

subject has been predicted to respond, or to be likely to respond, to administration of the checkpoint inhibitor by a method as described herein.

In another aspect, the disclosure refers to an *in vitro* or *in silico* method of designing or preparing a human subject-specific pharmaceutical composition or kit or panel of polypeptides for treatment of a specific human subject, the method comprising:

(i) selecting a fragment of a polypeptide, which fragment has been identified as immunogenic for the subject by the method as disclosed herein;

(ii) selecting a first sequence of up to 50 consecutive amino acids of the polypeptide, which consecutive amino acids comprise the amino acid sequence of the fragment selected in step (i);

(iii) repeating steps (i) and (ii) to select a second amino acid sequence of up to 50 consecutive amino acids of the same or a different polypeptide to the first amino acid sequence;

(iv) optionally further repeating steps (i) and (ii) to select one or more additional amino acid sequences of up to 50 consecutive amino acids of the same or different polypeptides to the first and second amino acid sequences; and

(vi) designing or preparing a subject-specific pharmaceutical composition, kit or panel of polypeptides having as active ingredients one or more polypeptides that together have all of the amino acid sequences selected in the preceding steps, optionally wherein one or more or each sequence is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the polypeptides.

In yet another aspect, the disclosure refers to use of a pharmaceutical composition, kit or panel of polypeptides in the manufacture of a medicament for treating cancer in a specific human subject, wherein the pharmaceutical composition, kit or panel has been designed or prepared for treatment of the subject according to the method as described in the disclosure, and wherein each of the fragments selected in a step as disclosed herein is a fragment of a different cancer-associated antigen, optionally wherein the medicament is for prophylactic use in the subject.

In yet another aspect, the disclosure refers to a method of designing or preparing a polypeptide for inducing an immune response in a specific human subject the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules or at least three HLA class II molecules of the subject, and designing or
5 preparing a polypeptide comprising the selected amino acid sequence; optionally wherein the method is (a) a method of designing or preparing a polypeptide for inducing a cytotoxic T cell response in a specific human subject, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject,
10 and designing or preparing a polypeptide comprising the selected amino acid sequence; or

(b) a method of designing or preparing a polypeptide for inducing a helper T cell response, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence.

In some embodiments the methods of the disclosure comprise the step of determining or obtaining the HLA class I genotype and/or the HLA class II genotype of the specific human subject.

CLAIMS

1. A method of identifying a fragment of a polypeptide as immunogenic for a specific human subject or predicting whether a polypeptide or a fragment of a polypeptide is immunogenic for a specific human subject, the method comprising the steps of
 - (i) determining whether the polypeptide comprises:
 - (a) an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or
 - (b) an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject; and
 - (ii) (I) identifying said sequence as a fragment of the polypeptide that is immunogenic for the subject; or
 - (II) predicting
 - A. that the polypeptide is immunogenic for the subject if the polypeptide comprises at least one sequence that meets the requirements of step (i); or
 - B. that the polypeptide is not immunogenic for the subject if the polypeptide does not comprise at least one sequence that meets the requirements of step (i).
2. The method of claim 1, wherein:
 - (a) the T cell epitope is capable of binding to at least three HLA class I molecules of the subject and consists of 9 consecutive amino acids of the polypeptide, or wherein the T cell epitope is capable of binding to at least three HLA class II molecules of the subject and consists of 15 consecutive amino acids of the polypeptide;
 - (b) step (i) comprises determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject;
 - (c) step (i) comprises determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject; and/or
 - (d) step (i) comprises determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject and identifying a fragment of the polypeptide that is a T cell epitope capable of binding to at

least one HLA class II molecule of the subject, wherein the HLA class II-binding epitope comprises the amino acid sequence of the HLA class I-binding T cell epitope.

3. The method of claim 1 or claim 2, wherein the polypeptide

(a) is expressed by a pathogenic organism, a virus or a cancer cell, is associated with an autoimmune disorder, or is an allergen or an ingredient of a pharmaceutical composition;

(b) is selected from the antigens listed in Tables 2 to 6;

(c) is an antigen or neoantigen expressed by a cancer cell, optionally wherein the cancer cell, the antigen or the neoantigen is in a sample taken from the subject;

(d) is a mutational neoantigen, optionally wherein

A. the neoantigen is present in a sample obtained from the subject; and/or

B. the immunogenic fragment comprises a neoantigen specific mutation; or

(e) is associated with or suspected of being associated with an autoimmune disorder or an autoimmune response in the subject, wherein determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject identifies the polypeptide and/or the fragment as immunogenic or associated with the autoimmune disorder or autoimmune response in the subject.

4. The method of any one of claims 1 to 3, wherein all of the fragments of the polypeptide that are a T cell epitope capable of binding to at least three HLA class I molecules and/or all of the fragments of the polypeptide that are a T cell epitope capable of binding to at least three HLA class II molecules of the subject are identified, optionally wherein the method is repeated for each polypeptide that is an active ingredient of a specific pharmaceutical composition.

5. The method of any one of the preceding claims, further comprising predicting whether the subject will have a cytotoxic T cell response or a helper T cell response to administration of one or more polypeptide or a pharmaceutical composition or kit comprising one or more polypeptides as active ingredients, wherein

A. a cytotoxic T cell response is predicted if the polypeptide(s) comprises at least one amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject;

B. a helper T cell response is predicted if the polypeptide(s) comprises at least one amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject;

C. no cytotoxic T cell response is predicted if the polypeptide(s) does not comprise any amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or

D. no helper T cell response is predicted if the polypeptide(s) does not comprise any amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject.

6. The method of claim 5, wherein the method further comprises determine the likelihood that the subject will have a cytotoxic T cell response and/or a helper T cell response that targets a polypeptide antigen that is expressed in the subject, the method comprising

(i) identifying one or more polypeptide antigens that comprises an amino acid sequence that is

(a) a T cell epitope capable of binding to at least three HLA class I or at least three HLA class II molecules of the subject; and

(b) comprised in the amino acid sequence of the polypeptide(s); and

(ii) using population expression frequency data for the one or more polypeptide antigens identified in step (i) to determine the likelihood that the subject will have a cytotoxic T cell response and/or a helper T cell response that targets a polypeptide antigen that is expressed in the subject.

7. The method of claim 5, wherein the polypeptide is a component of a pharmaceutical composition and the method comprises determining the likelihood that the subject will develop anti-drug antibodies (ADA) following administration of the polypeptide, wherein a predicted T helper cell response corresponds to a higher likelihood of ADA and no predicted T helper cell response corresponds to a lower likelihood of ADA, optionally wherein the polypeptide is a checkpoint inhibitor.

8. The method of any one of claims 1 to 6 further comprising predicting whether the subject will have a clinical response to administration of a pharmaceutical composition, kit or panel of polypeptides comprising one or more polypeptides as active ingredients, the method comprising

determining whether the one or more active ingredient polypeptides together comprise at least two 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 predicting

A. that the subject will have a clinical response to administration of the pharmaceutical composition, kit or panel of polypeptides if the one or more active ingredient polypeptides together comprise at least two different sequences each of which is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or

B. that the subject will not have a clinical response to administration of the pharmaceutical composition, kit or panel of polypeptides if the one or more active ingredient polypeptides together comprise no more than one sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject, optionally 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).

9. The method of any one of claims 1 to 6 and 8, further comprising determining the likelihood that the specific human subject will have a clinical response to administration of a pharmaceutical composition, kit or panel of polypeptides comprising one or more polypeptides as active ingredients, wherein one or more of the following factors corresponds to a higher likelihood of a clinical response:

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

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

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject; optionally wherein the target polypeptide antigens are expressed in the subject, further optionally wherein the target polypeptides antigens are in one or more samples obtained from the subject;

(c) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens and/or optionally target

polypeptide antigens that have been determined to comprise at least one amino acid sequence that is both

A. comprised in 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 an active ingredient polypeptide; and

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

10. The method of any one of claims 1 to 6, 8 and 9, further comprising determining the likelihood that the specific human subject will have a clinical response to administration of a pharmaceutical composition, kit or panel of polypeptides comprising one or more polypeptides as active ingredients, 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), optionally using population expression data for each antigen identified in step (i) to determine the probability that the subject expresses two or more of the antigens identified in step (i) that together comprise at least two different amino acid sequences of step (i), further optionally 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); 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.

11. The method of claim 9 or claim 10, wherein one or more of the following factors correspond 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 II 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 II of the subject, optionally wherein the target polypeptide antigens are expressed in the subject, optionally wherein the target polypeptides antigens are in one or more samples obtained from the subject;

(c) a higher number of target polypeptide antigens comprising

i. 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; and

ii. 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 II of the subject;

(d) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens, that have been 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 II of the subject

(e) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens, that have been determined to comprise

i. 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; and

ii. 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 II of the subject;

(f) 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 that have been 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 II of the subject; and/or

(g) 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 that have been determined to comprise

i. 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; and

ii. 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 II of the subject.

12. The method of claim 6, further comprising repeating the method for one or more further pharmaceutical compositions, kits or panels of polypeptides and ranking the compositions, kits or panels of polypeptides by their likelihood to induce a clinical response in the subject.

13. The method of any one of claims 1 to 12, further comprising predicting whether administration of the polypeptide, pharmaceutical composition, kit or panel of polypeptides will induce a toxic immune response in the subject, wherein

(a) the polypeptide(s) comprises at least one amino acid sequence that

- i. is capable of binding to at least three HLA class I of the subject; and
- ii. corresponds to a fragment of a human polypeptide expressed in healthy cells;

and a toxic immune response is predicted; or

(b) the polypeptide(s) do not comprise any amino acid sequence that

- i. is capable of binding to at least three HLA class I of the subject; and
- ii. corresponds to a fragment of a human polypeptide expressed in healthy cells;

and no toxic immune response is predicted.

14. The method of any one of the preceding claims further comprising selecting or recommending for treatment of the specific human subject administration to the subject of a polypeptide that comprises a polypeptide fragment that is identified as immunogenic for the subject, or of a polypeptide that is predicted to be immunogenic, or to induce a cytotoxic T cell or helper T cell response, or of a pharmaceutical composition, kit or panel of polypeptides that is predicted to induce a clinical response, or of a polypeptide or pharmaceutical composition that is predicted not to induce a toxic immune response or not to induce ADA in the subject.

15. The method of any one of claims 1 to 4 further comprising predicting whether the subject will have a clinical response to administration of a checkpoint inhibitor to treat cancer, the method comprising determining whether one or more cancer associated antigens together comprise at least two 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 predicting

A. that the subject will have a clinical response to administration of a checkpoint inhibitor if the one or more cancer associated antigens together comprise at least two different sequences each of which is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or

B. that the subject will not have a clinical response to administration of a checkpoint inhibitor if the one or more cancer associated antigens together comprise no more than one

sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject.

16. The method of any one of claims 1 to 4 and 15, wherein the method further comprises determining the likelihood that the subject will have a clinical response to administration of a checkpoint inhibitor to treat cancer, the method comprising

- (i) selecting a plurality of polypeptide antigens that are associated with the cancer type of the subject;
- (ii) identifying which of said cancer associated antigens comprise an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and
- (iii) using population expression data for each cancer associated antigen identified in step to determine the likelihood that the subject will have a clinical response to administration of a checkpoint inhibitor to treat cancer, wherein a higher probability that the subject expresses one or more of the cancer associated antigens identified in step (ii) that together comprise at least two 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 corresponds to a more likely clinical response.

17. The method of claim 16 further comprising selecting or recommending administration of a checkpoint inhibitor for treatment of the subject.

18. Use of a checkpoint inhibitor in the preparation of a medicament for treating cancer in a human subject in need thereof, wherein the subject has been predicted to respond, or to be likely to respond, to administration of the checkpoint inhibitor by a method according to claim 16.

19. The method of any one of claims 5, 7, 8 and 15, wherein the subject has been predicted to have a toxic immune response or ADA development, or not to have a cytotoxic T cell or helper T cell or clinical response, or not to respond to treatment with a checkpoint inhibitor and the method further comprises selecting or recommending a different treatment for the subject.

20. An *in vitro* or *in silico* method of designing or preparing a human subject-specific pharmaceutical composition or kit or panel of polypeptides for treatment of a specific human subject, the method comprising:

- (i) selecting a fragment of a polypeptide, which fragment has been identified as immunogenic for the subject by the method of any one of claims 1 to 3;
- (ii) selecting a first sequence of up to 50 consecutive amino acids of the polypeptide, which consecutive amino acids comprise the amino acid sequence of the fragment selected in step (i);
- (iii) repeating steps (i) and (ii) to select a second amino acid sequence of up to 50 consecutive amino acids of the same or a different polypeptide to the first amino acid sequence;
- (iv) optionally further repeating steps (i) and (ii) to select one or more additional amino acid sequences of up to 50 consecutive amino acids of the same or different polypeptides to the first and second amino acid sequences; and
- (v) designing or preparing a subject-specific pharmaceutical composition, kit or panel of polypeptides having as active ingredients one or more polypeptides that together have all of the amino acid sequences selected in the preceding steps, optionally wherein one or more or each sequence is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the polypeptides.

21. The method of claim 20, wherein the the fragment selected in step (i) is an HLA class I-binding epitope, and the method further comprises selecting a longer fragment of the polypeptide, which longer fragment

- a. comprises the fragment selected in step (i); and
 - b. is a T cell epitope capable of binding to at least three or to the most possible HLA class II molecules of the subject;
- and wherein the sequence of up to 50 consecutive amino acids selected in step (ii) comprises said longer sequence.

22. The method of claim 20 or 21, wherein each polypeptide either consists of one of the selected amino acid sequences, or comprises or consists of two or more of the selected amino acid sequences arranged end to end or overlapping in a single peptide, optionally wherein all of the neoepitopes formed at the join between any two of the selected amino acid sequences

arranged end to end in a single polypeptide have been screened to eliminate polypeptides comprising a neoepitope amino acid sequence that

- (a) corresponds to a fragment of a human polypeptide expressed in healthy cells;
- (b) is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or
- (c) meets both requirements (a) and (b).

23. The method of any one of claims 20 to 22, wherein the one or more polypeptides have been screened to eliminate polypeptides comprising an amino acid sequence that

- (a) corresponds to a fragment of a human polypeptide expressed in healthy cells; or
- (b) corresponds to a fragment of a human polypeptide expressed in healthy cells and is a T cell epitope capable of binding to at least three HLA class I molecules of the subject.

24. Use of a pharmaceutical composition, kit or panel of polypeptides in the manufacture of a medicament for treating cancer in a specific human subject, wherein the pharmaceutical composition, kit or panel has been designed or prepared for treatment of the subject according to the method of any one of claims 20 to 23, and wherein each of the fragments selected in step (i) is a fragment of a different cancer associated antigen, optionally wherein the medicament is for prophylactic use in the subject.

25. The use of claim 24, wherein the polypeptide(s) of the medicament together comprises at least seven different T cell epitopes, each of which is capable of binding to at least three HLA class I molecules and/or at least three HLA class II molecules of the subject.

26. The use according to claim 24 or claim 25, wherein the cancer-associated antigens are selected from the antigens listed in Table 2.

27. The use according to any one of claims 24 to 26, wherein one or more or each of the cancer-associated antigen sequences is flanked at the N and/or C terminus by additional amino acids that are not part of the cancer-associated antigen amino acid sequence .

28. The use according to any one of claims 24 to 27, wherein one or more or each active ingredient peptide or polypeptide of the pharmaceutical composition, kit or panel of polypeptides either

(i) consists of a sequence of up to 50 consecutive amino acids from one of the cancer-associated antigens; or

(ii) comprises or consist of two or more of said sequences of up to 50 consecutive amino acids of the cancer-associated antigens arranged end to end or overlapping in a single peptide.

29. The use according to claim 28, wherein the one or more peptides or polypeptides do not comprise any neoepitopes that span a join between any two of said amino acid sequences that are arranged end to end in a single peptide and that

(1) corresponds to a fragment of a human polypeptide expressed in healthy cells;

(2) is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or

(3) meets both requirements (1) and (2).

30. The use according to any one of claims 24 to 29, wherein the medicament is for administration in combination with chemotherapy, targeted therapy or a checkpoint inhibitor.

31. A method of designing or preparing a polypeptide for inducing an immune response in a specific human subject the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules or at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence; optionally wherein the method is

(a) a method of designing or preparing a polypeptide for inducing a cytotoxic T cell response in a specific human subject, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence; or

(b) a method of designing or preparing a polypeptide for inducing a helper T cell response, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence.