WHAT IS CLAIMED IS:

1. A method for identifying a tumor antigen candidate in a tumor cell sample, the method comprising:

(a) generating a tumor-specific proteome database by:

(i) extracting a set of subsequences (k-mers) comprising at least 33 base pairs from tumor RNA-sequences;

(ii) comparing the set of tumor subsequences of (i) to a set of corresponding control subsequences comprising at least 33 base pairs extracted from RNA-sequences from normal cells;

(iii) extracting the tumor subsequences that are absent in the corresponding control subsequences, thereby obtaining tumor-specific subsequences; and

(iv) *in silico* translating the tumor-specific subsequences, thereby obtaining the tumor-specific proteome database;

(b) generating a personalized tumor proteome database by:

(i) comparing the tumor RNA-sequences to a reference genome sequence to identify single-base mutations in said tumor RNA-sequences;

(ii) inserting the single-base mutations identified in (i) in the reference genome sequence, thereby creating a personalized tumor genome sequence;

(iii) *in silico* translating the expressed protein-coding transcripts from said personalized tumor genome sequence, thereby obtaining the personalized tumor proteome database;

(c) comparing the sequences of major histocompatibility complex (MHC)-associated peptides (MAPs) from said tumor with the sequences of the tumor-specific proteome database of (a) and the personalized tumor proteome database of (b) to identify the MAPs; and

(d) identifying a tumor antigen candidate among the MAPs identified in (c), wherein a tumor antigen candidate is a peptide whose sequence and/or encoding sequence is overexpressed or overrepresented in tumor cells relative to normal cells.

2. The method of claim 1, wherein the above-noted method further comprises (1) isolating and sequencing major histocompatibility complex (MHC)-associated peptides

(MAPs) from the tumor cell sample, and/or (2) performing whole transcriptome sequencing on the tumor cell sample, to obtain the tumor RNA-sequences.

3. The method of claim 2, wherein said isolating MAPs comprises (i) releasing said MAPs from said cell sample by mild acid treatment; and (ii) subjecting the released MAPs to chromatography.

4. The method of claim 3, wherein said method further comprises filtering the released peptides with a size exclusion column prior to said chromatography.

5. The method of any one of claims 1 to 4, wherein said subsequences comprises from 33 to 54 base pairs.

6. The method of any one of claims 1 to 5, further comprising assembling overlapping tumor-specific subsequences into longer tumor subsequences (contigs).

7. The method of claim 6, wherein said size exclusion column has a cut-off of about 3000 Da.

8. The method of any one of claims 1 to 7, wherein said sequencing of MAPs comprises subjecting the isolated MAPs to mass spectrometry (MS) sequencing analysis.

9. The method of any one of claims 1 to 8, wherein said method further comprises generating a personalized normal proteome database using corresponding normal cells.

10. The method of claim 9, wherein said identifying in (d) comprises excluding said MAP if its sequence is detected in the normal personalized proteome database.

11. The method of any one of claims 1 to 10, wherein the method further comprises generating 24- or 39-nucleotide k-mer databases from said tumor RNA-sequences and from RNA-sequences from normal cells to obtain a tumor k-mer database and a normal k-mer database; and comparing the tumor k-mer database and a normal k-mer database to 24- or 39-nucleotide k-mer derived from the MAP encoding sequence, wherein an overexpression or overrepresentation of the k-mer derived from the MAP encoding sequence in said tumor k-mer database relative to said normal k-mer database is indicative that the corresponding MAP is a tumor antigen candidate.

12. The method of claim 11, wherein the k-mer derived from the MAP encoding sequence is overexpressed or overrepresented by at least 10-fold in said tumor k-mer database relative to said normal k-mer database.

13. The method of claim 11 or 12, wherein the k-mer derived from the MAP encoding sequence is absent from said normal k-mer database.

14. The method of any one of claims 1 to 13, wherein said method comprises:

(a) isolating and sequencing MAPs in a tumor cell sample;

(b) performing whole transcriptome sequencing on said tumor cell sample, thereby obtaining tumor RNA-sequences;

(c) generating a tumor-specific proteome database by:

(i) extracting a set of subsequences comprising at least 33 nucleotides from said tumor RNA-sequences;

(ii) comparing the set of tumor subsequences of (i) to a set of corresponding control subsequences comprising at least 33 nucleotides extracted from RNA-sequences from normal cells;

(iii) extracting the tumor subsequences that are absent, or underexpressed by at least 4-fold, in the corresponding control subsequences, thereby obtaining tumor-specific subsequences; and

(iv) *in silico* translating the tumor-specific subsequences, thereby obtaining the tumor-specific proteome database;

(d) generating a personalized tumor proteome database by:

(i) comparing the tumor RNA-sequences to a reference genome sequence to identify single-base mutations in said tumor RNA-sequences;

(ii) inserting the single-base mutations identified in (i) in the reference genome sequence, thereby creating a personalized tumor genome sequence;

(iii) *in silico* translating the expressed protein-coding transcripts from said personalized tumor genome sequence, thereby obtaining the personalized tumor proteome database;

(e) generating a personalized normal proteome database by:

(i) comparing RNA-sequences from normal cells to a reference genome sequence to identify single-base mutations in said normal RNA-sequences;

(ii) inserting the single-base mutations identified in (i) in the reference genome sequence, thereby creating a personalized normal genome sequence;

(iii) *in silico* translating the expressed protein-coding transcripts from said personalized normal genome sequence, thereby obtaining the personalized normal proteome database;

(f) generating a normal and a tumor k-mer database by (i) extracting a set of subsequences comprising at least 24 nucleotides from said RNA-sequences from normal cells and said tumor RNA-sequences;

(g) comparing the sequences of the MAPs obtained in (a) with the sequences of the tumor-specific proteome database of (c) and the personalized tumor proteome database of (d) to identify the MAPs; and

(h) identifying a tumor antigen candidate among the MAPs identified in (f), wherein a tumor antigen candidate corresponds to a MAP (1) whose sequence is not present in the personalized normal proteome database; and (2) (i) whose sequence is present in the personalized tumor proteome database; and/or (ii) whose encoding sequence is overexpressed or overrepresented in said tumor k-mer database relative to said normal k-mer database.

15. The method of any one of claims 1 to 14, wherein said method further comprises selecting MAPs having a length of 8 to 11 amino acids.

16. The method of any one of claims 1 to 15, wherein said normal cells are thymic cells.

17. The method of claim 16, wherein said thymic cells are medullary thymic epithelial cells (mTEC).

18. The method of any one of claims 1 to 17, further comprising comparing the coding sequence of said tumor antigen candidate to sequences from normal tissues.

19. The method of any one of claims 1 to 18, wherein said MAPs have a length of 8 to 11 amino acids.

20. The method of any one of claims 1 to 19, further comprising assessing the binding of the tumor antigen candidate to an MHC molecule.

21. The method of claim 20, wherein said binding is assessed using an MHC binding prediction algorithm.

22. The method of any one of claims 1 to 21, further comprising assessing the frequency of T cells recognizing the tumor antigen candidate in a cell population.

23. The method of claim 22, wherein the frequency of T cells recognizing the tumor antigen candidate is assessed using multimeric MHC class I molecules comprising said tumor antigen candidate in their peptide binding groove.

24. The method of any one of claims 1 to 23, further comprising assessing the ability of the tumor antigen candidate to induce T cell activation.

25. The method of claim 24, wherein the ability of the tumor antigen candidate to induce T cell activation is assessed by measuring cytokine production by T cells contacted with cells having said tumor antigen candidate bound to MHC class I molecules at their cell surface.

26. The method of claim 25, wherein said cytokine production comprises interferon-gamma (IFN-γ) production.

27. The method of any one of claims 1 to 26, further comprising assessing the ability of said tumor antigen candidate to induce T-cell-mediated tumor cell killing and/or to inhibit tumor growth.