The inventors have developed a binary classifier that is able to separate subjects having cancer from a background population. Using this classifier, the inventors were able to demonstrate a clear association between HLA genotype and cancer risk. These findings confirm the central role of tumor specific T cell responses in the control of tumor growth and mean that HLA genotype analysis may be used to improve diagnostic tests for the early identification of subjects at a high risk of developing cancer.

Accordingly, in a first aspect, the disclosure provides a method for determining the risk that a human subject will develop a cancer, the method comprising quantifying the HLA triplets (HLAT) of the subject that are capable of binding to T cell epitopes in the amino acid sequence of tumor associated antigens (TAAs), wherein each HLA of a HLAT is capable of binding to the same T cell epitope, and determining the risk that the subject will develop a cancer, wherein, with respect to a TAA, a lower number of HLATs capable of binding to T cell epitopes of the TAA corresponds to a higher risk that the subject will develop cancer, and wherein an HLAT is any combination of three class I HLA alleles expressed by the subject.

Accordingly, in a further aspect, the disclosure provides a use of one or more peptides or one of more polynucleic acids or vectors in the manufacture of a medicament for treating cancer in a subject, wherein the subject has been determined to have an elevated risk of developing cancer using a method according to claim 1 or claim 2, and wherein the use comprises that the one or more peptides or one of more polynucleic acids or vectors that encode one or more peptides, that comprise an amino acid sequence that (i) is a fragment of a TAA; and (ii) comprises a T cell epitope capable of binding to an HLAT of the subject; optionally wherein the TAA fragment is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the TAA, are to be administered to the subject.

Accordingly, in a further aspect, the disclosure provides a system for determining the risk that a human subject will develop a cancer, the system comprising: (a) a storage module configured to store data comprising the HLA class I genotype of a subject and the amino acid sequences of TAAs; (b) a computation module configured to quantify the HLAT of the subject that are capable of binding to T cell epitopes in the amino acid sequence of the TAAs, wherein each HLA of a HLAT is capable of binding to the same T cell epitope, and wherein an HLAT is any combination of three class I HLA alleles expressed by the subject; and (c) an output module configured to display an indication of the risk that the subject will develop a cancer and/or a recommended treatment for the subject.

Accordingly, in a first example the disclosure provides a method for determining the risk that a human subject will develop a cancer, the method comprising quantifying the HLA triplets (HLAT) of the subject that are capable of binding to T cell epitopes in the amino acid sequence of tumor associated antigens (TAAs), wherein each HLA of a HLAT is capable of binding to the same T cell epitope, and determining the risk that the subject will develop a cancer, wherein, with respect to a TAA, a lower number of HLATs capable of binding to T cell epitopes of the TAA corresponds to a higher risk that the subject will develop cancer.

The findings described herein also suggest that the risk of cancer can be reduced by using vaccines that are personalised to effectively activate a subject's immune system to kill tumor cells.

Accordingly, in a further example the disclosure provides a method of treating cancer in a subject, wherein the subject has been determined to have an elevated risk of developing cancer using the method above, and wherein the method of treatment comprises administering to the subject one or more peptides or one of more polynucleic acids or vectors that encode one or more peptides, that comprise an amino acid sequence that (i) is a fragment of a TAA; and (ii) comprises a T cell epitope capable of binding to HLAT of the subject.

In further examples, the disclosure provides

- a peptide, or polynucleic acids or vectors that encode a peptide, for use in a method of treating cancer in a specific human subject, wherein the peptides comprises an amino acid sequence that (i) is a fragment of a TAA; and (ii) comprises a T cell epitope capable of binding to an HLAT of the subject ; and
- a peptide, or polynucleic acids or vectors that encode a peptide for use in the manufacture of a medicament for treating cancer in a specific human subject, wherein the peptides comprises an amino acid sequence that (i) is a fragment of a TAA; and (ii) comprises a T cell epitope capable of binding to an HLAT of the subject.

In a further example the disclosure provides a system for determining the risk that a human subject will develop a cancer, the system comprising:

(i) a storage module configured to store data comprising the HLA class I genotype of a subject and the amino acid sequences of TAAs;

(ii) a computation module configured to quantify the HLAT of the subject that are capable of binding to T cell epitopes in the amino acid sequence of the TAAs, wherein each HLA of a HLAT is capable of binding to the same T cell epitope; and

(iii) an output module configured to display an indication of the risk that the subject will develop a cancer and/or a recommended treatment for the subject.

(iv)

The methods and compositions of the present 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 examples of the disclosure set forth are considered to be illustrative and not limiting. Various changes to the described examples 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 examples and preferred features described except where such a combination is clearly impermissible or is stated to be expressly avoided. As used in this specification and the appended claims, the singular forms "a", "an", and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a peptide" includes two or more such peptides.

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

Description of the Figures

Fig. 1

ROC curve of HLA restricted PEPI biomarkers.

Fig. 2

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

CLAIMS

1. A method for determining the risk that a human subject will develop a cancer, the method comprising quantifying the HLA triplets (HLAT) of the subject that are capable of binding to T cell epitopes in the amino acid sequence of tumor associated antigens (TAAs), wherein each HLA of a HLAT is capable of binding to the same T cell epitope, and determining the risk that the subject will develop a cancer, wherein, with respect to a TAA, a lower number of HLATs capable of binding to T cell epitopes of the TAA corresponds to a higher risk that the subject will develop cancer, and wherein an HLAT is any combination of three class I HLA alleles expressed by the subject.

2. The method according to claim 1, wherein the cancer is selected from melanoma, lung cancer, renal cell cancer, colorectal cancer, bladder cancer, glioma, head and neck cancer, ovarian cancer, non-melanoma skin cancer, prostate cancer, kidney cancer, stomach cancer, liver cancer, cervix uteri cancer, oesophagus cancer, non-Hodgkin lymphoma, leukemia, pancreatic cancer, corpus uteri cancer, lip cancer, oral cavity cancer, thyroid cancer, brain cancer, nervous system cancer, gallbladder cancer, larynx cancer, pharynx cancer, myeloma, nasopharynx cancer, Hodgkin lymphoma, testis cancer, breast cancer, gastric cancer, bladder cancer, hepatocellular cancer, pediactric cancer and Kaposi sarcoma.

3. The method according to claim 1 or claim 2, wherein the subject is determined to have an elevated risk of developing cancer, and wherein the method further comprises selecting for the subject a treatment for cancer.

4. The method according to claim 3, wherein the treatment comprises that a peptide, or a polynucleic acid or vector that encodes a peptide, that comprises an amino acid sequence that
   (a) is a fragment of a TAA; and
   (b) comprises a T cell epitope capable of binding to an HLAT of the subject;

optionally wherein the TAA fragment is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the TAA, is to be administered to the subject.

5. The method according to claim 4, wherein the TAA is selected from those listed in Table 2 or Table 11.

6. Use of one or more peptides or one of more polynucleic acids or vectors in the manufacture of a medicament for treating cancer in a subject, wherein the subject has been determined to have an elevated risk of developing cancer using a method according to claim 1 or claim 2, and wherein the use comprises that the one or more peptides or one of more polynucleic acids or vectors that encode one or more peptides, that comprise an amino acid sequence that
   (i) is a fragment of a TAA; and
   (ii) comprises a T cell epitope capable of binding to an HLAT of the subject;
   optionally wherein the TAA fragment is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the TAA, are to be administered to the subject.

7. The use according to claim 6, wherein the TAA is selected from those listed in Table 2 or Table 11.

8. A system for determining the risk that a human subject will develop a cancer, the system comprising:
   (a) a storage module configured to store data comprising the HLA class I genotype of a subject and the amino acid sequences of TAAs;
   (b) a computation module configured to quantify the HLAT of the subject that are capable of binding to T cell epitopes in the amino acid sequence of the TAAs, wherein each HLA of a HLAT is capable of binding to the same T cell epitope, and wherein an HLAT is any combination of three class I HLA alleles expressed by the subject; and
   (c) an output module configured to display an indication of the risk that the subject will develop a cancer and/or a recommended treatment for the subject.