the infraclavicular (level III axillary lymph) nodes.

14 vaccine peptides were designed and prepared for patient BRC05 (Table 42). Peptides PBRC05-P01-P10 were made for this patient based on population expression data. The last 3 peptides in the Table 42 (SSX-2, MORC, MAGE-B1) were designed from antigens that expression was measured directly in the tumor of the patient.

Table 42 – Vaccine peptides for patient BRC05

BRC05 vaccine peptides	Target Antigen	Antigen Expression	20mer peptide	MAXHLA Class I	MAXHLA Class II
PBRC05_P1	SPAG9	88%	XXXXXXXXXXXXXXXXXXXX	3	4
PBRC05_P2	AKAP4	85%	XXXXXXXXXXXXXXXXXXXX	3	4
PBRC05_P3	MAGE-A11	59%	XXXXXXXXXXXXXXXXXXXX	3	3
PBRC05_P4	NY-SAR-35	49%	XXXXXXXXXXXXXXXXXXXX	3	3
PBRC05_P5	FSIP1	49%	XXXXXXXXXXXXXXXXXXXX	3	3
PBRC05_P6	NY-BR-1	47%	XXXXXXXXXXXXXXXXXXXX	3	4
PBRC05_P7	MAGE-A9	44%	XXXXXXXXXXXXXXXXXXXX	3	3
PBRC05_P8	SCP-1	38%	XXXXXXXXXXXXXXXXXXXX	3	6
PBRC05_P9	MAGE-A1	37%	XXXXXXXXXXXXXXXXXXXX	3	3
PBRC05_P10	MAGE-C2	21%	XXXXXXXXXXXXXXXXXXXX	3	3
PBRC05_P11	MAGE-A12	13%	XXXXXXXXXXXXXXXXXXXX	3	4
PBRC05_P12	SSX-2	6%	XXXXXXXXXXXXXXXXXXXX	3	1
PBRC05_P13	MORC	ND	XXXXXXXXXXXXXXXXXXXX	3	4
PBRC05_P14	MAGE-B1	ND	XXXXXXXXXXXXXXXXXXXX	3	3

Note: Bold red means CD8 PEPI, Underline means best binding CD4 allele.

T cell responses were measured cells in peripheral mononuclear cells 2 weeks after the 1st vaccination with the mix of peptides PBRC05_P1, PBRC05_P2, PBRC05_P3, PBRC05_P4, PBRC05_P5, PBRC05_P6, PBRC05_P7.

Table 43 - Antigen specific T cell responses: Number of spots / 300,000 PBMC

Antigen	Stimulant	Exp1	Exp2	Average
SPAG9	PBRC05_P1	2	1	1.5
AKAP4	PBRC05_P2	11	4	7.5
MAGE-A11	PBRC05_P3	26	32	29
NY-SAR-35	PBRC05_P4	472	497	484.5
FSIP1	PBRC05_P5	317	321	319
NY-BR-1	PBRC05_P6	8	12	10
MAGE-A9	PBRC05_P7	23	27	25
None	Negative Control (DMSO)	0	3	1.5

The results show that a single immunization with 7 peptides induced potent T cell responses against 3 out of the 7 peptides demonstrating potent MAGE-A11, NY-SAR-35, FSIP1 and MAGE-A9 specific T cell responses. There were weak responses against AKAP4 and NY-BR-1 and no response against SPAG9.

References

[1] Bagarazzi et al. Immunotherapy against HPV16/18 generates potent TH1 and cytotoxic cellular immune responses. Science Translational Medicine. 2012; 4(155):155ra138.

[2] Gudmundsdotter et al. Amplified antigen-specific immune responses in HIV-1 infected individuals in a double blind DNA immunization and therapy interruption trial. Vaccine. 2011; 29(33):5558-66.

[3] Bioley et al. HLA class I - associated immunodominance affects CTL responsiveness to an ESO recombinant protein tumor antigen vaccine. Clin Cancer Res. 2009; 15(1):299-306.

[4] Valmori et al. Vaccination with NY-ESO-1 protein and CpG in Montanide induces integrated antibody/Th1 responses and CD8 T cells through cross-priming. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(21):8947-52.

[5] Yuan et al. Integrated NY-ESO-1 antibody and CD8+ T-cell responses correlate with clinical benefit in advanced melanoma patients treated with ipilimumab.Proc Natl Acad Sci U S A. 2011;108(40):16723-16728.

[6] Kakimi et al. A phase I study of vaccination with NY-ESO-1f peptide mixed with Picibanil OK-432 and Montanide ISA-51 in patients with cancers expressing the NY-ESO-1 antigen.Int J Cancer. 2011;129(12):2836-46.
[7] Wada et al. Vaccination with NY-ESO-1 overlapping peptides mixed with Picibanil OK-432 and montanide ISA-51 in patients with cancers expressing the NY-ESO-1 antigen. J Immunother. 2014;37(2):84-92.
[8] Welters et al. Induction of tumor-specific CD4+ and CD8+ T-cell immunity in cervical cancer patients by a human papillomavirus type 16 E6 and E7 long peptides vaccine. Clin. Cancer Res. 2008; 14(1):178-87.
[9] Kenter et al. Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia. N Engl J Med. 2009; 361(19):1838-47.
[10] Welters et al. Success or failure of vaccination for HPV16-positive vulvar lesions correlates with kinetics and phenotype of induced T-cell responses. PNAS. 2010; 107(26):11895-9.
[11] http://www.ncbi.nlm.nih.gov/projects/gv/mhc/main.fcgi?cmd=initThe MHC database, NCBI (Accessed Mar 7, 2016).
[12] Karkada et al. Therapeutic vaccines and cancer: focus on DPX-0907. Biologics. 2014;8:27-38.
[13] Butts et al. Randomized phase IIB trial of BLP25 liposome vaccine in stage IIIB and IV non-small-cell lung cancer. J Clin Oncol. 2005;23(27):6674-81.
[14] Yuan et al.Safety and immunogenicity of a human and mouse gp100 DNA vaccine in a phase I trial of patients with melanoma.Cancer Immun. 2009;9:5.
[15] Kovjazin et al. ImMucin: a novel therapeutic vaccine with promiscuous MHC binding for the treatment of MUC1-expressing tumors. Vaccine. 2011;29(29-30):4676-86.
[16] Cathcart et al. Amultivalent bcr-abl fusion peptide vaccination trial in patients with chronic myeloid leukemia.Blood. 2004;103:1037-1042.
[17] Chapuis et al. Transferred WT1-reactive CD8+ T cells can mediate antileukemic activity and persist in post-transplant patients. Sci Transl Med. 2013;5(174):174ra27.
[18] Keilholz et al. A clinical and immunologic phase 2 trial of Wilms tumor gene product 1 (WT1) peptide vaccination in patients with AML and MDS. Blood; 2009; 113(26):6541-8.
[19] Walter et al. Multipeptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. Nat Med. 2012;18(8):1254-61.
[20] Phuphanich et al. Phase I trial of a multi-epitope-pulsed dendritic cell vaccine for patients with newly diagnosed glioblastoma. Cancer Immunol Immunother. 2013;62(1):125-35.
[21] Kantoff et al. Overall survival analysis of a phase II randomized controlled trial of a Poxviral-based PSA-targeted immunotherapy in metastatic castration-resistant prostate cancer. J Clin Oncol. 2010;28(7):1099-105.
[22] Tagawa et al. Phase I study of intranodal delivery of a plasmid DNA vaccine for patients with Stage IV melanoma. Cancer. 2003;98(1):144-54.
[23] Slingluff et al. Randomized multicenter trial of the effects of melanoma-associated helper peptides and cyclophosphamide on the immunogenicity of a multipeptide melanoma vaccine.J Clin Oncol. 2011;29(21):2924-32.

[24] Kaida et al. Phase 1 trial of Wilms tumor 1 (WT1) peptide vaccine and gemcitabine combination therapy in patients with advanced pancreatic or biliary tract cancer. J Immunother. 2011;34(1):92-9.

[25]Fenoglio et al. A multi-peptide, dual-adjuvant telomerase vaccine (GX301) is highly immunogenic in patients with prostate and renal cancer. Cancer Immunol Immunother; 2013; 62:1041–1052.

[26]Krug et al. WT1 peptide vaccinations induce CD4 and CD8 T cell immune responses in patients with mesothelioma and non-small cell lung cancer. Cancer Immunol Immunother; 2010; 59(10):1467-79.

[27]Slingluff et al. Clinical and immunologic results of a randomized phase II trial of vaccination using four melanoma peptides either administered in granulocyte-macrophage colony-stimulating factor in adjuvant or pulsed on dendritic cells. J Clin Oncol; 2003; 21(21):4016-26.

[28]Hodi et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med; 2010;363(8):711-23.

[29] Carmon et al. Phase I/II study exploring ImMucin, a pan-major histocompatibility complex, anti-MUC1 signal peptide vaccine, in multiple myeloma patients. Br J Hematol. 2014; 169(1):44-56.

[30]http://www.merckgroup.com/en/media/extNewsDetail.html?newsId=EB4A46A2AC4A52E7C1257AD9001F3186&newsType=1(Accessed Mar 28, 2016)

[31] Trimble et al. Safety, efficacy, and immunogenicity of VGX-3100, a therapeutic synthetic DNA vaccine targeting human papillomavirus 16 and 18 E6 and E7 proteins for cervical intraepithelial neoplasia 2/3: a randomised, double-blind, placebo-controlled phase 2b trial. Lancet. 2015;386(10008):2078-88.

[32]Cusi et al. Phase I trial of thymidylate synthase poly epitope peptide (TSPP) vaccine in advanced cancer patients. Cancer Immunol Immunother; 2015; 64:1159–1173.

[33]Asahara et al. Phase I/II clinical trial using HLA-A24-restricted peptide vaccine derived from KIF20A for patients with advanced pancreatic cancer. J Transl Med; 2013;11:291.

[34]Yoshitake et al. Phase II clinical trial of multiple peptide vaccination for advanced head and neck cancer patients revealed induction of immune responses and improved OS. Clin Cancer Res; 2014;21(2):312-21.

[35]Okuno et al. Clinical Trial of a 7-Peptide Cocktail Vaccine with Oral Chemotherapy for Patients with Metastatic Colorectal Cancer. Anticancer Res; 2014; 34: 3045-305.

[36]Rapoport et al. Combination Immunotherapy after ASCT for Multiple Myeloma Using MAGE-A3/Poly-ICLC Immunizations Followed by Adoptive Transfer of Vaccine-Primed and Costimulated Autologous T Cells. Clin Cancer Res; 2014; 20(5): 1355–1365.

[37]Greenfield et al. A phase I dose-escalation clinical trial of a peptidebased human papillomavirus therapeutic vaccine with Candida skin test reagent as a novel vaccine adjuvant for treating women with biopsy-proven cervical intraepithelial neoplasia 2/3. Oncoimmunol; 2015; 4:10, e1031439.

[38] Snyder et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med. 2014; 371(23):2189-99.

[39] Van Allen et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science; 2015; 350:6257.
[40] Li et al. Thrombocytopenia caused by the development of antibodies to thrombopoietin. Blood; 2001; 98:3241-3248
[41] Takedatsu et al. Determination of Thrombopoietin-Derived Peptides Recognized by Both Cellular and Humoral Immunities in Healthy Donors and Patients with Thrombocytopenia. 2005; 23(7): 975-982
[42] Eisenhauer et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer; 2009; 45(2):228-47.
[43] Therasse et al. New guidelines to evaluate the response to treatment in solid tumors: European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst; 2000; 92:205-216.
[44] Tsuchida & Therasse. Response evaluation criteria in solid tumors (RECIST): New guidelines. Med Pediatr Oncol. 2001; 37:1-3.
[45] Durie et al. International uniform response criteria for multiple myeloma. Leukemia; 2006;20:1467–1473.

CLAIMS

1. A polypeptide that comprises a fragment of up to 50 consecutive amino acids of

(a) a colorectal cancer-associated antigen selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, LEMD1, MAGE-A8, MAGE-A6 and MAGE-A3, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 21 to 40 and 234 to 250;

(b) an ovarian cancer-associated antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN, and AKAP-3 wherein the fragment comprises the amino acid sequence of any one of SEQ ID NOs: 272 to 301; and/or

(c) a breast cancer associated antigen selected from PIWIL-2, AKAP-4, EpCAM, BORIS, HIWI, SPAG9, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, PRAME, NY-SAR-35, MAGE-A9, NY-BR-1, SURVIVIN, MAGE-A11, HOM-TES-85 and NY-ESO-1 wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194;

optionally wherein the fragment is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the breast, ovarian or colorectal cancer-associated antigen.

2. The polypeptide of claim 1, wherein the polypeptide

a. is a fragment of a colorectal cancer-associated antigen selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, MAGE-A6, MAGE-A3 and LEMD1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 21 to 40 and 234 to 250; or

b. comprises or consists of two or more fragments of one or more colorectal cancer associated antigens selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, MAGE-A6, MAGE-A3 and LEMD1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID

NOs: 21 to 40 and 234 to 250, optionally wherein the fragments overlap or are arranged end to end in the polypeptide; or

c. is a fragment of a ovarian cancer-associated antigen selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN and AKAP-3, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 272 to 301; or

d. comprises or consists of two or more fragments of one or more ovarian cancer associated antigens selected from PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN and AKAP-3, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 272 to 301, optionally wherein the fragments overlap or are arranged end to end in the polypeptide; or

e. is a fragment of a breast cancer associated antigen selected from SPAG9, AKAP-4, BORIS, NY-SAR-35, NY-BR-1, SURVIVIN, MAGE-A11, PRAME, MAGE-A9, HOM-TES-85, PIWIL-2, EpCAM, HIWI, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, wherein the fragment comprises the amino acid sequence from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194; or

f. comprises or consists of two or more fragments of one or more breast cancer associated antigens selected from SPAG9, AKAP-4, BORIS, NY-SAR-35, NY-BR-1, SURVIVIN, MAGE-A11, PRAME, MAGE-A9, HOM-TES-8, PIWIL-2, EpCAM, HIWI, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 1 to 20, 24 and 172 to 194; optionally wherein the fragments overlap or are arranged end to end in the polypeptide and.

3. The polypeptide according to claim 1 or claim 2, wherein the polypeptide comprises or consists of fragments of at least two different cancer-associated antigens, wherein the cancer-associated antigens are selected from

(a) TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, MAGE-A8, MAGE-A6, MAGE-A3 and LEMD1;

(b) PIWIL-4, WT1, EpCAM, BORIS, AKAP-4, OY-TES-1, SP17, PIWIL-2, PIWIL-3, SPAG9, PRAME, HIWI, SURVIVIN and AKAP-3; and/or

(c) SPAG9, AKAP-4, BORIS, NY-SAR-35, NY-BR-1, SURVIVIN, MAGE-A11, PRAME, MAGE-A9, HOM-TES-8, PIWIL-2, EpCAM, HIWI, PLU-1, TSGA10, ODF-4, SP17, RHOXF-2;

wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 21 to 40 and 234 to 250; SEQ ID NOs: 272 to 301; and/or SEQ ID NOs: 1 to 20, 24 and 172 to 194.

4. The polypeptide according to any one of claims 1 to 3, comprising or consisting of one or more amino acid sequences selected from SEQ ID NOs: 41-80, 251 to 271, 302 to 331 and 196 to 233.

5. The polypeptide according to any one of claims 1 to 4 comprising or consisting of the amino acid sequence of any one of SEQ ID NOs: 81 to 142, 332 to 346 and 435-449.

6. A panel of two or more polypeptides according to any one of claims 1 to 5, wherein

(a) each polypeptide comprises a different amino acid sequence selected from SEQ ID NOs: 21 to 40 and 234 to 250; or

(b) each polypeptide comprises a different amino acid sequence selected from SEQ ID NOs: 272 to 301; or

(c) each peptide comprises a different amino acid sequence selected from SEQ ID NOs: 1 to 20, 24 and 172 to 194; or(c) each peptide comprises a different amino acid sequence selected from SEQ ID NOs: 1 to 40, 234 to 250, 272 to 301 and 172 to 194.

7. The panel of polypeptides according to claim 6 comprising six peptides having the amino acid sequences of SEQ ID NOs: 130, 121, 131, 124, 134, 126.

8. A pharmaceutical composition or kit having one or more polypeptides according to any one of claims 1 to 5, or a panel of polypeptides according to claim 6 or claim 7, or a polypeptide comprising at least two amino acid sequences selected SEQ ID NOs: 21 to 40 and 234 to 250; SEQ ID NOs: 272 to 301; and/or SEQ ID NOs: 1 to 20, 24 and 172 to 194 as an active ingredient.

9. A method of vaccination, providing immunotherapy or inducing a cytotoxic T cell response in a subject, the method comprising administering to the subject a pharmaceutical composition according to claim 8.

10. The method of claim 9 that is a method of treating cancer, optionally colorectal cancer, ovarian cancer or breast cancer.

11. A method of identifying a human subject who will likely have a cytotoxic T cell response to administration of a pharmaceutical composition according to claim 8, the method comprising

(i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition comprise a sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(ii) identifying the subject as likely to have a cytotoxic T cell response to administration of the pharmaceutical composition.

12. The method of claim 11 further comprising using population expression data for each antigen that

(a) is selected from TSP50, EpCAM, SPAG9, CAGE1, FBXO39, SURVIVIN, LEMD1, MAGE-A8, MAGE-A6, MAGE-A3, PIWIL-4, WT1, BORIS, AKAP-4, OY-TES-1,

SP17, PIWIL-2, PIWIL-3, PRAME, HIWI, PLU-1, TSGA10, ODF-4, RHOXF-2, NY-SAR-35, MAGE-A9, NY-BR-1, MAGE-A11, HOM-TES-85, NY-ESO-1 and AKAP-3; and

(b) comprises an amino acid sequence that is

i. a fragment of an active ingredient peptide of the pharmaceutical composition; and

ii. a T cell epitope capable of binding to at least three HLA class I molecules of the subject;

to determine the likelihood that the subject will have a cytotoxic T cell response that targets one or more polypeptide antigens that are expressed by cancer cells of the subject.

13. A method of identifying a subject who will likely have a clinical response to a method of treatment according to claim 10, the method comprising

(i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition comprise two or more different amino acid sequences each of which is

a. a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

b. a fragment of a cancer-associated antigen expressed by cancer cells of the subject, optionally wherein the cancer-associated antigen is present in a sample obtained from the subject; and

(ii) identifying the subject as likely to have a clinical response to the method of treatment.

14. A method of determining the likelihood that a specific human subject will have a clinical response to a method of treatment according to claim 10, wherein one or more of the following factors corresponds to a higher likelihood of a clinical response:

(a) presence in the active ingredient polypeptide(s) of a higher number of amino acid sequences and/or different amino acid sequences that are each a T cell epitope capable of binding to at least three HLA class I of the subject;

(b) a higher number of target polypeptide antigens, comprising at least one amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject;

optionally wherein the target polypeptide antigens are expressed in the subject, further optionally wherein the target polypeptides antigens are in one or more samples obtained from the subject;

(c) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens and/or optionally target polypeptide antigens that have been determined to comprise at least one amino acid sequence that is both

A. comprised in in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject; and/or

(d) a higher number of target polypeptide antigens that the subject is predicted to express, optionally a higher number of target polypeptide antigens that the subject expresses with a threshold probability, and/or optionally the target polypeptide antigens that have been determined to comprise at least one amino acid sequence that is both

A. comprised in in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject.

15. The method of claim 14, wherein the method comprises

(i) identifying which polypeptide antigens targeted by the active ingredient polypeptide(s) comprise an amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject;

(ii) using population expression data for each antigen identified in step (i) to determine the probability that the subject expresses one or more of the antigens

identified in step (i) that together comprise at least two different amino acid sequences of step (i); and

(iii) determining the likelihood that the subject will have a clinical response to administration of the pharmaceutical composition, kit or panel of polypeptides, wherein a higher probability determined in step (ii) corresponds to a more likely clinical response.

16. The method of claim 15, wherein the at least two different amino acid sequences are comprised in the amino acid sequence of two different polypeptide antigens targeted by the active ingredient polypeptide(s).

17. The method of any one of claims 13 to 16 further comprising selecting or recommending administration of the pharmaceutical composition as a method of treatment for the subject, and optionally further treating the subject by administering the pharmaceutical composition.

18. A method of treatment according to claim 10, wherein the subject has been identified as likely to have a clinical response or as having above a threshold minimum likelihood of having a clinical response to the treatment by a method according to any one of claims 13 to 16.

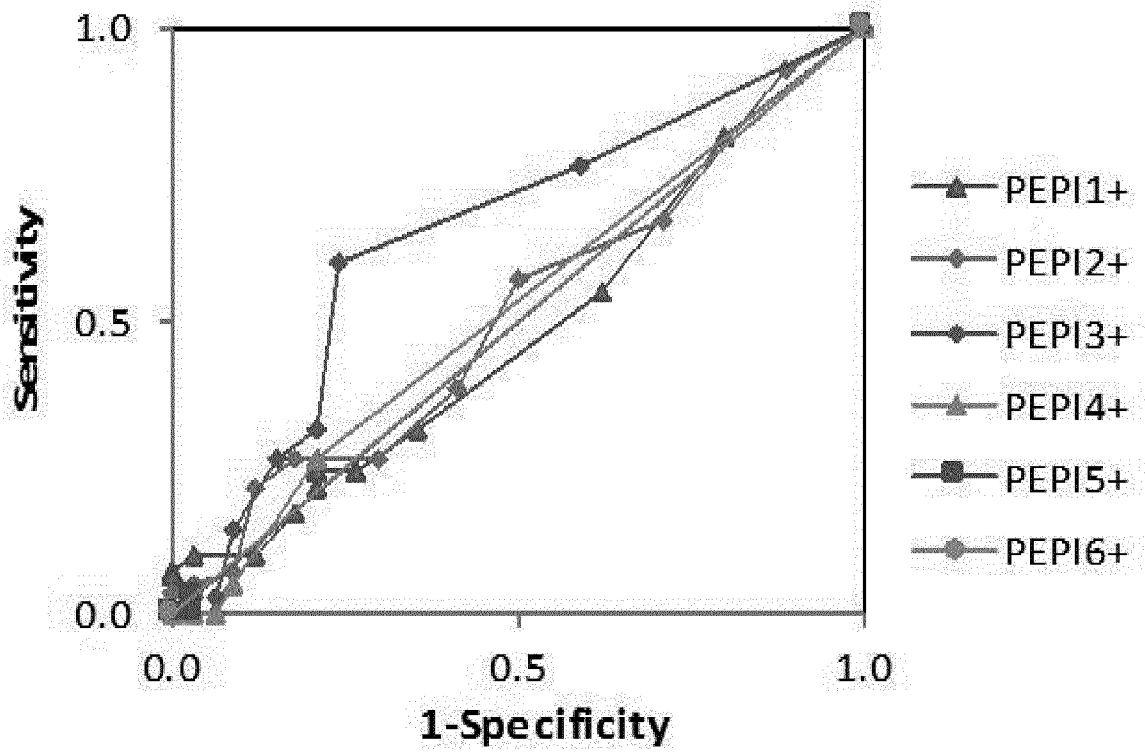
19. The method of any one of claims 9, 10, 17 and 18 wherein the treatment is administered in combination with chemotherapy, targeted therapy or a checkpoint inhibitor.

20. A method of identifying a human subject who will likely not have a clinical response to a method of treatment according to claim 10, the method comprising

(i) determining that the active ingredient peptide(s) of the pharmaceutical composition do not comprise two or more different amino acid sequences each of which is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(iii) identifying the subject as likely not to have a clinical response to the method of treatment.

1/25
Figure 1


1.0
0.5
0.0
Sensitivity
0.0
0.5
1.0
1-Specificity
PEPI1+
PEPI2+
PEPI3+
PEPI4+
PEPI5+
PEPI6+

2/25
Figure 2

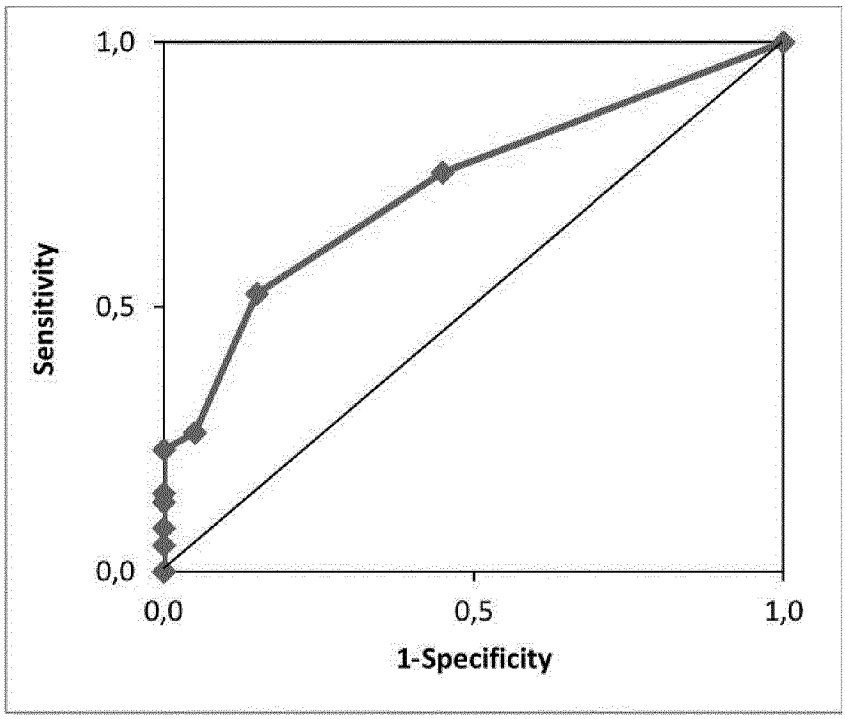


Fig. 3A

Patient ID	#epitope / HPV-16 E6 Pools															#epitope / HPV-16 E17 Pools									
	1-19	11-29	21-39	31-49	41-59	51-69	61-79	71-89	81-99	91-109	101-119	111-129	121-139	131-149	141-158	1-19	11-29	21-39	31-49	41-59	51-69	61-79	71-89	81-98
1	TN	TN	TN	FN	TN	FP	TN	TN	TN	TN	FN	TN	TN	TN	TN	FP	TN	TN	TN	FP	TN	TN	TN	TN
2	FN	FN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	FP	TN	TN	TN	TN
3	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
6	TN	TN	TN	FN	TN	FP	TN	TN	TN	TN	FN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	FP
7	TN	TN	TN	TN	FN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
8	TN	FN	TN	FN	TN	FN	TN	FP	TN	TN	FN	TN	TN	TN	FN	TN	TN	TN	TN	TP	TN	TN	TN	TN
9	TN	TN	TN	TN	FN	FN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
10	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	FP	TN
11	TN	FN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
13	TN	TN	TN	TN	TN	FN	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
16	TN	TN	TN	TN	TN	TP	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
18	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN
22	TN	FN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	FP	TN	TN	TN	TN
23	TN	TN	TN	FN	TN	TN	TN	TN	TN	FN	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
27	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
28	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN
29	FN	TN	FN	TN	TN	TN	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
30	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
100	TN	TN	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
102	TN	TN	TN	TN	TN	TP	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	TN	TN
103	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
105	TN	TN	TN	TN	TN	TP	FN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	FP	TN	TN	TN	TN
107	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN

Fig. 3B

Patient ID	#epitope / HPV-16 E6 Pools																#epitope / HPV-16 E17 Pools									
	1-19	11-29	21-39	31-49	41-59	51-69	61-79	71-89	81-99	91-109	101-119	111-129	121-139	131-149	141-158	1-19	11-29	21-39	31-49	41-59	51-69	61-79	71-89	81-98
1	TN	FP	FP	TP	FP	FP	FP	FP	FP	FP	TP	FP	TN	TN	FP	FP	FP	TN	TN	FP	FP	FP	FP	FP
2	TP	TP	FP	FP	FP	FP	FP	FP	FP	FP	TN	FP	TN	FP	FP	FP	FP	TN	TN	FP	TN	FP	FP	FP
3	FP	FP	TN	FP	FP	TP	TN	FP	FP	FP	FP	TP	TN	FP	TP	FP	TN	TN	TN	FP	TN	TN	FP	FP
6	FP	FP	FP	FN	FP	TP	TN	FP	FP	FP	FN	TN	TN	TN	TN	FP	FP	TN	FP	FP	TN	TN	FP	FP
7	FP	FP	FP	FP	FP	TP	TN	FP	FP	FP	FP	FP	TN	FP	FP	FP	TN	TN	TN	FP	TN	TN	FP	FP
8	FP	TP	TN	TP	FP	TP	FP	FP	FP	FP	FP	FP	FP	FP	TP	FP	FP	TN	TN	TP	TN	FP	FP	FP
9	FP	FP	TN	FP	TP	TP	TN	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	TN	TN	FP	TN	FP	FP	FP
10	FP	FP	TN	FP	FP	TP	FP	FP	FP	FP	FP	FP	TN	FP	FP	FP	TN	TN	TN	FP	TN	TN	FP	TN
11	FP	TP	FP	FP	FP	FP	TN	TN	FP	FP	FP	FP	TN	FP	FP	FP	FP	TN	TN	FP	TN	TN	FP	FP
13	FP	FP	FP	FP	FP	TP	FP	FP	TP	TP	FP	FP	TN	FP	FP	FP	FP	TN	TN	FP	FP	TN	FP	FP
16	FP	FP	FP	FP	TN	TP	TN	FP	FP	TP	FP	TN	TN	FP	FP	FP	FP	TN	FP	FP	TN	FP	FP	FP
18	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	TN	FP	FP	FP	FP	TN	TN	TP	FP	FP	FP	FP
22	FP	TP	FP	FP	FP	TP	FP	FP	FP	FP	FP	FP	TN	FP	FP	FP	FP	TN	TN	FP	TN	FP	FP	FP
23	TN	FP	FP	TP	FP	FP	FP	FP	FP	TP	FP	FP	TN	TP	FP	FP	FP	TN	FP	FP	FP	TN	FP	FP
27	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	TN	FP	FP	FP	FP	TN	TN	FP	TN	FP	FP	FP
28	FP	FP	TN	TN	FP	FP	FP	FP	FP	FP	TN	FP	TN	FP	FP	FP	TN	TN	TN	FP	FP	TN	FP	FP
29	TP	FP	TP	FP	FP	FP	FP	FP	FP	TP	FP	FP	FP	FP	FP	FP	TN	TN	FP	FP	FP	FP	FP	FP
30	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	TN	TN	TN	FP	FP	TN	TN	TN	FP	FP	TN	FP	FP
100	FP	FP	FP	FP	FP	TP	FP	FP	FP	FP	TN	TN	FP	FP	FP	FP	FP	FP	FP	FP	TN	FP	FP	TN
102	FP	FP	FP	FP	FP	TP	FP	FP	FP	FP	TN	FP	FP	FP	FP	FP	FP	TN	TN	FP	FN	TN	FP	TN
103	FP	TP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	TN	FP	FP	FP	FP	TN	TN	FP	TN	TN	FP	FP
105	FP	FP	FP	FP	FP	TP	TP	FP	FP	FP	TN	FP	TN	FP	FP	FP	TN	TN	TN	FP	TN	FP	FP	FP
107	FP	FP	FP	FP	FP	FP	FP	FP	FP	TP	TN	FP	FP	FP	FP	FP	FP	TN	FP	FP	TN	FP	TP	FP

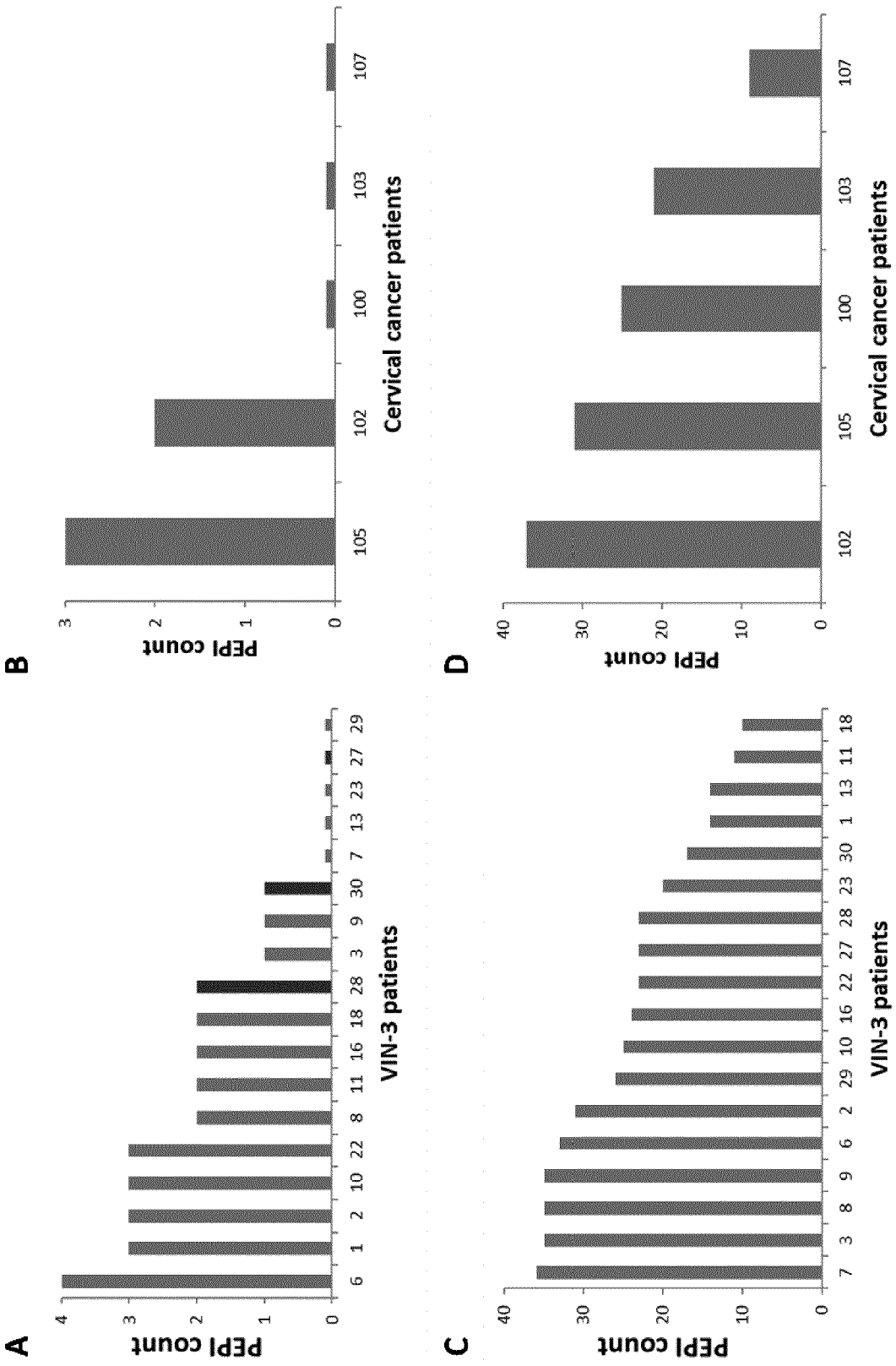
Fig. 4A

	#PEPI / HPV-16 E6				E7	
Patient ID	E6.1	E6.2	E6.3	E6.4	E7.1	E7.2
1	FP	TP	TP	FN	TN	FN
2	FP	TP	TP	TN	TN	TP
3	TP	TP	TP	TP	TN	TP
6	TP	TP	TP	FN	TP	TP
7	TP	TP	TP	TP	TN	TP
8	TP	TP	TP	FN	FP	TP
9	TP	TP	TP	FN	FP	TP
10	FP	TP	TP	FN	TN	TP
11	TP	TP	FN	FN	TN	TP
13	TP	TP	FN	FN	TN	FN
16	TP	TP	TP	FN	FN	TP
18	FP	TP	FN	FN	FN	TP
22	TP	TP	TP	FN	FN	FN
23	FP	TP	TP	FN	TN	TN
27	TP	TP	TP	FN	FN	FN
28	TP	TP	TP	FN	TN	FN
29	FP	TP	TP	FN	FP	TP
30	FP	FP	TN	TN	FN	FN
100	TP	TP	TP	FN	TN	TP
102	TP	TP	TP	FN	FN	TP
103	TP	TP	TP	FN	TN	TN
105	TP	TP	TP	TP	TN	TN
107	TP	TP	FP	FN	TN	FN

Fig. 4B

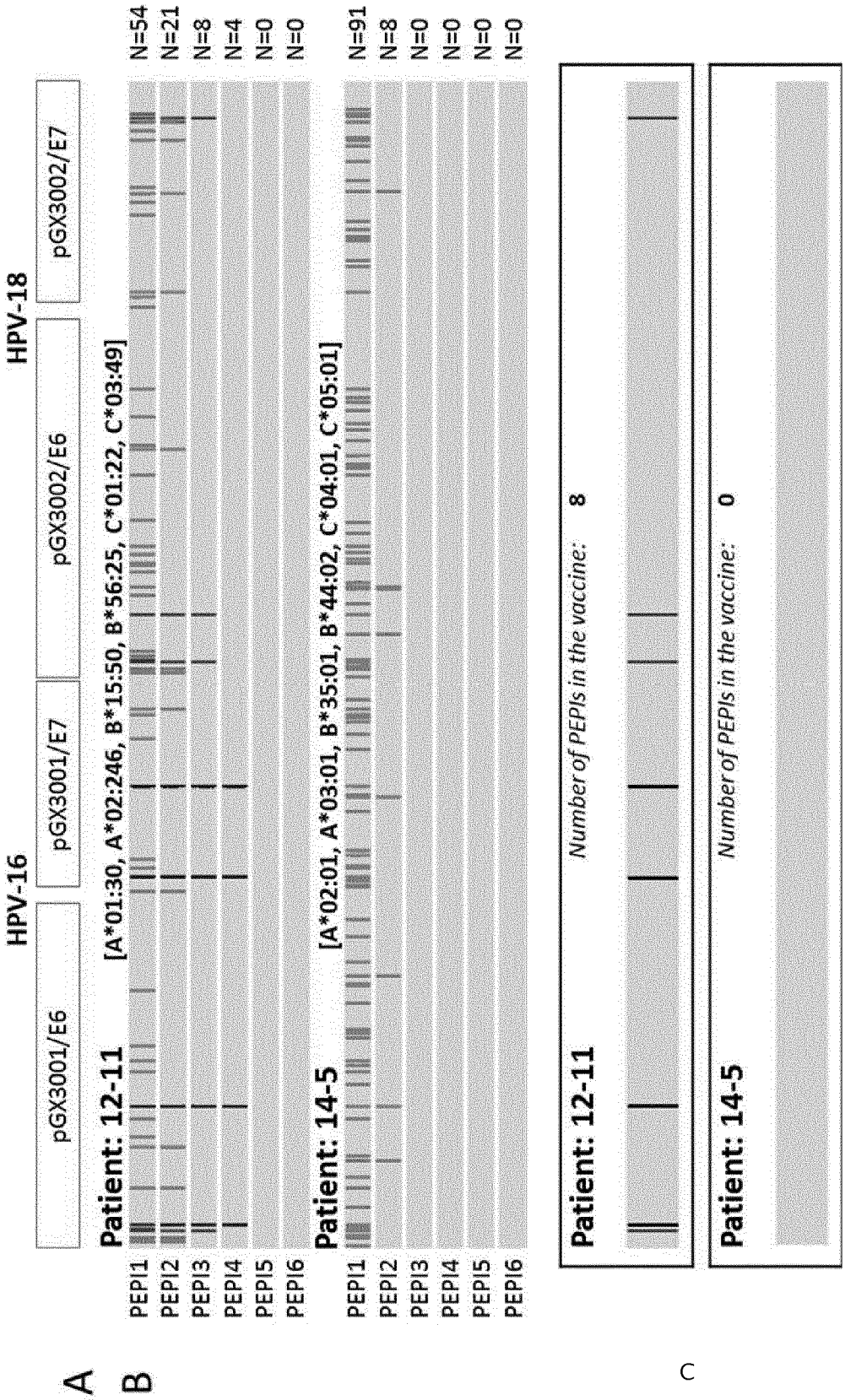
	#epitope / HPV-16 E6&E7 Pools					
Patient ID	E6.1	E6.2	E6.3	E6.4	E7.1	E7.2
1	FP	TP	TP	FN	FP	TP
2	FP	TP	TP	TN	TN	TP
3	TP	TP	TP	FN	FP	TP
6	TP	TP	TP	FN	TP	TP
7	TP	TP	TP	FN	FP	TP
8	TP	TP	TP	FN	FP	TP
9	TP	TP	TP	FN	FP	TP
10	FP	TP	TP	FN	FP	TP
11	TP	TP	TP	FN	FP	TP
13	TP	TP	TP	FN	TN	TP
16	TP	TP	TP	FN	TN	TP
18	FP	TP	TP	FN	FN	TP
22	TP	TP	TP	FN	TP	TP
23	FP	TP	TP	FN	FP	FP
27	TP	TP	TP	FN	TP	TP
28	TP	TP	TP	FN	FP	TP
29	FP	TP	TP	FN	FP	TP
30	FP	FP	FP	TN	FN	TP
100	TP	TP	TP	FN	FP	TP
102	TP	TP	TP	FN	TP	TP
103	TP	TP	TP	FN	TN	FP
105	TP	TP	TP	FN	TN	FP
107	TP	TP	TP	FN	FP	TP

6/25
Figure 5


A
PEPI count
VIN-3 patients
B
PEPI count
Cervical cancer patients
C
PEPI count
VIN-3 patients
D
PEPI count
Cervical cancer patients

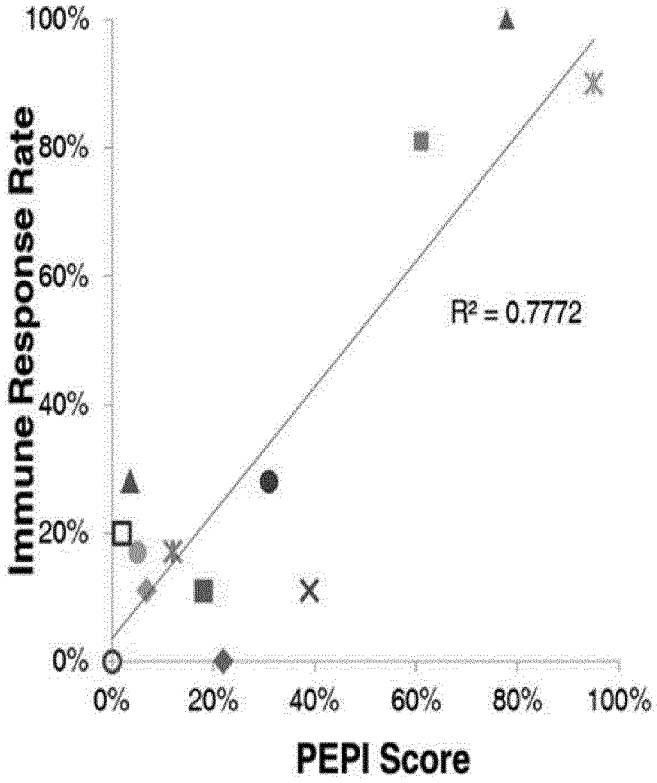
7/25

Figure 6


A
B
C
HPV-16
HPV-18
pGX3001/E6
pGX3001/E7
pGX3002/E6
pGX3002/E7
Patient: 12-11
[A*01:30, A*02:246, B*15:50, B*56:25, C*01:22, C*03:49]
Patient: 14-5
[A*02:01, A*03:01, B*35:01, B*44:02, C*04:01, C*05:01]
PEPI1
PEPI2
PEPI3
PEPI4
PEPI5
PEPI6
N=54
N=21
N=8
N=4
N=0
N=0
N=91
N=8
N=0
N=0
N=0
N=0
Patient: 12-11
Number of PEPIs in the vaccine: 8
Patient: 14-5
Number of PEPIs in the vaccine: 0

8/25

Figure 7



- MMNLMQPKTQQTYTYD (JUP)
- GRGSTTTNYLLDRDDYRNTSD (ADA17)
- LKKGAADGGKLDGNAKLNRSLK (BAP31)
- FPPKDDHTLKFLYDDNQRPYPP (TOP2A)
- QRPPFSQLHRFLADALNT (DDR1)
- RYRKPDYTLDDGHGLLRFKST (Abl-2)
- ALDQCKTSCALMQQHYDQTSCFSSP (ITGB8)
- STAPPAHGVTSAPDTRPAPGSTAPP (Muc-1)
- YLEPGPVTA (gp-100)
- MTPGTQSPFFLLLLLTVLTVV (Muc-1)
- SSKALQRPV (Bcr-abl)
- RMFPNAPYL (WT1)
- RMFPNAPYL (WT1, HLA-A*0201)

Fig. 8

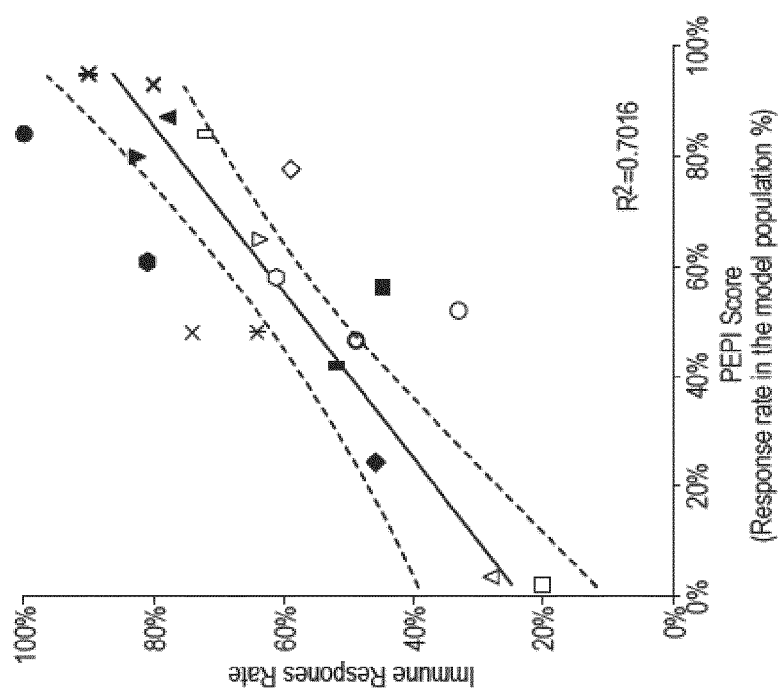
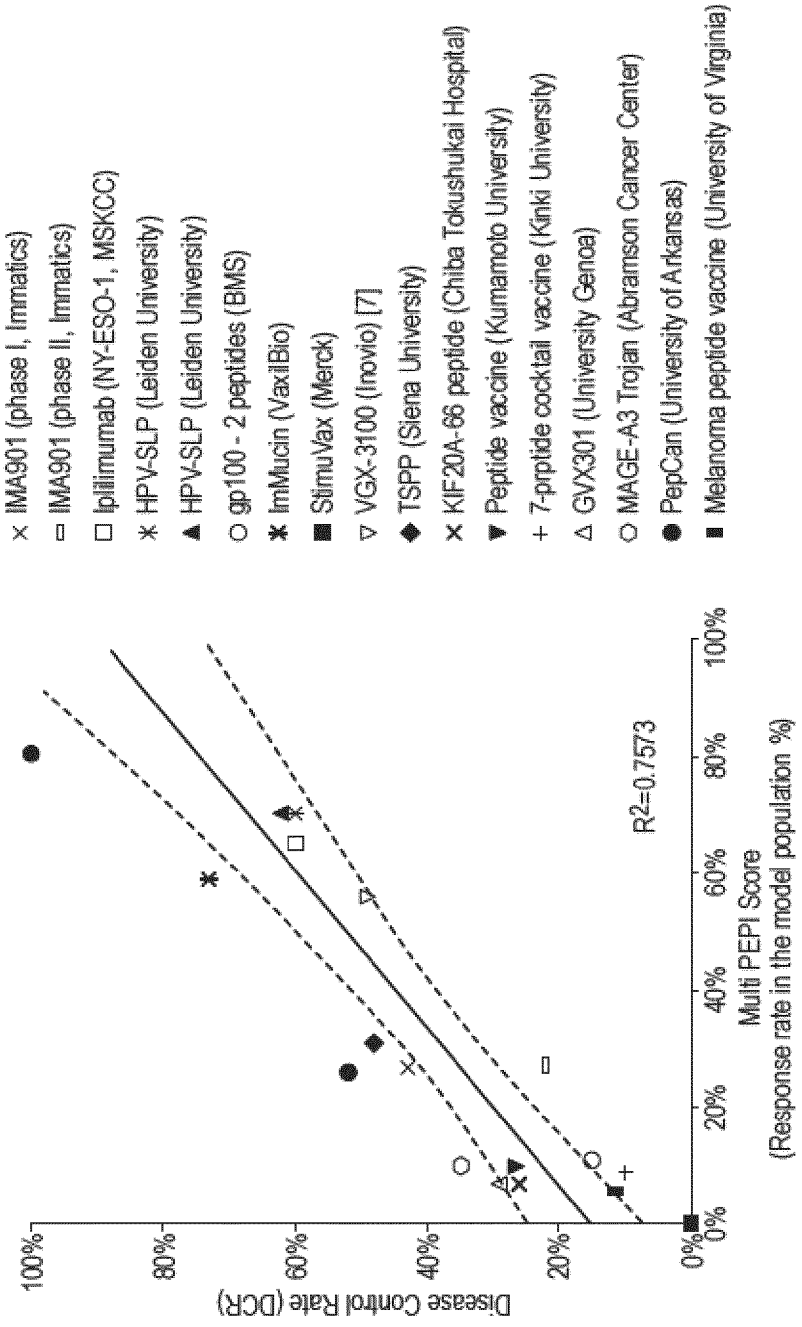

□ StimuVax (Merck)
△ gp100 vaccine (Sloan-Kettering Institute)
× IMA901 (phase I, Immatics)
* IMA901 (phase II, Immatics)
○ ICT107 (Cedars-Sinai Medical Center)
■ ProstVac (Bavarian Nordic)
◆ Synchrotope TA2M (Mannkind Co.)
O MELITAC 12.1 (Fred Hutchinson CRC)
◇ WT1 vaccine (NCRC Hospital, Tokyo)
□ Ipilimumab (NY-ESO-1, MSKCC)
▲ VGX-3100 (Inovio)
× HIVIS-1 (Karolinska)
✱ ImMucin (VaxilBio)
● NY-ESO-1 OLP (Osaka University)
▽ GVX301 (University Genoa)
▼ WT1 vaccine (MSKCC)
● WT1 vaccine (CBF)
○ DPX-0907 (ImmunoVaccine Inc.)
■ Melanoma peptide vaccine (University of Virginia)
Immune Responses Rate
100%
80%
60%
40%
20%
0%
0%
20%
40%
60%
80%
100%
PEPI Score
(Response rate in the model population %)
R²=0.7016

Fig. 9


× IMA901 (phase I, Immatics)
▫ IMA901 (phase II, Immatics)
□ Ipilimumab (NY-ESO-1, MSKCC)
✳ HPV-SLP (Leiden University)
▲ HPV-SLP (Leiden University)
○ gp100 - 2 peptides (BMS)
✱ ImMucin (VaxilBio)
■ StimuVax (Merck)
▽ VGX-3100 (Inovio) [7]
◆ TSPP (Siena University)
✕ KIF20A-66 peptide (Chiba Tokushukai Hospital)
▼ Peptide vaccine (Kumamoto University)
+ 7-prptide cocktail vaccine (Kinki University)
△ GVX301 (University Genoa)
○ MAGE-A3 Trojan (Abramson Cancer Center)
● PepCan (University of Arkansas)
▮ Melanoma peptide vaccine (University of Virginia)
Disease Control Rate (DCR)
100%
80%
60%
40%
20%
0%
0%
20%
40%
60%
80%
100%
R2=0.7573
Multi PEPI Score
(Response rate in the model population %)

11/25
Figure 10

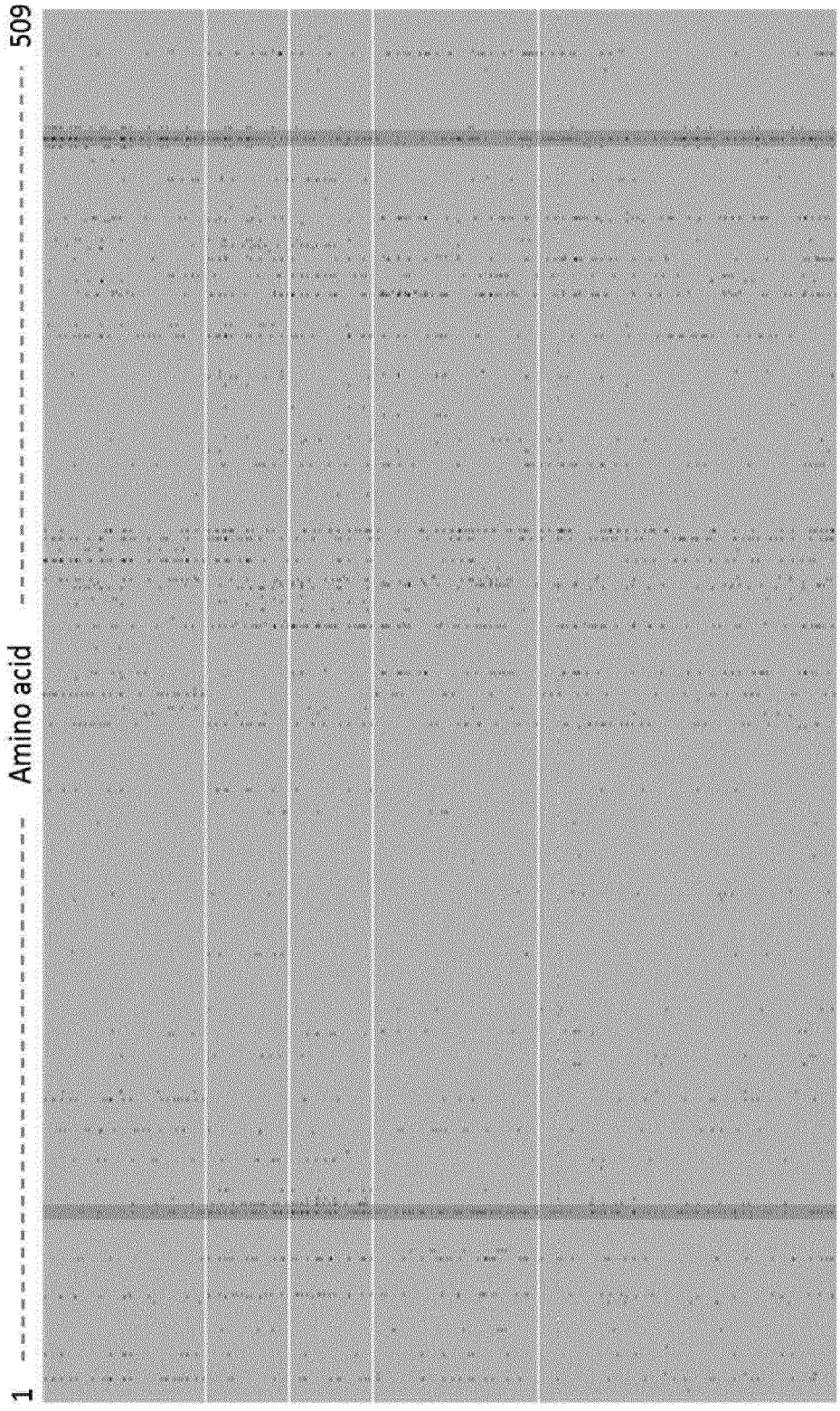
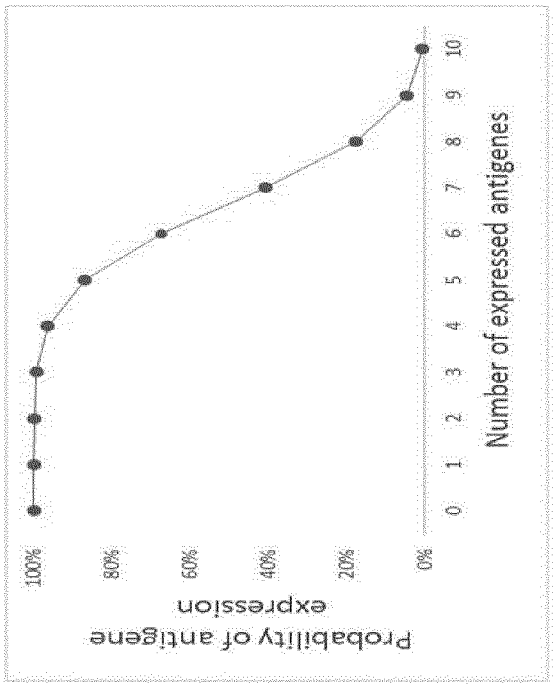


Figure 11

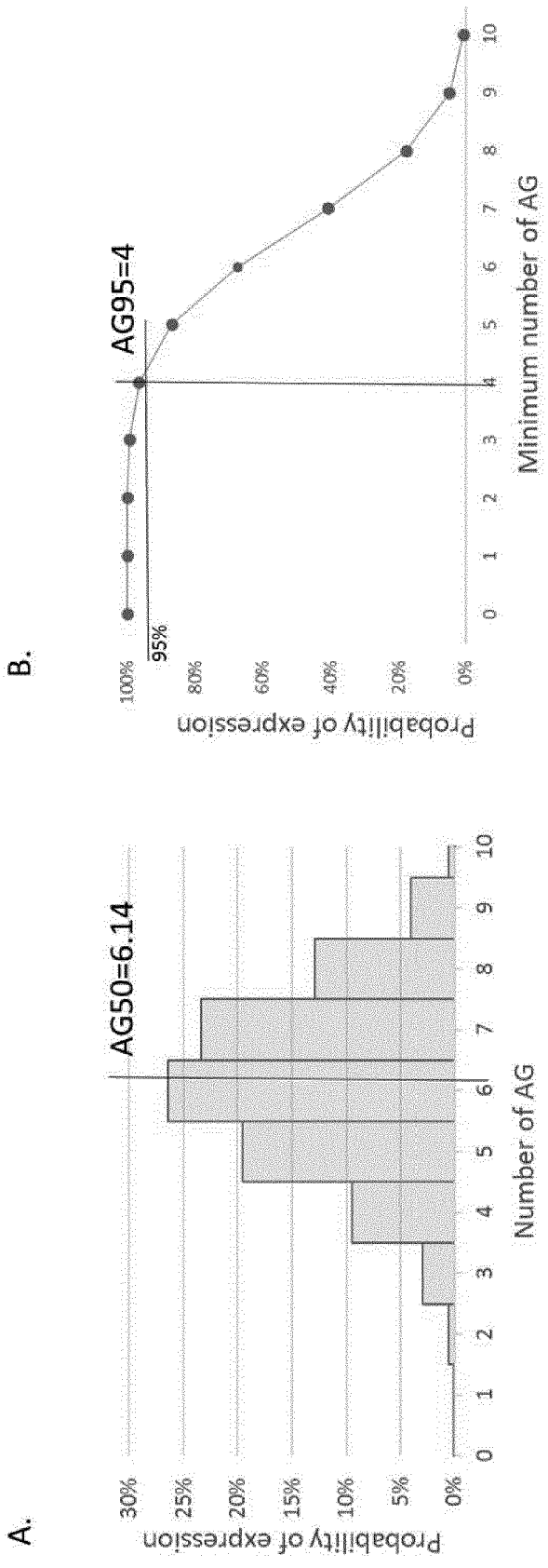


Antigens	Expression rate
Cancer testis antigens	Determined from 1053 tumor samples
AKAP-4	85%
BORIS	71%
SPAG9	88%
PRAME	55%
NY-SAR-35	48%
MAGE-A9	44%
NY-BR-1	47%
SURVIVIN	71%
MAGE-A11	59%
HOM-TES-85	47%

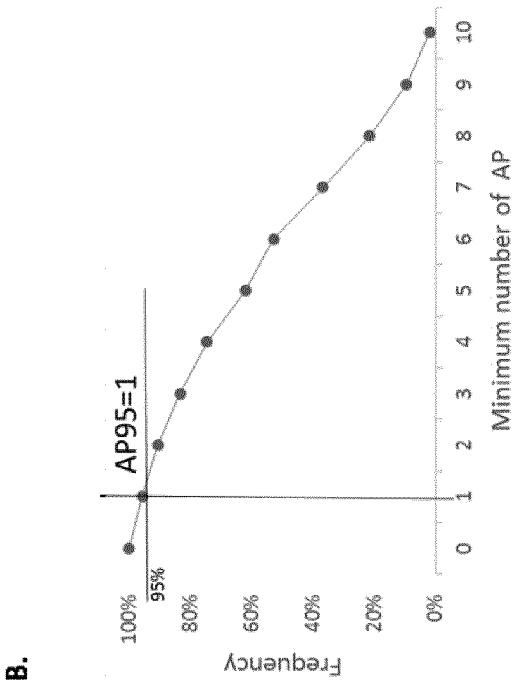
Conclusion

1. 99% of breast tumors express ≥3 antigens
2. No biopsy is needed

13/25
Figure 12



14/25
Figure 13



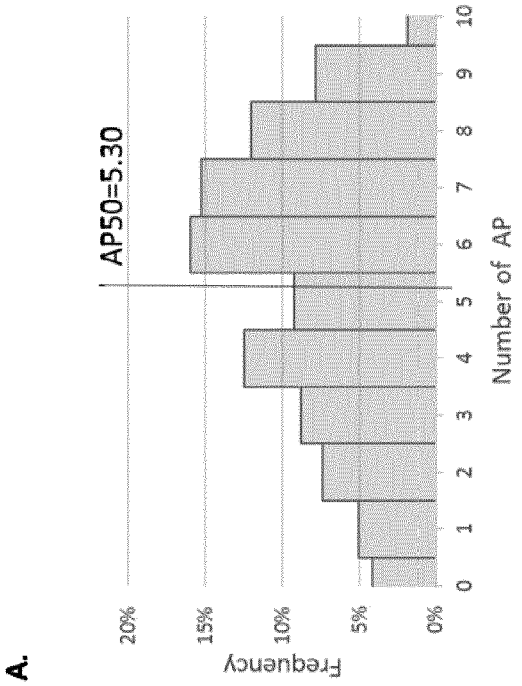
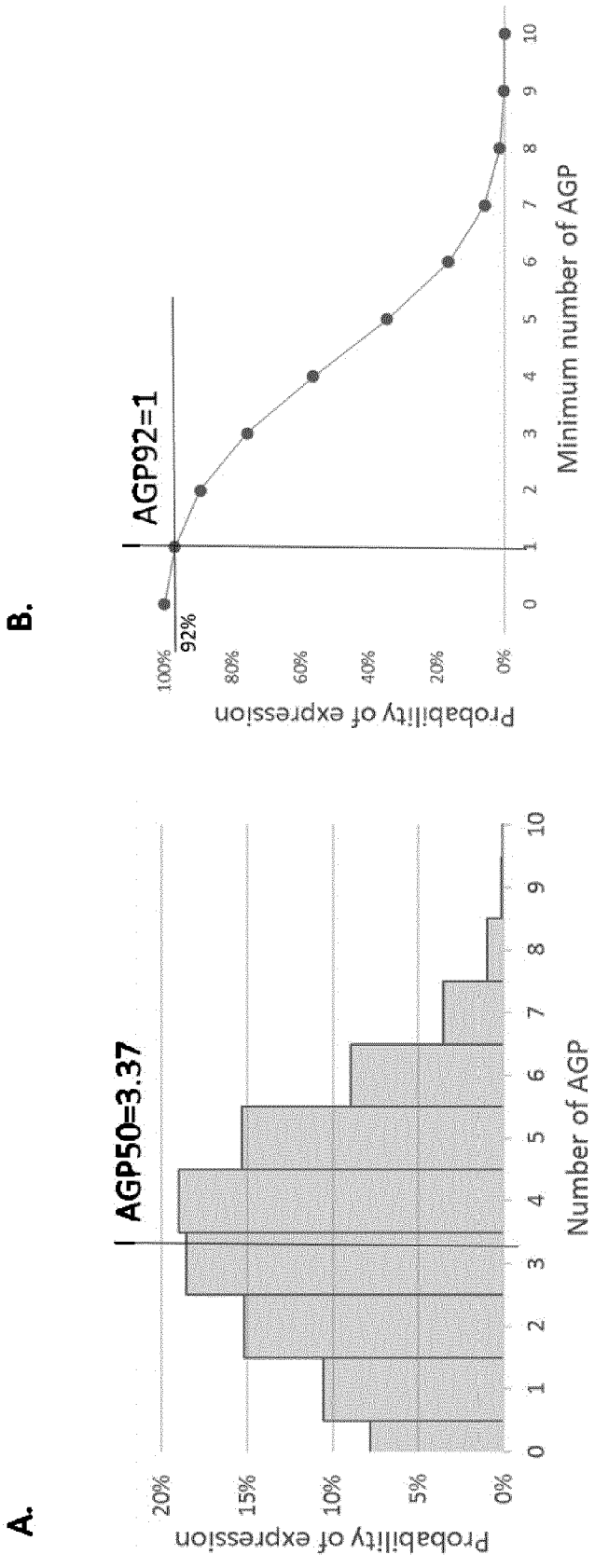


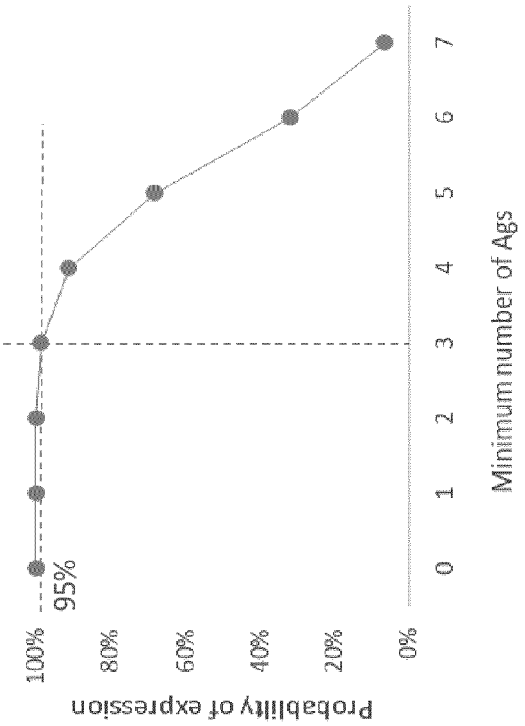
Figure 14


B.
AGP92=1
92%
100%
80%
60%
40%
20%
0%
Probability of expression
0
1
2
3
4
5
6
7
8
9
10
Minimum number of AGP
A.
AGP50=3.37
20%
15%
10%
5%
0%
Probability of expression
0
1
2
3
4
5
6
7
8
9
10
Number of AGP

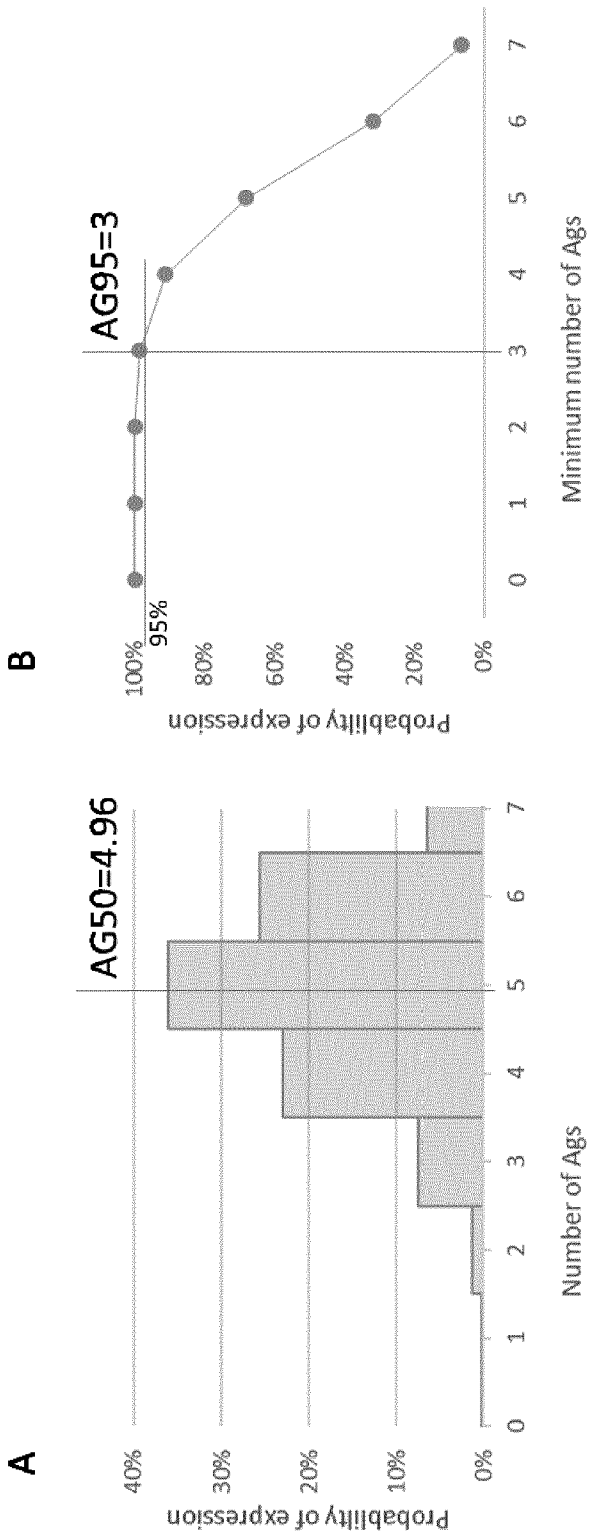
16/25
Figure 15

EXPRESSION RATES OF POLYPEPI1018 TARGET ANTIGENS IN CRC

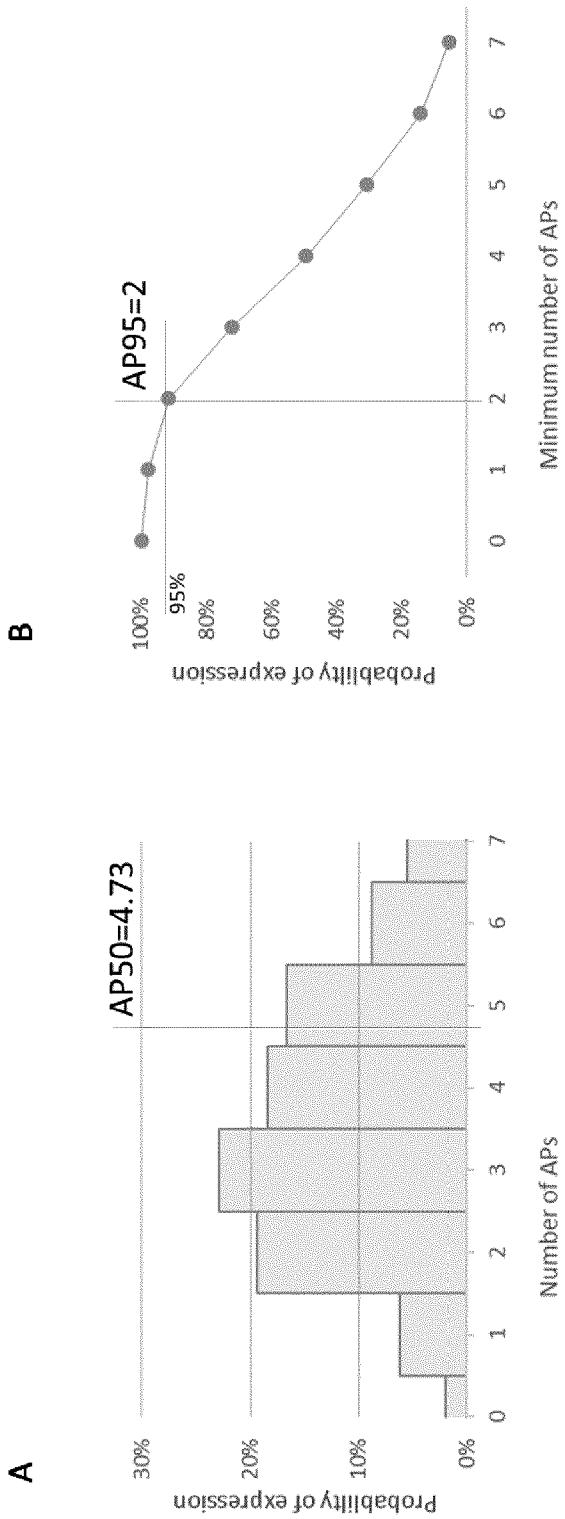
TSP50	89.47%
EpCAM	88.35%
Survivin	87.28%
CAGE1	74.47%
SPAG9	74.36%
FBXO39	38.60%
MAGE-A8	43.75%



17/25
Figure 16



18/25
Figure 17



19/25
Figure 18

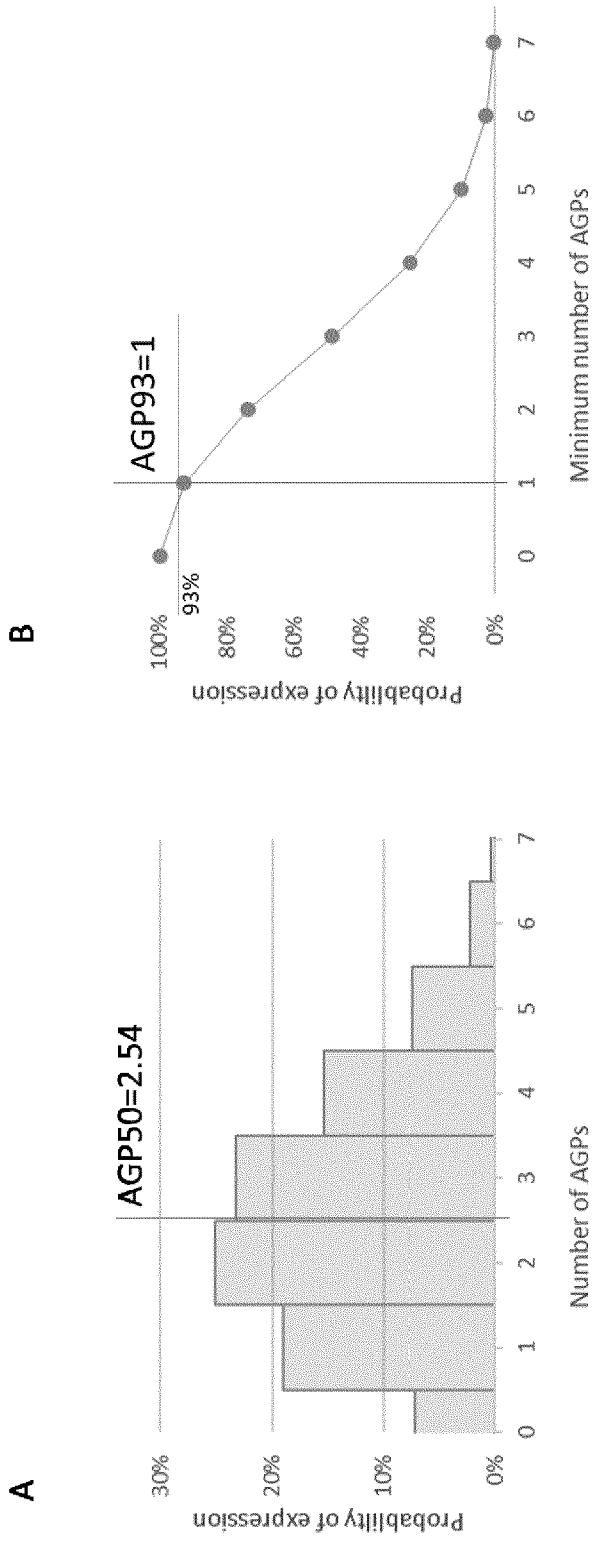
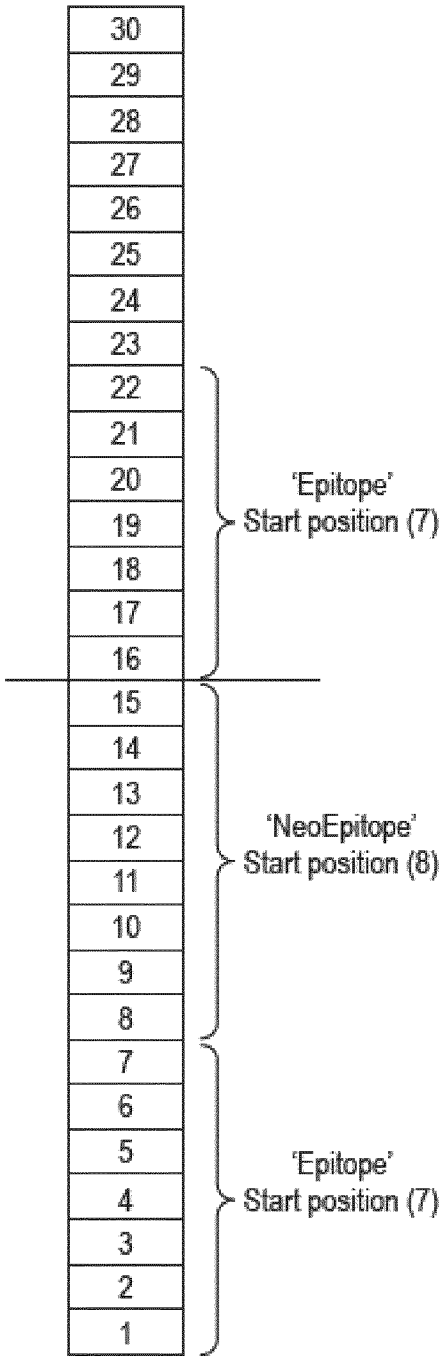
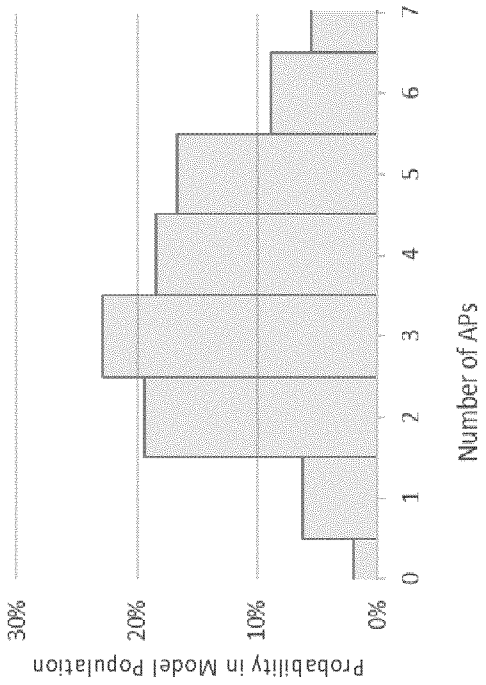


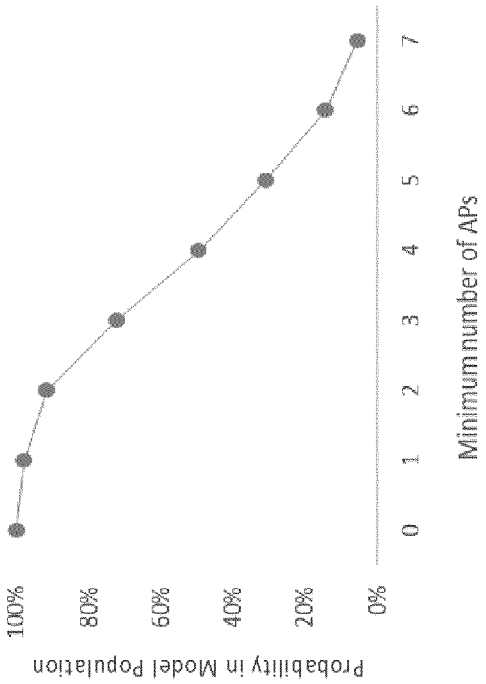
Fig. 19


30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
'Epitope'
Start position (7)
'NeoEpitope'
Start position (8)
'Epitope'
Start position (7)

21/25

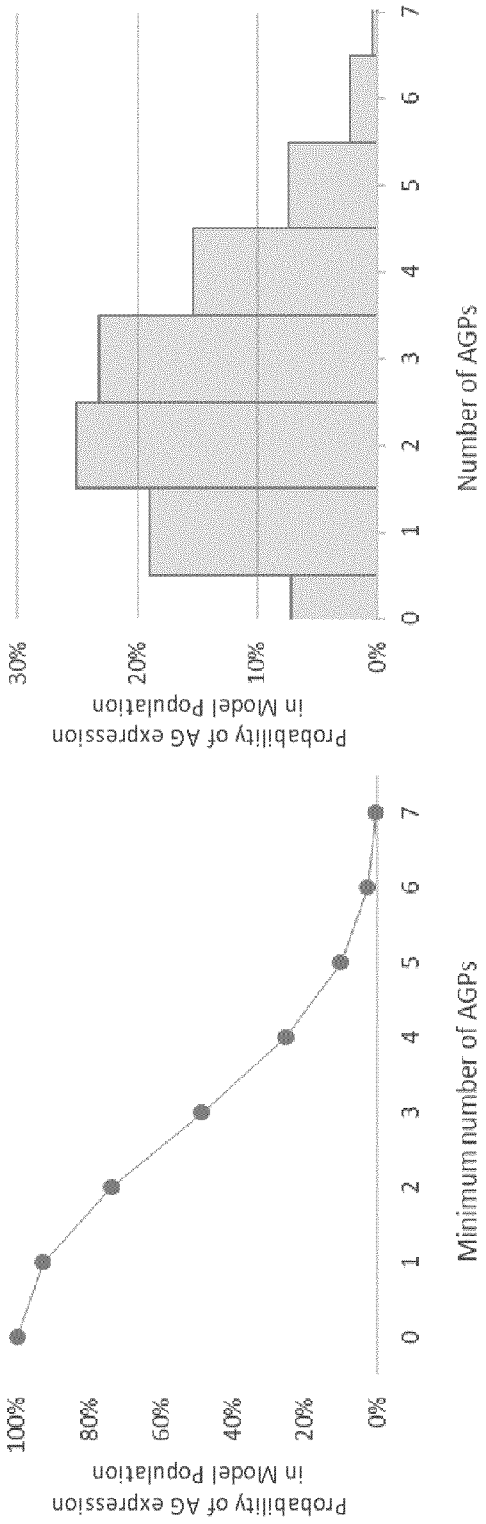
Figure 20





22/25

Figure 21



23/25
Figure 22

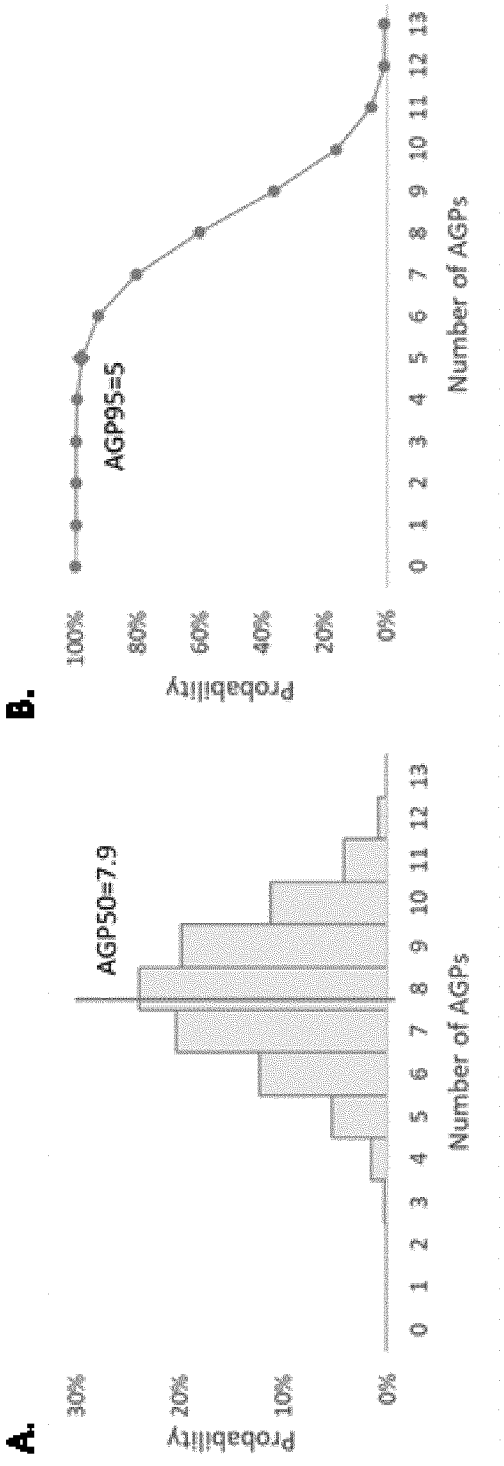
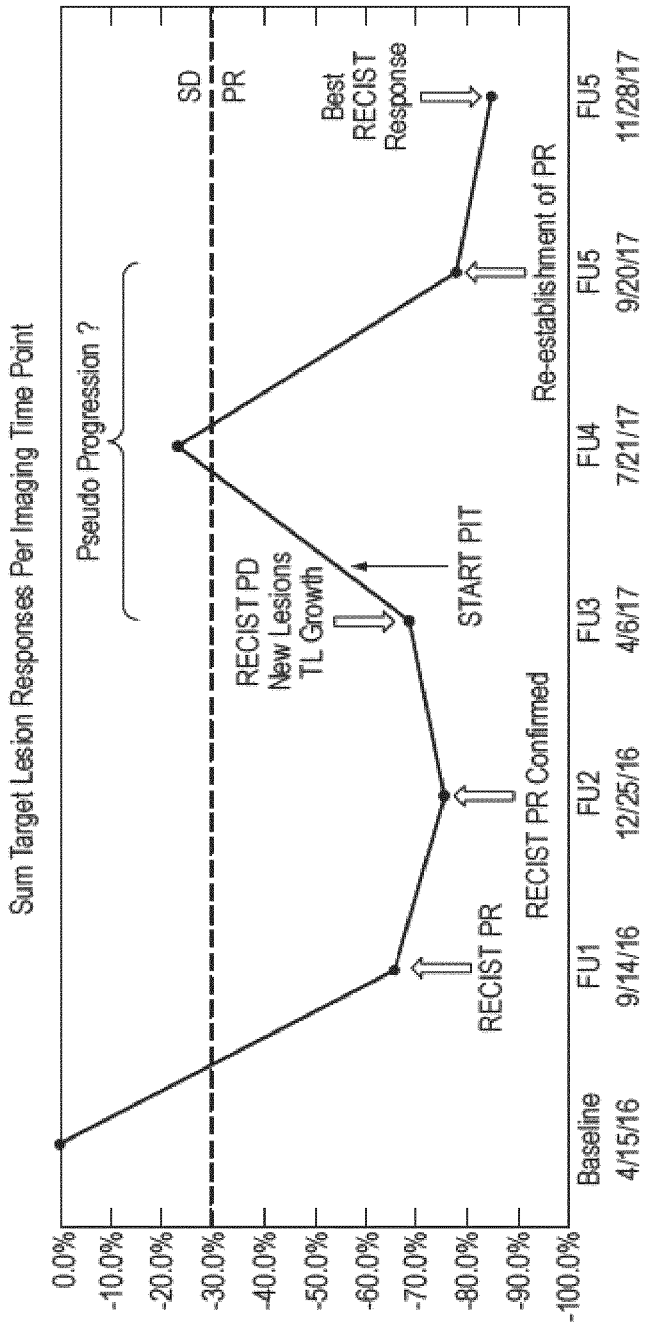
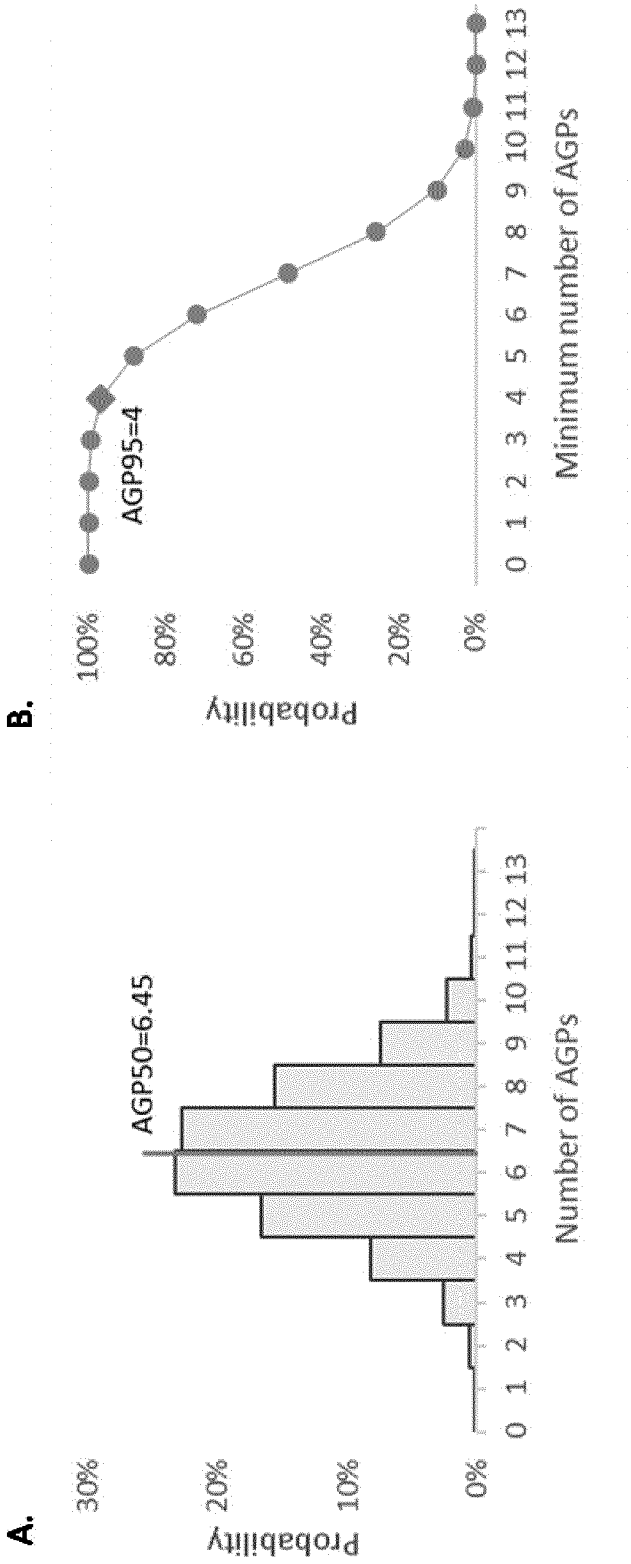


Fig. 23


Sum Target Lesion Responses Per Imaging Time Point
0.0%
-10.0%
-20.0%
-30.0%
-40.0%
-50.0%
-60.0%
-70.0%
-80.0%
-90.0%
-100.0%
SD
PR
Pseudo Progression ?
RECIST PD
New Lesions
TL Growth
START PIT
Best
RECIST
Response
RECIST PR
RECIST PR Confirmed
Re-establishment of PR
Baseline
4/15/16
FU1
9/14/16
FU2
12/25/16
FU3
4/6/17
FU4
7/21/17
FU5
9/20/17
FU5
11/28/17

25/25
Figure 24



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2018/055230

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/569 A61K38/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data, EMBASE, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2016/172722 A1 (NANTOMICS LLC [US]; NANT HOLDINGS IP LLC [US]) 27 October 2016 (2016-10-27) page 5, paragraph 0016 - paragraph 0017 page 9, paragraph 0029 page 11, paragraph 0033 - page 14, paragraph 0040; sequence 1048484 -----	1-3,7-20
Y	WEI JINGYAN ET AL: "Screening of single-chain variable fragments against TSP50 from a phage display antibody library and their expression as soluble proteins", JOURNAL OF BIOMOLECULAR SCREENING, vol. 11, no. 5, August 2006 (2006-08), pages 546-552, XP002779536, ISSN: 1087-0571 page 547, right-hand column, paragraph 3 ----- -/--	1-3,7-20

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 March 2018	08/06/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Montero Lopez, B

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2016/090177 A1 (VERIK BIO INC [US]) 9 June 2016 (2016-06-09) page 5, line 8 - line 23 page 7, line 15 - page 8, line 7 page 22, line 10 - page 23, line 9 page 36, line 25 - page 40, line 2 page 55, line 7 - page 57, line 3 page 59, line 5 - page 75 -----	1-3,7-20
A	US 2010/074925 A1 (CARMON LIOR [IL]) 25 March 2010 (2010-03-25) page 3, paragraph 0033 - page 4, paragraph 0050; table 1 page 6, paragraph 00787 - page 7, paragraph 110 page 8, paragraph 0131 - paragraph 0135 page 11, paragraph 0159 - paragraph 0167 -----	1-3,7-20
A	WO 00/18238 A1 (LONG ISLAND JEWISH RES INST [US]) 6 April 2000 (2000-04-06) page 7, line 5 - line 7 page 38, line 27 - page 40, line 4 -----	1-3,7-20
A	ZHENG LEI ET AL: "High Expression of Testes-Specific Protease 50 Is Associated with Poor Prognosis in Colorectal Carcinoma", PLOS ONE, vol. 6, no. 7, July 2011 (2011-07), XP002779537, ISSN: 1932-6203 abstract page 6, right-hand column, paragraph 2 -----	1-3,7-20

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2018/055230

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-3, 7-20(all partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-3, 7-20(all partially)

Polypeptide comprising a fragment of up to 50 consecutive aminoacids of a TSP50 colorectal cancer cancer-associated antigen, wherein the fragment comprises SEQ ID NO:21 or 130, a panel of at least two polypeptides comprising a polypeptide comprising SEQ ID NO:21 or 130, pharmaceutical composition or kit comprising such a polypeptide or panel and its use for vaccination; method for identifying a subject likely or not likely to have a T cell response or clinical response to administration of the pharmaceutical composition

---

2-72. claims: 1-4, 7-20(all partially)

Polypeptide comprising a fragment of up to 50 consecutive aminoacids of a colorectal cancer-associated antigen, wherein the fragment comprises either SEQ ID NO:22-40, 234-250, 251-271, 121, 131, 124, 134, or 126, a panel of at least two polypeptides comprising a polypeptide comprising of either SEQ ID NO:22-40, 234-250, 251-271, 121, 131, 124, 134, or 126, pharmaceutical composition or kit comprising such a polypeptide or panel and its use for vaccination; method for identifying a subject likely or not likely to have a T cell response or clinical response to administration of the pharmaceutical composition

---

73-147. claims: 1-6, 8-20(all partially)

Polypeptide comprising a fragment of up to 50 consecutive aminoacids of an ovarian-cancer associated antigen, wherein the fragment comprises either SEQ ID NO:272-301, 302-331 or 332-346, a panel of at least two polypeptides comprising a polypeptide comprising either SEQ ID NO:272-301, 302-331 or 332-346, pharmaceutical composition or kit comprising such a polypeptide or panel and its use for vaccination; method for identifying a subject likely or not likely to have a T cell response or clinical response to administration of the pharmaceutical composition

---

200-346. claims: 1-6, 8-20(all partially)

Polypeptide comprising a fragment of up to 50 consecutive aminoacids of a breast cancer-associated antigen, wherein the fragment comprises either SEQ ID NO:1-20, 24, 172-194, 41-80, 81-142, 196-233, or 435-449, a panel of at least two polypeptides comprising a polypeptide comprising either SEQ ID NO:1-20, 24, 172-194, 41-80, 41-80, 81-142, 196-233, or 435-449, pharmaceutical composition or kit comprising such a

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

polypeptide or panel and its use for vaccination; method for identifying a subject likely or not likely to have a T cell response or clinical response to administration of the pharmaceutical composition

---

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2018/055230

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2016172722 A1	27-10-2016	AU 2016253145 A1	04-01-2018
		CA 2988388 A1	27-10-2016
		EP 3286361 A1	28-02-2018
		JP 2018513187 A	24-05-2018
		US 2018141998 A1	24-05-2018
		WO 2016172722 A1	27-10-2016
WO 2016090177 A1	09-06-2016	AU 2015358387 A1	15-06-2017
		CA 2969231 A1	09-06-2016
		EP 3227679 A1	11-10-2017
		US 2016161486 A1	09-06-2016
		WO 2016090177 A1	09-06-2016
US 2010074925 A1	25-03-2010	AU 2007298494 A1	27-03-2008
		CA 2665816 A1	27-03-2008
		EP 2089423 A1	19-08-2009
		US 2010074925 A1	25-03-2010
		US 2017072036 A1	16-03-2017
		WO 2008035350 A1	27-03-2008
WO 0018238 A1	06-04-2000	AU 6273199 A	17-04-2000
		US 6617434 B1	09-09-2003
		US 2005100914 A1	12-05-2005
		WO 0018238 A1	06-04-2000