

PEPTIDE VACCINES

Field

The disclosure relates to polypeptides and vaccines that find use in the prevention or treatment of cancer, in particular cancers that express certain antigens expressed in gastric cancers, lung
5 cancers, melanomas and bladder 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
10 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
15 (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, including tumour associated antigens
20 (TAA), are processed into peptides. These peptides bind to HLA molecules 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. The inventors have demonstrated that an epitope that binds to a single HLA class I allele expressed in a subject is essential, but not sufficient to induce tumor specific T cell responses. Instead tumour specific T
25 cell responses are optimally activated when an epitope is recognised and presented by the HLA molecules encoded by at least three HLA class I genes of an individual (WO/2018/158456, WO/2018/158457, WO/2018/158455, EP 3370065 and EP 3369431).

Based on this discovery the inventors have identified the T cell epitopes from certain gastric and/or lung and/or melanoma and/or bladder cancer associated-polypeptide antigens
30 (tumor specific antigens (TSA) and/or 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

- 5 (a) a gastric cancer-associated antigen selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 30;
- 10 (b) a lung cancer-associated antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 90 to 119;
- 15 (c) a melanoma-associated antigen selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, NY-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 178 to 207; and/or
- 20 (d) a bladder cancer-associated antigen selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 268 to 297.

In some specific cases the disclosure provides a polypeptide that

- 25 (a) is a fragment of a gastric cancer-associated antigen selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 30; or
- 30 (b) comprises or consists of two or more fragments of one or more gastric cancer associated antigens selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1, wherein each fragment comprises a different amino acid

sequence selected from any one of SEQ ID NOs: 1 to 30, optionally wherein the fragments overlap or are arranged end to end in the polypeptide; or

(c) is a fragment of a lung cancer-associated antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 90 to 119; or

(d) comprises or consists of two or more fragments of one or more lung cancer associated antigens selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 90 to 119, optionally wherein the fragments overlap or are arranged end to end in the polypeptide; or

(e) is a fragment of a melanoma-associated antigen selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 178 to 207; or

(f) comprises or consists of two or more fragments of one or more melanoma associated antigens selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 178 to 207, optionally wherein the fragments overlap or are arranged end to end in the polypeptide; or

(g) is a fragment of a bladder cancer-associated antigen selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 268 to 297; or

(h) comprises or consists of two or more fragments of one or more bladder cancer associated antigens selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3,

MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 268 to 297, optionally wherein the fragments overlap or are arranged end to end in the polypeptide.

- 5 In some specific cases the polypeptide comprises or consists of fragments of at least two different cancer-associated antigens, wherein the cancer-associated antigens are selected from
- (a) DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1 and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 1 to 30;
- 10 and/or
- (b) BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1; and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 90 to 119; and/or
- 15 (c) PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1; and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 178 to 207 and/or
- 20 (d) PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12; and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 268 to 297.

In some cases the polypeptide comprises or consists of one or more amino acid sequences selected from SEQ ID NOs: 31 to 60, 90 to 119, 178 to 207 and/or 268 to 297.

In some cases the polypeptide comprises or consists of one or more amino acid sequences selected from SEQ ID NOs: 61 to 75, 120 to 149, 208 to 237, and/or 298 to 327.

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: 1 to 30; and/or SEQ ID NOs: 31 to 60; and/or SEQ ID NOs: 61 to 75; or selected from SEQ ID NOs: 90 to 119; and/or SEQ ID NOs: 120 to 149; and/or SEQ ID NOs:

150 to 164; or selected from SEQ ID NOs: 178 to 207; and/or SEQ ID NOs: 208 to 237; and/or
SEQ ID NOs: 238 to 252; or selected from SEQ ID NOs: 268 to 297; and/or SEQ ID NOs: 298
to 327; and/or SEQ ID NOs: 328 to 342.

In a further aspect the disclosure provides a pharmaceutical composition or kit having one or
5 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: 1 to 30;
SEQ ID NOs: 31 to 60; and/or SEQ ID NOs: 61-75; and/or SEQ ID NOs: 90 to 119; SEQ ID
NOs: 120 to 149; and/or SEQ ID NOs: 150 to 164; and/or SEQ ID NOs: 178 to 207; and/or SEQ
ID NOs: 208 to 237; and/or SEQ ID NOs: 238 to 252; and/or SEQ ID NOs: 268 to 297; and/or
10 SEQ ID NOs: 298 to 327; and/or SEQ ID NOs: 328 to 342 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 CD8+ 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
15 gastric cancer, lung cancer, melanoma and bladder 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 gastric cancer, lung cancer, melanoma and bladder cancer; and
- 20 - 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 gastric
cancer, lung cancer, melanoma and bladder cancer.

In a further aspect the disclosure provides a method of identifying a human subject who will
likely have a CD8+ T cell response to administration of a pharmaceutical composition as
25 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 CD8+ T cell response to administration
30 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

- 5 (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
- 10 (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 described above, wherein one or more of the following factors corresponds to a higher likelihood of a clinical
15 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
20 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
25 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
30 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

5 determined to comprise 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.

In some cases the cancer-associated antigens may be a gastric cancer antigen selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-
10 A3, LAGE-1, MAGE-A10, MAGE-A1, SSX1, a lung cancer antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2NY -SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1, MAGE-A1, a melanoma antigen selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10, MAGE-A1, and/or a bladder cancer antigen
15 selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8 and HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE A12 . 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
20 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.

25 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
30 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
- 5 (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 and/or class II genotype of the subject.

The present disclosure includes methods of treating a human subject likely to respond to a polypeptide(s) or pharmaceutical composition comprising the polypeptide(s) of the current disclosure comprising: (a) determining that a polypeptide(s) or pharmaceutical composition comprising the polypeptide(s) comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the human subject; and (b) administering to the human subject the polypeptide or pharmaceutical composition. In certain embodiments, the method further comprises using population expression data for each antigen to determine the likelihood that the human subject will have a CD8+ T cell response that targets one or more polypeptide antigens that are expressed by cancer cells of the human subject, wherein the antigen: (a) is selected from gastric cancer antigens DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 SSX1, and/or lung cancer antigens BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1, MAGE-A1, and/or melanoma cancer antigens PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, NY-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10, MAGE-A1, and/or bladder cancer antigens PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8 and HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12; and (b) comprises an amino acid sequence that is a fragment of an active ingredient peptide of the pharmaceutical composition, and a T cell epitope capable of binding to at least three HLA class I molecules of the human subject.

30 The present disclosure includes methods of treating a human subject likely to respond to a polypeptide(s) or pharmaceutical composition comprising the polypeptide(s) of the current

disclosure comprising: (a) determining that a polypeptide(s) or pharmaceutical composition comprising the polypeptide(s) 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 human subject; and a fragment of a cancer-associated antigen expressed by cancer cells of the human subject, optionally wherein the cancer-associated antigen is present in a sample obtained from the human subject; and (b) administering to the human subject the polypeptide or pharmaceutical composition.

The present disclosure includes methods of treating a human subject likely to respond to a polypeptide(s) or pharmaceutical composition comprising the polypeptide(s) of the current disclosure comprising determining any one of the following: (a) presence in a polypeptide(s) or pharmaceutical composition comprising the 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 human subject; (b) a higher number of target polypeptide antigens in the human subject or a sample from the human subject, comprising at least one amino acid sequence that is both comprised in an active ingredient polypeptide, and a T cell epitope capable of binding to at least three HLA class I of the human subject; (c) a higher probability that the human 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 comprised in an active ingredient polypeptide, and a T cell epitope capable of binding to at least three HLA class I of the human subject; and/or (d) a higher number of target polypeptide antigens that the human subject is predicted to express, optionally a higher number of target polypeptide antigens that the human 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 comprised in an active ingredient polypeptide, and a T cell epitope capable of binding to at least three HLA class I of the human subject; and (e) administering to the human subject the polypeptide or pharmaceutical composition. In certain embodiments, the method further comprises identifying which polypeptide antigens targeted by the polypeptide(s) or pharmaceutical composition comprising the polypeptide(s) comprise an amino acid sequence that is both comprised in an active ingredient polypeptide, and a T cell epitope capable of binding to at least three HLA class I of the human subject using population expression data for each antigen

identified to determine the probability that the human subject expresses one or more of the antigens identified that together comprise at least two different amino acid sequences; and administering to the human subject the polypeptide or pharmaceutical composition. In certain embodiments, 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).

The 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.

20

Description of the Figures

Fig. 1

ROC curve of HLA restricted PEPI biomarkers.

Fig. 2

25 ROC curve of ≥ 1 PEPI3+ Test for the determination of the diagnostic accuracy. AUC = 0.73 classifies a fair diagnostic value for the PEPI biomarker.

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

30

induced T cell response set of individuals ($p < 0.001$). B: Class I HLA restricted epitopes (PEPI1+). The OPA between predicted epitopes and CD8+ T cell responses was 25% (not statistically significant). True positive (TP), both peptide and T cell responses were detected (shaded); True negative (TN): neither peptides nor T cell responses were detected (shaded);

5 False negative (FN), only T cell responses were detected; False positive (FP), only peptide were detected.

Fig. 4

Correlation between PEPI Test predicted CD4 peptides and T-cell reactivity measured with peptide pools in patients treated with SLP vaccine. A: ≥ 3 HLA class II allele-binding

10 PEPIs; B: single HLA class II allele-binding epitopes. Gray: true positive (TP) and true negative (TN) responses; White: false negative (FN) and false positive (FP) responses. TP: both peptide and T cell responses were detected; TN: neither peptides nor T cell responses were detected; FN: only T cell responses were detected; FP: only peptides were detected.

Fig. 5

15 Multiple HLA binding peptides that define the HPV-16 LPV vaccine specific T cell response set of 20 VIN-3 and 5 cervical cancer patients. PEPI counts were compared to clinical responses after treatment with LPV. Predicted CD8+ T cell responders according to HLA class I PEPIs (A) and CD4+ T cell responders according to HLA class II PEPIs (B). Correlation between HLA class I (C) and class II (D) PEPI count and clinical response at 3 months follow-up in VIN-3

20 patients. Predicted T cell responders: PEPI count ≥ 1 . Gray column, patient with HPV16 E6- and/or E7-specific T cell response; Dashed column, patient without T cell responses. CR, complete clinical responder; PR, partial clinical responder; NR, clinical non-responder.

Fig. 6

The multiple HLA class I binding peptides that define the HPV vaccine specific T cell response

25 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

30 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

TSA expression probability targeted by IMA901 vaccine.

Fig. 8

HLA Class I allele binding properties of TUMAPs of IMA901 peptide vaccine for 2,915 common alleles. (A) and for the Class I genotype (6 alleles) of 51 HLA-A*02+ RCC patients. Percentages at the bottom indicate the proportion of HLAs the TUMAPs can bind to. Lines in darker grey indicate binding HLA alleles. (B) Probability indicates the proportion of patients who can present the indicated number of TUMAPs with their three or more HLAs. AP indicates number of antigens which can generate at least one PEPI. In this case, since both the antigens and the predicted PEPIs are 9mers, AP=TUMAP=PEPI.

Fig. 9

Correlation between immune response measured for any TUMAP and immune response against expressed antigen on the tumor (AGP)

Fig. 10

Correlation study between immune response rates (IRR) and PEPI Score, between objective response rates (ORR) and MultiPEPI Scores and between objective response rates (ORR) and MultiAg PEPI Scores. A: Preliminary experiment to explore the relationship between PEPI Score and immune response rate of therapeutic vaccines ($r^2=0.7$, $p=0.001$) B: IRR - PEPI Score plot. ($r^2=0.47$, $p=0.001$). C: MultiPEPI Score and clinical response rate of therapeutic vaccines ($r^2=0.75$, $p=0.001$). D: ORRs plotted against the MultiPEPI Score ($r^2=0.12$, $p=0.124$). E: ORRs plotted against the MultiAg PEPI Score for vaccines with multiple antigens ($r^2=0.64$; $p=0.009$). F: ORRs plotted against the MultiPEPI Score for vaccines with multiple antigens ($r^2=0.87$; $p=0.0002$). G: ORRs plotted against the MultiPEPI Score in patients with target antigen positive

disease ($r^2=0.56$ and $p=0.005$). Dark grey dashed lines indicate the 95% confidence interval; grey dashed line indicates the trendline.

Fig. 11

OBERTO trial design (NCT03391232)

5 Fig. 12

Antigen expression in CRC cohort of OBERTO trial (n=10). A: Expression frequencies of PolyPEPI1018 source antigens determined based on 2391 biopsies. B: PolyPEPI1018 vaccine design specified as 3 out of 7 TSAs are expressed in CRC tumors with above 95% probability. C: In average, 4 out of the 10 patients had pre-existing immune responses against each target
10 antigens, referring to the real expression of the TSAs in the tumors of the patients. D: 7 out of the 10 patients had pre-existing immune responses against minimum of 1 TSA, in average against 3 different TSAs.

Fig. 13

Immunogenicity of PolyPEPI1018 in CRC patients confirms proper target antigen and target
15 peptide selection. Upper part: target peptide selection and peptide design of PolyPEPI1018 vaccine composition. Two 15mers from CRC specific CTA (TSA) selected to contain 9mer PEPI3+ predominant in representative Model population. Table: PolyPEPI1018 vaccine has been retrospectively tested during a preclinical study in a CRC cohort and was proven to be immunogenic in all tested individuals for at least one antigen by generating PEPI3+s. Clinical
20 immune responses were measured specific for at least one antigen in 90% of patients, and multi-antigen immune responses were also found in 90% of patients against at least 2, and in 80% of patients against at least 3 antigens as tested with IFNy fluorospot assay specifically measured for the vaccine-comprising peptides.

Fig. 14

25 Clinical response for PolyPEPI1018 treatment. A: Swimmer plot of clinical responses of OBERTO trial (NCT03391232). B: Association progression free survival (PFS) and AGP count. C: Association tumour volume and AGP count.

Fig. 15

Peptide hotspot analysis example: PRAME antigen hotspot on 433 patients of the Model
30 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).

5 Fig. 16

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

Fig. 17

Antigen expression distribution for gastric cancer based on the calculation of multi-antigen
 10 responses from expression frequencies of the selected 14 different CTAs. A: non-cumulative distribution to calculate the expected value for the number of expressed antigens (AG50). This value shows that probably 7.18 vaccine antigens will be expressed by gastric tumor cells. B: cumulative distribution curve of the minimum number of expressed antigens (CTA expression curve). This shows that minimum 5 vaccine antigens will be expressed with 95% probability in
 15 gastric cancer cell (AG95).

Fig. 18

PEPI representing antigens: gastric cancer vaccine-specific CTA antigens with ≥ 1 PEPI, called as "AP") distribution within the Model Population (n=433) for gastric cancer vaccine. A: non-cumulative distribution of AP where the average number of APs is: AP50=7.98, meaning that in
 20 average almost 8 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 three vaccine antigen will have PEPIs in 95% of the Model Population (n=433) (AP95=3).

Fig. 19

PEPI represented expressed antigen (gastric cancer vaccine-specific CTA antigens expressed by
 25 the tumor, for which ≥ 1 PEPI is predicted, called as "AGP") distribution within the model population (n=433) calculated with CTA expression rates for gastric cancer. A: non-cumulative distribution of AGP where the expected value for number expressed CTAs represented by PEPI is AGP50=3.86. AGP50 is a measure of the effectiveness of the disclosed gastric cancer vaccine in attacking gastric tumor in an unselected patient population. AGP50 = 3.86 means that at least
 30 3 CTAs from the vaccine will probably be expressed by the gastric 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 PEPs in 95% of the population and the remaining 5% of the population will likely have no AGP at all (AGP95=1).

5 Fig. 20

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

Fig. 21

Antigen expression distribution for lung cancer based on the calculation of multi-antigen
10 responses from expression frequencies of the selected 13 different CTAs. A: non-cumulative distribution to calculate the expected value for the number of expressed antigens (AG50). This value shows that probably 4.76 vaccine antigens will be expressed by lung tumor cells. B: cumulative distribution curve of the minimum number of expressed antigens (CTA expression curve). This shows that minimum 2 vaccine antigens will be expressed with 95% probability in
15 lung cancer cell (AG95).

Fig. 22

PEPI representing antigens: lung cancer vaccine-specific CTA antigens with ≥ 1 PEPI, called as "AP") distribution within the Model Population (n=433) for lung cancer vaccine. A: non-cumulative distribution of AP where the average number of APs is: AP50=7.6, meaning that in
20 average almost 8 CTAs will have PEPs 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 two vaccine antigen will have PEPs in 95% of the Model Population (n=433) (AP95=2).

Fig. 23

PEPI represented expressed antigen (lung cancer vaccine-specific CTA antigens expressed by
25 the tumor, for which ≥ 1 PEPI is predicted, called as "AGP") distribution within the model population (n=433) calculated with CTA expression rates for lung cancer. A: non-cumulative distribution of AGP where the expected value for number expressed CTAs represented by PEPI is AGP50=2.77. AGP50 is a measure of the effectiveness of the disclosed lung cancer vaccine in attacking lung tumor in an unselected patient population. AGP50 = 2.77 means that at least 3
30 CTAs from the vaccine will probably be expressed by the lung tumor cells and present PEPs 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 PEPs in 91% of the population and the remaining 9% of the population will likely have no AGP at all (AGP95=0).

Fig. 24

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

Fig. 25

- Antigen expression distribution for melanoma based on the calculation of multi-antigen responses from expression frequencies of the selected 15 different CTAs. A: non-cumulative distribution to calculate the expected value for the number of expressed antigens (AG50). This value shows that probably 7.62 vaccine antigens will be expressed by gastric tumor cells. B: cumulative distribution curve of the minimum number of expressed antigens (CTA expression curve). This shows that minimum 5 vaccine antigens will be expressed with 95% probability in melanoma cell (AG95).
- 10

- 15 Fig. 26

- PEPI representing antigens: melanoma vaccine-specific CTA antigens with ≥ 1 PEPI, called as "AP") distribution within the Model Population (n=433) for melanoma vaccine. A: non-cumulative distribution of AP where the average number of APs is: AP50=8.29, meaning that in average 8 CTAs will have PEPs 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 three vaccine antigen will have PEPs in 95% of the Model Population (n=433) (AP95=2).
- 20

Fig. 27

- PEPI represented expressed antigen (melanoma 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 melanoma. A: non-cumulative distribution of AGP where the expected value for number expressed CTAs represented by PEPI is AGP50=4.22. AGP50 is a measure of the effectiveness of the disclosed melanoma vaccine in attacking gastric tumor in an unselected patient population. AGP50 = 4.22 means that at least 3 CTAs from the vaccine will probably be expressed by the gastric tumor cells and present PEPs 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 PEPs in
- 25
- 30

95% of the population and the remaining 5% of the population will likely have no AGP at all (AGP95=1).

Fig. 28

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

Fig. 29

Antigen expression distribution for bladder cancer based on the calculation of multi-antigen responses from expression frequencies of the selected 17 different CTAs. A: non-cumulative distribution to calculate the expected value for the number of expressed antigens (AG50). This value shows that probably 8.85 vaccine antigens will be expressed by bladder 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 bladder cancer cell (AG95).

Fig. 30

PEPI representing antigens: bladder cancer vaccine-specific CTA antigens with ≥ 1 PEPI, called as "AP") distribution within the Model Population (n=433) for bladder cancer vaccine. A: non-cumulative distribution of AP where the average number of APs is: AP50=9.44, meaning that in average almost 8 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 three vaccine antigen will have PEPIs in 95% of the Model Population (n=433) (AP95=3).

Fig. 31

PEPI represented expressed antigen (bladder 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 bladder cancer. A: non-cumulative distribution of AGP where the expected value for number expressed CTAs represented by PEPI is AGP50=3.90. AGP50 is a measure of the effectiveness of the disclosed bladder cancer vaccine in attacking bladder tumor in an unselected patient population. AGP50 = 3.90 means that at least 3 CTAs from the vaccine will probably be expressed by the bladder 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 95% of the population and the remaining 5% of the population will likely have no AGP at all ($AGP_{95}=1$).

Fig. 32

Probability of vaccine antigen expression in the Patient-A's tumor cells. There is over 95% probability that 5 out of the 13 target antigens in the vaccine regimen is expressed in the patient's tumor. Consequently, the 13 peptide vaccines together can induce immune responses against at least 5 ovarian cancer antigens with 95% probability (AGP_{95}). It has 84% probability that each peptide will induce immune responses in the Patient-A. AGP_{50} is the mean (expected value) =7.9 (it is a measure of the effectiveness of the vaccine in attacking the tumor of Patient-A).

Fig. 33

Treatment schedule of Patient-A.

Fig. 34

T cell responses of patient-A. A. Left: Vaccine peptide-specific T cell responses (20-mers). right: CD8+ cytotoxic T cell responses (9-mers). Predicted T cell responses are confirmed by bioassay.

Fig. 35

MRI findings of Patient-A 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.

Fig. 36

Probability of vaccine antigen expression in the Patient-B's tumor cells and treatment schedule of Patient-B. A: There is over 95% probability that 4 out of the 13 target antigens in the vaccine is expressed in the patient's tumor. B: Consequently, the 12 peptide vaccines together can induce immune responses against at least 4 breast cancer antigens with 95% probability (AGP_{95}). It has 84% probability that each peptide will induce immune responses in the Patient-B. $AGP_{50} = 6.45$; it is a measure of the effectiveness of the vaccine in attacking the tumor of Patient-B. C: Treatment schedule of Patient-B.

Fig. 37

T cell responses of Patient-A. Left: Vaccine peptide-specific T cell responses (20-mers) of P. Right: Kinetic of vaccine-specific CD8+ cytotoxic T cell responses (9-mers). Predicted T cell responses are confirmed by bioassay.

Fig. 38

Treatment schedule of Patient-C.

Fig. 39

- T cell responses of Patient-C. A: Vaccine peptide-specific T cell responses (20-mers). B: Vaccine peptide-specific CD8+ T cell responses (9-mers). C-D: Kinetics of vaccine-specific CD4+ T cells and CD8+ cytotoxic T cell responses (9-mers), respectively. Long lasting immune responses both CD4 and CD 8 T cell specific are present after 14 months.

Fig. 40

Treatment schedule of Patient-D.

- 10 Fig. 41

Immune responses of Patient-D for PIT treatment. A: CD4+ specific T cell responses (20mer) and B: CD8+ T cell specific T cell responses (9mer). 0.5-4 months refer to the timespan following the last vaccination until PBMC sample collection.

Fig. 42

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

Description of the Sequences

- SEQ ID NOs: 1 to 30 set forth 9 mer T cell epitopes described in Table 20a.
- 20 SEQ ID NOs: 31 to 60 set forth 15 mer T cell epitopes described in Table 20a.
- SEQ ID NOs: 61 to 75 set forth gastric cancer vaccine peptides described in Table 21a.
- SEQ ID NOs: 76 to 89 set forth gastric cancer associated antigens.
- SEQ ID NOs: 90 to 119 set forth 9 mer T cell epitopes described in Table 20b.
- SEQ ID NOs: 120 to 149 set forth 15 mer T cell epitopes described in Table 20b.
- 25 SEQ ID NOs: 150 to 164 set forth lung cancer vaccine peptides described in Table 21b.
- SEQ ID NOs: 165 to 177 set forth lung cancer associated antigens.
- SEQ ID NOs: 178 to 207 set forth 9 mer T cell epitopes described in Table 20c.
- SEQ ID NOs: 208 to 237 set forth 15 mer T cell epitopes described in Table 20c.
- SEQ ID NOs: 238 to 252 set forth melanoma vaccine peptides described in Table 21c.
- 30 SEQ ID NOs: 253 to 267 set forth melanoma associated antigens.
- SEQ ID NOs: 268 to 297 set forth 9 mer T cell epitopes described in Table 20d.

SEQ ID NOs: 298 to 327 set forth 15 mer T cell epitopes described in Table 20d.

SEQ ID NOs: 328 to 342 set forth bladder cancer vaccine peptides described in Table 21d.

SEQ ID NOs: 343 to 359 set forth bladder cancer associated antigens.

5 SEQ ID NOs: 360 to 372 set forth sequences of personalized vaccine of Patient-A and are described in Table 22.

SEQ ID NOs: 373 to 384 set forth sequences of personalized vaccine of Patient-B and are described in Table 24.

SEQ ID NO: 385 sets forth the 30 amino acid CRC_P3 peptide, Figure 13.

SEQ ID NOs: 386 to 394 set forth the 9mer sequences shown in Figure 8.

10

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*,
15 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.

The nomenclature used to designate the amino acid sequence of the HLA molecule is as
20 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
25 sequence determining the binding specificity to non-self antigenic peptides (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
30 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. In some embodiments, the HLA genotype of an individual is obtained or determined by assaying a biological sample from the individual. The biological sample typically contains subject DNA. The biological sample may be, for example, a blood, serum, plasma, saliva, urine, expiration, cell or tissue sample. In some embodiments the biological sample is a saliva sample. In some embodiments the biological sample is a buccal swab sample. An HLA genotype may be obtained or 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. In some embodiments, the HLA genotype is determined using sequence specific primer (SSP) technologies. In some embodiments, the HLA genotype is determined using sequence specific oligonucleotide (SSO) technologies. In some embodiments, the HLA genotype is determined using sequence based typing (SBT) technologies. In some embodiments, the HLA genotype is determined using next generation sequencing. 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”.

HLA-epitope binding

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.

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.

The term “personal epitope”, or “PEPI” as used herein distinguishes a subject-specific epitope from an HLA specific epitope. 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 (or II) molecule of an individual.

5 “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.

10 “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.

15 “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.

20 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
25 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.

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,
30 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/EpitopePrediction.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/>

HLA molecules regulate T cell responses. Until recently, the triggering of an immune response to individual epitopes was thought to be determined by recognition of the epitope by the product of single HLA allele, i.e. HLA-restricted epitopes. However, HLA-restricted
5 epitopes induce T cell responses in only a fraction of individuals. 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.

The inventors discovered that multiple HLA expressed by an individual need to present
10 the same peptide in order to trigger a T cell response. Therefore the fragments of a polypeptide antigen (epitopes) that are immunogenic for a specific individual (PEPIs) are those that can bind to multiple class I (activate CD8+ T cells eg, cytotoxic T cells) or class II (activate CD4+ T cells, eg. helper T cells or CD4+ killer cells) HLAs expressed by that individual. This discovery is described in PCT/EP2018/055231, PCT/EP2018/055232, PCT/EP2018/055230, EP 3370065 and
15 EP 3369431.

Polypeptides

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

20 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 polypeptides are typically about 9 to 50 or 15 to 40
25 or 20 to 30 amino acids long.

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
 5 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.

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
 10 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.
 15 Springer New York. 2012).

The peptides of the disclosure may comprise or consist of one or more fragments of one or more gastric cancer-associated antigens selected from DPPA2 (SEQ ID NO: 76, Q7Z7J5.1), CAGE-1 (SEQ ID NO: 77, Q8TC20.1), TSP50 (SEQ ID NO: 78, Q9UI38.1), HIWI (SEQ ID NO: 79, Q96J94.1), SURVIVIN (SEQ ID NO: 80, O15392.1), 5T4 (SEQ ID NO: 81, Q13641.1),
 20 PRAME (SEQ ID NO: 82, P78395.1), KK-LC-1 (SEQ ID NO 83, Q5H943.1), MAGE-A2 (SEQ ID NO: 84, P43356.1), MAGE-A3 (SEQ ID NO: 85, P43357.1), LAGE-1 (SEQ ID NO: 86, O75638.1), MAGE-A10 (SEQ ID NO: 87, P43363.1), MAGE-A1 (SEQ ID NO: 88, P43355.1) and SSX1 (SEQ ID NO: 89, Q16384.1); and/or one or more lung cancer-associated antigens selected from BRDT (SEQ ID NO: 165, Q58F21.1), PRAME (SEQ ID NO: 166, P78395.1),
 25 NALP4 (SEQ ID NO: 167, Q96MN2.1), MAGE-A12 (SEQ ID NO: 168, P43365.1), MAGE-A2 (SEQ ID NO: 169, P43356.1), SURVIVIN (SEQ ID NO: 170, O15392.1), DPPA2 (SEQ ID NO: 171, Q7Z7J5.1), NY-SAR-35 (SEQ ID NO 172, Q8N0W7.1), LDHC (SEQ ID NO: 173, P07864.1), MAGE-C2 (SEQ ID NO: 174, Q9UBF1.1), MAGE-A3 (SEQ ID NO: 175, P43357.1), KK-LC-1 (SEQ ID NO: 176, Q5H943.1) and MAGE-A1 (SEQ ID NO: 177, P43355.1); and/or one or more melanoma cancer-associated antigens selected from PRAME
 30 (SEQ ID NO: 253, P78395.1), MAGE-A2 (SEQ ID NO: 254, P43356.1), MAGE-C1 (SEQ ID

NO: 255, P43355.1), SURVIVIN (SEQ ID NO: 256, O15392.1), MAGE-A12 (SEQ ID NO: 257, P43365.1), Ny-ESO-1 (SEQ ID NO: 258, P78358.1), MAGE-C2 (SEQ ID NO: 259, Q9UBF1.1), MAGE-A6 (SEQ ID NO 260, P43360.1), BORIS (SEQ ID NO: 261, Q8NI51.1), LAGE-1 (SEQ ID NO: 262, O75638.1), MAGE-A11 (SEQ ID NO: 263, P43364.1), SSX-1 (SEQ ID NO: 264, Q16384.1), MAGE-A3 (SEQ ID NO: 265, P43357.1) MAGE-A10 (SEQ ID NO: 266, P43363.1) and MAGE-A1 (SEQ ID NO: 267, P43355.1); and/or one or more bladder cancer-associated antigens selected from PIWIL2 (SEQ ID NO: 343, Q8TC59.1), CTAGE1 (SEQ ID NO: 344, Q96RT6.1), MAGE-A9 (SEQ ID NO: 345, P43362.1), EpCAM (SEQ ID NO: 346, P16422.1), OY-TES-1 (SEQ ID NO: 347, Q8NEB7.1), NY-ESO-1 (SEQ ID NO: 348, P78358.1), SURVIVIN (SEQ ID NO: 349, O15392.1), MAGE-C1 (SEQ ID NO 350, O60732.1), MAGE-A2 (SEQ ID NO: 351, P43356.1), LAGE-1 (SEQ ID NO: 352, O75638.1), MAGE-A3 (SEQ ID NO: 353, P43357.1), MAGE-A8 (SEQ ID NO: 354, P43361.1), HAGE (SEQ ID NO: 355, Q9NXZ2.1), MAGE-A1 (SEQ ID NO: 356, P43355.1), MAGE-C2 (SEQ ID NO: 357, Q9UBF1.1), MAGE-A10 (SEQ ID NO: 358, P43363.1), and MAGE-A12 (SEQ ID NO: 359, P43365.1).

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 PEPs in the same polypeptide are within 50 amino acids of each other). Typically each polypeptide may comprise at least one fragment of a target polypeptide antigen wherein the fragment comprises an amino acid sequence that is a T cell epitope capable of binding to at least three or at least four HLA class II alleles in some subjects or a high proportion of subjects or the maximum proportion of subjects.

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 (Figure 42). 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 PEPs 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 bestEPs in a single treatment (administration of one or more pharmaceutical compositions that together comprise multiple PEPs), can be illustrated by the personalised vaccine polypeptides described in Example 21 below. Exemplary CTA expression probabilities in ovarian cancer are as follows: BAGE: 30%; MAGE A9: 37%; MAGE A4: 34%; MAGE A10: 52%. If patient-A were treated with a vaccine comprising PEPs in only BAGE and MAGE A9, then the probability of having a mAGP (multiple expressed antigens with PEPI) would be 11%. If patient-A were treated with a vaccine comprising only PEPs 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-A's PIT vaccine described in Example 21 contains a further 9 PEPs, 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-B with a

vaccine comprising PEPs in only MAGE C2: 21% and MAGE A1 has a mAGP probability of 7%. Treatment of patient-B with a vaccine comprising PEPs in only SPC1: 38%; MAGE A9 has a mAGP probability of 11%. Treatment of patient-B with a vaccine comprising PEPs in MAGE C2: 21%; MAGE A1: 37%; SPC1: 38%; MAGE A9 has a mAGP probability of 44% (44 > 7 + 11). Patient's PIT vaccine described in Example 21 contains a further 8 PEPs, 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 or more of the cancer associated antigens, or CTAs, such as the CTA discussed above. In some cases each fragment may comprise or consist of a different target epitope having an amino acid sequence selected from SEQ ID NOs: 1-30, SEQ ID NOs: 31-60, SEQ ID NOs: 90 to 119, SEQ ID NOs: 120 to 149, SEQ ID NOs: 178 to 207, SEQ ID NOs: 208 to 237, SEQ ID NOs: 268 to 297, and/or SEQ ID NOs: 298 to 327; 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 20, or 21, or 22, or 23, or 24, or 25, or 26, or 27, or 28, or 29; or SEQ ID NOs: 31 to 32, or to 33, or 34, or 35, or 36, or 37, or 38, or 39, or 40, or 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; or selected from SEQ ID NOs: 90 to 91, or to 92, or 93 or 94, or 95, or 96, or 97, or 98, or 99, or 100, or 101, or 102, or 103, or 104, or 105, or 106, or 107, or 108, or 109, or 110, or 111, or 112, or 113, or 114, or 115, or 116, or 117, or 118; or SEQ ID NOs: 120 to 121, or to 122, or 123, or 124, or 125, or 126, or 127, or 128, or 129; or 130; or 131; or 132; or 133; or 134; or 135; or 136; or 137; or 138; or 139; or 140; or 141; or 142; or 143; or 144; or 145; or 146; or 147; or 148; or 149; or selected from SEQ ID NOs: 178 to 179, or to 180, or 181, or 182, or 183, or 184, or 185, or 186, or 187, or 188, or 189, or 190, or 191, or 192, or 193, or 194, or 195, or 196, or 197, or 198, or 199, or 200, or 201, or 202, or 203, or 204, or 205, or 206; or SEQ ID NOs: 208 to 209, or 210, or 211, or 212, or 213, or 214, or 215, or 216, or 217, or 218, or 219, or 220, or 221, or 222, or 223, or 224, or 225, or 226, or 227, or 228, or 229, or 230, or 231, or 232, or 233, or 234, or 235, or 236; or selected from SEQ ID NOs: 268 to 269, or to 270, or 271, or 272, or 273, or 274, or 275, or 276, or 277, or 278, or 279, or 280, or 281, or

282, or 283, or 284, or 285, or 286, or 287, or 288, or 289, or 290, or 291, or 292, or 293, or 294, or 295, or 296; or selected from SEQ ID NOs: 298 to 299, or to 300, or 301, or 302, or 303, or 304, or 305, or 306, or 307, or 308, or 309, or 310, or 311, or 312, or 313, or 314, or 315, or 316, or 317, or 318, or 319, or 320, or 321, or 322, or 323, or 324, or 325, or 326; or selected from
5 any of these groups of sequences but excluding any specific combinations of sequences that are within 50-60 amino acids of each other in any one or more of the antigens of SEQ ID NOs: 76 to 89, 165 to 177, 253 to 267 and/or 343 to 359, such as any combination of SEQ ID NOs: 3 and 8; SEQ ID NOs: 9 and 10; SEQ ID NOs: 12 and 16; SEQ ID NOs: 13, 18 and 25; SEQ ID NOs: 21 and 24; SEQ ID NOs: 23 and 30; SEQ ID NOs: 93 and 94; SEQ ID NOs: 99, 100 and 102; SEQ
10 ID NOs: 109 and 111; SEQ ID NOs: 96, 101 and 113; SEQ ID NOs: 188 and 190; and/or SEQ ID NOs: 201 and 203; SEQ ID NOs: 268 and 270; SEQ ID NOs: 271 and 281; SEQ ID NOs: 272 and 275; and/or SEQ ID NOs: 288 and 291. In some cases each fragment may comprise or consist of a different amino acid sequence selected from SEQ ID NOs: 60 or 61 to 75; SEQ ID NOs: 150 to 164; SEQ ID NOs: 238 to 252; and/or SEQ ID NOs: 328 to 342.

15 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.

Pharmaceutical Compositions, Methods of Treatment and Modes of Administration

20 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
25 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.

30 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 PEPs. 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. imiquimod 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, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosine, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's 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, 5 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. 10 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.

15 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, 20 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 25 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 30 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

5 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 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

10 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

15 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.

20 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

25 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

30 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, intraperitoneally, intrathecally, intratracheally, intracardially, intralobally, intramedullary, 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, Nivolumab, Pembrolizumab, Atezolizumab and Durvalumab. Examples of chemotherapy agents include alkylating agents including nitrogen mustards such as mechlorethamine (HN2), cyclophosphamide, ifosfamide, melphalan (L-sarcosine) and chlorambucil; anthracyclines; epothilones; nitrosoureas such as carmustine (BCNU), lomustine (CCNU), semustine (methyl-CCNU) and streptozocin (streptozotocin); triazines 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 IFN α , 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
5 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 alphas.

In some cases the method of treatment is a method of vaccination or a method of
10 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
15 one or more PEPs to an individual to elicit a T cell response to recognize and kill cells that display the one or more PEPs on their cell surface in conjunction with a class I HLAs. In another specific embodiment, immunotherapy comprises the administration of one or more PEPs to an individual to elicit a cytotoxic T cell response against cells that display tumor associated antigens (TAAs), tumor specific antigens (TSA) or cancer testis antigens (CTAs)
20 comprising the one or more PEPs 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 PEPs presented by class II HLAs to an individual to elicit a CD4⁺ T helper or CD4⁺ killer cell response to provide co-stimulation to cytotoxic T cells that recognize and kill diseased cells that display the one or more PEPs on their cell surface in conjunction with a class
25 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 PEPs, 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
30 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 gastric cancer, lung cancer, melanoma and/or bladder 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
5 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). The cancer
10 may or may not be one that is associated with or caused by a viral infection and/or viral TAAs.

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
15 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
20 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 CD8+ or 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
25 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
30 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 30. 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 gastric cancer antigens DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1, SSX1, and/or lung cancer antigens BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1, MAGE-A1, and/or melanoma cancer antigens PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, NY-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10, MAGE-A1, and/or bladder cancer antigens PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3,

MAGE-A8 and HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12 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
5 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 different HLA class I molecules that bind to the same sequence T cell epitope of the target antigen. Moreover, as described above, at least two epitopes that binds to at least 3
10 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
15 class I molecules of the subject.

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
20 frequencies in the disease (e.g. MAGE-A3 in a tumor type like gastric cancer, lung cancer, melanoma or bladder 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 PEPs across the population (BestEPs) in multiple frequently expressed target antigens in the
25 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
30 more multiple HLA-binding PEPs in the active ingredient polypeptides; (ii) the presence of PEPs 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 gastric cancer, lung cancer, melanoma or bladder 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 gastric cancer, can be determined by detecting the antigen in tumor, e.g. gastric 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 or AG50 – potency of a vaccine:** The number of antigens in a cancer vaccine that a specific tumor type expresses with 95% or 50% probability. AG95 and AG50 are indicators of the vaccine's potency, and are independent of the immunogenicity of the vaccine antigens. AG95 and AG50 are calculated from the tumor antigen expression rate data. Such data may be obtained from experiments published in peer reviewed scientific journals. Technically, AG95 and AG50 are 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 at least 3 HLAs of a subject and induce T cell responses. PEPI3+ can be determined using the PEPI3+ Test in subjects whose 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: $0 \leq \text{AGP50} \leq \text{AG50}$.

- **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 PEPs 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.

20 Further embodiments of the disclosure – (1A) – Gastric 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: 61 to 75.
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, further comprising at least one additional peptide comprising a fragment of an antigen selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1.

4. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 30.
5. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 31-60.
- 5 6. The pharmaceutical composition of item 1, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
7. The pharmaceutical composition of item 6, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosine, dinitrochlorobenzene
10 (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
8. 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
15 any one of SEQ ID NOs: 61 to 75.
9. The pharmaceutical composition of item 8, 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.
- 20 10. The pharmaceutical composition of item 8, wherein the one or more nucleic acid molecules encode at least one additional peptide comprising a fragment of an antigen selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1.
11. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 30.
- 25 12. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 31 -60.
13. The pharmaceutical composition of item 8, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
- 30 14. The pharmaceutical composition of item 13, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin

(BCG), corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

- 5 15. 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;
 - 10 (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

15 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.
16. The method of item 15, wherein the subject has gastric cancer.
17. The method of item 15, wherein the pharmaceutical composition comprises 2 or more

20 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.
18. The method of item 15, wherein the pharmaceutical composition further comprises comprising at least one additional peptide comprising a fragment of an antigen selected from

25 DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1.
19. The method of item 18, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 30.
20. The method of item 18, wherein the fragment of an antigen comprises an amino acid

30 sequence selected from any one of SEQ ID NOs: 31-60.

21. The method of item 15, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
22. The method of item 21, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG),
5 corynebacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
23. The method of item 15, further comprising administering a chemotherapeutic agent, a
10 checkpoint inhibitor, a targeted therapy, radiation therapy, another immunotherapy, neoadjuvant therapy or combination thereof to the identified subject.
24. The method of item 15, 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
15 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
20 (ii) confirming the subject as likely to have a clinical response to the method of treatment.
25. 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
25 (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
30 (iii) administering the composition of item 1 to the identified subject.
26. 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: 61 to 75; 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: 61 to 75.

27. 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 DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 1 to 30.

Further embodiments of the disclosure – (1B) – Lung 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: 150 to 164.
- 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, further comprising at least one additional peptide comprising a fragment of an antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1.
- 4. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 90 to 119.
- 5. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 120 to 149.
- 6. The pharmaceutical composition of item 1, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
- 7. The pharmaceutical composition of item 6, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin

- (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
- 5 8. 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: 150 to 164.
9. The pharmaceutical composition of item 8, 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
10 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.
10. The pharmaceutical composition of item 8, wherein the one or more nucleic acid molecules encode at least one additional peptide comprising a fragment of an antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-
15 C2, MAGE-A3, KK-LC-1 and MAGE-A1.
11. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 90 to 119.
12. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 120 to 149.
- 20 13. The pharmaceutical composition of item 8, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
14. The pharmaceutical composition of item 13, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene
25 (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
15. 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
30 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.
16. The method of item 15, wherein the subject has lung cancer.
17. The method of item 15, 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.
18. The method of item 15, wherein the pharmaceutical composition further comprises comprising at least one additional peptide comprising a fragment of an antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1.
19. The method of item 18, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 90 to 119.
20. The method of item 18, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 120 to 149.
21. The method of item 15, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
22. The method of item 21, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysollecithin, pluronic

polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

23. The method of item 15, further comprising administering a chemotherapeutic agent, a checkpoint inhibitor, a targeted therapy, radiation therapy, another immunotherapy, neoadjuvant therapy or combination thereof to the identified subject.

24. The method of item 15, 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.

25. 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.

26. 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: 150 to 164; 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: 150 to 164.

27. 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 BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 90 to 119.

Further embodiments of the disclosure – (1C) - melanoma

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: 238 to 252.
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, further comprising at least one additional peptide comprising a fragment of an antigen selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1.
4. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 178 to 207.
5. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 208 to 237.
6. The pharmaceutical composition of item 1, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
7. The pharmaceutical composition of item 6, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

8. 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:61 to 75.
9. The pharmaceutical composition of item 8, wherein the one or more nucleic acid molecules
5 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.
10. The pharmaceutical composition of item 8, wherein the one or more nucleic acid molecules encode at least one additional peptide comprising a fragment of an antigen selected from
10 PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1.
11. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 178 to 207.
12. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an
15 amino acid sequence selected from any one of SEQ ID NOs: 208 to 237.
13. The pharmaceutical composition of item 8, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
14. The pharmaceutical composition of item 13, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin
20 (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosine, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
15. A method of identifying and treating a human subject having cancer who will likely have a
25 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
30 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
- 5 (iv) administering the composition of item 1 to the identified subject.
16. The method of item 15, wherein the subject has melanoma.
17. The method of item 15, 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
- 10 more peptides, or 12 or more peptides.
18. The method of item 15, wherein the pharmaceutical composition further comprises comprising at least one additional peptide comprising a fragment of an antigen selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and
- 15 MAGE-A1.
19. The method of item 18, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs:178 to 207.
20. The method of item 18, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs:208 to 237.
- 20 21. The method of item 15, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
22. The method of item 21, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG),
- 25 corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

23. The method of item 15, further comprising administering a chemotherapeutic agent, a checkpoint inhibitor, a targeted therapy, radiation therapy, another immunotherapy, neoadjuvant therapy or combination thereof to the identified subject.
24. The method of item 15, further comprising prior to the administering step,
- 5 (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
- c. a fragment of a cancer-associated antigen expressed by cancer cells of the subject as determined in step (i); and
- 10 d. 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.
25. A method of identifying and treating a human subject having cancer who will likely have
- 15 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
- 20 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.
26. A kit comprising:
- a. a first pharmaceutical composition comprising one or more peptides, wherein
- 25 each peptide comprises a different one of the amino acid sequence of any one of SEQ ID NOs: 238 to 267; 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: 238 to 267.
- 30 27. 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 PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, NY-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 178 to 207.

5

Further embodiments of the disclosure – (1D) – bladder 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: 328 to 342.
2. The pharmaceutical composition of item 1, comprising 2 or more peptides, 3 or more peptides,
10 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, further comprising at least one additional peptide comprising a fragment of an antigen selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8,
15 HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12.
4. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 268 to 297.
5. The pharmaceutical composition of item 3, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 298 to 327.
- 20 6. The pharmaceutical composition of item 1, further comprising a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.
7. The pharmaceutical composition of item 6, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene
25 (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
8. 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
30 any one of SEQ ID NOs: 328 to 342.

9. The pharmaceutical composition of item 8, 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.
- 5 10. The pharmaceutical composition of item 8, wherein the one or more nucleic acid molecules encode at least one additional peptide comprising a fragment of an antigen selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12.
- 10 11. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 268 to 297.
12. The pharmaceutical composition of item 10, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 298 to 327.
13. The pharmaceutical composition of item 8, further comprising a pharmaceutically acceptable
- 15 adjuvant, diluent, carrier, preservative, or combination thereof.
14. The pharmaceutical composition of item 13, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinisone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant
- 20 (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.
15. 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
- 25 (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;
- 30 (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.

16. The method of item 15, wherein the subject has bladder cancer.

5 17. The method of item 15, 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.

10 18. The method of item 15, wherein the pharmaceutical composition further comprises comprising at least one additional peptide comprising a fragment of an antigen selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12.

15 19. The method of item 18, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 268 to 297.

20. The method of item 18, wherein the fragment of an antigen comprises an amino acid sequence selected from any one of SEQ ID NOs: 298 to 327.

20 21. The method of item 15, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable adjuvant, diluent, carrier, preservative, or combination thereof.

25 22. The method of item 21, wherein the adjuvant is selected from the group consisting of Montanide ISA-51, QS-21, GM-CSF, cyclophosphamide, bacillus Calmette-Guerin (BCG), corynebacterium parvum, levamisole, azimezone, isoprinosone, dinitrochlorobenzene (DNCB), keyhole limpet hemocyanins (KLH), Freund's adjuvant (complete), Freund's adjuvant (incomplete), mineral gels, aluminum hydroxide (Alum), lysolecithin, pluronic polyols, polyanions, oil emulsions, dinitrophenol, diphtheria toxin (DT), and combinations thereof.

30 23. The method of item 15, further comprising administering a chemotherapeutic agent, a checkpoint inhibitor, a targeted therapy, radiation therapy, another immunotherapy, neoadjuvant therapy or combination thereof to the identified subject.

24. The method of item 15, further comprising prior to the administering step,

- (v) 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
 - e. a fragment of a cancer-associated antigen expressed by cancer cells of the subject as determined in step (i); and
 - f. a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and
 - (vi) confirming the subject as likely to have a clinical response to the method of treatment.
- 25. 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.
- 26. 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: 328 to 342; 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: 328 to 342.
- 27. 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 PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12, wherein each

fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 268 to 297.

Further embodiments of the disclosure – (2B) - Lung

- 5 1. A polypeptide that comprises a fragment of up to 50 consecutive amino acids of a lung cancer-associated antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 90 to 119, optionally wherein the fragment is flanked at the N and/or C terminus by additional
10 amino acids that are not part of the sequence of the lung cancer-associated antigen.
2. The polypeptide of item 1, wherein the polypeptide
 - a. is a fragment of a lung cancer-associated antigen selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1, wherein the fragment comprises
15 an amino acid sequence selected from any one of SEQ ID NOs: 90 to 119; or
 - b. comprises or consists of two or more fragments of one or more lung cancer associated antigens selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1, wherein each fragment comprises a different amino acid sequence
20 selected from any one of SEQ ID NOs: 90 to 119, optionally wherein the fragments overlap or are arranged end to end in the polypeptide.
3. The polypeptide according to item 1 or item 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 BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2,
25 SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1; and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 90 to 119.
4. The polypeptide according to any one of items 1 to 3, comprising or consisting of one or more amino acid sequences selected from SEQ ID NOs: 120 to 149.
- 30 5. The polypeptide according to any one of items 1 to 4 comprising or consisting of the amino acid sequence of any one of SEQ ID NOs: 150 to 164.

6. A panel of two or more polypeptides according to any one of items 1 to 5, wherein each polypeptide comprises a different amino acid sequence selected from SEQ ID NOs: 90 to 119.
7. A pharmaceutical composition or kit comprising one or more polypeptides according to any one of items 1 to 5, or a panel of polypeptides according to item 6, or a polypeptide
5 comprising at least two amino acid sequences selected SEQ ID NOs: 90 to 119, or one or more polynucleic acids or vectors encoding said one or more polypeptides.
8. 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 or the peptides, polynucleic acids or vectors of a kit according to item 7.
- 10 9. The method of item 8 that is a method of treating cancer, optionally lung cancer.
10. A method of identifying a human subject who will likely have a cytotoxic T cell response to administration of a pharmaceutical composition or the peptides, polynucleic acids or vectors of a kit according to item 7, the method comprising
- (i) determining that the active ingredient polypeptide(s) of the pharmaceutical
15 composition or kit 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 or the peptides, polynucleic acids or vectors of the kit.
- 20 11. The method of item 10 further comprising using population expression data for each antigen that
- (a) is selected from BRDT, PRAME, NALP4, MAGE-A12, MAGE-A2, SURVIVIN, DPPA2, NY-SAR-35, LDHC, MAGE-C2, MAGE-A3, KK-LC-1 and MAGE-A1; and
- (b) comprises an amino acid sequence that is
- 25 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 CD8+ T cell response that targets
30 one or more polypeptide antigens that are expressed by cancer cells of the subject.

12. A method of identifying a subject who will likely have a clinical response to a method of treatment according to item 9, 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
 - 5 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
 - 10 (ii) identifying the subject as likely to have a clinical response to the method of treatment.
13. A method of determining the likelihood that a specific human subject will have a clinical response to a method of treatment according to item 9, 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;
 - 15 (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
 - 20 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
 - 25 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;
 - 30 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
- 5 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.
14. The method of item 13, wherein the method comprises
- (i) identifying which polypeptide antigens targeted by the active ingredient polypeptide(s) comprise an amino acid sequence that is both
 - 10 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
 - 15 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.
- 20 15. The method of item 14, 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).
16. The method of any one of items 12 to 15 further comprising selecting or recommending administration of the pharmaceutical composition or the peptides, polynucleic acids or vectors of
- 25 the kit as a method of treatment for the subject, and optionally further treating the subject by administering the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit.
17. A method of treatment according to item 9, wherein the subject has been identified as likely to have a clinical response or as having above a threshold minimum likelihood of having a
- 30 clinical response to the treatment by a method according to any one of items 12 to 15.

18. The method of any one of items 8, 9, 16 and 17 wherein the treatment is administered in combination with chemotherapy, targeted therapy or a checkpoint inhibitor.

19. A method of identifying a human subject who will likely not have a clinical response to a method of treatment according to item 9, the method comprising

- 5 (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
- (ii) identifying the subject as likely not to have a clinical response to the method of treatment.

10

Further embodiments of the disclosure – (3B) - Melanoma

1. A polypeptide that comprises a fragment of up to 50 consecutive amino acids of a melanoma-associated antigen selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1,
 - 15 MAGE-A3, MAGE-A10 and MAGE-A1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 178 to 207, 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 melanoma-associated antigen.
2. The polypeptide of item 1, wherein the polypeptide
 - 20 c. is a fragment of a melanoma-associated antigen selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 178 to 207; or
 - 25 d. comprises or consists of two or more fragments of one or more melanoma associated antigens selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 178 to 207,
 - 30 optionally wherein the fragments overlap or are arranged end to end in the polypeptide.

3. The polypeptide according to item 1 or item 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 PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1; and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 178 to 207.
4. The polypeptide according to any one of items 1 to 3, comprising or consisting of one or more amino acid sequences selected from SEQ ID NOs: 208 to 237.
5. The polypeptide according to any one of items 1 to 4 comprising or consisting of the amino acid sequence of any one of SEQ ID NOs: 238 to 252.
6. A panel of two or more polypeptides according to any one of items 1 to 5, wherein each polypeptide comprises a different amino acid sequence selected from SEQ ID NOs: 178 to 207.
7. A pharmaceutical composition or kit comprising one or more polypeptides according to any one of items 1 to 5, or a panel of polypeptides according to item 6, or a polypeptide comprising at least two amino acid sequences selected SEQ ID NOs: 178 to 207, or one or more polynucleic acids or vectors encoding said one or more polypeptides.
8. 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 or the peptides, polynucleic acids or vectors of a kit according to item 7.
9. The method of item 8 that is a method of treating cancer, optionally melanoma.
10. A method of identifying a human subject who will likely have a cytotoxic T cell response to administration of a pharmaceutical composition or the peptides, polynucleic acids or vectors of a kit according to item 7, the method comprising
- (i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition or kit comprise a sequence that 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 to have a cytotoxic T cell response to administration of the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit.
11. The method of item 10 further comprising using population expression data for each antigen that

(a) is selected from PRAME, MAGE-A2, MAGE-C1, SURVIVIN, MAGE-A12, Ny-ESO-1, MAGE-C2, MAGE-A6, BORIS, LAGE-1, MAGE-A11, SSX-1, MAGE-A3, MAGE-A10 and MAGE-A1; and

(b) comprises an amino acid sequence that is

- 5 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 CD8+ T cell response that targets one or more polypeptide antigens that are expressed by cancer cells of the subject.

10

12. A method of identifying a subject who will likely have a clinical response to a method of treatment according to item 9, 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

15

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

20

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

13. A method of determining the likelihood that a specific human subject will have a clinical response to a method of treatment according to item 9, 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;

25

(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

30

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.

14. The method of item 13, wherein the method comprises

(iv) 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;

(v) 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

(vi) 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.

15. The method of item 14, 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).
16. The method of any one of items 12 to 15 further comprising selecting or recommending
5 administration of the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit as a method of treatment for the subject, and optionally further treating the subject by administering the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit.
17. A method of treatment according to item 9, wherein the subject has been identified as
10 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 items 12 to 15.
18. The method of any one of items 8, 9, 16 and 17 wherein the treatment is administered in combination with chemotherapy, targeted therapy or a checkpoint inhibitor.
19. A method of identifying a human subject who will likely not have a clinical response to a
15 method of treatment according to item 9, 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
 - (ii) identifying the subject as likely not to have a clinical response to the method of
20 treatment.

Further embodiments of the disclosure – (4B) - Bladder

1. A polypeptide that comprises a fragment of up to 50 consecutive amino acids of a bladder cancer-associated antigen selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1,
25 NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 268 to 297, 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 bladder cancer-associated antigen.
- 30 2. The polypeptide of item 1, wherein the polypeptide

- e. is a fragment of a bladder cancer-associated antigen selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 268 to 297; or
- 5 f. comprises or consists of two or more fragments of one or more bladder cancer associated antigens selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12, wherein each
- 10 fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 268 to 297, optionally wherein the fragments overlap or are arranged end to end in the polypeptide.
3. The polypeptide according to item 1 or item 2, wherein the polypeptide comprises or consists of fragments of at least two different cancer-associated antigens, wherein the cancer-
- 15 associated antigens are selected from PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-A1, MAGE-C2, MAGE-A10 and MAGE-A12; and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 268 to 297.
4. The polypeptide according to any one of items 1 to 3, comprising or consisting of one or
- 20 more amino acid sequences selected from SEQ ID NOs: 298 to 327.
5. The polypeptide according to any one of items 1 to 4 comprising or consisting of the amino acid sequence of any one of SEQ ID NOs: 328 to 342.
6. A panel of two or more polypeptides according to any one of items 1 to 5, wherein each polypeptide comprises a different amino acid sequence selected from SEQ ID NOs: 268 to 297.
- 25 7. A pharmaceutical composition or kit comprising one or more polypeptides according to any one of items 1 to 5, or a panel of polypeptides according to item 6, or a polypeptide comprising at least two amino acid sequences selected SEQ ID NOs: 268 to 297, or one or more polynucleic acids or vectors encoding said one or more polypeptides.
8. A method of vaccination, providing immunotherapy or inducing a cytotoxic T cell
- 30 response in a subject, the method comprising administering to the subject a pharmaceutical composition or the peptides, polynucleic acids or vectors of a kit according to item 7.

9. The method of item 8 that is a method of treating cancer, optionally bladder cancer.
10. A method of identifying a human subject who will likely have a cytotoxic T cell response to administration of a pharmaceutical composition or the peptides, polynucleic acids or vectors of a kit according to item 7, the method comprising
- 5 (i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition or kit comprise a sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and
- (iv) identifying the subject as likely to have a cytotoxic T cell response to administration of the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit.
- 10 11. The method of item 10 further comprising using population expression data for each antigen that
- (a) is selected PIWIL2, CTAGE1, MAGE-A9, EpCAM, OY-TES-1, NY-ESO-1, SURVIVIN, MAGE-C1, MAGE-A2, LAGE-1, MAGE-A3, MAGE-A8, HAGE, MAGE-
- 15 A1, MAGE-C2, MAGE-A10 and MAGE-A12; 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
- 20 the subject;
- to determine the likelihood that the subject will have a CD8⁺ T cell response that targets one or more polypeptide antigens that are expressed by cancer cells of the subject.
12. A method of identifying a subject who will likely have a clinical response to a method of treatment according to item 9, the method comprising
- 25 (i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition comprise two or more different amino acid sequences each of which is
- e. a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and
- f. a fragment of a cancer-associated antigen expressed by cancer cells of the subject,
- 30 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.

13. A method of determining the likelihood that a specific human subject will have a clinical response to a method of treatment according to item 9, wherein one or more of the following factors corresponds to a higher likelihood of a clinical response:

- 5 (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
 - 10 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;
- 15 (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
 - 20 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
 - 25 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.
14. The method of item 13, wherein the method comprises
 - (vii) identifying which polypeptide antigens targeted by the active ingredient
 - 30 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;
- (viii) 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
- (ix) 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.
15. The method of item 14, 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).
16. The method of any one of items 12 to 15 further comprising selecting or recommending administration of the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit as a method of treatment for the subject, and optionally further treating the subject by administering the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit.
17. A method of treatment according to item 9, 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 items 12 to 15.
18. The method of any one of items 8, 9, 16 and 17 wherein the treatment is administered in combination with chemotherapy, targeted therapy or a checkpoint inhibitor.
19. A method of identifying a human subject who will likely not have a clinical response to a method of treatment according to item 9, 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
- (ii) identifying the subject as likely not to have a clinical response to the method of treatment.

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 class 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.

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

The accuracy of the prediction of multiple HLA binding epitopes was also determined (Table 3). 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.

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

This study investigates whether the presentation of one or more epitopes of a polypeptide antigen by one or more HLA class I molecule of an individual is predictive for a CTL response.

The study was carried out by retrospective analysis of six clinical trials, conducted on 71
10 cancer patients and 9 HIV-infected patients (Table 4). 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
15 cases, correlation between CTL responses and clinical responses were reported.

No patient was excluded from the retrospective 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
20 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
1	VGX-3100	HPV16-E6 HPV16-E7	Cervical cancer	17/18	5 x 17	IFN- γ ELISPOT	High Resolution SBT

		HPV18-E6 HPV18-E7 HPV16/18					
2	HIVIS vaccine	HIV-1 Gag HIV-1 RT	AIDS	9/12	2 x 9	IFN- γ ELISPOT	Low-Medium Resolution SSO
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
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	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	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
Total	6	7		80	157		

The reported CD8+ T cell responses of the training dataset were compared with the HLA class I restriction profile of epitopes (9 mers) of the vaccine antigens. The antigen sequences and the HLA class 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 CD8+ T cell response data where CD8+ T cells are IFN- γ producing CTL specific for vaccine peptides (9 mers). 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 the CD8+ T cell response or 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 alleles 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% true positive rate (diagnostic sensitivity) (Table 6).

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 –Retrospective Validation of the ≥ 1 PEPI3+ threshold as novel biomarker for PEPI test

In a retrospective analysis, the test cohort of 81 datasets from 51 patients was used to validate the ≥ 1 PEPI3+ threshold to predict an antigen-specific CD8+ T cell response or 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 CD8+ T cell responses (CTL responses) reported from the clinical trials (Table 7).

The retrospective validation demonstrated that a PEPI3+ peptide induces CD8+ T cell response (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)	The likelihood that an individual that meets the ≥ 1 PEPI3+ threshold has antigen-specific CTL responses after treatment with immunotherapy.	84%
Sensitivity	The proportion of subjects with antigen-specific CTL responses after treatment with immunotherapy who meet the ≥ 1 PEPI3+ threshold.	75%
Specificity	The proportion of subjects without antigen-specific CTL responses after treatment with	55%

		immunotherapy who do not meet the ≥ 1 PEPI3+ threshold.	
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 value.

- 5 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

10 **Example 4 – Clinical Validation of the ≥ 1 PEPI3+ threshold as novel biomarker for PEPI test**

- 15 The PEPI3+ biomarker-based vaccine design has been tested first time in a phase I clinical trial in metastatic colorectal cancer (mCRC) patients in the OBERTO phase I/II clinical trial (NCT03391232). In this study, we evaluated the safety, tolerability and immunogenicity of a single or multiple dose(s) of PolyPEPI1018 as an add-on to maintenance therapy in subjects with mCRC. PolyPEPI1018 is a peptide vaccine containing 12 unique epitopes derived from 7 conserved TSAs frequently expressed in mCRC (WO2018158455 A1). These epitopes were designed to bind to at least three autologous HLA alleles that are more likely to induce T-cell responses than epitopes presented by a single HLA (See Examples 2 & 3). mCRC patients in the first line setting received

the vaccine (dose: 0.2 mg/peptide) just after the transition to maintenance therapy with a fluoropyrimidine and bevacizumab. Vaccine-specific T-cell responses were first predicted by identification of PEPI3+-s *in silico* (using the patient's complete HLA genotype and antigen expression rate specifically for CRC) and then measured by ELISpot after one cycle of vaccination (phase I part of the trial).

Seventy datasets from 10 patients (Phase 1 cohort and dataset of OBERTO trial) was used to prospectively validate that PEPI3+ biomarker predicts antigen-specific CTL responses. For each dataset, predicted PEPI3+-s were determined *in silico* and compared to the vaccine-specific immune responses measured by ELISPOT assay from the patients' blood. Diagnostic characteristics (positive predictive value, negative predictive value, overall percent agreement) determined this way were then compared with the retrospective validation results described in Example 3.

The overall percent agreement was 64%, with high positive predictive value of 79%, representing 79% probability that the patient with predicted PEPI3+ will produce CD8 T cell specific immune response against the analyzed antigen. Clinical trial data were significantly correlated with the retrospective trial results ($p=0.01$) and provides evidence for the PEPI3+ calculation with PEPI test to predict antigen-specific T cell responses based on the complete HLA-genotype of patients (Table 9).

Table 9. Prospective validation of the ≥ 1 PEPI3+ and PEPI test

Parameter	Definition	Retrospective validation n = 81*	Prospective validation (OBERTO) n = 70**
PPV Positive Predictive Value	The likelihood that an individual with a positive PEPI test result has antigen-specific T cell responses	84%	79%
NPV Negative Predictive Value	The likelihood that an individual with a negative PEPI test result does not have antigen-specific T cell responses	42%	51%
OPA Overall Percent Agreement	The percentage of results that are true results, whether positive or negative	70%	64%
Fisher's exact probability test (p)		0.01	0.01

*51 patients; 6 clinical trials; 81 dataset **10 patients; Treos phase I clinical trial (OBERTO); 70 datasets

Example 5 – The ≥ 1 PEPI3+ test predicts CD8+ T cell reactivities

Supporting data were obtained to show that the ≥ 1 PEPI3+ correlates with clinical immunogenicity data but the state-of-art mono-HLA specific epitope determination does not show correlation with vaccine-specific immunogenicity.

The ≥ 1 PEPI3+ calculation was compared with a state-of-art 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 HPV-16 synthetic long peptide vaccine (LPV) in two different clinical trials were determined from DNA samples. 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 M. J. 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* 14, 178-187 (2008)., G. G. Kenter, et al. Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia. *N Engl J Med* 361, 1838-1847 (2009). M. J. Welters, et al. Success or failure of vaccination for HPV16-positive vulvar lesions correlates with kinetics and phenotype of induced T-cell responses. *Proc Natl Acad Sci U S A* 107, 11895-11899 (2010). The publications also report the T cell responses of each vaccinated patient to pools of overlapping peptides of the vaccine. 25 (20 having VIN-3 and 5 having cervical cancer) patients had immune response data available, and 25 had clinical response data available.

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 CD8+ T cell response. Peptides that comprised no PEPI3+ were predicted not to induce a CD8+ T cell response.

The ≥ 1 PEPI3+ threshold correctly predicted 529 out of 555 negative CD8+ T cell responses (95% true negative (TN) rate) and 9 out of 45 positive CD8+ T cell responses (20% true positive (TP) rate) measured after vaccination (FIG. 3A). Overall, the agreement between

the ≥ 1 PEPI3+ threshold 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. Forty-two HLA class I-binding epitopes predicted 45 CD8+ T cell responses (93% TP rate). In contrast, of the 555 negative T cell responses, only 105 were ruled out by HLA-binding epitopes (19% TN rate) (FIG. 3B). Overall, the agreement between a single HLA class I allele-binding epitope and CD8+ T cell response was 25%, which was not statistically significant.

Example 6 - 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 5) were investigated for CD4+ T helper responses following LPV vaccination (FIG. 4). The TP rate of the prediction of HLA class II restricted epitopes was 95%, since the State of Art tool predicted 112 positive responses (positive CD4+ T cell reactivity to a peptide pool for a person's HLA class II alleles) out of 117. The TN rate was 0% since it could rule out 0 of 33 negative T cell responses. Overall, the agreement between HLA-restricted class II PEPI prediction and CD4+ T cell reactivity was 75% (not significant).

The HLA class II-binding PEPI3+-s predicted 86 of 117 positive CD4+ T- cell responses (73% TP rate) and ruled out 17 of 33 negative T cell responses (52% TN rate). Overall, the agreement between HLA class II PEPI3+-s and CD4+ T-cell response was 69% ($p = 0.005$) (FIG. 4A).

Example 7 - The ≥ 1 PEPI3+ test predicts T cell responses to full length LPV polypeptides

Using the same studies reported in Examples 5 and 6, 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 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 70%, PPV 85%)

and of 2 out of 5 cervical cancer patients (sensitivity 40%, PPV 100%) (FIG. 5A). The CD4+ T cell reactivities (PEPI3+) were correctly predicted 100% both of VIN-3 and cervical cancer patients (FIG. 5B).

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. There was also a correlation between the number of PEPIs and clinical response in VIN-3 patients for HLA class II PEPIs but not HLA class I PEPIs, confirming the post-hoc analysis results from the clinical trial (FIG. 5C and 5D).

Example 8 – Case Study, PEPI3+ correlation with vaccine-specific immunogenicity

“Vaccine-1” is an HPV16 based DNA vaccine containing full length E6 and E7 antigens with a linker in between. “Vaccine-2” is an HPV18 based DNA vaccine containing full length E6 and E7 antigens with a linker in between (FIG 6A). A Phase II clinical trial investigated the cell responses of 17 HPV-infected patients with cervical cancer who were vaccinated with both Vaccine-1” and “Vaccine-2” (“Vaccine-3” vaccination, Bagarazzi *et al.* Science Translational Medicine. 2012; 4(155):155ra138.).

FIG. 6B 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 (FIG. 6C).

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 9 – 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 10) (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 10. 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

5 **Example 10 – *In silico* trial based on the identification of multiple HLA binding epitopes in a multi-peptide vaccine IMA901 predict the reported clinical trial immune response rate**

Probability of targeting multiple antigens in the tumor of RCC patients

IMA901 is a therapeutic vaccine for renal cell cancer (RCC) comprising 9 peptides
 10 derived from tumor-associated antigens (TUMAPs). It was demonstrated that TUMAPs are naturally presented in human cancer tissue, they are overexpressed antigens shared by a subset of patients with the given cancer entity (Table 11). We estimated the probability that a TSA is expressed in a subject treated with IMA901 vaccine using available data from the scientific literature (FIG. 7). We used the Bayesian convention assuming that the expression probabilities
 15 follow a Beta-distribution.

We defined AG50 as the number of TSAs (AG) in the cancer vaccine that a specific tumor type expresses with 50% probability. The AG50 modelling of cancer vaccines assumes that each AG produces an effect proportional to the expression rate of the AG in the tumor type (if each AG in the vaccine is immunogenic).

20 For IMA901 vaccine targeting 9 antigens (9 TUMAPs), the AG50 value is 4.7, meaning that about half of the antigens are overexpressed in 50% of patients' tumor. Moreover, the probability of targeting 2 expressed antigens is 100% and 3 antigens is 96%. These results suggest high potency of IMA901 vaccine based on target antigen selection.

Table 11. Overexpression of TAAs in RCC tumors selected for IMA901 vaccine

TAA (AG)	Published expression rate in RCC tumors*	Estimated expression rate (95% CI)
ADF-001	5/11 ¹	46% (21%, 72%)

ADF-002	5/11	46% (21%, 72%)
APO-001	9/11 ¹	77% (52%, 95%)
CCN-001	4/11	38% (15%, 65%)
GUC-001	0/2 ²	25% (1%, 71%)
K67-001	2/2	75% (29%, 99%)
MET-001	11/11	92% (74%, 100%)
MUC-001	0/11	8% (0%, 26%)
RGS-001	7/11	62% (35%, 85%)

*expression is defined as overexpression in tumors compared to healthy tissues provided in the source publications

¹Walter S et al, Multi-peptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival, *Nature Medicine*, (2012), 18, 1254–1261

²Krüger T et al, Lessons to be learned from primary renal cell carcinomas, *Cancer Immunol, Immunother*, 2005, 54, 826-836

Probability of inducing immune responses against multiple antigens in the tumor of RCC patients

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) (Walter S et al, Multi-peptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival, *Nature Medicine*, (2012), 18, 1254–1261). Each of the 9 peptides in IMA901 have been identified as HLA-A*02-restricted epitopes. Based on currently accepted standards, all 9 peptides are strong candidates to boost T cell responses against renal cancer since their presence has been detected in renal cancer patients, and because the trial patients were specifically selected to have at least one HLA molecule (HLA-A*02) capable of presenting each of the peptides. Despite this restriction the immune response rate of the phase I and phase II clinical trials measured for at least one peptide of the vaccine was 74% and 64%, respectively.

We analyzed by *in silico* prediction the HLA binding properties of each TUMAP in IMA901 and found that 8 out of the 9 TUMAPs can bind to many HLA-A*02 alleles confirming the identification process (FIG. 8). However, we found that each TUMAP can bind to many other HLA-B* and HLA-C* alleles (FIG. 8A).

Since the complete 4-digit HLA genotype of subjects who participated in IMA901 clinical trials were not available, we used the genotype data of 51 HLA-A*02 selected RCC subjects from another clinical trial, to characterize the immunogenicity of IMA901 vaccine (REF: Chowell D, Morris LGT, Grigg CM, Weber JK, Samstein RM, et al. Patient HLA class I

genotype influences cancer response to checkpoint blockade immunotherapy. Science. 2018; 359 (6375): 582-587.). As presented on FIG. 8B, only few TUMAPs are able to bind to multiple HLAs of the same subject. The most immunogenic peptide in this context turned to be MET-001 capable of generating PEPI in 35% of RCC patients. However, CCN-001 could not generate
 5 PEPI in any of the patients, in agreement with FIG. 8A; CCN-001 can bind only to HLA-A*02 alleles. Based on FIG. 8A, MUC-001 is theoretically able to bind other alleles, too (both HLA-B and HLA-C), however those alleles were not present in the patients of our model population, therefore this peptide could not generate PEPI, either.

The immunogenicity of IMA901 vaccine determined in the 2 clinical trials was compared
 10 with the PEPI response rate determined using the PEPI test in our RCC model population. We found 67% (CI95 53-78%) immune response to at least one peptide of the IMA901 vaccine. According to PEPI test, 33% (CI95 22-47%) of these HLA-A*02+ subjects did not have 3 HLAs binding to any TUMAPs. Interestingly, IMA901 did not induce T cell responses in 25% and 36% of HLA-A*02 selected subjects in the Phase I and Phase II clinical trials, respectively.

15 Furthermore, PEPI test predicted 30% (CI95 19-43%) of subjects with 1 PEPI to one TUMAP, and 37% (CI95 25-51%) have ≥ 2 PEPIs to at least two IMA901 peptides, which is in agreement with the average 40% and 27% immune response to 1 or ≥ 2 TUMAPs in both clinical trials (Table 12). The differences between the immunogenicity found in the 3 cohorts can be explained by the differences in the HLA genotype of the study subjects as well as the potential errors in
 20 measuring T cell responses and in determining PEPIs with the PEPI test (see Example 1). The phase I and phase II study results show the variability of the immune response rates of the same vaccine in different trial cohorts. However, the agreements between PEPI response rates and immunogenicity of peptide vaccines are determined by the host HLA sequences.

**Table 12. Immunogenicity of IMA901 vaccine is determined by the host HLA genotype (multiple
 25 HLAs)**

Immune responses to TUMAPs	Phase I (n=27)*	Phase II (n=61)*	Phase I + II (n=88)	RCC model population (n=51)** (CI95%)
No peptide	25%	36%	33%	33% (22-47%)
≥ 1 peptide	74%	64%	67%	67% (53-78%)
1 peptide	44%	38%	40%	30% (19-43%)
≥ 2 peptides	29%	26%	27%	37% (25-51%)

*reported immunological data for the trials conducted with IMA901 vaccine (REF: Walter Nat Med 2012);

**Predicted by PEPI test

Similarly to the AG50, we defined AP50 as the average number of antigens with PEPI of a vaccine which shows how the vaccine can induce immune response against the antigens targeted by the composition (cancer vaccine specific immune response). AP, therefore is depending of the HLA heterogeneity of the analyzed population and is independent on the expression of the antigen on the tumor. The IMA901 composition can induce immune response against an average of 1.06 vaccine antigens (AP50=1.06) meaning that in the HLA-A*02 selected RCC model population it can induce immune response against at least one vaccine antigen. This result is far less compared to the designed intention of immunogenicity (HLA-matched patients treated with 9 peptides).

Comparison of Immunogenicity and Clinical Response of TUMAPs in the IMA901 Peptide Vaccine

An immune response induced by a vaccine against a single antigen might not be sufficient for clinical activity, as the given antigen might not be expressed in the patient. Therefore, we defined AGP as the immune response which targets an expressed antigen, taking into account both the immunogenicity and expression probability of the vaccine antigen on the tumor, presented above. AGP depends on the antigen (AG) expression rate in the indicated tumor and the HLA genotype of subjects capable to make PEPI (P) in the study population.

Therefore, we investigated the correlation between immune responses against different number of antigens (TUMAPs) and the immune responses against likely expressed antigens (AGP). We found that an immune response elicited by one peptide (1 TUMAP) corresponds to 0.98 AGP, meaning that there is 98% probability that the immune response induced by any peptide of the IMA901 vaccine will target an expressed antigen on the tumor (FIG. 9). However, immune responses elicited by 2 or 3 TUMAPs correspond to only 1.44 and 2.21 AGP, respectively. 0.35 AGP corresponding to 0 TUMAP indicates the cumulated error of PEPI test prediction (see Example 1).

To characterize the potency of a cancer vaccine we defined AGP50, a parameter showing the number of antigens that the vaccine induced CTLs can recognize in a tumor with 50%

probability. The computation is similar to the AG50 but in addition to the expression, the occurrence of the PEPI presentation on certain vaccine antigen is also considered. AGP50 for IMA901 vaccine for the RCC model population is 1.10.

In a retrospective analysis, IMA901 clinical trial investigators found that significantly more subjects who responded to multiple TUMAPs of IMA901 experienced disease control (DC, stable disease or partial response) compared with subjects who had no response or responded to only 1 TUMAP (Table 13). Since the presence of PEPIs accurately predicted the responders to TUMAPs, we investigated the relationship between disease control rate in the TUMAP responder subpopulation and AGP. Similarly, to the investigators we analyzed the percentage of patients who are likely to have immune response against an expressed antigen (i.e.: ≥ 1 AGP) for the subpopulations predicted to have immune response to 0, 1 or 2 TUMAPs using our RCC model population. Interestingly, percentage of patients with 1 AGP is similar to the percentage of patients with disease control in the subpopulations: i.e.: 33% of patients had disease control vs 47% (CI95 23-67%) had 1 AGP and considerably more patients had disease control and AGP in the subgroup with immune response to 2 TUMAPs 75% vs 90 % (CI95 70, 97%), respectively. These results suggest that only those patients are likely to experience clinical benefit, who have immune response against at least one expressed tumor antigen. Moreover, the percentage of patients with 1 AGP in our RCC model population is similar to the disease control rate of the phase I and phase II trials conducted with IMA901 vaccine (Table 12).

Table 13. Correlation between clinical benefit and AGP

Subpopulation	% of pts with DC in the clinical subpopulation	% pts with 1 AGP in the model subpopulation (CI95)
No IR	14%	5% (0%,18%)
IR to 1TUMAP	33%	47% (23%,67%)
IR to ≥ 2 TUMAPs	75%	90% (70%,97%)
Phase I	40%	49% (35%,61%)
Phase II	31%	

Analysis of IMA901 vaccine potency in multiple populations

As shown in Table 14, AG50 value of 4.7 was observed for IMA901 vaccine, suggesting high potency based on target antigen selection. However, AP50 for IMA901 in both the unselected general population and HLA-A*02 selected subjects were only 0.75 and 1.12,

respectively. Similar results were obtained for unselected RCC model population and HLA-A*02 selected populations. This results demonstrate that HLA-A*02 enrichment improved the antigenicity of IMA901, however did not ensure the immunogenicity of the vaccine. Consequently, the AGP50 values describing the potency of the vaccine are low in each population.

Table 14. Potency of IMA901 vaccine in in unselected population and HLA-A*02 selected subjects

Model Population	AG50	AP50	AGP50
All Subjects (n=433)	4.7	0.75	0.49
HLA-A*02 Subjects (n=180)	4.7	1.12	0.81
RCC population (n=129)	4.7	0.61	0.70
RCC subpopulation A*02 (n=51)	4.7	1.06	1.10

Example 11 – *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 9, may be used to predict CTL reactivity rates of vaccines, i.e. used in an *in silico* efficacy trial and to determine the correlation between the clinical outcome of vaccine trials and PEPI.

Published clinical trial results were collected from studies with therapeutic vaccines, which included 1,790 subjects in 64 clinical studies, treated with 42 therapeutic vaccines covering 61 different antigens (Table 15). The same vaccines used in those clinical trials were used to perform *in silico* trials with the model population of 433 human leukocyte antigen (HLA)-genotyped subjects (described in Example 9). No subjects were excluded for reasons other than data availability. IRR was defined as the proportion of subjects in the study population with T cell responses induced by the study vaccine. ORR was defined as the proportion of subjects in the study population with objective response (complete and partial response) after vaccination. The proportion of subjects with PEPIs (personal epitopes that bind to 3 HLA alleles of a subject), multiple PEPIs, and PEPIs in multiple antigens were computed in the *in silico* trials

to obtain the PEPI Score, MultiPEPI Score, and MultiAgPEPI Score, respectively. The immune and objective response rates (IRR and ORR) from the published clinical trials were compared with the PEPI Score, MultiPEPI Score, and MultiAgPEPI Score. All reported and calculated scores are summarized in Table 16.

5 **Table 15. Summary of patient demographics in the published clinical trials**

Characteristic	Count	Percentage
Total subjects	1,790	
Total studies	64	
Subjects with HIV infection	12	1%
Subjects with neoplasia or dysplasia	172	9%
Subjects with cancer	1606	90%
Subjects with solid tumors	1503	84%
Subjects with liquid tumors	103	6%
Subjects with metastatic tumors	788	44%
Subjects with non-metastatic tumors	818	46%
HLA selected subjects	918	51%
Non- HLA selected subjects	872	49%
Trials with HLA selected subjects	32	50%
Trials without HLA selected subjects	32	50%

Table 16. Response rates and PEPI Scores

Immunotherapy	IRR	ORR	PEPI Score	Multi PEPI Score	MultiAg PEPI Score
PSMA-Survivin pulsed DC vaccine	---	18%	3%	0%	0%
Peptide vaccine	---	3%	10%	0%	0%
HPV-SLP	83%	60%	73%	70%	34%
	100%	60%	73%	70%	34%
VGX-3100	78%	50%	87%	56%	64%
Melanoma peptide vaccine	52%	12%	42%	6%	6%
GAA peptides vaccine	55%	15%	18%	0%	0%
KRM-20 peptide vaccine	40%	13%	36%	15%	15%
Peptide vaccine	100%	25%	81%	3%	1%

S-288310 peptide vaccine	67%	17%	44%	8%	8%
KIF20A-66 peptide	70%	26%	38%	7%	---
PepCan	65%	52%	62%	26%	---
Iplilimumab (NYESO-1 specific response)	72%	25%	84%	65%	---
p53 SLP70-248	88%	0%	77%	52%	---
	100%	0%	77%	52%	---
	0%	---	77%	52%	---
p53 SLP70-235	21%	---	75%	52%	---
GVX301	64%	0%	65%	7%	---
TroVax vaccine (OXB-301)	65%	0%	94%	83%	---
	57%	0%	94%	83%	---
StimuVax	21%	---	2%	---	---
IMA901	74%	---	48%	27%	27%
	64%	---	48%	27%	27%
ICT107	33%	---	52%	---	---
ProstVac	67%	---	50%	23%	---
	45%	---	50%	23%	---
	76%	---	50%	23%	---
	67%	---	50%	23%	---
	50%	---	50%	23%	---
	72%	---	50%	23%	---
Synchrotope TA2M	46%	---	24%	7%	---
MELITAC 12·1	49%	---	47%	19%	---
HIVIS	50%	---	88%	---	---
	80%	---	93%	---	---
ImMucin	90%	---	95%	70%	---
	100%	47%	95%	70%	---
NY-ESO-1 OLP	71%	---	84%	65%	---
	82%	0%	84%	65%	---
WT1 vaccine	83%	---	80%	77%	---
WT1 peptide vaccine	72%	6%	86%	---	---
RHAMM-R3 peptide vaccine	44%	0%	0%	---	---
GMMG-MM5 peptides	35%	---	86%	21%	21%
INGN-225 p53 vaccine	58%	4%	82%	61%	---
HR2822	8%	---	3%	---	---
GV1001	17%	---	3%	---	---
	45%	---	3%	---	---
Vx-001	51%	---	33%	---	---
	66%	7%	33%	---	---
	58%	4%	33%	---	---
	71%	0%	33%	---	---
NY-ESO-1f	90%	0%	55%	18%	---

GL-0817 (MAGE-A3 Trojan)	33%	---	29%	3%	---
	57%	0%	29%	3%	---
	0%	0%	29%	3%	---
DPX0907 (per peptide)	0%	---	22%	---	---
	11%	---	18%	---	---
	11%	---	7%	---	---
	11%	---	39%	---	---
	17%	---	12%	---	---
	17%	---	5%	---	---
	22%	---	31%	---	---
CV9103 mRNA vaccine	80%	---	100%	---	---
TG4010 vaccine	38%	13%	43%	6%	---
	26%	0%	43%	6%	---
	21%	---	43%	6%	---
	---	0%	---	---	---
SVN-2B peptide vaccine	60%	---	35%	---	---
TSPP peptide vaccine	---	5%	72%	31%	---
Her2/neu peptide vaccine (p369)	62%	---	4%	---	---
Her2/neu peptide vaccine (p688)	31%	---	1%	---	---
Her2/neu peptide vaccine (p971)	54%	---	0%	---	---
MART-1 Peptide Vaccine	15%	---	0%	---	---

We investigated the correlation between ≥ 1 PEPI3+ Score and immune response rate in a previous study of 12 peptide vaccines derived from cancer antigens that induced T cell responses in a subpopulation of 172 subjects from 19 clinical trials, that were identified from peer reviewed publications. The experimentally determined response rates reported from the trials were compared with the ≥ 1 PEPI3+ Scores and linear correlation between ≥ 1 PEPI3+ Score and response rate ($R^2 = 0.70$) was found ($p=0.001$) (FIG. 10A). The correlation between ≥ 1 PEPI3+ Score and immune response rate was then confirmed by the analysis of 59 clinical trials involving 1,343 subjects who were treated with 40 different vaccines. Each vaccine was analyzed by comparing the published IRR from the clinical trial to the PEPI Score from the model population. (FIG. 10B). The correlation between the IRR and PEPI Score was significant ($r^2=0.465$ and $p=0.001$). This result demonstrated that the PEPI Score determined by *in silico* trials in the MP accurately predicts the IRRs observed in clinical trials.

To test whether polyclonal T cell response increases the likelihood of tumor shrinkage, ORR and MultiPEPI Score were compared. Preliminary experiments analyzed the relationship

between clinical response (either ORR or DCR) and MultiPEPI Score in 17 clinical trials conducted with peptide- and DNA-based immunotherapy vaccines. The results from these experiments demonstrated a significant correlation between clinical response rate and MultiPEPI Score ($r^2 = 0.75$, $p < 0.001$). To confirm these findings, ORR data from 27 clinical trials with 21
5 different vaccines, involving 600 subjects, were collected and analyzed (FIG. 10C). The MultiPEPI Score was calculated as the percentage of subjects in the model population with multiple PEPIs from the study vaccine. The results from this experiment demonstrated that ORR does not correlate with MultiPEPI Score (FIG. 10D).

Results from previous studies suggested that T cell responses against multiple antigens
10 were associated with longer progression free- and overall survival. Consequently, we hypothesized that the induction of T cell responses against multiple tumor antigens increases the likelihood of tumor shrinkage. To test this hypothesis, ORR data from 10 clinical trials conducted with 9 different vaccines, involving 263 subjects, that were treated with multiantigen-targeting vaccine were collected and analyzed. The MultiAg PEPI Score was calculated as the
15 percentage of subjects with vaccine-specific PEPIs on at least two antigens. The results from this experiment demonstrated a significant correlation between ORR and MultiAg PEPI Score ($r^2 = 0.64$; $p = 0.01$), and ORR and MultiPEPI Score ($r^2 = 0.88$ and $p = 0.001$) (FIG. 10E and F, respectively). These results suggest that T cell responses against multiple tumor antigens may recognize a larger tumor cell population, thereby increasing the likelihood of tumor shrinkage.

20 The next analysis explored whether PEPI-specific T cell responses against antigens expressed in the tumor of interest, increase the likelihood of tumor shrinkage. A total of 15 clinical trials enrolled subjects with target antigen positive disease and 11 clinical trials had no subject preselection based on antigen expression. The proportion of subjects with objective response was significantly higher in CTs with target antigen-positive subjects compared with
25 CTs without target pre-selection (21.0% vs 3.6%, respectively, $p = 0.03$).

The correlation between ORR and MultiPEPI Score was statistically significant in subjects with confirmed expression of target antigens ($r^2 = 0.56$, $p = 0.005$) (FIG. 10G). These results emphasize the importance of the presence of cognate PEPI in the tumor, and also that the presence of the cognate PEPI in the tumor increases the likelihood of tumor shrinkage.

This study demonstrated that the link between a subject's HLA genotype and PEPI is the most important factor in predicting clinical response to a vaccine. This study also showed that the PEPI Score can predict the clinical outcome of therapeutic vaccines.

5 **Example 12 - Study design of OBERTO Phase I/II Clinical Trial and preliminary safety data**

OBERTO trial is a Phase I/II trial of PolyPEPI1018 Vaccine and CDx for the Treatment of Metastatic Colorectal Cancer (NCT03391232). Study design is shown in FIG. 11.

10 Enrollment criteria

- Histologically confirmed metastatic adenocarcinoma originating from the colon or the rectum
- Presence of at least 1 measurable reference lesion according to RECIST 1.1
- PR or stable disease during first-line treatment with a systemic chemotherapy regimen and 1 biological therapy regimen
- 15 • Maintenance therapy with a fluoropyrimidine (5-fluorouracil or capecitabine) plus the same biologic agent (bevacizumab, cetuximab or panitumumab) used during induction, scheduled to initiate prior to the first day of treatment with the study drug
- Last CT scan at 3 weeks or less before the first day of treatment

Subject Withdrawal and Discontinuation.

20

- During the initial study period (12W), if a patient experiences disease progression and needs to start a second-line therapy, the patient will be withdrawn from the study.
- During the second part of the study (after 2nd dose) if a patient experiences disease progression and needs to start a second-line therapy, the patient will remain in the study, receive the third
- 25 vaccination as scheduled and complete follow-up.

- Transient local erythema and edema at the site of vaccination were observed as expected, as well as a flu-like syndrome with minor fever and fatigue. These reactions are already well-known for peptide vaccination and usually are associated with the mechanism of action,

because fever and flu-like syndrome might be the consequence and sign for the induction of immune responses (this is known as typical vaccine reactions for childhood vaccinations).

- Only one serious adverse event (SAE) “possibly related” to the vaccine was recorded (Table 17).
- 5 • One dose limiting toxicity (DLT) not related to the vaccine occurred (syncope).

Safety results are summarized in Table 17.

Table 17. Serious adverse events reported in the OBERTO clinical trial. No related SAE occurred (only 1 “possibly related”).

Patient ID	SAE	Relatedness
010001	Death due to disease progression	Unrelated
010004	Embolism	Unlikely Related
010004	Abdominal pain	Unrelated
010007	Bowel Obstruction	Unrelated
020004	Non-Infectious Acute Encephalitis	Possibly Related

10

Example 13 – Expression frequency based target antigen selection during vaccine design and it’s clinical validation for mCRC

Shared tumor antigens enable precise targeting of all tumor types – including the ones with low mutational burden. Population expression data collected previously from 2,391 CRC biopsies represents the variability of antigen expression in CRC patients worldwide (FIG. 12A).

15

PolyPEPI1018 is a peptide vaccine we designed to contain 12 unique epitopes derived from 7 conserved testis specific antigens (TSAs) frequently expressed in mCRC. In our model we supposed, that by selecting the TSA frequently expressed in CRC, the target identification will be correct and will eliminate the need for tumor biopsy. We have calculated that the probability of 3 out of 7 TSAs being expressed in each tumor is greater than 95%. (FIG. 12B)

20

In a phase I study we evaluated the safety, tolerability and immunogenicity of PolyPEPI1018 as an add-on to maintenance therapy in subjects with metastatic colorectal cancer (mCRC) (NCT03391232) (See also in Example 4).

Immunogenicity measurements proved pre-existing immune responses and indirectly confirmed target antigen expression in the patients. Immunogenicity was measured with enriched Fluorospot assay (ELISPOT) from PBMC samples isolated prior to vaccination and in different time points following a following single immunization with PolyPEPI1018 to confirm vaccine-induced T cell responses; PBMC samples were *in vitro* stimulated with vaccine-specific peptides (9mers and 30mers) to determine vaccine-induced T cell responses above baseline. In average 4, at least 2 patients had pre-existing CD8 T cell responses against each target antigen (FIG. 12C). 7 out of 10 patients had pre-existing immune responses against at least 1 antigen (average 3) (FIG 12D). These results provide proof for the proper target selection, because CD8+ T cell response for a CRC specific target TSA prior to vaccination with PolyPEPI1018 vaccine confirms the expression of that target antigen in the analyzed patient. Targeting the real (expressed) TSAs is the prerequisite for an effective tumor vaccine.

Example 14 - Pre-clinical and Clinical Immunogenicity of PolyPEPI1018 Vaccine proves proper peptide selection

PolyPEPI1018 vaccine contains six 30mer peptides, each designed by joining two immunogenic 15mer fragments (each involving a 9mer PEPI, consequently there are 2 PEPIs in each 30mer by design) derived from 7 TSAs (FIG 13). These antigens are frequently expressed in CRC tumors based on analysis of 2,391 biopsies (FIG 12).

Preclinical immunogenicity results calculated for the Model Population (n=433) and for a CRC cohort (n=37) resulted in 98% and 100% predicted immunogenicity based on PEPI test predictions and this was clinically proved in the OBERTO trial (n=10), with immune responses measured for at least one antigen in 90% of patients. More interestingly, 90% of patients had vaccine peptide specific immune responses against at least 2 antigens and 80% had CD8+ T cell response against 3 or more different vaccine antigens, showing evidence for appropriate target antigen selection during the design of PolyPEPI1018. CD4+ T cell specific and CD8+ T cell specific clinical immunogenicity is detailed in Table 18.

Table 18. Clinical immunogenicity results for PolyPEPI1018 in mCRC.

Immunological responses	% Patients (n)
CD4+ T cell responses	100% (10/10)

CD8+ T cell responses against ≥3 antigens	80% (8/10)
Both CD8+ and CD4+ T cell responses	90% (9/10)
Ex vivo detected CD8+ T cell response	71% (5/7)
Ex vivo detected CD4+ T cell response	86% (6/7)
Average increase of the fraction of polyfunctional (IFN- γ and TNF- α positive) CD8+ T cells compared to pre-vaccination	0.39%
Average increase of the fraction of polyfunctional (IL-2 and TNF- α positive) CD4+ T cells compared to pre-vaccination	0.066%

High immune response rates were found for both effector and memory effector T cells, both for CD4+ and CD8+ T cells, and 9 of 10 patients' immune responses were boosted or *de novo* induced by the vaccine. Also, the fractions of CRC-reactive, polyfunctional CD8+ and CD4+ T cells have been increased in patient's PBMC after vaccination by 2.5- and 13-fold, respectively.

Example 15 - Clinical response for PolyPEPI1018 treatment

The OBERTO clinical trial (NCT03391232), that has been further described in Examples 4, 12, 13 and 14 was analyzed for preliminary objective tumor response rates (RECIST 1.1) (FIG. 14). Of the eleven vaccinated patients on maintenance therapy, 5 had stable disease (SD) at the time point of the preliminary analysis (12 weeks), 3 experienced unexpected tumor responses (partial response, PR) observed on treatment (maintenance therapy + vaccination) and 3 had progressed disease (PD) according to RECIST 1.1 criteria. Stable disease as best response was achieved in 69% of patients on maintenance therapy (capecitabine and bevacizumab). Patient 020004 had durable treatment effect after 12 weeks, and patient 010004 had long lasting treatment effect, qualified for curative surgery. Following the 3rd vaccination this patient had no evidence of disease thus being complete responder, as shown on the swimmer plot on FIG. 14.

After one vaccination, ORR was 27%, DCR was 63%, and in patients receiving at least 2 doses (out of the 3 doses), 2 of 5 had ORR (40%) and DCR was as high as 80% (SD+PR+CR in 4 out of 5 patients) (Table 19).

Table 19. Clinical response for PolyPEPI1018 treatment after ≥ 1 and ≥ 2 vaccination dose

Number of vaccination dose	Objective Response Rate (CR+PR)	Disease Control Rate (SD + PR+CR)
≥ 1	27% (3/11)	63% (7/11)
≥ 2	40% (2/5)	80% (4/5)

Based on the data of the 5 patients receiving multiple doses of PolyPEPI1018 vaccine in the OBERTO-101 clinical trial, preliminary data suggests that higher AGP count (>2) is associated with longer PFS and elevated tumor size reduction (FIG.14B and C).

5

Example 16 – Gastric cancer peptide vaccine design for large population

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

10 Gastric cancer CTAs were identified and ranked based on the overall expression frequencies of antigens found in gastric cancer tumor samples as reported in peer reviewed publications.

Based on the ranked expression rate the most frequently expressed CTA were selected as target antigens for gastric cancer vaccine. The expression rates of the selected gastric cancer,
15 specific CTAs are illustrated in Figure 16.

To select immunogenic peptides from the target CTAs, the PEPI3+ Test and the Model Population described in Example 8 were used to identify the 9 mer epitopes (PEPI3+s) that are most frequently presented by at least 3 HLAs of the individuals in the Model Population. These epitopes are referred to herein as “bestEPIs”. An illustrative example of the “PEPI3+ hotspot”
20 analysis and bestEPI identification is shown in FIG. 15 for the PRAME antigen.

The reported expression frequency for each CTA was multiplied 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 gastric cancer antigens in the highest proportion of individuals (Table 20a). 15 mers were then selected encompassing each of the selected 9 mers
25 (Table 20a). The 15 mers were selected to bind to most HLA class II alleles of most subjects. These 15 mers can induce both CTL and T helper responses in the highest proportion of subjects.

Table 20a. BestEPI list (9-mers underlined) for selecting gastric cancer peptides for vaccine composition. N%: Antigen expression frequency in gastric 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%
1	31	DPPA2	100%	KRNK <u>KMMKRLMTVEK</u>	284	43%	12%	43%
2	32	CAGE-1	77%	ASQLASKM <u>HSLALM</u>	610	42%	94%	33%
3	33	TSP50	57%	GFSYEQDPTLRDPEA	105	51%	0%	29%
4	34	DPPA2	100%	KIEVYLRRLHRHAYPE	115	28%	98%	28%
5	35	HIWI	76%	GFTTSILQYENSIML	251	37%	86%	28%
6	36	SURVIVIN	100%	AKKVRRRAIEQLAAMD	128	26%	25%	26%
7	37	HIWI	76%	SIAGFVASINEGMTR	643	28%	44%	21%
8	38	TSP50	57%	STIMETQFPVSEKGV	84	36%	0%	21%
9	39	5T4	52%	RLELASNHFLYLPRD	214	40%	97%	21%
10	40	5T4	52%	SNHFLYLPRDVLAQL	219	40%	100%	21%
11	41	5T4	52%	SSASSFSSAPFLAS	36	33%	88%	17%
12	42	MAGE-A2	31%	REDSVFAHPRKLLMQ	234	55%	74%	17%
13	43	KK-LC-1	80%	RNTGEMSSNSTALAL	26	21%	0%	17%
14	44	CAGE-1	77%	NIENYSTNALIQPVD	97	21%	14%	16%
15	45	SURVIVIN	100%	KDHRISTFKNWPFLE	15	15%	83%	15%
16	46	MAGE-A2	31%	SFSTTINYTLWRQSD	70	37%	65%	12%
17	47	KK-LC-1	80%	SRDILNNFPHSIARQ	63	13%	11%	11%
18	48	MAGE-A3	37%	LTQHFEVQENYLEYRQ	246	27%	56%	10%
19	49	LAGE-1	14%	ITMPFSSPMEAEVR	92	65%	8%	9%
20	50	MAGE-A3	37%	KASSSLQLVFGIELM	153	23%	58%	9%
21	51	MAGE-A10	30%	RNYEDHFPLLFSEAS	166	27%	26%	8%
22	52	MAGE-A1	31%	ETSYVKVLEYVIKVS	273	26%	85%	8%
23	53	MAGE-A3	37%	QAALSRKVAELVHFL	106	21%	45%	8%
24	54	KK-LC-1	80%	SNTDNNLAVYDLSD	51	10%	0%	8%
25	55	PRAME	20%	RHSQTLKAMVQAWPF	64	37%	38%	7%
26	56	MAGE-A2	31%	SKASEYLQLVFGIEV	152	24%	69%	7%
27	57	MAGE-A1	31%	SAFPTTINFTRQRQP	62	24%	0%	7%
28	58	SSX1	13%	QVEHPQMTFGRLLHRI	93	55%	20%	7%

29	59	MAGE-A1	31%	PRALAETSYVKVLEY	268	16%	39%	5%
30	60	PRAME	20%	DQLLRHVMNPLETSL	314	22%	63%	4%

Fifteen 30 mer peptides were then designed (Table 21a). 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 2-a.

5

Table 21a – 30mer gastric cancer vaccine peptides of composition PolyPEPI1311

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
61	GC1311-01	CAGE-1/DPPA2	ASQLASKMHSLLALMKRNKKMMKRLMTVEK	64%	94%
62	GC1311-02	HIWI/MAGE-A10	GFTTSILQYENSIMLRNYEDHFPLLFSEAS	62%	86%
63	GC1311-03	MAGE-A2/SURVIVIN	SFSTTINYTLWRQSDAKKVRRAIEQLAAMD	49%	69%
64	GC1311-04	KK-LC-1/SURVIVIN	RNTGEMSSNSTALALKDHRISTFKNWPFLF	39%	83%
65	GC1311-05	KK-LC-1/HIWI	SRDILNNFPHSIARQSIAGFVASINEGMTR	36%	45%
66	GC1311-06	MAGE-A2/PRAME	REDSVFAHPRKLLMQDQLLRHVMNPLETSL	64%	91%
67	GC1311-07	5T4/5T4	RLELASNHFLYLPRDSNHFLYLPRDVLAQL	65%	100%
68	GC1311-08	5T4/TSP50	SSASSFSSSAPFLASSTTMTQFPVSEGKV	55%	88%
69	GC1311-09	TSP50/MAGE-A1	GFSYEQDPTLRDPEAPRALAETSYVKVLEY	67%	39%
70	GC1311-10	KK-LC-1/SSX1	SNTDNNLAVYDLSDRDQVEHPQMTFGR LHRI	58%	20%
71	GC1311-11	DPPA2/LAGE-1	KIEVYLRLHRHAYPEITMPFSSPMEAE LVR	82%	99%
72	GC1311-12	MAGE-A1/MAGE-A3	ETSYVKVLEYVIKVSLTQHFVQENYLEYRQ	41%	91%
73	GC1311-13	MAGE-A3/CAGE-1	QAALSRKVAELVHFLNIENYSTNALIQPVD	30%	49%
74	GC1311-14	PRAME/MAGE-A3	RHSQTLKAMVQAWPFKASSSLQLVFGIELM	53%	77%
75	GC1311-15	MAGE-A2/MAGE-A1	SKASEYLQLVFGIEVSAFPTTINFTRQRQP	41%	69%

* 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).

10

Characterization of PolyPEPI1311

Tumor heterogeneity can be addressed by including peptide sequences that target multiple CTAs in a vaccine or immunotherapy regime. The PolyPEPI1311, composition targets 14 different CTAs. Based on the antigen expression rates for these 14 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 5 of the 14 target antigens (AG95=5) as shown by the antigen expression curve in FIG. 17.

15

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. gastric 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.

5 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 (gastric cancer vaccine
10 specific immune response). The PolyPEPI1311 composition can induce immune response against an average of 8 vaccine antigens (AP50=7.98) and 95% of the Model Population can induce immune response against at least three vaccine antigens (AP95=3)(FIG. 18).

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
15 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.

20 Combining the data of AG of gastric cancer and AP in the Model Population we determined the AGP value of PolyPEPI1311 that represents the probability distribution of vaccine antigens that induce immune responses against antigens expressed in gastric tumors. For PolyPEPI1311, the AGP50 value in the Model Population is 3.86. The AGP95=1, means that 95% of the subjects in the Model Population induce immune responses against at least one
25 expressed vaccine antigen (FIG. 19).

Example 17 – Patient selection using companion diagnostic for gastric cancer, lung cancer, melanoma and bladder cancer vaccine

30 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).

10 **Example 18 – Lung cancer peptide vaccine design for large population**

The PEPI3+ Test described above was used to design 9-mer and 15-mer peptides for use in lung cancer vaccines, using the same method described in Example 16 above for gastric cancer (Tables 20b, 21b). The expression rates of the selected gastric cancer, specific CTAs are illustrated in Figure 20.

15

Table 20b. BestEPI list (9-mers underlined) for selecting lung cancer peptides for vaccine composition. N%: Antigen expression frequency in lung cancers; B%: bestEPI frequency, i.e. 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%.

20

SEQ ID NO. 9mer	SEQ ID NO. 15mer	Antigen		BestEPIs and Optimized 15 mer				
		Antigen	N%	Antigen	N%	Antigen	N%	Antigen
90	120	BRDT	48%	DAYKFAADVRLMFMN	333	58%	100%	28%
91	121	PRAME	63%	RHSQTLKAMVQAWPF	64	37%	38%	23%
92	122	BRDT	48%	DLWKHSFSWPFQRPV	42	43%	88%	21%
93	123	NALP4	38%	QSTTSVYSSFVFNLF	366	50%	75%	19%
94	124	NALP4	38%	KRAMEAFNLVRESEQ	319	48%	67%	18%
95	125	NALP4	38%	FVIDSFEEELQGGLNE	228	46%	47%	17%
96	126	MAGE-A12	36%	QVALSRKMAELVHFL	106	48%	63%	17%

97	127	MAGE-A2	27%	REDSVFAHPRKLLMQ	234	55%	74%	15%
98	128	SURVIVIN	57%	AKKVRRRAIEQLAAMD	128	26%	25%	15%
99	129	PRAME	63%	KEEQYIAQFTSQFLS	280	23%	99%	14%
100	130	PRAME	63%	DQLLRHVMNPLETLS	314	22%	63%	14%
101	131	MAGE-A12	36%	RNFQDFFPVIFSKAS	141	37%	90%	13%
102	132	PRAME	63%	RGRLDQLLRHVMNPL	310	21%	65%	13%
103	133	DPPA2	30%	KRNKKMMKRLMTVEK	284	43%	12%	13%
104	134	NY-SAR-35	28%	SSYFVLANGHILPNS	94	42%	85%	12%
105	135	BRDT	48%	ILKEMLAKKHFSYAW	279	23%	55%	11%
106	136	LDHC	29%	SVMDLVGSILKNLRR	255	38%	85%	11%
107	137	LDHC	29%	KLKGEMMDLQHGSFL	57	37%	1%	11%
108	138	MAGE-A2	27%	SSFSTTINYTLWRQS	69	37%	74%	10%
109	139	MAGE-C2	24%	MASESLVMSSNVSF	357	40%	15%	10%
110	140	MAGE-C2	24%	EHFVYGEPRELLTKV	263	40%	60%	10%
111	141	MAGE-C2	24%	SESLVMSSNVSFSE	359	38%	33%	9%
112	142	MAGE-A3	33%	LTQHVFQENYLEYRQ	246	27%	56%	9%
113	143	MAGE-A12	36%	SKASEYLQLVFGIEV	152	24%	69%	9%
114	144	SURVIVIN	57%	KDHRISTFKNWPFLE	15	15%	83%	9%
115	145	DPPA2	30%	KIEVYLRRLHRHAYPE	115	28%	98%	9%
116	146	MAGE-A3	33%	KASSSLQLVFGIELM	153	23%	58%	8%
117	147	KK-LC-1	35%	RNTGEMSSNSTALAL	26	21%	0%	7%
118	148	MAGE-C2	24%	SSFSTSSSLILGGPE	53	30%	44%	7%
119	149	MAGE-A1	28%	PRALAETSYVKVLEY	268	16%	39%	5%

Table 21b – 30mer lung cancer vaccine peptides of composition PolyPEPI821

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
150	LC821-01	BRDT/SURVIVIN	DAYKFAADVRLMFMNAKKVRRRAIEQLAAMD	64%	100%
151	LC821-02	PRAME/MAGE-A2	RHSQTLKAMVQAWPFREDSVFAHPRKLLMQ	67%	78%
152	LC821-03	BRDT/MAGE-A12	DLWKHSFSWPFQRPVSKASEYLQLVFGIEV	57%	91%
153	LC821-04	PRAME/LDHC	DQLLRHVMNPLETLSSVMDLVGSILKNLRR	53%	94%
154	LC821-05	NALP4/SURVIVIN	QSTTSVYSSVFVNLFKDHRISTFKNWPFLE	55%	92%
155	LC821-06	MAGE-C2/NALP4	SSFSTSSSLILGGPEFVIDSFEELQGGLNE	60%	80%
156	LC821-07	MAGE-A2/NALP4	SSFSTTINYTLWRQSKRAMEAFNLVRESEQ	70%	90%
157	LC821-08	NY-SAR-35/MAGE-A12	SSYFVLANGHILPNSRNFQDFFPVIFSKAS	67%	98%
158	LC821-09	MAGE-C2/PRAME	MASESLVMSSNVSFKEEQYIAQFTSQFLS	51%	99%
159	LC821-10	PRAME/MAGE-A3	RGRLDQLLRHVMNPLLTQHVFQENYLEYRQ	42%	85%
160	LC821-11	MAGE-C2/LDHC	SESLVMSSNVSFSEKLKGEMMDLQHGSFL	67%	33%
161	LC821-12	MAGE-A12/MAGE-A1	QVALSRKMAELVHFLPRALAETSYVKVLEY	63%	72%

162	LC821-13	MAGE-C2/BRDT	EHFVYGEPRELLTKVILKEMLAKKHFSYAW	59%	75%
163	LC821-14	DPPA2/MAGE-A3	KIEVYLRRLHRHAYPEKASSSLQLVFGIELM	53%	99%
164	LC821-15	KK-LC-1/DPPA2	RNTGEMSSNSTALALKRNKMMKRLMTVEK	60%	12%

* 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 PolyPEPI821

5 The PolyPEPI821 composition targets 13 different CTAs. Based on the antigen expression rates for these 13 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 2 of the 10 target antigens (AG95=2) as shown by the antigen expression curve in FIG. 21.

10 The PolyPEPI821 composition can induce immune response against an average of approaching 8 vaccine antigens (AP50=7.60) and 95% of the Model Population can induce immune response against at least two vaccine antigens (AP95=2)(FIG. 22).

For PolyPEPI821, the AGP50 value in the Model Population is 2.77. The AGP91=1, means that 91% of the subjects in the Model Population induce immune responses against at least one expressed vaccine antigen (FIG. 23).

Example 19 – Melanoma peptide vaccine design for large population

The PEPI3+ Test described above was used to design 9-mer and 15-mer peptides for use in melanoma vaccines, using the same method described in Example 16 above for gastric cancer (Tables 20c, 21c). The expression rates of the selected gastric cancer, specific CTAs are illustrated in Figure 24.

Table 20c. BestEPI list (9-mers underlined) for selecting melanoma peptides for vaccine composition. N%: Antigen expression frequency in melanomas; B%: bestEPI frequency, i.e. 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	Antigen	BestEPIs and Optimized 15 mer
--------	---------	-------------------------------

NO. 9mer	SEQ ID NO. 15mer							
		Antigen	N%	Opt. 15mer	Opt. Position	B%	HLAII** (CD4)	B%*N%
178	208	PRAME	90%	<u>LERLAYLHARLRELL</u>	457	52%	100%	47%
179	209	MAGE-A2	61%	<u>EDSVFAHPRKLLMQD</u>	235	55%	43%	34%
180	210	PRAME	90%	<u>RHSQTLKAMVQAWPF</u>	64	37%	38%	33%
181	211	MAGE-C1	45%	<u>SFSYTLTSLFQSSPE</u>	450	57%	99%	26%
182	212	Survivin	96%	<u>TAKKVRRAIEQLAAM</u>	127	26%	26%	25%
183	213	MAGE-C1	45%	<u>SPSFSTLVSFLQSS</u>	273	53%	97%	24%
184	214	MAGE-A2	61%	<u>SSFSTTINYTLWRQS</u>	69	37%	74%	23%
185	215	MAGE-A12	45%	<u>QVALSRKMAELVHFL</u>	106	48%	63%	22%
186	216	MAGE-C1	45%	<u>DDSYVFVNTLDLTSE</u>	971	48%	99%	22%
187	217	MAGE-C1	45%	<u>FSYTLASLLQSSHES</u>	783	47%	100%	21%
188	218	PRAME	90%	<u>KEEQYIAQFTSQFLS</u>	280	23%	99%	20%
189	219	MAGE-C1	45%	<u>QIPMTSSFSTLLSI</u>	339	45%	65%	20%
190	220	PRAME	90%	<u>DQLLRHVMNPLETLS</u>	314	22%	63%	20%
191	221	Ny-ESO-1	38%	<u>RGPE SRLLEFY LAMP</u>	81	52%	65%	20%
192	222	MAGE-C2	45%	<u>MASESL SVMSSNVSF</u>	357	40%	15%	18%
193	223	MAGE-C2	45%	<u>REHFVYGEPRELLTK</u>	262	40%	74%	18%
194	224	MAGE-A6	62%	<u>QYFVQENYLEYRQVP</u>	248	27%	93%	17%
195	225	BORIS	27%	<u>MFTSSRMSSFNRHMK</u>	263	57%	66%	15%
196	226	LAGE-1	35%	<u>DFTVSGNLLFMSVRD</u>	125	43%	82%	15%
197	227	MAGE-A2	61%	<u>SKASEYLQLVFGIEV</u>	152	24%	69%	15%
198	228	Survivin	96%	<u>KDHRISTFKNWPFLE</u>	15	15%	83%	14%
199	229	MAGE-A11	54%	<u>SHSYVLVTSLNLSYD</u>	286	26%	100%	14%
200	230	SSX-1	25%	<u>QVEHPQMTFGR LHRI</u>	93	55%	20%	14%
201	231	MAGE-A3	59%	<u>KASSSLQLVFGIELM</u>	153	23%	58%	14%
202	232	MAGE-C2	45%	<u>SSFSTSSSLILGGPE</u>	53	30%	44%	13%
203	233	MAGE-A3	59%	<u>QAALSRKVAELVHFL</u>	106	21%	45%	12%
204	234	MAGE-A11	54%	<u>SPTAMDAIFGSL SDE</u>	181	23%	0%	12%
205	235	MAGE-A10	44%	<u>RNYEDHFPLLFSEAS</u>	166	27%	26%	12%
206	236	MAGE-A11	54%	<u>YAGREHFLFGEPKRL</u>	344	18%	71%	9%
207	237	MAGE-A1	37%	<u>PRALAETSYVKVLEY</u>	268	16%	39%	6%

Fifteen 30 mer peptides were then designed (Table 21c). 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 (BestEPs) from Table 20c.

5

Table 21c – 30mer melanoma vaccine peptides of composition PolyPEPI621

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
238	MC621-01	PRAME/PRAME	LERLAYLHARLRELDDQLLRHVMNPLETLS	58%	100%
239	MC621-02	Survivin/MAGE-A2	TAKKVRRAIEQLAAMEDSVFAHPRKLLMQD	65%	54%
240	MC621-03	PRAME/MAGE-A3	RHSQTLKAMVQAWPFKASSSLQLVFGIELM	53%	77%
241	MC621-04	MAGE-C1 /MAGE-A2	DDSYVFNVTLDLTSESSFSTTINYTLWRQS	70%	99%
242	MC621-05	MAGE-C1 /MAGE-A2	SFSYTLTSLFQSSPESKASEYLQLVFGIEV	64%	99%
243	MC621-06	MAGE-A11/MAGE-C1	SHSYVLVTSLNLSYDSPFSSTLVSLFQSS	73%	100%
244	MC621-07	LAGE-1 /MAGE-C2	DFTVSGNLLFMSVRDMASESLSVMSNVSF	60%	83%
245	MC621-08	BORIS/MAGE-A12	MFTSSRMSSFNHRMKQVALSRKMAELVHFL	74%	81%
246	MC621-09	MAGE-C2/MAGE-C1	SSFSTSSSLILGGPEFSYTLASLLQSSHES	54%	100%
247	MC621-10	Survivin/MAGE-C1	KDHRISTFKNWPFLEQIPMTSSFSSTLLSI	51%	89%
248	MC621-11	Ny-ESO-1/MAGE-A10	RGPE SRLLEFYLAMPRNYEDHFPLLFSEAS	73%	66%
249	MC621-12	MAGE-A3 /MAGE-A6	QAALSRKVAELVHFLQYFVQENYLEYRQVP	48%	97%
250	MC621-13	MAGE-A11 /MAGE-A1	SPTAMDAIFGSLSDPEPRALAETSYVKVLEY	53%	39%
251	MC621-14	MAGE-C2/MAGE-A11	REHFVYGEPRELLTKYAGREHFLFGEPKRL	57%	83%
252	MC621-15	SSX-1/PRAME	QVEHPQMTFGRHLHRIKEEQYIAQFTSQFLS	67%	99%

* 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 PolyPEPI621

5 The PolyPEPI621 composition targets 15 different CTAs. Based on the antigen expression rates for these 15 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 5 of the 15 target antigens (AG95=5) as shown by the antigen expression curve in FIG. 25.

10 The PolyPEPI621 composition can induce immune response against an average of 8 vaccine antigens (AP50=8.29) and 95% of the Model Population can induce immune response against at least three vaccine antigens (AP95=2)(FIG. 26).

Combining the data of AG of melanoma and AP in the Model Population we determined the AGP value of PolyPEPI621 that represents the probability distribution of vaccine antigens that induce immune responses against antigens expressed in gastric tumors. For PolyPEPI621, 15 the AGP50 value in the Model Population is 2.77. The AGP95=1, means that 95% of the subjects in the Model Population induce immune responses against at least one expressed vaccine antigen (FIG. 27).

Example 20 – Bladder cancer peptide vaccine design for large population

The PEPI3+ Test described above was used to design 9-mer and 15-mer peptides for use in bladder cancer vaccines, using the same method described in Example 16 above for gastric cancer (Tables 120d, 21d). The expression rates of the selected gastric cancer, specific CTAs are illustrated in Figure 28.

Table 20d. BestEPI list (9-mers underlined) for selecting bladder cancer peptides for vaccine composition. N%: Antigen expression frequency in bladder 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%
268	298	PIWIL2	82%	<u>FVASINLT</u> LTWKYSR	760	67%	93%	55%
269	299	PIWIL2	82%	RNFYDPTSAMVL <u>QQH</u>	341	60%	49%	49%
270	300	PIWIL2	82%	YSRVVFQMPHQEIVD	772	40%	77%	33%
271	301	CTAGE1	53%	QNYIDQFLLTSFPTF	33	59%	90%	31%
272	302	MAGE-A9	61%	EFMFQEALKLKVAEL	101	49%	100%	30%
273	303	EpCAM	54%	RTYWIIIEELKHKARE	140	51%	100%	28%
274	304	OY-TES-1	28%	ESTPMIMENIQELIR	276	67%	82%	19%
275	305	MAGE-A9	61%	SSISVYYTLWSQFDE	67	30%	97%	19%
276	306	NY-ESO-1	31%	RGPE SRLLEFY LAMP	81	52%	65%	16%
277	307	SURVIVIN	62%	AKKVRRAIEQLAAMD	128	26%	25%	16%
278	308	MAGE-C1	27%	SSFSYTLLSLFQSSP	449	57%	98%	16%
279	309	CTAGE1	53%	SFVLFLFGGNNFIQN	14	29%	85%	15%
280	310	EpCAM	54%	QTLIYYVDEKAPDFS	246	28%	34%	15%
281	311	MAGE-A2	25%	REDSVFAHPRKLLMQ	234	55%	74%	14%
282	312	MAGE-C1	27%	DDSYVFNVTLDLTSE	971	48%	99%	13%
283	313	LAGE-1	30%	DFTVSGNLLFMSVRD	125	43%	82%	13%
284	314	MAGE-A3	42%	LTQH FVQENYLEYRQ	246	27%	56%	11%
285	315	MAGE-A8	57%	EEAIWEALSVMGLYD	221	20%	78%	11%
286	316	HAGE	24%	NDLQMSNFVNLKNIT	377	43%	67%	10%

287	317	MAGE-A8	57%	EKVAELVRFLLRKYQ	114	18%	99%	10%
288	318	MAGE-A3	42%	KASSSLQLVFGIELM	153	23%	58%	10%
289	319	SURVIVIN	62%	KDHRISTFKNWPFLE	15	15%	83%	9%
290	320	MAGE-A2	25%	SSFSTTINYTLWRQS	69	37%	74%	9%
291	321	MAGE-A3	42%	QAALSRKVAELVHFL	106	21%	45%	9%
292	322	MAGE-A1	34%	ETSYVKVLEYVIKVS	273	26%	85%	9%
293	323	MAGE-C2	19%	MASESLVMSSNVSF	357	40%	15%	8%
294	324	MAGE-C2	19%	REHFVYGEPRELLTK	262	40%	74%	8%
295	325	MAGE-A10	28%	RNYEDHFPLLFSEAS	166	27%	26%	7%
296	326	MAGE-A12	29%	KASEYLQLVFGIEVV	153	24%	83%	7%
297	327	LAGE-1	30%	DSRLLQLHITMPFSS	84	20%	99%	6%

Fifteen 30 mer peptides were then designed (Table 21d). 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 20d.

5

Table 21d – 30mer bladder cancer vaccine peptides of composition PolyPEPI1411

SEQID	TREOSID	Source Antigen	Peptide (30mer)	HLAI* (CD8)	HLAII** (CD4)
328	BLV1411-01	MAGE-C1/PIWIL2	DDSYVFNLTDLTSENFYDPTSAMVLQQH	81%	99%
329	BLV1411-02	MAGE-A9/OY-TES-1	EFMFQEALKLKVAELESTPMIMENIQELIR	77%	100%
330	BLV1411-03	PIWIL2/MAGE-A1	YSRVVFQMPHQEIIVDETSYVKVLEYVIKVS	54%	93%
331	BLV1411-04	CTAGE1/MAGE-A2	QNYIDQFLLTSTFPTFREDSTVFAHPRKLLMQ	82%	100%
332	BLV1411-05	HAGE/EpCAM	NDLQMSNFVNLKNITRTYWI I IELKHKARE	66%	100%
333	BLV1411-06	MAGE-A9/MAGE-A8	SSISVYYTLWSQFDEEEKVAELVRFLLRKYQ	42%	100%
334	BLV1411-07	NY-ESO-1/MAGE-A10	RGPESSLLEFYLAMPRNYEDHFPLLFSEAS	73%	66%
335	BLV1411-08	CTAGE1/MAGE-A8	SFVLFLFGGNNFIQNEEAIWEALSVMGLYD	46%	94%
336	BLV1411-09	EpCAM/MAGE-C2	QTLIYYVDEKAPEFSREHFVYGEPRELLTK	64%	80%
337	BLV1411-10	MAGE-C1/MAGE-A12	SSFSYTLTSLFQSSPKASEYLQLVFGIEVV	64%	98%
338	BLV1411-11	PIWIL2/LAGE-1	FVASINLTTLTKWYSRDFTVSGNLLFMSVRD	76%	94%
339	BLV1411-12	MAGE-A3/MAGE-A3	LTQHFVQENYLEYRQKASSSLQLVFGIELM	47%	71%
340	BLV1411-13	MAGE-A2/LAGE-1	SSFSTTINYTLWRQSDSRLLQLHITMPFSS	61%	99%
341	BLV1411-14	SURVIVIN/MAGE-C2	AKKVRRAIEQLAAMDMAESLSVMSSNVSF	52%	28%
342	BLV1411-15	MAGE-A3/SURVIVIN	QAALSRKVAELVHFLKDHRISTFKNWPFLE	30%	90%

* 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).

10 Characterization of PolyPEPI1411

The PolyPEPI1411 composition targets 17 different CTAs. Based on the antigen expression rates for these 17 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 17 target antigens (AG95=4) as shown by the antigen expression curve in FIG. 29.

The PolyPEPI1411 composition can induce immune response against an average of 9 vaccine antigens (AP50=9.44) and 95% of the Model Population can induce immune response against at least three vaccine antigens (AP95=3)(FIG. 30).

Combining the data of AG of bladder cancer and AP in the Model Population we determined the AGP value of PolyPEPI1411 that represents the probability distribution of vaccine antigens that induce immune responses against antigens expressed in bladder tumors. For PolyPEPI1411, the AGP50 value in the Model Population is 3.90. The AGP95=1, means that 95% of the subjects in the Model Population induce immune responses against at least one expressed vaccine antigen (FIG. 31).

Example 21 – Personalised Immunotherapy (PIT) design and treatment for ovarian-, breast- and colorectal cancer

This Example provides proof of concept data from 4 metastatic cancer patients treated with personalized immunotherapy vaccine compositions to support the principals of binding of epitopes by multiple HLAs of a subject to induce cytotoxic T cell responses, on which the present disclosure is partly based on.

Composition for Treatment of Ovarian Cancer with POC01 – PIT (Patient A)

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.

The HLA class I and class II genotype of a metastatic ovarian adenocarcinoma cancer patient (Patient-A) was determined from a saliva sample.

To make a personalized pharmaceutical composition for Patient-A 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-A (Table 22). In addition, each peptide is optimized to bind the maximum number of HLA class II of the patient.

Table 22: Personalized vaccine of Patient-A.

POC01 vaccine for Patient-A	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%	PSQASSGQARMFPNAPYLP	4	1
POC01_P7	HIWI	63%	RRSIAGFVASINEGMTRWFS	3	4
POC01_P8	PRAME	60%	MQDIKMILKMVQLDSIEDLE	3	4
POC01_P9	AKAP-3	58%	ANSVVSDDMMVSIMKTLKIQV	3	4
POC01_P10	MAGE-A4	37%	REALSNKVDELAHFLLRKYR	3	2
POC01_P11	MAGE-A9	37%	ETSYEKVINYLVMNLNAREPI	3	4
POC01_P12a	MAGE-A10	52%	DVKEVDPTGHSFVLVTSLGL	3	4
POC01_P12b	BAGE	30%	SAQLLQARLMKEESPVSWR	3	2

5

Eleven PEPI3 peptides in this immunotherapy composition can induce T cell responses in Patient-A 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-A 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 the 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 any expressed antigen) and maximum number of antigens (all antigens are expressed and present a PEPI).

15

The probability that Patient-A will express one or more of the 13 antigens is shown in Fig. 32. AGP95 (AGP with 95% probability) = 5, AGP50 (the mean -expected value- of the discrete probability distribution) = 7.9, mAGP (probability that AGP is at least 2) = 100%, AP = 13.

- 5 A pharmaceutical composition for Patient-A may be comprised of at least 2 from the 13 peptides (Table 22), because the presence in a vaccine or immunotherapy composition of at least two polypeptide fragments (epitopes) that can bind to at least three HLAs of an individual (≥ 2 PEPI3+) was determined to be predictive for a clinical response. The peptides are synthesized, dissolved in a pharmaceutically acceptable solvent and mixed with an adjuvant prior to injection.
- 10 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-A the 13 peptides were formulated as 4 x 3 or 4 peptide (POC01/1, POC01/2, POC01/3, POC01/4). One treatment cycle is defined as administration of
- 15 all 13 peptides within 30 days.

Patient history:

Diagnosis: Metastatic ovarian adenocarcinoma

Age: 51

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

20 **Tumor pathology:**

2011: first diagnosis of ovarian adenocarcinoma; Wertheim operation and chemotherapy; lymph node removal

2015: metastasis in pericardial adipose tissue, excised

2016: hepatic metastases

- 25 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)

- 30 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. FIG. 33.

2017-2018: Patient-A received 8 cycles of vaccination as add-on therapy, and lived 17 months (528 days) after start of the treatment. During this interval, after the 3rd and 4th vaccine treatment she experienced partial response as best response. She died in October 2018.

5 An interferon (IFN)- γ ELISPOT bioassay confirmed the predicted T cell responses of Patient-A to the 13 peptides. Positive T cell responses (defined as >5 fold above control, or >3 fold above control and >50 spots) were detected for all 13 20-mer peptides and all 13 9-mer peptides having the sequence of the PEPI of each peptide capable of binding to the maximum HLA class I alleles of Patient-A (FIG. 34).

10 **Patient's tumor MRI findings (Baseline April 15, 2016) (BL: baseline for tumor response evaluation on FIG. 35)**

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 (FU1: follow up 1 on FIG. 35)

15 Dec/25/2016 (before PIT vaccine treatment) There was dramatic reduction in tumor burden with confirmation of response obtained at (FU2: follow up 2 on FIG. 35)

Jan - Mar 2017 – TOPO protocol (topoisomerase)

April/6/2017 (FU3 on FIG. 35) demonstrated regrowth of existing lesions and appearance of new lesions leading to disease progression. Peritoneal carcinomatosis with increased amount of

20 ascites. Progressive hepatic tumor and lymph node

April 21 2017 START PIT

Jul/26/17 (after the 2nd Cycle of PIT): (FU4 on FIG. 35) Progression / Pseudo-Progression

Rapid progression in lymph nodes, hepatic, retroperitoneal and thoracic areas, significant pleural fluid and ascites. Initiate Carboplatin, Gemcitabine, Avastin.

25 Sep/20/17 (after 3 Cycles of PIT): (FU5 on FIG. 35) Partial Response

Complete remission in the pleural region/fluid and ascites

Remission in hepatic, retroperitoneal area and lymph nodes

The findings suggest pseudo progression.

Nov/28/17 (after 4 Cycles of PIT): (FU6 on FIG. 35) Partial Response

30 Complete remission in the thoracic region. Remission in hepatic, retroperitoneal area and lymph nodes

Apr/13/18: Progression

Complete remission in the thoracic and retroperitoneal regions. Progression in hepatic centers and lymph nodes

Jun/12/2018: Stable disease

- 5 Complete remission in the thoracic and retroperitoneal regions. Minimal regression in hepatic centers and lymph nodes

July 2018: Progression

October 2018: Patient-A died

Partial MRI data for Patient-A is shown in Table 23 and Figure 35.

10 **Table 23. 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

Design, safety and immunogenicity of Personalised Immunotherapy Composition (PBRC01) for treatment of metastatic breast cancer (Patient-B) PT9

- 15 The HLA class I and class II genotype of metastatic breast cancer Patient-B was determined from a saliva sample. To make a personalized pharmaceutical composition for Patient-B 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-B (Table 24). In addition, each peptide is optimized to bind the
- 20 maximum number of HLA class II of the patient. The twelve peptides target twelve breast cancer antigens. The probability that Patient-B will express one or more of the 12 antigens is shown in FIG. 36.

Table 24. 12 peptides for Patient-B breast cancer patient

BRC01 vaccine peptides	Target Antigen	Antigen Expression	20mer peptide	MAXHL A 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%	DVKEVDPTSHSYVLVTSNL	3	4
PBRC01_cP6	NY-SAR-35	49%	ENAHGQSLEEDSALEALLNF	3	2
PBRC01_cP7	HOM-TES-85	47%	MASFRKLTSEKVPNNHPSR	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 TSAs expressed in the breast cancer cells of BRC09. Additional efficacy parameters: AGP50 = 6.45, mAGP = 100%, AP = 12.

- 5 For treatment of Patient-B 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 (FIG 36C).

Patient history:

- 2013: Diagnosis: breast carcinoma diagnosis; CT scan and bone scan ruled out metastatic
10 disease.
2014: bilateral mastectomy, postoperative chemotherapy
2016: extensive metastatic disease with nodal involvement both above and below the diaphragm. Multiple liver and pulmonary metastases.

Therapy:

- 15 2013-2014: Adriamycin-Cyclophosphamide and Paclitaxel
2017: Letrozole, Palbociclib and Gosorelin and PIT vaccine
2018: Worsening conditions, patient died in January

PIT vaccine treatment began on 7 April 2017. Patient-B's treatment schedule and

main characteristics of disease are shown in Table 25.

Table 25 – Treatment and response of Patient-B

Date (2017)	Mar	May	Jun	Sep	Nov	Dec
Treatment regimen	PIT Vaccine					
	Palbocyclib Letrozole Gosorelin		Anticancer drug treatment interruption			
Neutrophils (1.7-3.5/mm ³)	ND	1.1	4.5	3.4	2.4	3
CEA (<5.0 ng/ml)	99	65	23	32	128	430
CA 15-3 (<31.3 U/ml)	322	333	138	76	272	230
T1: Right axillar lymph node	15 mm & 11.6 SUV _{max}	9 mm & 2.0 SUV _{max}	nd*	nd	nd	6 mm & 0 SUV _{max}
T2: Right lung metastasis	10 mm & 4.8 SUV _{max}	7 mm & 0 SUV _{max}	nd	nd	nd	4 mm & 0 SUV _{max}
Left iliac bone metastasis	Non measurable & 4.0 SUV _{max}	Regression & 0 SUV _{max}	nd	nd	nd	Regression & 0 SUV _{max}
Multiple liver metastases	Non measurable & 11.5 SUV _{max}	Partial regression & 6.1 SUV _{max}	nd	nd	nd	Progression & 16.8 SUV _{max}

*no data

It was predicted with 95% confidence that 8-12 vaccine peptides would induce T cell responses in Patient-B. Peptide-specific T cell responses were measured in all available PBMC samples using an interferon (IFN)- γ ELISPOT bioassay (FIG. 37). The results confirmed the prediction: Nine peptides reacted positive demonstrating that T cells can recognize Patient-B's tumor cells expressing FISPI, BORIS, MAGE-A11, HOM-TES-85, NY-BR-1, MAGE-A9, SCP1, MAGE-A1 and MAGE-C2 antigens. Some tumor specific T cells were present after the 1st vaccination and boosted with additional treatments (e.g. MAGE-A1) others induced after boosting (e.g. MAGE-A9). Such broad tumor specific T cell responses are remarkable in a late stage cancer patient.

Patient-B history and results

Mar 7, 2017: Prior PIT Vaccine treatment

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

Mar 2017: Treatment initiation - Letrozole, *Palbociclib*, Gosorelin & PIT Vaccine

May 2017: Drug interruption

May 26 2017: After 1 cycle of PIT

10 83% reduction of tumor metabolic activity (PET CT) liver, lung lymphnodes and other
 metastases.

June 2017: Normalized Neutrophils values indicate *Palbociclib* interruption as affirmed by the
 patient

4 Months Delayed Rebound of Tumor Markers

15 **Mar to May 2017:** CEA and CA remained elevated consistently with the outcome of her anti-
 cancer treatment (Ban, Future Oncol 2018)

June to Sept 2017: CEA and CA decreased consistently with the delayed responses to
 immunotherapies

Quality of life

20 **Feb to Mar 2017:** Poor, hospitalized with jaundice

April to Oct 2017: Excellent

Nov 2017: (Worsening conditions (tumor escape?))

Jan 2018: Patient-B died.

Immunogenicity results are summarized in FIG. 37.

25

Clinical outcome measurements of the patient: One month prior to the initiation of PIT
 vaccine treatment PET CT documented extensive DFG avid disease with nodal involvement both
 above and below the diaphragm (Table 25). She had progressive multiple hepatic, multifocal
 osseous and pulmonary metastases and retroperitoneal adenopathy. Her intrahepatic enzymes
 30 were elevated consistent with the damage caused by her liver metastases with elevated bilirubin
 and jaundice. She accepted Letrozole, *Palbociclib* and Gosorelin as anti-cancer treatment. Two

month after initiation of PIT vaccinations the patient felt very well and her quality of life normalized. In fact, her PET CT showed a significant morphometabolic regression in the liver, lung, bone and lymph node metastases. No metabolic adenopathy was identifiable at the supra-diaphragmatic stage.

5 The combination of Palbocyclob and the personalised vaccine was likely to have been responsible for the remarkable early response observed following administration of the vaccine. Palbocyclob 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 results of Patient-B's treatment suggest that PIT vaccine may be used as add-on to the
10 state-of-art therapy to obtain maximal efficacy.

 Patient-B's tumor biomarkers were followed to disentangle the effects of state-of-art therapy from those of PIT vaccine. Tumor markers were unchanged during the initial 2-3 months of treatment then sharply dropped suggesting of a delayed effect, typical of immunotherapies (Table 25). Moreover, at the time the tumor biomarkers dropped the patient had already
15 voluntarily interrupted treatment and confirmed by the increase in neutrophil counts.

 After the 5th PIT treatment the patient experienced symptoms. The levels of tumor markers and liver enzymes were increased again. 33 days after the last PIT vaccination, her PET CT showed significant metabolic progression in the liver, peritoneal, skeletal and left adrenal site confirming the laboratory findings. The discrete relapse in the distant metastases could be due to
20 potential immune resistance; perhaps caused by downregulation of both HLA expression that impairs the recognition of the tumor by PIT induced T cells. However, the PET CT had detected complete regression of the metabolic activity of all axillary and mediastinal axillary supra-diaphragmatic targets (Table 25). These localized tumor responses may be accounted to the known delayed and durable responses to immunotherapy, as it is unlikely that after anti-cancer
25 drug treatment interruption these tumor sites would not relapse.

Personalised Immunotherapy Composition for treatment of patient with metastatic breast carcinoma (Patient-C) PT13

 PIT vaccine similar in design to that described for Patient-A and Patient-B was prepared
30 for the treatment of a patient (Patient-C) with metastatic breast carcinoma. PIT vaccine

contained 12 PEPIs. The PIT vaccine has a predicted efficacy of AGP = 4. The patient's treatment schedule is shown in FIG.38.

Tumor Pathology

- 5 2011 Original tumor: HER2-, ER+, sentinel lymph node negative
 2017 Multiple bone metastases: ER+, cytokeratin 7+, cytokeratin 20-,CA125-, TTF1-, CDX2-

Treatments

- 2011 Wide local resection, sentinel lymph nodes negative; radiotherapy
 2017- Anti-cancer therapy (Tx): Letrozole (2.5 mg/day), Denosumab;
 10 Radiation (Rx): one bone
 PIT vaccine (3 cycles) as add-on to standard of care
 Bioassay confirmed positive T cell responses (defined as >5 fold above control, or >3 fold above control and >50 spots) to 11 out of the 12 20-mer peptides of the PIT vaccine and 11 out of 12 9-mer peptides having the sequence of the PEPI of each peptide capable of binding to
 15 the maximum HLA class I alleles of the patient (FIG. 39).

Long-lasting memory T-cell responses were detected after 14 months of the last vaccination (FIG. 24C-D).

Treatment Outcome

- Clinical results of treatment of Patient-C are shown in Table 26. Patient-C has partial response
 20 and signs of healing bone metastases.

Table 26 – Clinical results of treatment of breast cancer Patient-C

	Before PIT	+70 days * (10w)	+150 days * (21w)	+388 days* (55w)
Bone Biopsy	Met. breast Cancer DCIS	Not done	RIB5 is negative	Not done
PET CT	Multiple metastases	Only RIB5 is DFG avid	Not done	Not done
CT	Multiple metastases	Not done	Not done	Healing bone mets (sclerotic foci)
CA-15-3	87	50	32	24

*After 3rd cycle of PIT vaccination

Immune responses are shown on FIG. 39. Predicted Immunogenicity, PEPI = 12 (CI95% [8,12]
 Detected Immunogenicity: 11 (20-mers) & 11 (9-mers) antigen specific T cell responses
 following 3 PIT vaccinations (FIG. 39A, B). After 4.5, 11 or 14 months of the last vaccination,
 PIT vaccine-specific immune response could still be detected (FIG. 39C, D).

5

**Personalised Immunotherapy Composition for treatment of patient with metastatic
 colorectal cancer (Patient-D) PT16**

Tumor pathology

- 2017 (Feb) mCRC (MSS) with liver metastases, surgery of primer tumor (in sigmoid colon).
 10 pT3 pN2b (8/16) M1. KRAS G12D, TP53-C135Y, KDR-Q472H, MET-T1010I
 mutations. SATB2 expression. EGFR wt, PIK3CA-I391M (non-driver).
 2017 (Jun) Partial liver resection: KRAS-G12D (35G>A) NRAS wt,
 2018 (May) 2nd resection: SATB2 expression, lung metastases 3 → 21
 Treatments
 15 2017 FOLFOX-4 (oxaliplatin, Ca-folate, 5-FU) → allergic reaction during 2nd
 treatment
 DeGramont (5-FU + Ca-folate)
 2018 (Jun) → FOLFIRI plus ramucirumab, biweekly; chemoembolization
 2018 (Oct) PIT vaccination (13 patient-specific peptides, 4 doses) as add-on to standard of
 20 care.

The patient's treatment schedule is shown in FIG. 40.

Treatment outcome

Patient in good overall condition, disease progression in lungs after 8 months confirmed by CT.

- Both PIT induced and pre-existing T cell responses were measured by enriched
 25 Fluorospot from PBMC, using 9mer and 20mer peptides for stimulation (FIG. 41).

Summary of immune response rate and immunogenicity results prove the proper design
 for target antigen selection as well as for the induction of multi-peptide targeting immune
 responses, both CD4+ and CD8+ specific ones.

Table 27. Summary table of immunological analysis of Patient A-D

Patient ID	Measured immunogenicity for the different vaccine peptides*	
	CD4+ T cells	CD8+ T cells
Patient-A	13/13 (100%)	13/13 (100%)
Patient-B	9/12 (75%)	1/12 (8%)
Patient-C	11/12 (92%)	11/12 (92%)
Patient-D	7/13 (54%)	13/13 (100%)
IRR (ratio of immune responder patients)	4/4	4/4
Ratio of immunogenic peptides (median)	10/12-13	10/12-13

*Following 1-3 cycles of vaccination

CLAIMS

1. A polypeptide that comprises a fragment of up to 50 consecutive amino acids of a gastric cancer-associated antigen selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 30, 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 gastric cancer-associated antigen.
2. The polypeptide of claim 1, wherein the polypeptide
- a. is a fragment of a gastric cancer-associated antigen selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1, wherein the fragment comprises an amino acid sequence selected from any one of SEQ ID NOs: 1 to 30; or
- b. comprises or consists of two or more fragments of one or more gastric cancer associated antigens selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1, wherein each fragment comprises a different amino acid sequence selected from any one of SEQ ID NOs: 1 to 30, optionally wherein the fragments overlap or are arranged end to end in the polypeptide.
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 DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1; and wherein each fragment comprises a different amino acid sequence selected from SEQ ID NOs: 1 to 30.
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: 31 to 60.

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: 61 to 75.
6. A panel of two or more polypeptides according to any one of claims 1 to 5, wherein each polypeptide comprises a different amino acid sequence selected from SEQ ID NOs: 1 to 30.
7. A pharmaceutical composition or kit comprising one or more polypeptides according to any one of claims 1 to 5, or a panel of polypeptides according to claim 6, or a polypeptide comprising at least two amino acid sequences selected SEQ ID NOs: 1 to 30, or one or more polynucleic acids or vectors encoding said one or more polypeptides.
8. 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 or the peptides, polynucleic acids or vectors of a kit according to claim 7.
9. The method of claim 8 that is a method of treating cancer, optionally gastric cancer.
10. A method of identifying a human subject who will likely have a cytotoxic T cell response to administration of a pharmaceutical composition or the peptides, polynucleic acids or vectors of a kit according to claim 7, the method comprising
- (i) determining that the active ingredient polypeptide(s) of the pharmaceutical composition or kit 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 or the peptides, polynucleic acids or vectors of the kit.
11. The method of claim 10 further comprising using population expression data for each antigen that
- (a) is selected from DPPA2, CAGE-1, TSP50, HIWI, SURVIVIN, 5T4, PRAME, KK-LC-1, MAGE-A2, MAGE-A3, LAGE-1, MAGE-A10, MAGE-A1 and SSX1; 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
- 5 the subject;
- to determine the likelihood that the subject will have a CD8+ T cell response that targets one or more polypeptide antigens that are expressed by cancer cells of the subject.
12. A method of identifying a subject who will likely have a clinical response to a method of
- 10 treatment according to claim 9, 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
- 15 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.
- 20 13. A method of determining the likelihood that a specific human subject will have a clinical response to a method of treatment according to claim 9, 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
- 25 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.

14. The method of claim 13, 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.

15. The method of claim 14, 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).

- 5 16. The method of any one of claims 12 to 15 further comprising selecting or recommending administration of the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit as a method of treatment for the subject, and optionally further treating the subject by administering the pharmaceutical composition or the peptides, polynucleic acids or vectors of the kit.

10

17. A method of treatment according to claim 9, 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 12 to 15.

- 15 18. The method of any one of claims 8, 9, 16 and 17 wherein the treatment is administered in combination with chemotherapy, targeted therapy or a checkpoint inhibitor.

19. A method of identifying a human subject who will likely not have a clinical response to a method of treatment according to claim 9, the method comprising

- 20 (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
(ii) identifying the subject as likely not to have a clinical response to the method of treatment.

25