

## **RANDOMIZED PEPTIDE LIBRARIES PRESENTED BY HUMAN LEUKOCYTE ANTIGENS**

### **CROSS-REFERENCE TO RELATED APPLICATIONS**

**[0001]** The present application claims the benefit of United States Provisional Application No. 62/726,060, filed August 31, 2018, the contents of which are hereby incorporated by reference in their entirety.

### **BACKGROUND**

**[0002]** T cells are vital to the adaptive immune response, having roles in response to infection and cancer. T cells recognize proteins derived from foreign pathogens as well as self, such as in cases of autoimmunity. Fragments of these proteins (e.g., peptides) are presented by human leukocyte antigen (HLA) molecules and recognized by the T cell via the T cell receptor (TCR).

**[0003]** Major histocompatibility class (MHC) I HLA molecules display peptides generated largely from processing endogenous antigens produced by the cell, such as self-antigens, but also foreign intracellular antigens such as peptides derived from viral proteins, into smaller peptides. Once a peptide is bound into the HLA peptide binding cleft, MHC class I HLA molecules interact with and stimulate CD8<sup>+</sup> cytotoxic T cells. MHC class I has 3 main loci A, B, and C, with each loci divided into many alleles. Alleles refer to the DNA sequence of a gene at the given locus and is usually denoted by at least a four-digit number (e.g., A\*24:02) the first letter designating the locus, a first number defining an allele group (or type) and the second number defining a specific protein within the allele group. A second and third number can be appended indicating silent coding variants and non-coding variants respectively.

**[0004]** Upon recognition of a specific peptide-HLA complex (pHLA), the T cell becomes activated and can (1) become cytotoxic, (2) secrete cytokines, and/or (3) recruit other immune cells. This complex interaction between a foreign or self-peptide, HLA molecule, and TCR is central to identifying how the immune system responds to recognized pathogens at the molecular level. One of

the greatest difficulties in this complex interaction during an immune response is understanding the specificities of TCRs in terms of the identity of the peptides that are recognized. New methods of identifying TCRs and the pHLAs that they recognize are needed.

### SUMMARY

**[0005]** Provided herein in some embodiments are antigen screening libraries comprising a plurality of Human Leukocyte Antigen (HLA)-antigen polypeptide complexes, the HLA-antigen polypeptide complexes comprising (a) an HLA polypeptide, the HLA polypeptide comprising a peptide binding cleft, (b) a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209, wherein the randomized antigen polypeptide specifically binds to the peptide binding cleft of the HLA polypeptide, and (c) a Beta-2 ( $\beta$ 2) microglobulin polypeptide.

**[0006]** In some embodiments, the plurality of HLA-antigen complexes comprises an HLA polypeptide selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises at least five, ten, fifteen, twenty, or twenty-five different HLA polypeptides selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises all of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E HLA polypeptides.

**[0007]** In some embodiments, the plurality of HLA-antigen complexes comprises an HLA polypeptide comprising an amino acid sequence at least 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 427 to 455.

**[0008]** In some embodiments, the plurality of the HLA-antigen polypeptide complexes comprises

at least about  $10^5$  different HLA-antigen polypeptide complexes comprising at least about  $10^5$  different randomized antigen polypeptides.

**[0009]** In some embodiments, the HLA polypeptide, the randomized antigen polypeptide, and the  $\beta$ 2-microglobulin polypeptide comprise a single polypeptide. In some embodiments, the single polypeptide further comprises a first flexible polypeptide linker and a second flexible polypeptide linker. In some embodiments, the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide. In these embodiments, the first flexible polypeptide linker separates the HLA polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the HLA polypeptide from the  $\beta$ 2-microglobulin polypeptide. In some embodiments, the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide. In these embodiments, the first flexible polypeptide linker separates the HLA polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the HLA polypeptide from the  $\beta$ 2-microglobulin polypeptide. In some embodiments, the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide. In these embodiments, the first flexible polypeptide linker separates the randomized antigen polypeptide from the  $\beta$ 2-microglobulin polypeptide, and a second flexible polypeptide linker separates the  $\beta$ 2-microglobulin polypeptide from the HLA polypeptide. In some embodiments, the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide. In these embodiments, the first flexible polypeptide linker separates the HLA polypeptide from the  $\beta$ 2-microglobulin polypeptide, and a second flexible polypeptide linker separates the randomized antigen polypeptide

from the HLA polypeptide. In some embodiments, the  $\beta$ 2-microglobulin polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the randomized antigen polypeptide on the single polypeptide. In these embodiments, the first flexible polypeptide linker separates the HLA polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the randomized antigen polypeptide from the  $\beta$ 2-microglobulin polypeptide. In some embodiments, the randomized antigen polypeptide is C-terminal to the  $\beta$ 2-microglobulin on the single polypeptide, and the HLA polypeptide is C-terminal to the randomized antigen polypeptide on the single polypeptide. In these embodiments, the first flexible polypeptide linker separates the  $\beta$ 2-microglobulin polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the randomized antigen polypeptide from the HLA polypeptide.

**[0010]** In some embodiments, each of the HLA-antigen complexes of the plurality of the HLA-antigen complexes do not comprise an epitope tag. In some embodiments, at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes comprise an epitope tag. In some embodiments, at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does not comprise an epitope tag and at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does comprise an epitope tag. In some embodiments, the epitope tag comprises a FLAG tag, a c-MYC tag, a HIS-tag, a hemagglutinin (HA) tag, a VSVg tag, or a V5 tag.

**[0011]** In some embodiments, the HLA-antigen complexes each comprise a membrane tethering domain. In some embodiments, the membrane tethering domain comprises Aga2. In some embodiments, the antigen screening library is expressed on a plurality of cells.

**[0012]** In some embodiments, the plurality of cells are a plurality of yeast cells. In some embodiments, the plurality of yeast cells are a plurality of yeast cells of the EBY100 strain of *Saccharomyces cerevisiae*.

**[0013]** In some embodiments, each cell of the plurality of cells expresses a specific HLA-antigen complex.

**[0014]** Provided herein in some embodiments are antigen screening libraries comprising a plurality of Human Leukocyte Antigen (HLA)-antigen polypeptide complexes, the HLA-antigen polypeptide complexes comprising an HLA polypeptide, the HLA polypeptide comprising a peptide binding cleft, and a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209, wherein the randomized antigen polypeptide specifically binds to the peptide binding cleft of the HLA polypeptide.

**[0015]** In some embodiments, the plurality of HLA-antigen complexes comprises an HLA polypeptide selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises at least five, ten, fifteen, twenty, or twenty-five different HLA polypeptides selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises all of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E HLA polypeptides.

**[0016]** In some embodiments, the plurality of HLA-antigen complexes comprises an HLA polypeptide comprising an amino acid sequence at least 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 427 to 455.

**[0017]** In some embodiments, the plurality of the HLA-antigen polypeptide complexes comprises at least about  $10^5$  different HLA-antigen polypeptide complexes comprising at least about  $10^5$  different randomized antigen polypeptides.

**[0018]** In some embodiments, the HLA polypeptide, the randomized antigen polypeptide, and the

$\beta$ 2-microglobulin polypeptide comprise a single polypeptide. In some embodiments, the single polypeptide further comprises a first flexible polypeptide linker separating the HLA polypeptide from the randomized antigen polypeptide. In certain of these embodiments, the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide. In certain of these embodiments, the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide.

**[0019]** In some embodiments, each of the HLA-antigen complexes of the plurality of the HLA-antigen complexes do not comprise an epitope tag. In some embodiments, at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes comprise an epitope tag. In some embodiments, at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does not comprise an epitope tag and at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does comprise an epitope tag. In some embodiments, the epitope tag comprises a FLAG tag, a c-MYC tag, a HIS-tag, a hemagglutinin (HA) tag, a VSVg tag, or a V5 tag.

**[0020]** In some embodiments, the HLA-antigen complexes each comprise a membrane tethering domain. In some embodiments, the membrane tethering domain comprises Aga2. In some embodiments, the antigen screening library is expressed on a plurality of cells.

**[0021]** In some embodiments, the plurality of cells are a plurality of yeast cells. In some embodiments, the plurality of yeast cells are a plurality of yeast cells of the EBY100 strain of *Saccharomyces cerevisiae*.

**[0022]** In some embodiments, each cell of the plurality of cells expresses a specific HLA-antigen complex.

**[0023]** Provided herein in some embodiments are antigen screening libraries comprising a plurality of antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes, the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes. In these embodiments, the antigen

screening libraries further comprise a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209, wherein the randomized antigen polypeptide specifically binds to the peptide binding cleft of the HLA polypeptide; and a Beta-2 ( $\beta$ 2) microglobulin polypeptide. In these embodiments, the antigen screening libraries also further comprise a plurality of HLA polypeptides constitutively expressed by one or more yeast cells and comprising a peptide binding cleft.

**[0024]** In some embodiments, the plurality of HLA-antigen complexes comprises an HLA polypeptide selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises at least five, ten, fifteen, twenty, or twenty-five different HLA polypeptides selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises all of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E HLA polypeptides.

**[0025]** In some embodiments, the plurality of HLA-antigen complexes comprises an HLA polypeptide comprising an amino acid sequence at least 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 427 to 455.

**[0026]** In some embodiments, the plurality of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes comprises at least about  $10^5$  different antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes comprising at least about  $10^5$  different randomized antigen polypeptides.

**[0027]** In some embodiments, the randomized antigen polypeptide and the  $\beta$ 2-microglobulin polypeptide comprise a single polypeptide. In some embodiments, the single polypeptide further

comprises a first flexible polypeptide linker. In certain of these embodiments, the randomized antigen polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide. In certain of these embodiments, the randomized antigen polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

**[0028]** In some embodiments, each of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes of the plurality of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes do not comprise an epitope tag. In some embodiments, at least one of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes of the plurality of antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes comprise an epitope tag. In some embodiments, at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does not comprise an epitope tag and at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does comprise an epitope tag. In some embodiments, the epitope tag comprises a FLAG tag, a c-MYC tag, a HIS-tag, a hemagglutinin (HA) tag, a VSVg tag, or a V5 tag.

**[0029]** In some embodiments, the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes each comprise a membrane tethering domain. In some embodiments, the membrane tethering domain comprises Aga2. In some embodiments, the antigen screening library is expressed on a plurality of cells.

**[0030]** In some embodiments, the plurality of cells are a plurality of yeast cells. In some embodiments, the plurality of yeast cells are a plurality of yeast cells of the EBY100 strain of *Saccharomyces cerevisiae*.

**[0031]** In some embodiments, each cell of the plurality of cells expresses a specific antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complex.

**[0032]** Provided herein in some embodiments are a plurality of nucleic acids encoding the antigen screening libraries in accordance with the present technology.

[0033] In some embodiments, the HLA polypeptide of the HLA-antigen complex is encoded by a nucleic acid that is at least about 85%, 87.5%, 90%, 95%, 97%, 98%, or 99% homologous to any one of SEQ ID NOs: 456 to 484. In some embodiments, the randomized antigen polypeptide of the HLA-antigen complex is encoded by a nucleic acid set forth in any one of SEQ ID NOs: 210 to 426.

[0034] In some embodiments, the plurality of nucleic acids is expressed by a plurality of cells.

[0035] Provided herein in some embodiments are a plurality of cells expressing the antigen screening library in accordance with the present technology.

[0036] In some embodiments, the plurality of cells is a plurality of yeast cells. In some embodiments, the plurality of yeast cells is a plurality of cells of the EBY100 strain of *Saccharomyces cerevisiae*. In some embodiments, each cell of the plurality of cells comprises a nucleic acid of the plurality of nucleic acids encoding a specific of HLA-antigen complex.

[0037] Provided herein in some embodiments are methods of selecting an antigen comprising contacting the plurality of cells in accordance with the present technology with a T cell receptor (TCR).

[0038] In some embodiments, the TCR is immobilized on a substrate. In some embodiments, the TCR is expressed by a cell.

[0039] In some embodiments, the selection is repeated for 2, 3, 4, or 5 cycles.

[0040] In some embodiments, the antigen is a polypeptide antigen. In some embodiments, the antigen is a polypeptide antigen that does not naturally occur. In some embodiments, the antigen is a polypeptide antigen that does not naturally occur in a human.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0041] **FIG. 1A** illustrates a schematic of an HLA antigen polypeptide construct coupled to a yeast cell in accordance with some embodiments of the present technology.

[0042] **FIG. 1B** illustrates an exemplary, non-limiting, embodiment of an HLA antigen polypeptide construct tethered to a cell in accordance with some embodiments of the present

technology.

[0043] **FIG. 2** illustrates an exemplary, non-limiting, depiction of a process for selecting a specific randomized antigen polypeptide that interacts with a specific T cell receptor in accordance with some embodiments of the present technology.

[0044] **FIGS. 3A and 3B** are maps of an example pCT vector (**FIG. 3A**) and an example pYAL vector (**FIG. 3B**).

[0045] **FIG. 4** illustrates characterization by flow cytometry of peptide-HLA (pHLA) expression on yeast surface for a plurality of allotypes in accordance with some embodiments of the present technology.

#### DETAILED DESCRIPTION

[0046] Described herein are antigen screening libraries useful for selection and/or identification of polypeptide ligands for T cell receptors (TCRs). In many cases, the antigen screening libraries are useful to discover polypeptide antigens that are capable of interacting with and stimulating human T cells as TCR ligands, including both endogenous TCR antigens and non-endogenous TCR antigens which may be novel TCR antigens and/or novel epitopes. Such novel antigens and/or novel epitopes are useful, at least for example, to stimulate one or more TCRs on T cells that may have become exhausted or anergized, and revive immune responses against cancer, tumors, or chronic viral infections. Accordingly, the present disclosure includes peptide library display, such as randomized peptide antigen libraries, in the context of a given HLA to determine the specificities and general recognition properties of TCRs restricted to HLA-mediated peptide recognition.

[0047] Once expressed using the methodologies described herein, a randomized peptide antigen library may be displayed by HLA molecules that are expressed on the surface of cells. In general, the cells that display these HLA-antigen polypeptide complexes are not normal antigen presenting cells of a host's immune system but rather are cells that can easily be transformed, transfected, transduced,

and/or electroporated with a nucleic acid encoding an HLA-antigen polypeptide, including without limitation, insect cells, yeast cells, and bacterial cells. In some embodiments, the randomized peptide antigen library is expressed by yeast cells. A mixture of plasmids that encode at least  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ ,  $10^{10}$ ,  $10^{11}$ ,  $10^{12}$ ,  $10^{13}$ ,  $10^{14}$ , or  $10^{15}$  distinct polypeptide antigens, and either one or a plurality of different HLA molecules, are transformed into yeast cells. Following transformation with the randomized peptide antigen library, the yeast cells that express the HLA-antigen polypeptide complex library are then contacted by a TCR, or other macromolecule having one or more antigen binding domains, serving as a bait. The TCRs are either (1) expressed by a cell or (2) recombinantly produced and, optionally, multimerized and/or immobilized, on a solid structure, such as a bead, or via a protein scaffold such as streptavidin or streptavidin conjugated dextran (referenced as the selection reagent). The cells expressing HLA-antigen polypeptide complexes that interact with the TCR selection reagent can be selected by an appropriate modality, and after 2, 3, 4, 5, 6, 7 or more rounds of enrichment (e.g., cycles) the nucleic acids encoding the HLA-antigen polypeptide complexes can be extracted from the enriched cells and sequencing can be performed to determine the polypeptide antigens that have been enriched. The enriched polypeptide antigens define the structural attributes that interact with a given TCR.

**[0048]** In some embodiments, the present disclosure includes an antigen screening library which comprises a plurality of HLA-antigen polypeptide complexes. In some embodiments, the HLA-antigen polypeptide complexes comprise (a) an HLA polypeptide, the HLA polypeptide comprising a peptide binding cleft; (b) a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 194, wherein the randomized antigen polypeptide is selected to specifically bind to the peptide binding cleft of the HLA polypeptide; and (c) a beta-2 ( $\beta 2$ ) microglobulin polypeptide. Also provided herein are derivatives of randomized peptide antigens and libraries thereof, compositions thereof, pharmaceutical compositions thereof, and uses of the same.

Also provided herein are nucleic acid sequences encoding one or more randomized peptide antigen libraries disclosed herein and derivatives thereof, and methods for expressing the one or more randomized peptide antigen libraries, peptides thereof, and derivatives thereof in one or more cells.

[0049] As set forth in the examples provided herein, a randomized peptide antigen library was designed (Example 1) and includes nucleic acid constructs (FIG. 1A) and peptide constructs tethered to a cell, such as a yeast cell (FIG. 1B). Expression of pHLA was characterized and validated using a yeast display (YD) system (Example 2). These pHLAs can interact with a TCR and determining whether interaction occurs can be determined with one or more processes described herein, such as was performed using the process illustrated in FIG. 2. Expression of pHLA was validated by flow cytometry (Example 2, Level 1) and can further be functionally validated by screening the randomized peptide antigen library using a candidate allotype-matched TCR (Example 2, Level 2).

[0050] The following description of the invention is merely intended to illustrate various embodiments of the present disclosure. As such, the specific modifications discussed are not to be construed as limitations on the scope of the present disclosure. It will be apparent to one skilled in the art that various equivalents, changes, and modifications may be made without departing from the scope of the present disclosure, and it is understood that such equivalent embodiments are to be included herein.

[0051] All references listed herein are incorporated by reference, in their entirety. Methods and apparatuses are provided here by way of example and are not intended to be limiting to the present disclosure.

### **Certain Definitions**

[0052] In the following description, some specific details are set forth in order to provide a thorough understanding of various embodiments. However, one skilled in the art will understand that the embodiments provided may be practiced without these details. Unless the context requires

otherwise, throughout the specification and claims which follow, the word “comprise” and variations thereof, such as, “comprises” and “comprising” are to be construed in an open, inclusive sense, that is, as “including, but not limited to.” 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. It should also be noted that the term “or” is generally employed in its sense including “and/or” unless the content clearly dictates otherwise. Further, headings provided herein are for convenience only and do not interpret the scope or meaning of the claimed embodiments.

**[0053]** The terms “peptide,” “polypeptide,” and “protein” are used interchangeably to refer to a polymer of amino acid residues, and are not limited to a minimum length, though a number of amino acid residues may be specified (e.g., 9mer is nine amino acid residues). Polypeptides may include amino acid residues including natural and/or non-natural amino acid residues. The terms also include post-expression modifications of the polypeptide, for example, glycosylation, sialylation, acetylation, phosphorylation, and the like. In some embodiments, the polypeptides may contain modifications with respect to a native or natural sequence, as long as the protein maintains the desired activity. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through mutations of hosts which produce the proteins or errors due to PCR amplification.

**[0054]** The term “acidic residue” refers to amino acid residues in D- or L-form having sidechains comprising acidic groups. Exemplary acidic residues include D and E.

**[0055]** The term “amide residue” refers to amino acids in D- or L-form having sidechains comprising amide derivatives of acidic groups. Exemplary residues include N and Q.

**[0056]** The term “aromatic residue” refers to amino acid residues in D- or L-form having sidechains comprising aromatic groups. Exemplary aromatic residues include F, Y, and W.

**[0057]** The term “basic residue” refers to amino acid residues in D- or L-form having sidechains comprising basic groups. Exemplary basic residues include H, K, and R.

**[0058]** The term “hydrophilic residue” refers to amino acid residues in D- or L-form having sidechains comprising polar groups. Exemplary hydrophilic residues include C, S, T, N, and Q.

**[0059]** The term “nonfunctional residue” refers to amino acid residues in D- or L-form having sidechains that lack acidic, basic, or aromatic groups. Exemplary nonfunctional amino acid residues include M, G, A, V, I, L and norleucine (Nle).

**[0060]** The term “neutral hydrophobic residue” refers to amino acid residues in D- or L-form having sidechains that lack basic, acidic, or polar groups. Exemplary neutral hydrophobic amino acid residues include A, V, L, I, P, W, M, and F.

**[0061]** The term “polar hydrophobic residue” refers to amino acid residues in D- or L-form having sidechains comprising polar groups. Exemplary polar hydrophobic amino acid residues include T, G, S, Y, C, Q, and N.

**[0062]** The term “hydrophobic residue” refers to amino acid residues in D- or L-form having sidechains that lack basic or acidic groups. Exemplary hydrophobic amino acid residues include A, V, L, I, P, W, M, F, T, G, S, Y, C, Q, and N.

**[0063]** “Percent (%) sequence identity” with respect to a reference polypeptide sequence is the percentage of amino acid residues in a candidate sequence that is identical with the amino acid residues in the reference polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are known, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software, or other software appropriate for nucleic acid sequences. Appropriate parameters for aligning sequences are able to be determined, including algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity

values are generated using the sequence comparison computer program ALIGN-2. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc., and the source code has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available from Genentech, Inc., South San Francisco, Calif., or may be compiled from the source code. The ALIGN-2 program should be compiled for use on a UNIX operating system, including digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

**[0064]** In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a some % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows: 100 times the fraction  $X/Y$ , where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

**[0065]** As used herein, the terms "homologous," "homology," or "percent homology" when used herein to describe to a nucleic acid sequence, relative to a reference sequence, can be determined using the formula described by Karlin & Altschul 1990, modified as in Karlin & Altschul 1993. Such a formula is incorporated into the basic local alignment search tool (BLAST) programs of Altschul 1990. Percent homology of sequences can be determined using the most recent version of BLAST, as of the filing date of this application.

**[0066]** “T cell receptor” (TCR), refers to an antigen/MHC binding heterodimeric protein product of a vertebrate, e.g. mammalian, TCR gene complex, including the human TCR  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  chains. For example, the complete sequence of the human  $\beta$  TCR locus has been sequenced, as published by Rowen 1996; the human TCR locus has been sequenced and resequenced, for example see Mackelprang 2006; see a general analysis of the T-cell receptor variable gene segment families in Arden 1995; each of which is herein specifically incorporated by reference for the sequence information provided and referenced in the publication.

**[0067]** “Bait” refers to a TCR or “other macromolecule having one or more antigen binding domains” that binds to an antigen of the present technology. The other macromolecule having one or more antigen binding domains is an antibody, a DARPin, or a synthetic molecule, including aptamers. The antigen binding domain binds a peptide, such as one or more of the HLA-peptide complexes of the present technology, or a nucleic acid, such as DNA and RNA.

**[0068]** “Exogenous” with respect to a nucleic acid or polynucleotide indicates that the nucleic acid is part of a recombinant nucleic acid construct or is not in its natural environment. For example, an exogenous nucleic acid can be a sequence from one species introduced into another species, i.e., a heterologous nucleic acid. Typically, such an exogenous nucleic acid is introduced into the other species via a recombinant nucleic acid construct. An exogenous nucleic acid also can be a sequence that is native to an organism and that has been reintroduced into cells of that organism. An exogenous nucleic acid that includes a native sequence can often be distinguished from the naturally occurring sequence by the presence of non-natural sequences linked to the exogenous nucleic acid, e.g., non-native regulatory sequences flanking a native sequence in a recombinant nucleic acid construct. In addition, stably transformed exogenous nucleic acids typically are integrated at positions other than the position where the native sequence is found. The exogenous elements may be added to a construct, for example, using genetic recombination. Genetic recombination is the breaking and rejoining of

DNA strands to form new molecules of DNA encoding a novel set of genetic information.

[0069] As used herein the term “about” refers to an amount that is near the stated amount by 10%.

### **Structural Characteristics of the HLA-Antigen Polypeptide Complexes**

[0070] Disclosed herein are antigen screening libraries, such as randomized peptide antigen libraries, which include a plurality of HLA-antigen polypeptide complexes. The HLA-antigen polypeptide complexes of the current disclosure minimally comprise at least three constituents: (a) a randomized antigen polypeptide, (b) a major histocompatibility class I (MHC I) HLA molecule, and (c) a  $\beta$ 2-microglobulin. In some embodiments, the randomized antigen polypeptide of (a) is randomized having at least one or more residues conserved that serve as anchor residues to bind to an HLA molecule of a specific type. Exemplary, but not limiting, randomized antigen polypeptide antigens and the HLA type with which they associate are shown in **Table 1** and given by SEQ ID NOs: 1 to 194 and **Table 2** and given by SEQ ID NOs: 195 to 209. In some embodiments, the randomized polypeptide antigens comprises a sequence that is at least about 70%, 75%, 80%, 85%, 87%, 87.5%, 90%, 95%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5%, or 100% identical to any one of, but not limited to, the amino acid sequences set forth in any one of SEQ ID NOs: 1 to 194 and SEQ ID NOs: 195 to 209. In some embodiments, the randomized polypeptide antigens comprise a sequence identical to any one of those set forth in any one of SEQ ID NOs: 1 to 194 and SEQ ID NOs: 195 to 209. Also envisioned within the present disclosure are randomized polypeptide antigen truncations that have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 25 amino acids truncated from the N-terminus or truncated from the C-terminus of any one of SEQ ID NOs: 1 to 194 and SEQ ID NOs: 195 to 209. In some embodiments, the HLA molecule of (b) is a HLA polypeptide and comprises a peptide binding cleft. Once expressed, in some embodiments, the randomized antigen polypeptide of (a) binds the HLA polypeptide of (b) at the peptide binding cleft.

**Table 1: Polypeptide Antigen Sequences Sorted By HLA Type**

| HLA type | SEQ ID NO: | Sequence               |
|----------|------------|------------------------|
| HLA A3   | 1          | X(V/L/M)XXXXXXK        |
|          | 2          | X(V/L/M)XXXXXXXK       |
|          | 3          | X(V/L/M)XXXXXXXXK      |
|          | 4          | X(V/L/M)XXXXXXXXXK     |
|          | 5          | X(V/L/M)XXXXXXXXXXK    |
| HLA A11  | 6          | X(V/L/F)XXXXX(K/R)     |
|          | 7          | X(V/L/F)XXXXXX(K/R)    |
|          | 8          | X(V/L/F)XXXXXXX(K/R)   |
|          | 9          | X(V/L/F)XXXXXXXX(K/R)  |
|          | 10         | X(V/L/F)XXXXXXXXX(K/R) |
| HLA A23  | 11         | XXXXXXXXY              |
|          | 12         | XXXXXXXXXY             |
|          | 13         | XXXXXXXXXXY            |
|          | 14         | XXXXXXXXXXXY           |
|          | 15         | XXXXXXXXXXXY           |
| HLA A24  | 16         | X(F/Y)XXXXX(I/L/F)     |
|          | 17         | X(F/Y)XXXXXX(I/L/F)    |
|          | 18         | X(F/Y)XXXXXXX(I/L/F)   |
|          | 19         | X(F/Y)XXXXXXXX(I/L/F)  |
|          | 20         | X(F/Y)XXXXXXXXX(I/L/F) |
| HLA A26  | 21         | X(V/L/F)XXXXX(F/Y)     |
|          | 22         | X(I/T)XXXXX(F/Y)       |
|          | 23         | X(V/L/F)XXXXXX(F/Y)    |
|          | 24         | X(I/T)XXXXXX(F/Y)      |
|          | 25         | X(V/L/F)XXXXXXX(F/Y)   |
|          | 26         | X(I/T)XXXXXXX(F/Y)     |
|          | 27         | X(V/L/F)XXXXXXXX(F/Y)  |
|          | 28         | X(I/T)XXXXXXXX(F/Y)    |
|          | 29         | X(V/L/F)XXXXXXXXX(F/Y) |
|          | 30         | X(I/T)XXXXXXXXX(F/Y)   |
| HLA A30  | 31         | X(F/Y)XXXXX(L)         |
|          | 32         | X(F/Y)XXXXXX(L)        |
|          | 33         | X(F/Y)XXXXXXX(L)       |
|          | 34         | X(F/Y)XXXXXXXX(L)      |
|          | 35         | X(F/Y)XXXXXXXXX(L)     |

|         |    |                      |
|---------|----|----------------------|
| HLA A31 | 36 | XXXXXXXX(K/R)        |
|         | 37 | XXXXXXXX(K/R)        |
|         | 38 | XXXXXXXX(K/R)        |
|         | 39 | XXXXXXXXXX(K/R)      |
|         | 40 | XXXXXXXXXXXX(K/R)    |
| HLA A33 | 41 | XXXXXXX(K/R)         |
|         | 42 | XXXXXXX(K/R)         |
|         | 43 | XXXXXXXXXX(K/R)      |
|         | 44 | XXXXXXXXXXXX(K/R)    |
|         | 45 | XXXXXXXXXXXX(K/R)    |
| HLA A68 | 46 | X(V)XXXXX(K/R)       |
|         | 47 | X(V)XXXXX(K/R)       |
|         | 48 | X(V)XXXXXXX(K/R)     |
|         | 49 | X(V)XXXXXXXX(K/R)    |
|         | 50 | X(V)XXXXXXXXX(K/R)   |
|         | 51 | X(T)XXXXX(K/R)       |
|         | 52 | X(T)XXXXX(K/R)       |
|         | 53 | X(T)XXXXXXX(K/R)     |
|         | 54 | X(T)XXXXXXXX(K/R)    |
|         | 55 | X(T)XXXXXXXXX(K/R)   |
| HLA B7  | 56 | XPXXXXXL             |
|         | 57 | XPXXXXXL             |
|         | 58 | XPXXXXXXXL           |
|         | 59 | XPXXXXXXXXXL         |
|         | 60 | XPXXXXXXXXXXL        |
| HLA B8  | 61 | XX(K)X(K/R)XX(L)     |
|         | 62 | XX(K)X(K/R)XXX(L)    |
|         | 63 | XX(K)X(K/R)XXXX(L)   |
|         | 64 | XX(K)X(K/R)XXXXX(L)  |
|         | 65 | XX(K)X(K/R)XXXXXX(L) |
| HLA B15 | 66 | X(Q/L)XXXXX(F/Y)     |
|         | 67 | X(Q/L)XXXXXX(F/Y)    |
|         | 68 | X(Q/L)XXXXXXX(F/Y)   |
|         | 69 | X(Q/L)XXXXXXXX(F/Y)  |
|         | 70 | X(Q/L)XXXXXXXXX(F/Y) |
| HLA B27 | 71 | X(R)XXXXX(F/Y)       |
|         | 72 | X(R)XXXXXX(F/Y)      |
|         | 73 | X(R)XXXXXXX(F/Y)     |

|         |     |                      |
|---------|-----|----------------------|
|         | 74  | X(R)XXXXXXXXX(F/Y)   |
|         | 75  | X(R)XXXXXXXXX(F/Y)   |
| HLA B35 | 76  | X(P)XXXXXX(F/Y)      |
|         | 77  | X(P)XXXXXX(F/Y)      |
|         | 78  | X(P)XXXXXX(F/Y)      |
|         | 79  | X(P)XXXXXXXXX(F/Y)   |
|         | 80  | X(P)XXXXXXXXX(F/Y)   |
|         | 81  | X(P)XXXXXX(M/L/I)    |
|         | 82  | X(P)XXXXXX(M/L/I)    |
|         | 83  | X(P)XXXXXXXXX(M/L/I) |
|         | 84  | X(P)XXXXXXXXX(M/L/I) |
|         | 85  | X(P)XXXXXXXXX(M/L/I) |
| HLA B40 | 86  | X(E)XXXXXX(L)        |
|         | 87  | X(E)XXXXXX(L)        |
|         | 88  | X(E)XXXXXX(L)        |
|         | 89  | X(E)XXXXXXXXX(L)     |
|         | 90  | X(E)XXXXXXXXX(L)     |
| HLA B51 | 91  | X(A/G)XXXXXX(V/I)    |
|         | 92  | X(A/G)XXXXXX(V/I)    |
|         | 93  | X(A/G)XXXXXXXXX(V/I) |
|         | 94  | X(A/G)XXXXXXXXX(V/I) |
|         | 95  | X(A/G)XXXXXXXXX(V/I) |
|         | 96  | X(P)XXXXXX(V/I)      |
|         | 97  | X(P)XXXXXX(V/I)      |
|         | 98  | X(P)XXXXXX(V/I)      |
|         | 99  | X(P)XXXXXXXXX(V/I)   |
|         | 100 | X(P)XXXXXXXXX(V/I)   |
| HLA B44 | 101 | X(E)XXXXXX(F/Y)      |
|         | 102 | X(E)XXXXXX(F/Y)      |
|         | 103 | X(E)XXXXXXXXX(F/Y)   |
|         | 104 | X(E)XXXXXXXXX(F/Y)   |
|         | 105 | X(E)XXXXXXXXX(F/Y)   |
| HLA B53 | 106 | X(P)XXXXXX(W)        |
|         | 107 | X(P)XXXXXX(W)        |
|         | 108 | X(P)XXXXXX(W)        |
|         | 109 | X(P)XXXXXXXXX(W)     |
|         | 110 | X(P)XXXXXXXXX(W)     |
|         | 111 | X(P)XXXXXX(F/L)      |

|         |     |                         |
|---------|-----|-------------------------|
|         | 112 | X(P)XXXXXXX(F/L)        |
|         | 113 | X(P)XXXXXXX(F/L)        |
|         | 114 | X(P)XXXXXXX(F/L)        |
|         | 115 | X(P)XXXXXXXXXX(F/L)     |
| HLA B57 | 116 | X(A/T/S)XXXXXX(F/Y)     |
|         | 117 | X(A/T/S)XXXXXX(F/Y)     |
|         | 118 | X(A/T/S)XXXXXXX(F/Y)    |
|         | 119 | X(A/T/S)XXXXXXXXXX(F/Y) |
|         | 120 | X(A/T/S)XXXXXXXXXX(F/Y) |
|         | 121 | X(A/T/S)XXXXXX(W)       |
|         | 122 | X(A/T/S)XXXXXX(W)       |
|         | 123 | X(A/T/S)XXXXXXXX(W)     |
|         | 124 | X(A/T/S)XXXXXXXXXX(W)   |
|         | 125 | X(A/T/S)XXXXXXXXXX(W)   |
| HLA C1  | 126 | X(L)XXXXX(L)            |
|         | 127 | X(L)XXXXXX(L)           |
|         | 128 | X(L)XXXXXXX(L)          |
|         | 129 | X(L)XXXXXXXXXX(L)       |
|         | 130 | X(L)XXXXXXXXXX(L)       |
|         | 131 | X(A)XXXXXXX(L)          |
|         | 132 | X(A)XXXXXXX(L)          |
|         | 133 | X(A)XXXXXXX(L)          |
|         | 134 | X(A)XXXXXXXXXX(L)       |
|         | 135 | X(A)XXXXXXXXXX(L)       |
| HLA C2  | 136 | X(A)XXXXX(L/V)          |
|         | 137 | X(A)XXXXXX(L/V)         |
|         | 138 | X(A)XXXXXXX(L/V)        |
|         | 139 | X(A)XXXXXXXXXX(L/V)     |
|         | 140 | X(A)XXXXXXXXXX(L/V)     |
|         | 141 | X(A)XXXXXX(F/Y)         |
|         | 142 | X(A)XXXXXX(F/Y)         |
|         | 143 | X(A)XXXXXXX(F/Y)        |
|         | 144 | X(A)XXXXXXXXXX(F/Y)     |
|         | 145 | X(A)XXXXXXXXXX(F/Y)     |
| HLA C3  | 146 | XXXXXXXXX(L/F/M/I)      |
|         | 147 | XXXXXXXXX(L/F/M/I)      |
|         | 148 | XXXXXXXXXX(L/F/M/I)     |
|         | 149 | XXXXXXXXXX(L/F/M/I)     |

|        |     |                         |
|--------|-----|-------------------------|
|        | 150 | XXXXXXXXXXXX(L/F/M/I)   |
| HLA C4 | 151 | X(Y/F)XXXXXX(L/F/M/I)   |
|        | 152 | X(Y/F)XXXXXX(L/F/M/I)   |
|        | 153 | X(Y/F)XXXXXX(L/F/M/I)   |
|        | 154 | X(Y/F)XXXXXX(L/F/M/I)   |
|        | 155 | X(Y/F)XXXXXXXX(L/F/M/I) |
|        | 156 | X(P)XXXXXX(L/F/M/I)     |
|        | 157 | X(P)XXXXXX(L/F/M/I)     |
|        | 158 | X(P)XXXXXX(L/F/M/I)     |
|        | 159 | X(P)XXXXXXXX(L/F/M/I)   |
|        | 160 | X(P)XXXXXXXX(L/F/M/I)   |
| HLA C5 | 161 | XX(D)XXXX(L/V)          |
|        | 162 | XX(D)XXXX(L/V)          |
|        | 163 | XX(D)XXXX(L/V)          |
|        | 164 | XX(D)XXXX(L/V)          |
|        | 165 | XX(D)XXXXXXXX(L/V)      |
| HLA C6 | 166 | XXXXXX(L/V/I)           |
|        | 167 | XXXXXXXX(L/V/I)         |
|        | 168 | XXXXXXXX(L/V/I)         |
|        | 169 | XXXXXXXXXX(L/V/I)       |
|        | 170 | XXXXXXXXXX(L/V/I)       |
|        | 171 | XXXXXX(Y)               |
|        | 172 | XXXXXX(Y)               |
|        | 173 | XXXXXXXX(Y)             |
|        | 174 | XXXXXXXXXX(Y)           |
|        | 175 | XXXXXXXXXX(Y)           |
| HLA C7 | 176 | XXXXXX(F/L)             |
|        | 177 | XXXXXX(F/L)             |
|        | 178 | XXXXXX(F/L)             |
|        | 179 | XXXXXX(F/L)             |
|        | 180 | XXXXXXXXXX(F/L)         |
|        | 181 | XXXXXX(Y)               |
|        | 182 | XXXXXX(Y)               |
|        | 183 | XXXXXXXX(Y)             |
|        | 184 | XXXXXXXXXX(Y)           |
|        | 185 | XXXXXXXXXX(Y)           |
| HLA C8 | 186 | XX(D)XXX(L)             |
|        | 187 | XX(D)XXXX(L)            |

|       |     |                     |
|-------|-----|---------------------|
| HLA E | 188 | XX(D)XXXXXX(L)      |
|       | 189 | XX(D)XXXXXX(L)      |
|       | 190 | XX(D)XXXXXXXX(L)    |
|       | 191 | X(L/M)XXXXXX(L/V)   |
|       | 192 | X(L/M)XXXXXX(L/V)   |
|       | 193 | X(L/M)XXXXXXXX(L/V) |
|       | 194 | X(L/M)XXXXXXXX(L/V) |

**Table 2: Additional Polypeptide Antigen Sequences for HLA A11**

| SEQ ID NO: | Sequence              |
|------------|-----------------------|
| 195        | X(I/L/V)XXXXXX(K/R)   |
| 196        | X(I/L/V)XXXXXX(K/R)   |
| 197        | X(I/L/V)XXXXXXXX(K/R) |
| 198        | X(I/L/V)XXXXXXXX(K/R) |
| 199        | X(I/L/V)XXXXXXXX(K/R) |
| 200        | X(Y/F)XXXXXX(K/R)     |
| 201        | X(Y/F)XXXXXX(K/R)     |
| 202        | X(Y/F)XXXXXXXX(K/R)   |
| 203        | X(Y/F)XXXXXXXX(K/R)   |
| 204        | X(Y/F)XXXXXXXX(K/R)   |
| 205        | X(N/Y)XXXXXX(K/R)     |
| 206        | X(N/Y)XXXXXX(K/R)     |
| 207        | X(N/Y)XXXXXXXX(K/R)   |
| 208        | X(N/Y)XXXXXXXX(K/R)   |
| 209        | X(N/Y)XXXXXXXX(K/R)   |

**[0071]** In some embodiments, antigen screening libraries of the present disclosure include (b) randomized antigen polypeptides encoded at least by, but not limited to, nucleotide sequences SEQ ID NOs: 210 to 411 provided at least in **Table 4**. In some embodiments, antigen screening libraries of the present disclosure include (b) randomized antigen polypeptides encoded at least by, but not limited to, nucleotide sequences SEQ ID NOs: 412 to 426 provided at least in **Table 5**. Nucleic acids that encode the randomized antigen polypeptides of (b) are encoded by a degenerate base sequence, effectively allowing any amino acid to be encoded at a given position corresponding to the degenerate base sequence. Each randomized antigen polypeptide has at least one conserved anchor position that

is encoded by a restricted degenerate code, or a specific sequence, which allows the randomized antigen polypeptide to more efficiently interact with a certain HLA type. Having at least one conserved anchor position per randomized antigen polypeptide increases efficiency of formation of a randomized antigen polypeptide and HLA complex compared to formation of an HLA complex with a fully randomized antigen polypeptide. In some embodiments, 1, 2, or 3 of the amino acid residues of a randomized antigen polypeptide are constant. In some embodiments, the randomized antigen polypeptide antigens comprises a sequence that is at least about 70%, 75%, 80%, 85%, 87%, 87.5%, 90%, 95%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5%, or 100% identical to any one of, but not limited to, the amino acid sequences set forth in any one of SEQ ID NOs: 210 to 411 and SEQ ID NOs: 412 to 426. In some embodiments, the randomized antigen polypeptide antigens comprise a sequence identical to any one of those set forth in any one of SEQ ID NOs: 210 to 411 and SEQ ID NOs: 412 to 426. Also envisioned within the present disclosure are randomized antigen polypeptide antigen polypeptide truncations that have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 25 amino acids truncated from the N-terminus or truncated from the C-terminus of any one of SEQ ID NOs: 210 to 411 and SEQ ID NOs: 412 to 426.

**[0072]** In some embodiments, amino acid residues of a randomized antigen polypeptide vary by 2, 3, or 4 different amino acids. For example, referring to **Table 1**, the second and the last position of a randomized antigen polypeptide that binds to HLA-A2 will comprise leucine or methionine; and leucine, methionine, or valine, respectively.

**[0073]** The amino acid sequences in Tables 1 and 2 above include random amino acid residues ('X') and explicitly defined amino acids located at residues referred to collectively as anchor positions. The anchor positions specified in the library design can be altered, for example, based on amino acid substitutions set forth in Table 3. One of ordinary skill in the art would appreciate that possible substitutions for X residue in the amino acid sequences of Tables 1 and 2 are not limited and can

include additional substitutions without departing from the scope of the disclosure. For example, amino acid substitutions can be used to identify important residues of the peptide sequence that contribute to binding of the HLA or to constrain or expand the members of the library described herein.

**[0074]** Conservative modifications will produce peptides having functional and chemical characteristics similar to those of the peptide from which such modifications are made. In contrast, substantial modifications in the functional and/or chemical characteristics of the peptides may be accomplished by selecting substitutions in the amino acid sequence that differ significantly in their effect on maintaining (a) the structure of the molecular backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the size of the molecule.

**[0075]** For example, a “conservative amino acid substitution” may involve a substitution of a native amino acid residue with a nonnative residue such that there is little or no effect on the polarity or charge of the amino acid residue at that position. Furthermore, any native residue in the polypeptide may also be substituted with alanine, as has been previously described for “alanine scanning mutagenesis” (see, for example, MacLennan 1998 and Sasaki & Sutoh 1998, which discuss alanine scanning mutagenesis).

**[0076]** Desired amino acid substitutions (whether conservative or non-conservative) can be determined by those skilled in the art at the time such substitutions are desired. Exemplary amino acid substitutions are set forth in Table 3.

**Table 3: Amino Acid Substitutions**

| Original Residues | Exemplary Substitutions |
|-------------------|-------------------------|
| Ala (A)           | Val, Leu, Ile           |
| Arg (R)           | Lys, Gln, Asn           |
| Asn (N)           | Gln                     |
| Asp (D)           | Glu                     |

|         |                                              |
|---------|----------------------------------------------|
| Cys (C) | Ser, Ala                                     |
| Gln (Q) | Asn                                          |
| Glu (E) | Asp                                          |
| Gly (G) | Pro, Ala                                     |
| His (H) | Asn, Gln, Lys, Arg                           |
| Ile (I) | Leu, Val, Met, Ala, Phe, Norleucine (Nle)    |
| Leu (L) | Norleucine (Nle), Ile, Val, Met, Ala, Phe    |
| Lys (K) | Arg, 1,4 Diaminobutyric Acid (Dab), Gln, Asn |
| Met (M) | Leu, Phe, Ile                                |
| Phe (F) | Leu, Val, Ile, Ala, Tyr                      |
| Pro (P) | Ala                                          |
| Ser (S) | Thr, Ala, Cys                                |
| Thr (T) | Ser                                          |
| Trp (W) | Tyr, Phe                                     |
| Tyr (Y) | Trp, Phe, Thr, Ser                           |
| Val (V) | Ile, Met, Leu, Phe, Ala, Norleucine (Nle)    |

[0077] In certain embodiments, conservative amino acid substitutions also encompass non-naturally occurring amino acid residues which are typically incorporated by chemical peptide synthesis rather than by synthesis in biological systems.

[0078] As noted in the foregoing section “Certain Definitions,” naturally occurring residues may be divided into classes based on common sidechain properties that may be useful for modifications of sequence. For example, non-conservative substitutions may involve the exchange of a member of one of these classes for a member from another class. Such substituted residues may be introduced into regions of the peptide that are homologous with non-human orthologs, or into the non-homologous regions of the molecule. In addition, one may also make modifications using P or G for the purpose of influencing chain orientation.

[0079] In making such modifications, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics; these are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine

(−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

**[0080]** The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is understood in the art (Kyte & Doolittle 1982). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

**[0081]** It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. The greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, i.e. with a biological property of the protein.

**[0082]** The following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 $\pm$ 1); glutamate (+3.0 $\pm$ 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (−0.4); proline (−0.5 $\pm$ 1); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred. One may also identify epitopes from primary amino acid sequences on the basis of hydrophilicity. These regions are also referred to as “epitopic core regions.”

**[0083]** A skilled artisan will be able to determine suitable variants of the polypeptide as set forth in the foregoing sequences using well known techniques. For identifying suitable areas of the molecule that may be changed without destroying activity, one skilled in the art may target areas not believed to be important for activity. For example, when similar polypeptides with similar activities from the

same species or from other species are known, one skilled in the art may compare the amino acid sequence of a peptide to similar peptides. With such a comparison, one can identify residues and portions of the molecules that are conserved among similar polypeptides. It will be appreciated that changes in areas of a peptide that are not conserved relative to such similar peptides would be less likely to adversely affect the biological activity and/or structure of the peptide. One skilled in the art would also know that, even in relatively conserved regions, one may substitute chemically similar amino acids for the naturally occurring residues while retaining activity (conservative amino acid residue substitutions). Therefore, even areas that may be important for biological activity or for structure may be subject to conservative amino acid substitutions without destroying the biological activity or without adversely affecting the peptide structure.

**[0084]** Additionally, one skilled in the art can review structure-function studies identifying residues in similar peptides that are important for activity or structure. In view of such a comparison, one can predict the importance of amino acid residues in a peptide that correspond to amino acid residues that are important for activity or structure in similar peptides. One skilled in the art may opt for chemically similar amino acid substitutions for such predicted important amino acid residues of the peptides.

**[0085]** One skilled in the art can also analyze the three-dimensional structure and amino acid sequence in relation to that structure in similar polypeptides. In view of that information, one skilled in the art may predict the alignment of amino acid residues of a peptide with respect to its three dimensional structure. One skilled in the art may choose not to make radical changes to amino acid residues predicted to be on the surface of the protein, since such residues may be involved in important interactions with other molecules. Moreover, one skilled in the art may generate test variants containing a single amino acid substitution at each desired amino acid residue. The variants can then be screened using activity assays known to those skilled in the art. Such data could be used to gather

information about suitable variants. For example, if one discovered that a change to a particular amino acid residue resulted in destroyed, undesirably reduced, or unsuitable activity, variants with such a change would be avoided. In other words, based on information gathered from such routine experiments, one skilled in the art can readily determine the amino acids where further substitutions should be avoided, either alone or in combination with other mutations.

**[0086]** A number of scientific publications have been devoted to the prediction of secondary structure (see, e.g., Moult 1996; Chou & Fasman 1974a; Chou & Fasman 1974b; Chou & Fasman 1978a; Chou & Fasman 1978b; and Chou & Fasman 1979). Moreover, computer programs are currently available to assist with predicting secondary structure. One method of predicting secondary structure is based upon homology modeling. For example, two polypeptides or proteins which have a sequence identity of greater than 30%, or similarity greater than 40% often have similar structural topologies. The recent growth of the protein structural data base (PDB) has provided enhanced predictability of secondary structure, including the potential number of folds within a polypeptide's or protein's structure (Holm & Sander 1999). It has been suggested that there are a limited number of folds in a given polypeptide or protein and that once a critical number of structures have been resolved, structural prediction will gain dramatically in accuracy (Brenner 1997).

**[0087]** Additional methods of predicting secondary structure include “threading” (Jones 1997; Sippl & Flockner 1996), “profile analysis” (Bowie 1991; Gribskov 1987; Gribskov 1990), and “evolutionary linkage” (Holm & Sander 1999; Brenner 1997).

**Table 4: Nucleic Acid Sequences Encoding Randomized Polypeptide Antigens**

| HLA type | SEQ ID NO: | Sequence                       |
|----------|------------|--------------------------------|
| HLA A2   | 210        | nnkmtgnnknnknnknnknnkntg       |
|          | 211        | nnkmtgnnknnknnknnknnkntg       |
|          | 212        | nnkmtgnnknnknnknnknnknnkntg    |
|          | 213        | nnkmtgnnknnknnknnknnknnknnkntg |

|         |     |                                                        |
|---------|-----|--------------------------------------------------------|
| HLA A1  | 214 | nnknnkgaknnknnknnknnknnktat                            |
|         | 215 | nnknnkgaknnknnknnknnknnknnktat                         |
|         | 216 | nnknnkgaknnknnknnknnknnknnknnktat                      |
| HLA A3  | 217 | nnkvtgnnknnknnknnknnknnkaaa                            |
|         | 218 | nnkvtgnnknnknnknnknnknnknnkaaa                         |
|         | 219 | nnkvtgnnknnknnknnknnknnknnknnkaaa                      |
|         | 220 | nnkvtgnnknnknnknnknnknnknnknnknnkaaa                   |
|         | 221 | nnkvtgnnknnknnknnknnknnknnknnknnknnkaaa                |
| HLA A11 | 222 | nnkbttnnknnknnknnknnknnknnkara                         |
|         | 223 | nnkbttnnknnknnknnknnknnknnknnkara                      |
|         | 224 | nnkbttnnknnknnknnknnknnknnknnknnkara                   |
|         | 225 | nnkbttnnknnknnknnknnknnknnknnknnknnkara                |
|         | 226 | nnkbttnnknnknnknnknnknnknnknnknnknnknnkara             |
| HLA A23 | 227 | nnknnknnknnknnknnknnknnknnktay                         |
|         | 228 | nnknnknnknnknnknnknnknnknnknnktay                      |
|         | 229 | nnknnknnknnknnknnknnknnknnknnknnknnktay                |
|         | 230 | nnknnknnknnknnknnknnknnknnknnknnknnknnktay             |
|         | 231 | nnknnknnknnknnknnknnknnknnknnknnknnknnknnknnktay       |
| HLA A24 | 232 | nnkwttnnknnknnknnknnknnknnkhtt                         |
|         | 233 | nnkwttnnknnknnknnknnknnknnknnkhtt                      |
|         | 234 | nnkwttnnknnknnknnknnknnknnknnknnkhtt                   |
|         | 235 | nnkwttnnknnknnknnknnknnknnknnknnknnkhtt                |
|         | 236 | nnkwttnnknnknnknnknnknnknnknnknnknnknnkhtt             |
| HLA A26 | 237 | nnkbthnnknnknnknnknnknnknnktwy                         |
|         | 238 | nnkayynnknnknnknnknnknnknnknnktwy                      |
|         | 239 | nnkbthnnknnknnknnknnknnknnknnknnktwy                   |
|         | 240 | nnkayynnknnknnknnknnknnknnknnknnktwy                   |
|         | 241 | nnkbthnnknnknnknnknnknnknnknnknnknnktwy                |
|         | 242 | nnkayynnknnknnknnknnknnknnknnknnknnktwy                |
|         | 243 | nnkbthnnknnknnknnknnknnknnknnknnknnknnktwy             |
|         | 244 | nnkayynnknnknnknnknnknnknnknnknnknnknnktwy             |
|         | 245 | nnkbthnnknnknnknnknnknnknnknnknnknnknnknnktwy          |
|         | 246 | nnkayynnknnknnknnknnknnknnknnknnknnknnknnktwy          |
| HLA A30 | 247 | nnktwynnknnknnknnknnknnknnknnkctn                      |
|         | 248 | nnktwynnknnknnknnknnknnknnknnknnkctn                   |
|         | 249 | nnktwynnknnknnknnknnknnknnknnknnknnknnkctn             |
|         | 250 | nnktwynnknnknnknnknnknnknnknnknnknnknnknnkctn          |
|         | 251 | nnktwynnknnknnknnknnknnknnknnknnknnknnknnknnknnknnkctn |

|         |     |                                          |
|---------|-----|------------------------------------------|
| HLA A31 | 252 | nnknnknnknnknnknnknnkarr                 |
|         | 253 | nnknnknnknnknnknnknnkarr                 |
|         | 254 | nnknnknnknnknnknnknnknnkarr              |
|         | 255 | nnknnknnknnknnknnknnknnknnkarr           |
|         | 256 | nnknnknnknnknnknnknnknnknnknnkarr        |
| HLA A33 | 257 | nnknnknnknnknnknnknnkarr                 |
|         | 258 | nnknnknnknnknnknnknnknnkarr              |
|         | 259 | nnknnknnknnknnknnknnknnknnkarr           |
|         | 260 | nnknnknnknnknnknnknnknnknnknnkarr        |
|         | 261 | nnknnknnknnknnknnknnknnknnknnknnkarr     |
| HLA A68 | 262 | nnkgttnnknnknnknnknnkarr                 |
|         | 263 | nnkgttnnknnknnknnknnknnkarr              |
|         | 264 | nnkgttnnknnknnknnknnknnknnkarr           |
|         | 265 | nnkgttnnknnknnknnknnknnknnknnkarr        |
|         | 266 | nnkgttnnknnknnknnknnknnknnknnknnkarr     |
|         | 267 | nnkactnnknnknnknnknnknnkarr              |
|         | 268 | nnkactnnknnknnknnknnknnknnkarr           |
|         | 269 | nnkactnnknnknnknnknnknnknnknnkarr        |
|         | 270 | nnkactnnknnknnknnknnknnknnknnknnkarr     |
|         | 271 | nnkactnnknnknnknnknnknnknnknnknnknnkarr  |
| HLA B7  | 272 | nnkcetnnknnknnknnknnknnkett              |
|         | 273 | nnkcetnnknnknnknnknnknnknnkett           |
|         | 274 | nnkcetnnknnknnknnknnknnknnknnkett        |
|         | 275 | nnkcetnnknnknnknnknnknnknnknnknnkett     |
|         | 276 | nnkcetnnknnknnknnknnknnknnknnknnknnkett  |
| HLA B8  | 277 | nnknnkaaannkarannknnknnkett              |
|         | 278 | nnknnkaaannkarannknnknnknnkett           |
|         | 279 | nnknnkaaannkarannknnknnknnknnkett        |
|         | 280 | nnknnkaaannkarannknnknnknnknnknnkett     |
|         | 281 | nnknnkaaannkarannknnknnknnknnknnknnkett  |
| HLA B15 | 282 | nnkcwannknnknnknnknnknnkttwt             |
|         | 283 | nnkcwannknnknnknnknnknnknnkttwt          |
|         | 284 | nnkcwannknnknnknnknnknnknnknnkttwt       |
|         | 285 | nnkcwannknnknnknnknnknnknnknnknnkttwt    |
|         | 286 | nnkcwannknnknnknnknnknnknnknnknnknnkttwt |
| HLA B27 | 287 | nnkagannknnknnknnknnknnkttwt             |
|         | 288 | nnkagannknnknnknnknnknnknnkttwt          |
|         | 289 | nnkagannknnknnknnknnknnknnknnkttwt       |

[illegible]

|         |     |                                     |
|---------|-----|-------------------------------------|
|         | 328 | nnkcctnnknknknknknknknknytt         |
|         | 329 | nnkcctnnknknknknknknknknytt         |
|         | 330 | nnkcctnnknknknknknknknknknytt       |
|         | 331 | nnkcctnnknknknknknknknknknknytt     |
| HLA B57 | 332 | nnkdetnnknknknknknknkntwt           |
|         | 333 | nnkdetnnknknknknknknkntwt           |
|         | 334 | nnkdetnnknknknknknknknkntwt         |
|         | 335 | nnkdetnnknknknknknknknknkntwt       |
|         | 336 | nnkdetnnknknknknknknknknknkntwt     |
|         | 337 | nnkdetnnknknknknknknkntgg           |
|         | 338 | nnkdetnnknknknknknknknkntgg         |
|         | 339 | nnkdetnnknknknknknknknknkntgg       |
|         | 340 | nnkdetnnknknknknknknknknknkntgg     |
|         | 341 | nnkdetnnknknknknknknknknknknkntgg   |
| HLA C1  | 342 | nnkttannknknknknknknknkttta         |
|         | 343 | nnkttannknknknknknknknkttta         |
|         | 344 | nnkttannknknknknknknknknkttta       |
|         | 345 | nnkttannknknknknknknknknknkttta     |
|         | 346 | nnkttannknknknknknknknknknknkttta   |
|         | 347 | nnkgctnnknknknknknknknkttta         |
|         | 348 | nnkgctnnknknknknknknknknkttta       |
|         | 349 | nnkgctnnknknknknknknknknknkttta     |
|         | 350 | nnkgctnnknknknknknknknknknknkttta   |
|         | 351 | nnkgctnnknknknknknknknknknknknkttta |
| HLA C2  | 352 | nnkgctnnknknknknknknknkstc          |
|         | 353 | nnkgctnnknknknknknknknknkstc        |
|         | 354 | nnkgctnnknknknknknknknknknkstc      |
|         | 355 | nnkgctnnknknknknknknknknknknkstc    |
|         | 356 | nnkgctnnknknknknknknknknknknknkstc  |
|         | 357 | nnkgctnnknknknknknknknkntwt         |
|         | 358 | nnkgctnnknknknknknknknknkntwt       |
|         | 359 | nnkgctnnknknknknknknknknknkntwt     |
|         | 360 | nnkgctnnknknknknknknknknknknkntwt   |
|         | 361 | nnkgctnnknknknknknknknknknknknkntwt |
| HLA C3  | 362 | nnknknknknknknknknknknkhtk          |
|         | 363 | nnknknknknknknknknknknknkhtk        |
|         | 364 | nnknknknknknknknknknknknknkhtk      |
|         | 365 | nnknknknknknknknknknknknknknkhtk    |

|        |     |                                   |
|--------|-----|-----------------------------------|
|        | 366 | nnknnknnknnknnknnknnknnknnknnkhtk |
| HLA C4 | 367 | nnktwtnnknnknnknnknnknnkhtk       |
|        | 368 | nnktwtnnknnknnknnknnknnkhtk       |
|        | 369 | nnktwtnnknnknnknnknnknnkhtk       |
|        | 370 | nnktwtnnknnknnknnknnknnkhtk       |
|        | 371 | nnktwtnnknnknnknnknnknnkhtk       |
|        | 372 | nnkcetnnknnknnknnknnknnkhtk       |
|        | 373 | nnkcetnnknnknnknnknnknnkhtk       |
|        | 374 | nnkcetnnknnknnknnknnknnkhtk       |
|        | 375 | nnkcetnnknnknnknnknnknnkhtk       |
|        | 376 | nnkcetnnknnknnknnknnknnkhtk       |
| HLA C5 | 377 | nnknnkgatnnknnknnknnkstt          |
|        | 378 | nnknnkgatnnknnknnknnkstt          |
|        | 379 | nnknnknnkgatnnknnknnknnkstt       |
|        | 380 | nnknnknnknnkgatnnknnknnknnkstt    |
|        | 381 | nnknnknnknnkgatnnknnknnknnkstt    |
| HLA C6 | 382 | nnknnknnknnknnknnknnkvtt          |
|        | 383 | nnknnknnknnknnknnknnkvtt          |
|        | 384 | nnknnknnknnknnknnknnkvtt          |
|        | 385 | nnknnknnknnknnknnknnkvtt          |
|        | 386 | nnknnknnknnknnknnknnknnkvtt       |
|        | 387 | nnknnknnknnknnknnknnkntat         |
|        | 388 | nnknnknnknnknnknnknnkntat         |
|        | 389 | nnknnknnknnknnknnknnkntat         |
|        | 390 | nnknnknnknnknnknnknnkntat         |
|        | 391 | nnknnknnknnknnknnknnkntat         |
| HLA C7 | 392 | nnknnknnknnknnknnknnkytt          |
|        | 393 | nnknnknnknnknnknnknnkytt          |
|        | 394 | nnknnknnknnknnknnknnkytt          |
|        | 395 | nnknnknnknnknnknnknnkytt          |
|        | 396 | nnknnknnknnknnknnknnknnkytt       |
|        | 397 | nnknnknnknnknnknnknnkntat         |
|        | 398 | nnknnknnknnknnknnknnkntat         |
|        | 399 | nnknnknnknnknnknnknnkntat         |
|        | 400 | nnknnknnknnknnknnknnkntat         |
|        | 401 | nnknnknnknnknnknnknnkntat         |
| HLA C8 | 402 | nnknnkgatnnknnknnkett             |
|        | 403 | nnknnkgatnnknnknnkett             |

|       |     |                                            |
|-------|-----|--------------------------------------------|
| HLA E | 404 | nnknnkgatnnknnknnknnknnkctt                |
|       | 405 | nnknnkgatnnknnknnknnknnknnkctt             |
|       | 406 | nnknnkgatnnknnknnknnknnknnknnkctt          |
|       | 407 | nnkmtgnnknnknnknnknnknnknnkctt             |
|       | 408 | nnkmtgnnknnknnknnknnknnknnknnkctt          |
| HLA E | 409 | nnkmtgnnknnknnknnknnknnknnknnknnkctt       |
|       | 410 | nnkmtgnnknnknnknnknnknnknnknnknnknnkctt    |
|       | 411 | nnkmtgnnknnknnknnknnknnknnknnknnknnknnkctt |
|       |     |                                            |
|       |     |                                            |

**Table 5: Additional Nucleic Acid Sequences Encoding Randomized Polypeptide Antigens for HLA A11**

| SEQ ID NO: | Sequence                                      |
|------------|-----------------------------------------------|
| 412        | nnkvttnnknnknnknnknnknnkara                   |
| 413        | nnkvttnnknnknnknnknnknnknnkara                |
| 414        | nnkvttnnknnknnknnknnknnknnknnkara             |
| 415        | nnkvttnnknnknnknnknnknnknnknnknnkara          |
| 416        | nnkvttnnknnknnknnknnknnknnknnknnknnkara       |
| 417        | nnktwtnnknnknnknnknnknnknnknnkara             |
| 418        | nnktwtnnknnknnknnknnknnknnknnkara             |
| 419        | nnktwtnnknnknnknnknnknnknnknnknnkara          |
| 420        | nnktwtnnknnknnknnknnknnknnknnknnknnkara       |
| 421        | nnktwtnnknnknnknnknnknnknnknnknnknnknnkara    |
| 422        | nnkwatnnknnknnknnknnknnknnknnkara             |
| 423        | nnkwatnnknnknnknnknnknnknnknnknnkara          |
| 424        | nnkwatnnknnknnknnknnknnknnknnknnknnkara       |
| 425        | nnkwatnnknnknnknnknnknnknnknnknnknnknnkara    |
| 426        | nnkwatnnknnknnknnknnknnknnknnknnknnknnknnkara |

**[0088]** One advantage of a randomized antigen polypeptide is that a single nucleic acid with a degenerate base code can potentially express a large amount of different randomized antigen polypeptides, which increases the chances that any one screening experiment will identify one or more randomized antigen polypeptides that interact with a certain TCR. In some embodiments, the nucleic acid that encodes the randomized antigen polypeptide can encode at least  $1 \times 10^4$ , at least  $1 \times 10^5$ , at least  $1 \times 10^6$ , at least  $1 \times 10^7$ , at least  $1 \times 10^8$ , at least  $1 \times 10^9$ , at least  $1 \times 10^{10}$ , at least  $1 \times 10^{11}$ , at least  $1 \times 10^{12}$ , at

least  $1 \times 10^{13}$ , at least  $1 \times 10^{14}$ , or at least  $1 \times 10^{15}$  different randomized polypeptide antigens.

**[0089]** Peptide antigens that bind in the binding cleft of an HLA molecule are generally of a restricted length range. The majority of polypeptides that bind to class I HLA molecules are 8, 9, 10, or 11 amino acids in length. In some embodiments, the randomized antigen polypeptide which binds to an HLA molecule and forming the HLA-antigen polypeptide complexes of the present disclosure is between 8 and 11 amino acids in length. In some embodiments, the randomized antigen polypeptide is between 8 and 10 amino acids in length. In some embodiments, the randomized antigen polypeptide is 8 amino acids in length. In some embodiments, the randomized antigen polypeptide is 9 amino acids in length. In some embodiments, the randomized antigen polypeptide is 10 amino acids in length. In some embodiments, the randomized antigen polypeptide is 11 amino acids in length.

**[0090]** Another constituent of the HLA-antigen polypeptide complexes described herein is an HLA molecule, such as an HLA polypeptide. For the purposes of the current disclosure, the HLA molecule is a class I major histocompatibility molecule. In some embodiments, the plurality of HLA polypeptides of the HLA-antigen polypeptide complexes of the current disclosure (HLA-antigen complexes) can comprise any of the following loci and alleles: A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, each of the HLA-antigen complexes in the plurality of HLA-antigen complexes comprise an HLA polypeptide selected from the group of HLA polypeptides consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises at least five, ten, fifteen, twenty, or twenty-five different HLA polypeptides selected from the group of HLA polypeptides consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E. In some embodiments, the plurality of HLA-antigen complexes comprises all of the HLA polypeptides in the group of HLA polypeptides

consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E.

**[0091]** In some embodiments, the amino acid sequence of the HLA polypeptide of the HLA-antigen polypeptide complex can comprise any of the amino acid sequences set forth in **Table 6**. In some embodiments, the HLA polypeptide comprises an amino acid sequence that is at least about 70%, 75%, 80%, 85%, 87%, 87.5%, 90%, 95%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5%, or 100% identical to any one of, but not limited to, the amino acid sequences set forth in any one of SEQ ID NOs: 427 to 455. In some embodiments, the HLA polypeptide comprises an amino acid sequence identical to any one of those set forth in any one of SEQ ID NOs: 427 to 455. In some embodiments, a portion of the HLA polypeptide that comprises the peptide binding cleft is identical to any one of SEQ ID Nos: 251 to 279, and the non-peptide binding cleft residues are at least about 70%, 75%, 80%, 85%, 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% identical to any one of SEQ ID NOs: 427 to 455. Also envisioned within the present disclosure are HLA polypeptide truncations that have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 25 amino acids truncated from the N-terminus or truncated from the C-terminus of any one of SEQ ID NOs: 427 to 455.

**Table 6: HLA Allele Amino Acid Sequences**

| SEQ ID NO: | Amino Acid Sequence                                                                                                                                                                                                                                                                                          |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 427        | GSHSMRYFFTSVSRPGRGEPRFIAVG YVDDTQFVRFDSDAASQKMEPRAP<br>WIEQEGPEYWDQETRNMKAHSQTDRANLGTLRGAYNQSEDGSHTIQIM<br>YGCDVGPDGRFLRGYRQDAYDGKDYIALNEDLRSWTAADMAAQITKRK<br>WEAVHAAEQRRVYLEGRCVDGLRRYLENGKETLQRTDPPKTHMTHHPI<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWELSS  |
| 428        | GSHSMRYFFTSVSRPGRGEPRFIAVG YVDDTQFVRFDSDAASQRMPEPRAP<br>WIEQEGPEYWDQETRNVKAQSQTDRVDLGTLRGAYNQSEAGSHTIQIMY<br>GCDVGSDGRFLRGYRQDAYDGKDYIALNEDLRSWTAADMAAQITKRK<br>WEAAHEAEQLRAYLDGTCVEWLRRYLENGKETLQRTDPPKTHMTHHPI<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWELSS |

|     |                                                                                                                                                                                                                                                                                                                |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 429 | GSHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRA<br>PWIEQEGPEYWDQETRNVKAQSQTDRVDLGTLRGAYNQSEAGSHTIQIM<br>YGCDVGPDPGRFLRGYRQDAYDGKDYIALNEDLRSWTAADMAAQITKRK<br>WEAAHAAEQQRAYLEGRCVEWLRRYLENGKETLQRTDPPKTHMTHHPI<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWELSS    |
| 430 | GSHSMRYFSTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAP<br>WIEQEGPEYWDEETGKVK AHSQTDRENLRALRAYNQSEAGSHTLQMM<br>FGCDVGS DGRFLRGYHQYAYDGKDYIALKEDLRSWTAADMAAQITQRK<br>WEAARVAEQLRAYLEGTCVDGLRRYLENGKETLQRTDPPKTHMTHHPI<br>DHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPSS     |
| 431 | GSHSMRYFSTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAP<br>WIEQEGPEYWDEETGKVK AHSQTDRENLRALRAYNQSEAGSHTLQMM<br>FGCDVGS DGRFLRGYHQYAYDGKDYIALKEDLRSWTAADMAAQITKRK<br>WEAAHVAEQQRAYLEGTCVDGLRRYLENGKETLQRTDPPKTHMTHHPI<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPSS    |
| 432 | GSHSMRYFSTSVSRPGSGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAP<br>WIEQERPEYWDQETRNVKAQSQTDRVDLGTLRGAYNQSEAGSHTIQIMY<br>GCDVGS DGRFLRGYEQHAYDGKDYIALNEDLRSWTAADMAAQITQRK<br>WEAARWAEQLRAYLEGTCVEWLRRYLENGKETLQRTDPPKTHMTHHPI<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWELSS    |
| 433 | GSHSMRYFTTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAP<br>WIEQERPEYWDQETRNVKAHSQIDRVLDLGTLRGAYNQSEAGSHTIQMM<br>YGCDVGS DGRFLRGYQQDAYDGKDYIALNEDLRSWTAADMAAQITQRK<br>WEAARVAEQLRAYLEGTCVEWLRRYLENGKETLQRTDPPKTHMTHHAV<br>SDHEATLRCWALSFPYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WASVVVPSGQEQRYTCHVQHEGLPKPLTLRWEPSS  |
| 434 | GSHSMRYFTTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAP<br>WIEQEGPEYWDRNTRNVKAHSQIDRVLDLGTLRGAYNQSEAGSHTIQMM<br>YGCDVGS DGRFLRGYQQDAYDGKDYIALNEDLRSWTAADMAAQITQRK<br>WEAARVAEQLRAYLEGTCVEWLRRYLENGKETLQRTDPPKTHMTHHAV<br>SDHEATLRCWALSFPYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQK<br>WASVVVPSGQEQRYTCHVQHEGLPKPLTLRWEPSS  |
| 435 | GSHSMRYFYTSMSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRA<br>PWIEQEGPEYWDRNTRNVKAQSQTDRVDLGTLRGAYNQSEAGSHTIQR<br>MYGCDVGPDPGRFLRGYHQYAYDGKDYIALKEDLRSWTAADMAAQITTK<br>HKWEAAHVAEQWRAYLEGTCVEWLRRYLENGKETLQRTDAPKTHMTH<br>HAVSDHEATLRCWALSFPYPAEITLTWQRDGEDQTQDTEL VETRPAGDGT<br>FQKWVAVVVPSGQEQRYTCHVQHEGLPKPLTLRWEPSS |

|     |                                                                                                                                                                                                                                                                                                             |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 436 | GSHSMRYFYTSVSRPGRGEPFISVGYVDDTQFVRFDSDAASPREEPRAP<br>WIEQEGPEYWDNRNTQIYKAQAQTDRESLRNLRGAYNQSEAGSHTLQSM<br>YGCDVGPDPGRLLRGHDQYAYDGKDIALNEDLRSWTAADTAAQITQRK<br>WEAAREAEQRRAYLEGECEVWLRRYLENGKDKLERADPPKTHVTHHPIS<br>DHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPS  |
| 437 | GSHSMRYFDTAMSRPGRGEPFISVGYVDDTQFVRFDSDAASPREEPRAP<br>WIEQEGPEYWDNRNTQIFKTNTQTDRESLRNLRGAYNQSEAGSHTLQSMY<br>GCDVGPDPGRLLRGHNQYAYDGKDIALNEDLRSWTAADTAAQITQRKW<br>EAARVAEQDRAYLEGTCVEWLRRYLENGKDTLERADPPKTHVTHHPISD<br>HEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQKW<br>AAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPS  |
| 438 | GSHSMRYFYTAMSRPGRGEPFIAVGYVDDTQFVRFDSDAASPRMAPRA<br>PWIEQEGPEYWDRETQISKNTQTYRESLRNLRGAYNQSEAGSHTLQRM<br>YGCDVGPDPGRLLRGHDQSA YDGKDIALNEDLSSWTAADTAAQITQRK<br>WEAAREAEQWRAYLEGLCVEWLRRYLENGKETLQRADPPKTHVTHHPISD<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPS |
| 439 | GSHSMRYFHTAMSRPGRGEPFITVGYVDDTLFVRFDSDATSPRKEPRAP<br>WIEQEGPEYWDRETQISKNTQTYRESLRNLRGAYNQSEAGSHTLQRM<br>YGCDVGPDPGRLLRGHNQYAYDGKDIALNEDLRSWTAADTAAQISQRKL<br>EAARVAEQLRAYLEGECEVWLRRYLENGKDKLERADPPKTHVTHHPISD<br>HEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQKW<br>AAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPS    |
| 440 | GSHSMRYFHTSVSRPGRGEPFITVGYVDDTLFVRFDSDATSPRKEPRAP<br>WIEQEGPEYWDRETQISKNTQTYRESLRNLRGAYNQSEAGSHTLQSMY<br>YGCDVGPDPGRLLRGHNQYAYDGKDIALNEDLRSWTAADTAAQITQRKW<br>EAARVAEQLRAYLEGECEVWLRRYLENGKETLQRADPPKTHVTHHPISD<br>HEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQKW<br>AAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPS   |
| 441 | GSHSMRYFYTAMSRPGRGEPFITVGYVDDTLFVRFDSDATSPRKEPRAP<br>WIEQEGPEYWDRETQISKNTQTYRENLRALRAYNQSEAGSHIIQRMYG<br>CDVGPDPGRLLRGYDQDAYDGKDIALNEDLSSWTAADTAAQITQRKWE<br>AARVAEQDRAYLEGLCVESLRRYLENGKETLQRADPPKTHVTHHPISDH<br>EVTLCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQKWA<br>AVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPS      |
| 442 | GSHSMRYFYTAMSRPGRGEPFITVGYVDDTLFVRFDSDATSPRKEPRAP<br>WIEQEGPEYWDRETQISKNTQTYRENLRALRAYNQSEAGSHIIQRMYG<br>CDVGPDPGRLLRGYDQDAYDGKDIALNEDLSSWTAADTAAQITQRKWE<br>AARVAEQLRAYLEGLCVESLRRYLENGKETLQRADPPKTHVTHHPISDH<br>EVTLCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQKWA<br>AVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPS      |

|     |                                                                                                                                                                                                                                                                                                                  |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 443 | GSHSMRYFYTAMSRPGRGEPRIAVGYVDDTQFVRFDSDAASPRTEPRAP<br>WIEQEGPEYWDRNTQIFKTNTQTYRENLRALRAYNQSEAGSHTWQTM<br>GCDVGPDPGRLLRGHNQYAYDGKDYIALNEDLSSWTAADTAAQITQRKW<br>EAAAREAEQLRAYLEGLCWEVLRRHLENGKETLQRADPPKTHVTHHPVSD<br>HEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQKW<br>AAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPSS       |
| 444 | GSHSMRYFYTAMSRPGRGEPRIAVGYVDDTQFVRFDSDAASPRTEPRAP<br>WIEQEGPEYWDRNTQIFKTNTQTYRENLRALRAYNQSEAGSHIIQRM<br>CDLGPDPGRLLRGHDQSAAYDGKDYIALNEDLSSWTAADTAAQITQRKWE<br>AARVAEQLRAYLEGLCWEVLRRYLENGKETLQRADPPKTHVTHHPVSD<br>HEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQKW<br>AAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPSS        |
| 445 | GSHSMRYFYTAMSRPGRGEPRIAVGYVDDTQFVRFDSDAASPRTEPRAP<br>WIEQEGPEYWDGETRNMKASAQTYRENLRALRAYNQSEAGSHIIQRM<br>GCDLGPDPGRLLRGHDQSAAYDGKDYIALNEDLSSWTAADTAAQITQRKW<br>EAAARVAEQLRAYLEGLCWEVLRRYLENGKETLQRADPPKTHVTHHPVS<br>DHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDRTFQK<br>WAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPSS      |
| 446 | CSHSMKYFFTSVSRPGRGEPRIAVGYVDDTQFVRFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQKYKRQAQTDREVSLRNLRGAYNQSEAGSHTLQW<br>MCGCDLGPDPGRLLRGYDQYAYDGKDYIALNEDLRSWTAADTAAQITQR<br>KWEAAREAEQRRAYLEGTCWEVLRRYLENGKETLQRAEHPKTHVTHHP<br>VSDHEATLRCWALGFYPAEITLTWQWDGEDQTQDTEL VETRPAGDGT<br>FQKWAAMVPSGEEQRYTCHVQHEGLPEPLTLRWEPSS      |
| 447 | CSHSMRYFYTAVSRPSRGEPHFIAVGYVDDTQFVRFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQKYKRQAQTDREVNLRLKLRGAYNQSEAGSHTLQR<br>MYGCDLGPDPGRLLRGYDQSAAYDGKDYIALNEDLRSWTAADTAAQITQR<br>KWEAAREAEQWRAYLEGTCWEVLRRYLENGKETLQRAEHPKTHVTHHP<br>PVSDHEATLRCWALGFYPTTEITLTWQRDGEDQTQDTEL VETRPAGDGT<br>FQKWAAMVPSGEEQRYTCHVQHEGLPEPLTLRWEPSS |
| 448 | GSHSMRYFYTAVSRPGRGEPHFIAVGYVDDTQFVRFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQKYKRQAQTDREVSLRNLRGAYNQSEAGSHIIQRM<br>YGCDVGPDPGRLLRGYDQYAYDGKDYIALNEDLRSWTAADTAAQITQRK<br>WEAAREAEQLRAYLEGLCWEVLRRYLKNGKETLQRAEHPKTHVTHHPV<br>SDHEATLRCWALGFYPAEITLTWQWDGEDQTQDTEL VETRPAGDGT<br>FQKWAAMVPSGEEQRYTCHVQHEGLPEPLTLRWEPSS     |
| 449 | GSHSMRYFSTSVSWPGRGEPRIAVGYVDDTQFVRFDSDAASPRGEPREP<br>WVEQEGPEYWDRETQKYKRQAQADRVNLRLKLRGAYNQSEAGSHTLQR<br>MFGCDLGPDPGRLLRGYNQFAYDGKDYIALNEDLRSWTAADTAAQITQR<br>KWEAAREAEQRRAYLEGTCWEVLRRYLENGKETLQRAEHPKTHVTHHP<br>VSDHEATLRCWALGFYPAEITLTWQWDGEDQTQDTEL VETRPAGDGT<br>FQKWAAMVPSGEEQRYTCHVQHEGLPEPLTLRWKPSS      |
| 450 | CSHSMRYFYTAVSRPGRGEPRIAVGYVDDTQFVQFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQKYKRQAQTDREVNLRLKLRGAYNQSEAGSHTLQR<br>MYGCDLGPDPGRLLRGYNQFAYDGKDYIALNEDLRSWTAADKAAQITQR<br>KWEAAREAEQRRAYLEGTCWEVLRRYLENGKKTQRAEHPKTHVTHHP                                                                                                   |

|     |                                                                                                                                                                                                                                                                                                                  |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     | VSDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQFQ<br>KWA AAVVPSGEEQRYTCHVQHEGLPEPLTLRWGPSS                                                                                                                                                                                                                     |
| 451 | CSHSMRYFDTAVSRPGRGEPFISVGYVDDTQFVRFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQKYKRQAQADRVNLRKLRGAYNQSEDGSHTLQW<br>MYGCDLGP DGRLLRGYDQSA YDGKD YIALNEDLRSWTAADTAAQITQR<br>KWEAAREAEQWRAYLEGTCVEWLRRYLENGKETLQRAEHPKTHVTHH<br>PVS DHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQFQ<br>QKWA AAVVPSGEEQRYTCHVQHEGLPEPLTLRWEPS |
| 452 | CSHSMRYFDTAVSRPGRGEPFISVGYVDDTQFVRFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQNYKRQAQADRVSLRNLRGAYNQSEDGSHTLQR<br>MYGCDLGP DGRLLRGYDQSA YDGKD YIALNEDLRSWTAADTAAQITQR<br>KLEAARAEQLRAYLEGTCVEWLRRYLENGKETLQRAEPPKTHVTHHPL<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQFQK<br>WA AAVVPSGQEQRYTCHMQHEGLQEPLTLSWEPSS   |
| 453 | CSHSMRYFDTAVSRPGRGEPFISVGYVDDTQFVRFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQKYKRQAQADRVSLRNLRGAYNQSEDGSHTLQR<br>MSGCDLGP DGRLLRGYDQSA YDGKD YIALNEDLRSWTAADTAAQITQR<br>KLEAARAEQLRAYLEGTCVEWLRRYLENGKETLQRAEPPKTHVTHHPL<br>SDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQFQK<br>WA AAVVPSGQEQRYTCHMQHEGLQEPLTLSWEPSS   |
| 454 | CSHSMRYFYTAVSRPGRGEPFI AVGYVDDTQFVRFDSDAASPRGEPRAP<br>WVEQEGPEYWDRETQKYKRQAQTD RVSLRNLRGAYNQSEAGSHTLQW<br>MYGCDLGP DGRLLRGYDQSA YDGKD YIALNEDLRSWTAADTAAQITQR<br>KWEAARAEQQRAYLEGTCVEWLRRYLENGKETLQRAEHPKTHVTHHL<br>VSDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTEL VETRPAGDGTQFQ<br>KWA AAVVPSGEEQRYTCHVQHEGLPEPLTLRWEPS  |
| 455 | GSHSLKYFHTSVSRPGRGEPFISVGYVDDTQFVRFNDAA SPRMVPRAP<br>WMEQEGSEYWDRETR SARDTAQIFRVNLR TLRGAYNQSEAGSHTLQWM<br>HGCELGP DGRFLRGYEQFAYDGKD YLTNEDLRSWTA VDTAAQISEQKS<br>NDASEAEHQRAYLEDTCVEWLHKYLEKGKETLLHLEPPKTHVTHHPISD<br>HEATLRCWALGFYPAEITLTWQQDGEGHTQDTEL VETRPAGDGTQFQKW<br>AAVVPSGEEQRYTCHVQHEGLPEPVT LRWKPAS  |

[0092] The HLA polypeptide of the HLA-antigen polypeptide complex can be encoded by a nucleic acid of any set forth in **Table 7**. In some embodiments, the HLA polypeptide is encoded by a nucleic acid sequence that is at least about 90%, 95%, 97%, 98%, 99%, or 100% homologous to at least, but not limited to, any one of the nucleic acid sequences listed in **Table 7**, such as SEQ ID NOs: 456 to 484. In some embodiments, the HLA polypeptide is encoded by a nucleic acid sequence identical to that set forth in any one of SEQ ID NOs: 456 to 484.

Table 7: HLA Allele Nucleic Acid Sequences

| SEQ ID NO: | DNA sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 456        | gggtccactccatgaggtatttctcacatccgtgtcccggcccggcgcggggagccccgttcacgcctgtgggct<br>acgtggacgacacgcagttcgtgcggttcgacagcgcgcgcgagccagaagatggagccgcgggcccgtgg<br>atagagcaggaggggcccggagtattgggaccaggagacacggaatgtgaaggccactcacagactgaccgagcg<br>aacctggggaccctgcgcggcgctacaaccagagcgcaggacggttctcacaccatccagataatgtatggctgcga<br>cgtggggccggacggcgcttctccgcggtaccggcaggacgcctacgacggcaaggattacatgcctgaac<br>gaggacctgcgctcttgaccgcggcgacatggcagctcagatcaccaagcgcaagtgggaggcggtccatgcg<br>gcgagcagcggagagtctacctggaggcgcggtgcgtggacgggtccgcagatactggagaacgggaaggga<br>gacactgcagcgcacggaccccccaagacacatatgaccaccacccatctctgacctgaggccaccctgaggt<br>gctgggcccctgggcttctacctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacac<br>ggagctcgtggagaccaggcctgcaggggatggaaccttcagaagtggcgcggtgtgtgtgcttctggagag<br>gagcagagatacacctgccatgtgcagcatgagggtctgcccaagcccctcacctgagatgggagctgtcttc   |
| 457        | gggtccactccatgaggtatttctcacatccgtgtcccggcccggcgcggggagccccgttcacgcctgtgggct<br>acgtggacgacacgcagttcgtgcggttcgacagcgcgcgcgagccagaaggatggagccgcgggcccgtgg<br>atagagcaggaggggcccggagtattgggaccaggagacacggaatgtgaaggccactcacagactgaccgagtg<br>gacctggggaccctgcgcggcgctacaaccagagcgcaggcgggttctcacaccatccagataatgtatggctgcga<br>cgtggggtcggacggcgcttctccgcggtaccggcaggacgcctacgacggcaaggattacatgcctgaac<br>gaggacctgcgctcttgaccgcggcgacatggcgggtcagatcaccaagcgcaagtgggaggcggtccatgag<br>gcgagcagtgagagcctacctggatggcagctgcgtggagtgggtccgcagatactggagaacgggaaggag<br>aactgcagcgcacggaccccccaagacacatatgaccaccacccatctctgacctgaggccaccctgaggtg<br>ctgggcccctgggcttctacctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacacg<br>gagctcgtggagaccaggcctgcaggggatggaaccttcagaagtggcgcggtgtgtgtgcttctggagagg<br>agcagagatacacctgccatgtgcagcatgagggtctgcccaagcccctcacctgagatgggagctgtcttc    |
| 458        | gggtccactccatgaggtatttctcacactccgtgtcccggcccggcgcggggagccccgttcacgcctgtggg<br>tacgtggacgacacgcagttcgtgcggttcgacagcgcgcgcgagccagaaggatggagccgcgggcccgtgg<br>atagagcaggaggggcccggagtattgggaccaggagacacggaatgtgaaggccactcacagactgaccgagtg<br>gacctggggaccctgcgcggcgctacaaccagagcgcaggacggttctcacaccatccagataatgtatggctgcga<br>cgtggggccggacggcgcttctccgcggtaccggcaggacgcctacgacggcaaggattacatgcctgaac<br>gaggacctgcgctcttgaccgcggcgacatggcagctcagatcaccaagcgcaagtgggaggcggtccatgcg<br>gcgagcagcagagagcctacctggaggcgcggtgcgtggagtgggtccgcagatactggagaacgggaaggga<br>gacactgcagcgcacggaccccccaagacacatatgaccaccacccatctctgacctgaggccaccctgaggt<br>gctgggcccctgggcttctacctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacac<br>ggagctcgtggagaccaggcctgcaggggatggaaccttcagaagtggcgcggtgtgtgtgcttctggagag<br>gagcagagatacacctgccatgtgcagcatgagggtctgcccaagcccctcacctgagatgggagctgtcttc |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 459 | <p>ggctccactccatgaggtatttccacatccgtgtcccggcccggccgcggggagccccgcttcacgcgcgtgggc<br/> tacgtggacgacacgcagttcgtgcggttcgacagcgacgcccgagccagaggatggagccgcgggcgccgtgg<br/> atagagcaggaggggcccggagtattgggacgagagacagggaagtgaaggccactcacagactgaccgaga<br/> gaacctgcggatcgcgctccgcgcctacaaccagagcgaggccggttctcacacctccagatgatgttggctgcga<br/> cgtggggtcggacggcgcttctccgcgggtaccaccagtacgcctacgacggcaaggattacatcgccctgaaag<br/> aggacctgcgctcttgaccgcggcgacatggcggtcagatcacccagcgcaagtgggagggcgcccggtgtgg<br/> cggagcagttgagagcctacctggagggcacgtgcgtggacgggtccgcagatacctggagaacgggaaggaga<br/> cactgcagcgacggaccccccaagacacatatgaccaccacccatctctgacctgaggccactctgagatgct<br/> gggcccgtggcttctacctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacacgg<br/> agcttggtagaccaggcctgcaggggatggaaccttcagaagtgggcagctgtggtggtaccttctggagaggag<br/> cagagatacacctgccatgtgcagcatgagggtctgcccaagcccctaccctgagatgggagccatcttcc</p>    |
| 460 | <p>ggctccactccatgaggtatttccacatccgtgtcccggcccggccgcggggagccccgcttcacgcgcgtgggc<br/> tacgtggacgacacgcagttcgtgcggttcgacagcgacgcccgagccagaggatggagccgcgggcgccgtgg<br/> atagagcaggaggggcccggagtattgggacgagagacagggaagtgaaggccactcacagactgaccgaga<br/> gaacctgcggatcgcgctccgcgcctacaaccagagcgaggccggttctcacacctccagatgatgttggctgcga<br/> cgtggggtcggacggcgcttctccgcgggtaccaccagtacgcctacgacggcaaggattacatcgccctgaaag<br/> aggacctgcgctcttgaccgcggcgacatggcggtcagatcaccaagcgcaagtgggagggcgcccgatgtgg<br/> cggagcagcagagagcctacctggagggcacgtgcgtggacgggtccgcagatacctggagaacgggaaggag<br/> aactgcagcgacggaccccccaagacacatatgaccaccacccatctctgacctgaggccactctgagatg<br/> ctgggcccgtggcttctacctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacacg<br/> gagcttggtagaccaggcctgcaggggatggaaccttcagaagtgggcagctgtggtggtaccttctggagaggga<br/> gcagagatacacctgccatgtgcagcatgagggtctgcccaagcccctaccctgagatgggagccatcttcc</p>    |
| 461 | <p>ggctccactccatgaggtatttccacatccgtgtcccggcccggcagtgagagccccgcttcacgcagtgggct<br/> acgtggacgacacgcagttcgtgcggttcgacagcgacgcccgagccagaggatggagccgcgggcgccgtgg<br/> atagagcaggagagggcctgagtattgggaccaggagacacggaatgtgaaggccactcacagactgaccgagtg<br/> gacctggggaccctgcgcggcgctacaaccagagcgaggccggttctcacacctccagataatgtatggctgcga<br/> cgtggggtcggacggcgcttctccgcgggtatgaacagcagcgctacgacggcaaggattacatcgccctgaa<br/> gaggacctgcgctcttgaccgcggcgacatggcggtcagatcacccagcgcaagtgggagggcgcccggttgg<br/> gcggagcagttgagagcctacctggagggcacgtgcgtggagtggctccgcagatacctggagaacgggaaggag<br/> aactgcagcgacggaccccccaagacacatatgaccaccacccatctctgacctgaggccacctgaggtg<br/> ctgggcccgtggcttctacctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacacg<br/> gagctcgtggagaccaggcctgcaggggatggaaccttcagaagtgggcggctgtggtggtccttctggagagg<br/> agcagagatacacctgccatgtgcagcatgagggtctgcccaagcccctaccctgagatgggagctgttctcc</p>      |
| 462 | <p>ggctccactccatgaggtatttaccacatccgtgtcccggcccggccgcggggagccccgcttcacgcgcgtgggc<br/> tacgtggacgacacgcagttcgtgcggttcgacagcgacgcccgagccagaggatggagccgcgggcgccgtgg<br/> atagagcaggagagggcctgagtattgggaccaggagacacggaatgtgaaggccactcacagattgaccgagtgg<br/> acctggggaccctgcgcggcgctacaaccagagcgaggccggttctcacacctccagatgatgtatggctgcgac<br/> gtggggtcggacggcgcttctccgcgggtaccagcagcgcctacgacggcaaggattacatcgccctgaaag<br/> aggacctgcgctcttgaccgcggcgacatggcggtcagatcacccagcgcaagtgggagggcgcccggtgtgg<br/> cggagcagttgagagcctacctggagggcacgtgcgtggagtggctccgcagatacctggagaacgggaaggaga<br/> cactgcagcgacggaccccccaagacacatatgactcaccacgctgtctctgacctgaggccacctgaggtgc<br/> tgggcccgtgagcttctacctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacacgg<br/> agctcgtggagaccaggcctgcaggggatggaaccttcagaagtgggcgctgtggtggtccttctggagaggag<br/> cagagatacacctgccatgtgcagcatgagggtctgcccaagcccctaccctgagatgggagccgttctcc</p> |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 463 | <p>ggctcccactccatgaggtatttcaccacatccgtgtcccggcccggccgaggagccccgcttcacgcgcgtgggc<br/> tacgtggacgacacgcagttcgtgctggttcgacagcgcgcgcgcgagccagaggatggagccgaggcgccgtgg<br/> atagagcaggaggggcccggagtattgggaccggaacacacggaatgtgaaggccactcacagattgaccgagt<br/> gacctggggaccctgcgcggcgccctacaaccagagcgcgagccggttctcacaccatccagatgatgatggctgcga<br/> cgtgggggtcggacggcgcttctccgcggtaccagcaggacgcctacgagcgcaaggattacatcgccctgaac<br/> gaggacctgcgctcttggaccgcgcgacatggcggtcagatcacccagcgcaagtgggaggcgcccgctgtg<br/> gcggagcagttgagagcctacgtggaggcacgtgcgtggagtggctccgagatacctggagaacgggaaggag<br/> aactgcagcgcacggaccccccaagacgcatatgactcaccacgctgtctctgacctgaggccacctgaggtg<br/> ctgggcccctgagcttctaccctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggacacg<br/> gagctcgtggagaccaggcctgcaggggatggaaccttccagaagtggcgctgtgtggtggtccttctggacagga<br/> gcagagatacacctgccatgtgcagcatgagggtctccccagccccctaccctgagatgggagccgtcttcc</p>    |
| 464 | <p>ggctcccactccatgaggtatttctacacctccatgtcccggcccggccgaggagccccgcttcacgcgcgtgggc<br/> tacgtggacgacacgcagttcgtgctggttcgacagcgcgcgcgcgagccagaggatggagccgaggcgccgtgg<br/> atagagcaggaggggcccggagtattgggaccggaacacacggaatgtgaaggccagtcacagactgaccgagt<br/> gacctggggaccctgcgcggcgccctacaaccagagcgcgagccggttctcacaccatccagaggatgatggctgcg<br/> acgtggggccggacggcgcttctccgcggtaccaccagtagcctacgagcgcaaggattacatcgccctgaa<br/> agaggacctgcgctcttggaccgcgcgacatggcagctcagaccaccaagcacaagtgggaggcgcccatgt<br/> ggcgaggagcagtgagagcctacctggaggcgacgtgcgtggagtggctccgagatacctggagaacgggaagg<br/> agacactgcagcgcacggacgccccaaaacgcatatgactcaccacgctgtctctgacctgaagccacctgagg<br/> tgctgggcccctgagcttctaccctgcggagatcacactgacctggcagcgggatggggaggaccagaccaggaca<br/> cggagctcgtggagaccaggcctgcaggggatggaaccttccagaagtgggtggctgtgtggtggtccttctggacag<br/> gagcagagatacacctgccatgtgcagcatgagggttggccaagccccctaccctgagatgggagccgtcttcc</p> |
| 465 | <p>ggctcccactccatgaggtatttctacacctccgtgtcccggcccggccgaggagccccgcttcacgcgcgtgggc<br/> acgtggacgacacccagttcgtgaggttcgacagcgcgcgcgcgagtcgagagaggagccgaggcgccgtgg<br/> atagagcaggaggggcccggagtattgggaccggaacacacagatctacaaggccacgacagactgaccgagag<br/> agcctgcggaaacctgcgcggcgccctacaaccagagcgcgagccggttctcacacctccagagcatgtacggctgcg<br/> acgtggggccggacggcgcttctccgcggtaccagtagcctacgagcgcaaggattacatcgccctgaa<br/> cgaggacctgcgctcttggaccgcccggacacggcggtcagatcacccagcgcaagtgggaggcgcccgctg<br/> aggcgaggacagcggagacctacctggaggcgagtgctggtggagtggctccgagatacctggagaacgggaag<br/> gacaagctggagcgcgctgacccccaaagacacagtgaccaccacccatctctgacctgaggccacctga<br/> ggtgctgggcccctgggttctaccctgcggagatcacactgacctggcagcgggatggcgaggacaaactcaggac<br/> actgagcttggagaccagaccagcaggagatagaaccttccagaagtgggcagctgtgtggtggtccttctggaga<br/> agagcagagatacacatgccatgtacagcatgaggggctgccgaagccccctaccctgagatgggagccgtcttcc</p>          |
| 466 | <p>ggctcccactccatgaggtatttgcacaccgcatgtcccggcccggccgaggagccccgcttcacgcgcgtgggc<br/> tacgtggacgacacgcagttcgtgaggttcgacagcgcgcgcgcgagtcgagagaggagccgaggcgccgtgg<br/> atagagcaggaggggcccggagtattgggaccggaacacacagatcttcaagaccaacacacagactgaccgagag<br/> agcctgcggaaacctgcgcggcgccctacaaccagagcgcgagccggttctcacacctccagagcatgtacggctgcg<br/> acgtggggccggacggcgcttctccgcggtacacagtagcctacgagcgcaaggattacatcgccctgaa<br/> cgaggacctgcgctcttggaccgcccggacacggcggtcagatcacccagcgcaagtgggaggcgcccgctg<br/> ggcgaggacagcagagcctacctggaggcgacgtgcgtggagtggctccgagatacctggagaacgggaagg<br/> acacgctggagcgcgagcccccaagacacagtgaccaccacccatctctgacctgaggccacctgag<br/> gtgctgggcccctgggttctaccctgcggagatcacactgacctggcagcgggatggcgaggacaaactcaggac<br/> actgagcttggagaccagaccagcaggagatagaaccttccagaagtgggcagctgtgtggtggtccttctggaga<br/> agagcagagatacacatgccatgtacagcatgaggggctgccgaagccccctaccctgagatgggagccgtcttcc</p>             |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 467 | <p>ggctcccactccatgaggtattttacaccgccatgtcccggcccggccgaggagccccgcttcacgcagtgggc<br/> tacgtggacgacacccagttcgtgaggttcgacagcgacgcccgagtcaggatggcgccccggcgccatgg<br/> atagagcaggagggggccggagtattggaccgggagacacagatctccaagaccaacacacagacttaccgagag<br/> agcctgcggaacctgcgcggcgctacaaccagagcgaggccgggtctcacaccctccagaggatgtacggctgc<br/> gacgtggggccggacggcgctcctcccgggcatgaccagtccgcctacgacggcaaggattacatcgccctga<br/> acgaggacctgagctcctggaccgcggcgacacggcgctcagatcacccagcgcaagtgggagggcgcccggt<br/> gaggcgagcagtgagagcctacctggaggcgctgtgcgtggagtggctccgagatacctggagaacgggaag<br/> gagacactgcagcgcgcgacccccaaagacacatgtgaccaccaccccatctctgacctgaggccacctga<br/> ggtgctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggcgaggacaaactcagga<br/> caccgagcttgggagaccagaccagcaggagatagaaccttcagaagtgggcagctgtggtggtgccttctggag<br/> aagagcagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccatcttc</p>  |
| 468 | <p>ggctcccactccatgaggtatttcacaccgccatgtcccggcccggccgaggagccccgcttcacccgtggg<br/> ctacgtggacgacacgctgttcgtgaggttcgacagcgacgccacgagtcagggaaggagccggcgccatg<br/> gatagagcaggagggggccggagtattggaccgggagacacagatctccaagaccaacacacagacttaccgaga<br/> gagcctgcggaacctgcgcggcgctacaaccagagcgaggccgggtctcacaccctccagaggatgtacggctg<br/> cgacgtggggccggacggcgcgctcctcccgggcgataaccagtacgcctacgacggcaaggattacatcgccctg<br/> aacgaggacctgcgctcctggaccgcggcgacacggcgctcagatctccagcgcaagtgggagggcgcccggt<br/> gtggcgagcagctgagagcctacctggaggcgagtgctggtgagtggtccgagatacctggagaacgggaag<br/> gacaagctggagcgcgctgacccccaaagacacacgtgaccaccaccccatctctgacctgaggccacctga<br/> ggtgctggggccctgggttctaccctgcggagatcacactgacctggcagcgggatggcgaggacaaactcaggac<br/> actgagcttgggagaccagaccagcaggagatagaaccttcagaagtgggcagctgtggtggtgccttctggaga<br/> agagcagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccgtcttc</p> |
| 469 | <p>ggctcccactccatgaggtatttcacacctcgtgtcccggcccggccgaggagccccgcttcacccgtggg<br/> tacgtggacgacacgctgttcgtgaggttcgacagcgacgccacgagtcagggaaggagccggcgccatgg<br/> atagagcaggagggggccggagtattggaccgggagacacagatctccaagaccaacacacagacttaccgagag<br/> agcctgcggaacctgcgcggcgctacaaccagagcgaggccgggtctcacaccctccagaggatgtacggctgcg<br/> acgtggggccggacggcgctcctcccgggcgataaccagtacgcctacgacggcaaggattacatcgccctgaa<br/> cgaggacctgcgctcctggaccgcggcgacacggcgctcagatcacccagcgcaagtgggagggcgcccggtg<br/> ggcgaggcagctgagagcctacctggaggcgagtgctggtgagtggtccgagatacctggagaacgggaagg<br/> agacactgcagcgcgcgacccccaaagacacacgtgaccaccaccccatctctgacctgaggccacctgag<br/> gtgctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggcgaggacaaactcaggac<br/> actgagcttgggagaccagaccagcaggagatagaaccttcagaagtgggcagctgtggtggtgccttctggaga<br/> agagcagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccgtcttc</p>    |
| 470 | <p>ggctcccactccatgaggtattttacaccgccatgtcccggcccggccgaggagccccgcttcacccgtggg<br/> tacgtggacgacacgctgttcgtgaggttcgacagcgacgccacgagtcagggaaggagccggcgccatgg<br/> atagagcaggagggggccggagtattggaccgggagacacagatctccaagaccaacacacagacttaccgagag<br/> aacctgcgcaccgcgctccgcgctacaaccagagcgaggccgggtctcacatcatccagaggatgtacggctgcg<br/> acgtggggccggacggcgctcctcccggggatgaccaggacgcctacgacggcaaggattacatcgccctgaa<br/> cgaggacctgagctcctggaccgcggcgacacggcgctcagatcacccagcgcaagtgggagggcgcccggtg<br/> ggcgaggcaggacagagcctacctggaggcgctgtgctggtgagtcgctccgagatacctggagaacgggaagg<br/> gacactgcagcgcgcgacccccaaagacacatgtgaccaccaccccatctctgacctgagggtcacctgaggt<br/> gctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggcgaggacaaactcaggacac<br/> cgagcttgggagaccagaccagcaggagatagaaccttcagaagtgggcagctgtggtggtgccttctggagaag<br/> agcagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccgtcttc</p>  |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 471 | <p>ggctccactccatgaggtatttctacaccgcatgtcccggcccggccgaggagccccgcttcatcaccgtgggc<br/> tacgtggacgacacgctgttcgtgaggttcgacagcgacgccacgagtcaggaggaaggagccgaggcccatgg<br/> atagagcaggaggggcccggagtattgggaccgggagacacagatctccaagaccaacacacagacttaccgagag<br/> aacctgcgcaccgcgctccgcgcctacaaccagagcgaggccgggtctcatcatccagaggatgtacggctgcg<br/> acgtggggccggacggcgccctctccgagggtatgaccaggacgacctacgacggcaaggattacatgccctgaa<br/> cgaggacctgagctcctggaccgcgccggacaccgcggtcagatcacccagcgcaagtgggagggcgccctgtg<br/> ggcggagcagctgagagcctacctggaggccctgtgctggagtcgctccgagatacctggagaacgggaaggga<br/> gacactgcagcgcgccgacccccaaagacacatgtgaccaccacccatctctgacctgaggtcacctgaggt<br/> gctgggcccctgggcttctacctgaggatcacactgacctggcagcgggatggcgaggacaaaactcaggacac<br/> cgagcttgggagaccagaccagcaggagatagaaccttccagaagtgggcagctgtggtggtgccttctggagaag<br/> agcagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccgtcttc</p> |
| 472 | <p>ggctccactccatgaggtatttctacaccgcatgtcccggcccggccgaggagccccgcttcatcagtgaggc<br/> tacgtggacgacacccagttcgtgaggttcgacagcgacgcccgagtcaggagcggagccccggcgccatgg<br/> atagagcaggaggggcccggagtattgggaccggaacacacagatcttcaagaccaacacacagacttaccgagaga<br/> acctgcggatcgcgctccgcgcctacaaccagagcgaggccgggtctcacacttgccagacgatgtatggctcgac<br/> gtggggccggacggcgccctctccggggcataaccagtacgcctacgacggcaagattacatgccctgaacg<br/> aggacctgagctcctggaccgcgccggacaccgcggtcagatcacccagcgcaagtgggagggcgccctgtgag<br/> gaggagcagctgagagcctacctggaggccctgtgctggagtggtccgcagacacctggagaacgggaaggag<br/> aactgcagcgcgccgacccccaaagacacacgtgaccaccacccagttctgacctgagggccacctgaggt<br/> gctgggcccctgggcttctacctgaggatcacactgacctggcagcgggatggcgaggacaaaactcaggacact<br/> gagcttgggagaccagaccagcaggagatagaaccttccagaagtgggcagctgtggtggtgccttctggagaaga<br/> gcagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccatcttc</p>      |
| 473 | <p>ggctccactccatgaggtatttctacaccgcatgtcccggcccggccgaggagccccgcttcatcagtgaggc<br/> tacgtggacgacacccagttcgtgaggttcgacagcgacgcccgagtcaggagcggagccccggcgccatgg<br/> atagagcaggaggggcccggagtattgggaccggaacacacagatcttcaagaccaacacacagacttaccgagaga<br/> acctgcggatcgcgctccgcgcctacaaccagagcgaggccgggtctcatcatccagaggatgtatggctcgac<br/> ctggggcccggacggcgccctctccggggcatgaccagtccgctacgacggcaagattacatgccctgaacg<br/> aggacctgagctcctggaccgcgccggacaccgcggtcagatcacccagcgcaagtgggagggcgccctgtgg<br/> cggagcagctgagagcctacctggaggccctgtgctggagtggtccgcagatacctggagaacgggaaggaga<br/> cactgcagcgcgccgacccccaaagacacacgtgaccaccacccagttctgacctgagggccacctgaggtg<br/> ctgggcccctgggcttctacctgaggatcacactgacctggcagcgggatggcgaggacaaaactcaggacactg<br/> agcttgggagaccagaccagcaggagatagaaccttccagaagtgggcagctgtggtggtgccttctggagaagag<br/> cagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccatcttc</p>         |
| 474 | <p>ggctccactccatgaggtatttctacaccgcatgtcccggcccggccgaggagccccgcttcatcagtgaggc<br/> tacgtggacgacacccagttcgtgaggttcgacagcgacgcccgagtcaggagcggagccccggcgccatgg<br/> atagagcaggaggggcccggagtattgggaccgggagacacggaacatgaaggcctccgcgagacttaccgaga<br/> gaacctgcggatcgcgctccgcgcctacaaccagagcgaggccgggtctcatcatccagaggatgtatggctgcg<br/> acctggggcccggacggcgccctctccggggcatgaccagtccgctacgacggcaaggattacatgccctgaa<br/> cgaggacctgagctcctggaccgcgccggacaccgcggtcagatcacccagcgcaagtgggagggcgccctgtg<br/> ggcggagcagctgagagcctacctggaggccctgtgctggagtggtccgcagatacctggagaacgggaaggga<br/> gacactgcagcgcgccgacccccaaagacacacgtgaccaccacccagttctgacctgagggccacctgaggtg<br/> tgctgggcccctgggcttctacctgaggatcacactgacctggcagcgggatggcgaggacaaaactcaggacac<br/> tgagcttgggagaccagaccagcaggagatagaaccttccagaagtgggcagctgtggtggtgccttctggagaag<br/> agcagagatacacatgccatgtacagcatgaggggctgccgaagcccctcacctgagatgggagccatcttc</p>    |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 475 | <p>tgctccactccatgaagtattttctacatccgtgtcccgccctggccgaggagccccgttcatctcagtgggcta<br/> cgtggagacacgcagttcgtgcggttcgacagcgacgccgcgagtcagagaggggagccgaggcgccgtggg<br/> tgagcagggaggggcccggagtattgggaccgggagacacagaagtacaagcgccaggcacagactgaccgagt<br/> agcctgcggaacctgcgcggcgcttacaaccagagcgaggccgggtctcacacctccagtgatgtgtgctgcg<br/> acctggggcccgcggcgccctctccgcgggtatgaccagtacgctacgacggaaggattacatcgccctgaa<br/> cgaggacctgcgctcctggaccgccgagacaccgcggtcagatcacccagcgcaagtgggagggcgccctg<br/> aggcgaggagcgcggagagcctacctggaggcgacgtgctggagtggctccgagatacctggagaacgggaag<br/> gagacactgcagcgcggaacacccaaagacacacgtgacccaccatccagtctctgacctgaggccacctga<br/> gggtgctgggcccctgggcttctacctgcggagatcacactgacctggcagtgggatggggaggacaaactcaggac<br/> accgagcttgggagaccaggccagcaggagatggaaccttcagaagtgggcagctgtggtgcttctggaga<br/> agagcagagatacacgtgccatgtgcagcacgaggggctgccggagcccctcacctgagatgggagccgtcttc</p>            |
| 476 | <p>tgctccactccatgaggtattttctacaccgtgtgtcccgcccagccgaggagccccacttcatcgagtgaggct<br/> acgtggagacacgcagttcgtgcggttcgacagcgacgccgcgagtcgaagaggggagccgaggcgccgtgg<br/> gtggagcagggaggggcccggagtattgggaccgggagacacagaagtacaagcgccaggcacagactgaccgagt<br/> gaacctgcggaactacgcggcgcttacaaccagagcgaggccgggtctcacacctccagaggatgtacgctgc<br/> gacctggggcccgcggcgccctctccgcgggtatgaccagtccgctacgacggcaaggattacatcgccctga<br/> acgaggacctgcgctcctggaccgccgagacacagcggtcagatcacccagcgcaagtgggagggcgccctg<br/> gaggcgaggcagtgagagcctacctggaggcgagtgctggagtggctccgagatacctggagaacgggaa<br/> ggagacactgcagcgcggaacacccaaagacacacgtgacccaccatccagtctctgacctgaggccacctg<br/> agggtgctgggcccctgggcttctacctacggagatcacactgacctggcagcgggatggcgaggacaaactcagg<br/> acaccgagcttgggagaccaggccagcaggagatggaaccttcagaagtgggcagctgtggtgcttctggaga<br/> gaagagcagagatacacgtgccatgtgcagcacgaggggctgccggagcccctcacctgagatgggagccatctt<br/> cc</p> |
| 477 | <p>ggctccactccatgaggtattttctacaccgtgtgtcccgcccggccgaggagccccacttcatcgagtgaggc<br/> tacgtggagacacgcagttcgtgcggttcgacagcgacgccgcgagtcagagaggggagccgaggcgccgtgg<br/> gtggagcagggaggggcccggagtattgggaccgggagacacagaagtacaagcgccaggcacagactgaccgagt<br/> gagcctgcggaacctgcgcggcgcttacaaccagagcgaggccgggtctcacatcatccagaggatgtatgctgc<br/> gacgtggggcccgcggcgccctctccgcgggtatgaccagtacgctacgacggaaggattacatcgccctga<br/> acgaggatctgcgctcctggaccgccgagacacgcggtcagatcacccagcgcaagtgggagggcgccctg<br/> aggcgaggcagctgagagcctacctggaggcctgtgctggagtggctccgagatacctgagaatgggaagg<br/> agacactgcagcgcggaacacccaaagacacacgtgacccaccatccagtctctgacctgaggccacctgag<br/> gtgctgggcccctgggcttctacctgcggagatcacactgacctggcagtgggatggggaggacaaactcaggaca<br/> ctgagcttgggagaccaggccagcaggagatggaaccttcagaagtgggcagctgtggtgcttctggaga<br/> gagcagagatacacgtgccatgtgcagcacgaggggctgccggagcccctcacctgagatgggagccgtcttc</p>             |
| 478 | <p>ggctccactccatgaggtattttctccacatccgtgtcttggcccggccgaggagccccgttcatcgagtgaggct<br/> acgtggagacacgcagttcgtgcggttcgacagcgacgccgcgagtcgaagaggggagccgaggcgccgtgg<br/> gtggagcagggaggggcccggagtattgggaccgggagacacagaagtacaagcgccaggcacaggtgaccgagt<br/> gaacctgcggaactgcgcggcgcttacaaccagagcgaggacgggtctcacacctccagaggatgtttgctgc<br/> gacctggggcccgcggcgccctctccgcgggtataaccagttcgctacgacggcaaggattacatcgccctgaa<br/> cgaggatctgcgctcctggaccgccgagacacggcggtcagatcacccagcgcaagtgggagggcgccctg<br/> ggcgaggcagcggagagcctacctggaggcgacgtgctggagtggctccgagatacctggagaacgggaagg<br/> agacactgcagcgcggaacacccaaagacacacgtgacccaccatccagtctctgacctgaggccacctgag<br/> gtgctgggcccctgggcttctacctgcggagatcacactgacctggcagtgggatggggaggacaaactcaggaca<br/> ccgagcttgggagaccaggccagcaggagatggaaccttcagaagtgggcagctgtggtgcttctggaga<br/> gagcagagatacacgtgccatgtgcagcacgaggggctgccggagcccctcacctgagatggaagccgtcttc</p>            |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 479 | <p>tgctcccactccatgaggtatttctacaccgccgtgtcccggccccggccgaggagagccccgcttcacgcagtgggc<br/> tacgtggacgacacgcagttcgtgcggttcgacagcgacgccgcgagtcgaagaggggagccgcggcgccgtgg<br/> gtggagcaggagggggccggagtattgggaccgggagacacagaagtacaagcgccaggcacagactgaccgagt<br/> gaacctgcggaaactgcgcggcgctacaaccagagcgaggccgggtctcacacctccagaggatgtatggctgc<br/> gacctggggccccacggcgccctcctccggtataaccagttcgctacgagcgcaaggattacatcgccctgaa<br/> tgaaggacctgcgctcctggaccgcccgacaaggcggtcagatcaccagcgcaagtgaggagcgggccctga<br/> ggcggagcagcgagagcctacctggagggcacgtgcgtggagtggctccgcagatacctggagaacgggaaga<br/> agacgctgcagcgcgcggaacacccaaagacacacgtgacccaccatccagtctctgacatgagggccacctgag<br/> gtgctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggcgaggaccaaactcaggac<br/> accgagcttgggagaccaggccagcaggagatggaacctccagaagtgggcagctgtggtggtgccttctggaga<br/> agagcagagatacacgtgccatgtgcagcacgaggggctgccagagcccctcacctgagatgggggccatcttcc</p>            |
| 480 | <p>tgctcccactccatgaggtatttgcacaccgccgtgtcccggccccggccgaggagagccccgcttcacgtcagtgggc<br/> tacgtggacgacacgcagttcgtgcggttcgacagcgacgccgcgagtcgagagggggagccccggcgccgtgg<br/> gtggagcaggagggggccggagtattgggaccgggagacacagaagtacaagcgccaggcacaggctgaccgagt<br/> gaacctgcggaaactgcgcggcgctacaaccagagcgaggacgggtctcacacctccagtggtatgtatggctgc<br/> gacctggggccccacggcgccctcctccggtatgacctccgctacgagcgcaaggattacatcgccctga<br/> acgaggacctgcgctcctggaccgcccgagacacggcggtcagatcaccagcgcaagtgaggagcgggccctga<br/> gaggcgagcagtgaggagcctacctggagggcacgtgcgtggagtggctccgcagatacctggagaacgggaag<br/> ggagacactgcagcgcgcggaacacccaaagacacacgtgacccaccatccagtctctgacatgagggccacctg<br/> agggtgctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggcgaggaccaaactcagg<br/> acaccgagcttgggagaccaggccagcaggagatggaacctccagaagtgggcagctgtggtggtgccttctgga<br/> gaagagcagagatacacgtgccatgtgcagcacgaggggctgccagagcccctcacctgagatgggagccatctt<br/> cc</p> |
| 481 | <p>Tgctcccactccatgaggtatttgcacaccgccgtgtcccggccccggccgaggagagccccgcttcacgtcagtggg<br/> ctacgtggacgacacgcagttcgtgcggttcgacagcgacgccgcgagtcgagagggggagccgcggcgccgtg<br/> ggtggagcaggagggggccggagtattgggaccgggagacacagaagtacaagcgccaggcacaggctgaccgag<br/> tgagcctgcggaaactgcgcggcgctacaaccagagcgaggacgggtctcacacctccagaggatgtatggctg<br/> cgacctggggccccacggcgccctcctccggtatgacctccgctacgacggcaaggattacatcgccctg<br/> aacgaggacctgcgctcctggaccgcccgagacaccgggtcagatcaccagcgcaagtgaggcgggccctga<br/> gcgcgaggagcagctgagagcctacctggagggcacgtgcgtggagtggctccgcagatacctggagaacgggaag<br/> gagacactgcagcgcgagaaaccccaaaagacacacgtgacccaccaccccctctctgacctgagggccacctga<br/> ggtgctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggggaggaccagaccagga<br/> caccgagcttgggagaccaggccagcaggagatggaacctccagaagtgggcagctgtggtggtgccttctggac<br/> aagagcagagatacacgtgccatgtgcagcacgaggggctgcaagagcccctcacctgagctgggagccatcttcc</p>            |
| 482 | <p>tgctcccactccatgaggtatttgcacaccgccgtgtcccggccccggccgaggagagccccgcttcacgtcagtgggc<br/> tacgtggacgacacgcagttcgtgcggttcgacagcgacgccgcgagtcgagagggggagccgcggcgccgtgg<br/> gtggagcaggagggggccggagtattgggaccgggagacacagaagtacaagcgccaggcacaggctgaccgagt<br/> gagcctgcggaaactgcgcggcgctacaaccagagcgaggacgggtctcacacctccagaggatgtctggctgc<br/> gacctggggccccacggcgccctcctccggtatgacctccgctacgagcgcaaggattacatcgccctga<br/> acgaggacctgcgctcctggaccgcccgagacaccgggtcagatcaccagcgcaagtgaggcgggccctga<br/> cgcgaggagcagctgagagcctacctggagggcacgtgcgtggagtggctccgcagatacctggagaacgggaag<br/> agacactgcagcgcgagaaaccccaaaagacacacgtgacccaccaccccctctctgacctgagggccacctga<br/> gtgctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggggaggaccagaccagga<br/> accgagcttgggagaccaggccagcaggagatggaacctccagaagtgggcagctgtggtggtgccttctggac<br/> agagcagagatacacgtgccatgtgcagcacgaggggctgcaagagcccctcacctgagctgggagccatcttcc</p>                 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 483 | tgctccactccatgaggtattttacaccgccgtgtcccggcccgccgaggagagccccgcttcacgcagtgggc<br>tacgtggacgacacgcagttcgtgcggttcgacagcgacgcccgcagtcgaagaggggagccgcccggcgccgtgg<br>gtggagcaggaggggcccggagtattgggaccgggagacacagaagtacaagcgccaggcacagactgaccgagt<br>gaggctgcggaacctgcgcggcgctacaaccagagcgaggccgggtctcacacctccagtggtatgtatggctgc<br>gacctggggcccgcggcgccctcctccgcggtatgaccagtcgcctacgacggcaaggattacatcgccctga<br>acgaggacctgcgctcctggaccgcccggacacggcggtcagatcaccagcgcaagtgaggagggcccgtg<br>gcggcgaggagcagagagcctacctggagggcacgtgcgtggagtggctccgagatacctggagaacgggaa<br>ggagacactgcagcgcgcggaacacccaaagacacacgtgacccaccatttggtctctgacctgaggccacctga<br>ggtgctggggccctgggcttctaccctgcggagatcacactgacctggcagcgggatggcgaggaccaaactcagga<br>caccgagcttgaggagaccagggcagcaggagatggaacctccagaagtgggcagctgtggtggtccttctggag<br>aagagcagagatacacgtgccatgtgcagcacgagggggtgccggagcccctaccctgagatgggagccatcttc<br>c |
| 484 | ggctccactcctgaagtatttcacacttccgtgtcccggcccgccgaggagagccccgcttcacgtctgtgggcta<br>cgtggacgacaccagttcgtgcgcttcgacaacgacgcccgcagtcaggatggtgccgcccggcgccgtggatg<br>gagcaggaggggtcagagtattgggaccgggagacacggagcgccaggacaccgcacagatttccgagtgaac<br>ctgcggacgctgcgcggcgctacaatcagagcgaggccgggtctcacacctgcagtggatgcatggctgcgagc<br>tggggcccgcggcgcttctcgcgggtatgaacagttcgcctacgacggcaaggattatctcacctgaatgagg<br>acctgcgctcctggaccgcggtggacacggcggtcagatctccgagcaaaagtcaaatgatcctctgaggcggag<br>caccagagagcctacctggaagacacatgcgtggagtggctccacaaatacctggagaaggggaaggagacactgc<br>ttcacctggagccccaaagacacacgtgactcaccacccatcttgacctgaggccacctgaggtgctggggccc<br>tgggcttctacctgaggatcacactgacctggcagcaggatggggagggccataccagacacggagctcgt<br>ggagaccaggcctgcaggggatggaacctccagaagtgggcagctgtggtggtccttctggagaggagcagaga<br>tacacgtgccatgtgcagcatgaggggctaccgagcccgtcacctgagatggaagccggcttcc                    |

[0093] In some embodiments, the plurality of the HLA-antigen polypeptide complexes of the randomized peptide antigen libraries comprise at least about  $10^5$  different HLA-antigen polypeptide complexes. Components of the  $10^5$  different HLA-antigen polypeptide complexes include, collectively, at least about  $10^5$  different randomized antigen polypeptides. In some embodiments, the plurality of the HLA-antigen polypeptide complexes of the randomized peptide antigen libraries comprise at least about  $10^7$  different HLA-antigen polypeptide complexes. Components of the  $10^7$  different HLA-antigen polypeptide complexes include, collectively, at least about  $10^7$  different randomized antigen polypeptides. In some embodiments, the plurality of the HLA-antigen polypeptide complexes of the randomized peptide antigen libraries comprise at least about  $10^9$  different HLA-antigen polypeptide complexes. Components of the  $10^9$  different HLA-antigen polypeptide complexes include, collectively, at least about  $10^9$  different randomized antigen polypeptides. In some embodiments, the plurality of

the HLA-antigen polypeptide complexes of the randomized peptide antigen libraries comprise at least about  $10^{11}$  different HLA-antigen polypeptide complexes. Components of the  $10^{11}$  different HLA-antigen polypeptide complexes include, collectively, at least about  $10^{11}$  different randomized antigen polypeptides.

**[0094]** In some embodiments, the plurality of the HLA-antigen polypeptide complexes of the randomized peptide antigen libraries further comprise a  $\beta$ 2-microglobulin polypeptide, which interacts with and stabilizes the HLA-antigen polypeptide complexes on the surface of the cell. The amino acid sequence of human  $\beta$ 2-microglobulin polypeptide is set forth in NCBI Seq. Ref. NP\_004039. In some embodiments, the human  $\beta$ 2-microglobulin polypeptide amino acid sequence of the present disclosure is a functional naturally occurring variant of the human  $\beta$ 2-microglobulin polypeptide having an amino acid sequence at least about 90%, 95%, 97%, 98%, or 99% identical to the human  $\beta$ 2-microglobulin polypeptide disclosed as NCBI Seq. Ref. NP\_004039.

**[0095]** The present disclosure also includes antigen screening libraries of a plurality HLA-antigen polypeptide where the  $\beta$ 2-microglobulin is constitutively expressed by a cell. In some embodiments, the  $\beta$ 2-microglobulin is encoded by a first nucleic acid, the randomized antigen polypeptide encoded by a second nucleic acid, and the HLA polypeptide is encoded by a third nucleic acid. In other embodiments, the  $\beta$ 2-microglobulin is encoded by a first nucleic acid and the randomized antigen polypeptide and the HLA polypeptide is encoded by a second nucleic acid. When encoded by the first nucleic acid, the  $\beta$ 2-microglobulin can be transduced, transfected, or transformed into a cell before or after the second nucleic acid or the third nucleic acid.

**[0096]** In some embodiments of the present disclosure, the  $\beta$ 2-microglobulin is fused to at least one of the randomized antigen polypeptides of the antigen screening library using techniques known to those of ordinary skill in the art. In these embodiments, the HLA polypeptides may or may not be a component of the antigen screening library. In other embodiments of the present disclosure, at least

one of the HLA polypeptides is fused to at least one of the randomized antigen polypeptides of the antigen screening library using techniques known to those of ordinary skill in the art. In these embodiments, the  $\beta$ 2-microglobulin can be expressed by a cell that is transduced, transfected, or transformed to express other components of the antigen screening library, such as the randomized antigen polypeptides and the HLA polypeptides. Similar to other embodiments described herein, the  $\beta$ 2-microglobulin is constitutively expressed by the cell. In certain of these embodiments, the cell is a yeast cell. In other embodiments, the  $\beta$ 2-microglobulin is not expressed by the cell that is transduced, transfected, or transformed to express other components of the antigen screening library, such as the randomized antigen polypeptides and the HLA polypeptides. In certain of these embodiments, the cell is a mammalian cell.

**[0097]** In addition to the (a) randomized antigen polypeptide, (b) MHC I HLA molecule, and (c)  $\beta$ 2-microglobulin features of the HLA-antigen polypeptide complex of the randomized peptide antigen libraries, the HLA-antigen polypeptide complexes of the present disclosure can further include (d) a signal sequence, (e) polypeptide linkers between any or all of (a), (b), or (c), (f) a membrane tethering domain, and, optionally, (g) an epitope tag, such as a FLAG tag, a c-Myc tag, a His-tag, a hemagglutinin (HA) tag, a VSVg tag, a V5 tag, an AU1 tag, an AU5 tag, a Glu-Glu tag, an OLLAS tag, a T7 tag, an S-TagHSV tag, a KT3 tag, a TK15 tag, an Fc tag, an Xpress tag, a Ty tag, a Strep tag, an NE tag, an E tag, a C-tag, and/or an AviTag. In some embodiments, the HLA-antigen complexes do not comprise an epitope tag. However, in some embodiments, at least one or more of each of the plurality of HLA-antigen complexes of the randomized peptide antigen libraries comprise the epitope tag which allows for confirmation of expression of at least one of the HLA-antigen complexes using an antibody specific for the epitope. In some embodiments, each of the plurality of HLA-antigen complexes of the randomized peptide antigen libraries comprise the epitope tag.

**[0098]** In some embodiments, the membrane tethering domain comprises a polypeptide linker

separating the membrane tethering domain from one or more other features ((a)-(e) and (g)) of the HLA-antigen polypeptide complex. In some embodiments, the features ((a)-(g)) of the HLA-antigen polypeptide complex are expressed as a single polypeptide. In some embodiments, the (b) HLA molecule (e.g., HLA polypeptide), the (a) randomized antigen polypeptide, and the (c)  $\beta$ 2-microglobulin polypeptide comprise a single polypeptide. In some embodiments, the (b) HLA polypeptide and the (a) randomized antigen polypeptide are expressed as a single polypeptide, while, the (c)  $\beta$ 2-microglobulin is expressed separately. For example, the (c)  $\beta$ 2-microglobulin can be supplied from a separate polypeptide encoded by the same nucleic acid that expresses the (a) randomized antigen polypeptide and the (b) HLA polypeptide, a separate nucleic acid, or endogenously produced by the cell. In some embodiments, the randomized antigen polypeptide is N-terminal to the HLA polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide. In some embodiments, the randomized antigen polypeptide is C-terminal to the HLA polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide. In some embodiments, the randomized antigen polypeptide is N-terminal to the HLA polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide. In some embodiments, the randomized antigen polypeptide is C-terminal to the HLA polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide.

**[0099]** The (a) a randomized antigen polypeptide, (b) a major histocompatibility class I (MHC I) HLA molecule, and (c) a  $\beta$ 2-microglobulin, can be separated by at least one flexible polypeptide linker, such as a first flexible polypeptide linker, a second flexible polypeptide linker, a third flexible polypeptide linker, a fourth flexible polypeptide linker, a fifth flexible polypeptide linker, or more flexible polypeptide linkers. In some embodiments, the at least one flexible polypeptide linker can range between about 3 and about 100 amino acid residues in length, between about 5 and about 80 amino acid residues in length, between about 10 and about 70 amino acid residues in length, between

about 3 and about 100 amino acid residues in length, between about 20 and about 60 amino acid residues in length. In some embodiments, the linker can be about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length. In some embodiments, the linker can be a glycine linker, or a Gly-Ser linker of the formula (GGGGS)<sub>x</sub>, wherein X is 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10. In some embodiments, the linker can suitably comprise a protease cleavage site such as a thrombin cleavage site.

**[00100]** In some embodiments, the HLA-antigen polypeptide complexes of the randomized peptide antigen libraries comprise a signal polypeptide which directs the HLA-antigen polypeptide complex to the cell surface via the secretory pathway. This signal peptide is cleaved in the endoplasmic reticulum and is not expressed by the HLA-antigen polypeptide complex when located on the cell-surface. The signal sequence can be any suitable sequence such as an endogenous HLA leader sequence, or a heterologous leader sequence imported from a different secretory or transmembrane molecule, such as an immunoglobulin leader sequence.

**[00101]** The HLA-antigen polypeptide complexes further comprise a membrane tethering domain, such as an anchor domain from a glycosylphosphatidylinositol (GPI) protein and/or a domain from yeast proteins having internal repeats (PIR protein). This membrane tethering domain can comprise a transmembrane domain or a domain that interacts with a cell surface protein. In some embodiments, the membrane tethering domain comprise at least one anchor domain of a GPI protein selected from the group consisting of yeast Aga2, Cwp1p, Cwp2p, Aga1p, Tip1p, Flo1p, Sed1p, YCR89w, and Tir1p and/or a PIR protein selected from the group consisting of yeast Pir1p, Pir2p, Pir3p, Pir4p, and Pir5p. A non-limiting example of membrane domain tethering is provided in **FIG. 1B**.

**[00102]** In other embodiments, components of the antigen screening libraries of a plurality HLA-antigen polypeptide complexes are expressed as more than one polypeptide and include a cleavage sequence which separates components of the antigen screening libraries of a plurality HLA-antigen

polypeptide complexes from one another. For example, the randomized peptide antigen is separated from the HLA polypeptide and/or from the Beta-2 ( $\beta$ 2) microglobulin polypeptide by the cleavage sequence. As another example, the HLA peptide is separated from the Beta-2 ( $\beta$ 2) microglobulin polypeptide by the nucleotide encoded cleavage sequence. In some embodiments, the components of the antigen screening libraries are separated by more than one cleavage sequence. Suitable cleavage sequences are known to those of ordinary skill in the art and include, but are not limited to, self-cleaving peptides (P2A, T2A, F2A, and E2A), proteolytic cleavage sites (a 3C site, a thrombin site, a TEV site, a Factor Xa site, and an EKT site) and an internal ribosome entry sequence (IRES).

**[00103]** In some embodiments, the antigen screening library and/or the HLA-antigen polypeptide complexes can be expressed by one or more cells that can easily be transfected, transduced, electroporated, or transformed with the nucleic acids described herein. In some embodiments, the antigen screening library and/or the HLA-antigen polypeptide complexes are expressed on a plurality of cells. In some embodiments, each cell of the plurality of cells expresses a specific HLA-antigen complex of the HLA-antigen polypeptide complexes and/or another component of the antigen screening library. In some embodiments, a nucleic acid or a plurality of nucleic acids encode the antigen screening library and/or the HLA-antigen polypeptide complexes. In some embodiments, the antigen screening library and/or the HLA-antigen polypeptide complexes comprise prokaryotic cells. In some embodiments, the cell expressing the HLA-antigen polypeptide complexes comprise eukaryotic cells. In some embodiments, the eukaryotic cells comprise yeast cells. In some embodiments, the yeast cells are a cell of *Saccharomyces cerevisiae*. In some embodiments, the *Saccharomyces cerevisiae* is of the strain EBY100. Transforming *Saccharomyces cerevisiae* with nucleic acids can be achieved by standard methods as long as the efficiency is sufficient to produce at least  $10^7$ ,  $10^8$ ,  $10^9$ , or  $10^{10}$  transformants.

**[00104]** In addition to the plurality of HLA-antigen polypeptide complexes of the antigen screening

libraries described above, the present technology also includes at least two or more antigen screening libraries having HLA-antigen polypeptide complexes that differ from those described above. In some embodiments, the HLA-antigen polypeptide complexes have fewer components and/or at least one different component than the plurality of HLA-antigen polypeptide complexes described above. For example, in some embodiments, HLA-antigen polypeptide complexes can also comprise (a) an HLA polypeptide having a peptide binding cleft; and (b) a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209 that specifically binds to the peptide binding cleft of the HLA polypeptide. In these embodiments, the HLA polypeptide, and the randomized antigen polypeptide comprise a single polypeptide. Also in these embodiments, the single polypeptide further comprises a first flexible polypeptide linker separating the HLA polypeptide from the randomized antigen polypeptide. When expressed on a single polypeptide separated by the first flexible polypeptide linker, the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide or the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide.

**[00105]** As another example, in some embodiments, antigen screening libraries of the present technology comprise (a) an HLA polypeptide constitutively expressed by one or more yeast cells, the HLA polypeptide comprising a peptide binding cleft, and (b) a plurality of Beta-2 ( $\beta$ 2) microglobulin polypeptide-antigen polypeptide complexes. In these embodiments, the plurality of Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes include a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209, wherein the randomized antigen polypeptide specifically binds to the peptide binding cleft of the HLA polypeptide; and (c) a Beta-2 ( $\beta$ 2) microglobulin polypeptide. In these embodiments, the randomized antigen polypeptide and the  $\beta$ 2-microglobulin polypeptide comprise a single polypeptide. Also, in these embodiments, the single polypeptide further comprises a first flexible polypeptide linker separating the Beta-2 ( $\beta$ 2)

microglobulin polypeptide from the randomized antigen polypeptide. When expressed on a single polypeptide separated by the first flexible polypeptide linker, the randomized antigen polypeptide is N-terminal to the Beta-2 ( $\beta$ 2) microglobulin polypeptide on the single polypeptide or the randomized antigen polypeptide is C-terminal to the Beta-2 ( $\beta$ 2) microglobulin polypeptide on the single polypeptide.

#### **Nucleic Acids Encoding HLA-Antigen Polypeptide Complexes**

**[00106]** Also disclosed herein are nucleic acids that encode HLA-antigen polypeptide complexes of the antigen screening libraries. Nucleic acids that encode the HLA-antigen polypeptide complexes of the current disclosure minimally encode: (a) a randomized antigen polypeptide, (b) an MHC I HLA molecule, and a (c)  $\beta$ 2-microglobulin. In addition to the (a) randomized antigen polypeptide, (b) MHC I HLA molecule, and (c)  $\beta$ 2-microglobulin features of the HLA-antigen polypeptide complex of the randomized peptide antigen libraries encoded by one or more nucleic acids, the HLA-antigen polypeptide complexes of the present disclosure further include nucleic acids which encode (d) a signal sequence, (e) polypeptide linkers between any or all of (a), (b), or (c), (f) a membrane tethering domain, and, optionally, (g) an epitope tag, such as a FLAG tag, a c-MYC tag, a HIS-tag, a hemagglutinin tag, a VSVg tag, a V5 tag, an AU1 tag, an AU5 tag, a Glu-Glu tag, an OLLAS tag, a T7 tag, an S-Tag, an HSV tag, a KT3 tag, a TK15 tag, an Fc tag, an Xpress tag, a Ty tag, a Strep tag, an NE tag, an E tag, a C-tag, and/or an AviTag (**FIG. 1A** and **FIG. 1B**).

**[00107]** In some embodiments, the nucleic acid encoding the (f) membrane tethering domain may further encode (e) one or more polypeptide linkers separating the membrane tethering domain from other features of the HLA-antigen polypeptide complex. In some embodiments, the nucleic acid encodes one or more flexible polypeptide linkers which separate the (a) HLA polypeptide from the (b) randomized antigen polypeptide and the (c)  $\beta$ 2-microglobulin polypeptide when all three features are encoded on the single nucleic acid.

**[00108]** In some embodiments, the nucleic acid encoding the single polypeptide further comprises nucleotides which encode a first flexible polypeptide linker and a second flexible polypeptide linker, wherein the nucleotide sequence encoding the first flexible polypeptide linker separates the nucleotide sequence encoding the HLA polypeptide from the nucleotide sequence encoding the randomized antigen polypeptide, and the nucleotide sequence encoding the second flexible polypeptide linker separates the nucleotide sequence encoding the randomized antigen polypeptide from the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide. In some embodiments, once expressed, the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

**[00109]** In some embodiments, the nucleotide sequence encoding the first flexible polypeptide linker separates the nucleotide sequence encoding the HLA polypeptide from the nucleotide sequence encoding the randomized antigen polypeptide, and the nucleotide sequence encoding the second flexible polypeptide linker separates the nucleotide sequence encoding the HLA polypeptide from the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide. In some embodiments, once expressed, the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

**[00110]** In some embodiments, the nucleotide sequence encoding the first flexible polypeptide linker separates the nucleotide sequence encoding the HLA polypeptide from the nucleotide sequence encoding the randomized antigen polypeptide, and the nucleotide sequence encoding the second flexible polypeptide linker separates the nucleotide sequence encoding the HLA polypeptide from the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide. In some embodiments, once expressed, the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single

polypeptide.

**[00111]** In some embodiments, the nucleotide sequence encoding the first flexible polypeptide linker separates the nucleotide sequence encoding the randomized antigen polypeptide from the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide, and the nucleotide sequence encoding the second flexible polypeptide linker separates the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide from the nucleotide sequence encoding the HLA polypeptide. In some embodiments, once expressed the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

**[00112]** In some embodiments, the nucleotide sequence encoding the first flexible polypeptide linker separates the nucleotide sequence encoding the HLA polypeptide from the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide, and the nucleotide sequence encoding the second flexible polypeptide linker separates the nucleotide sequence encoding the randomized antigen polypeptide from the nucleotide sequence encoding the HLA polypeptide. In some embodiments, once expressed, the  $\beta$ 2-microglobulin polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the randomized antigen polypeptide on the single polypeptide.

**[00113]** In some embodiments, the nucleotide sequence encoding the first flexible polypeptide linker separates the nucleotide sequence encoding the HLA polypeptide from the nucleotide sequence encoding the randomized antigen polypeptide, and the nucleotide sequence encoding the second flexible polypeptide linker separates the nucleotide sequence encoding the randomized antigen polypeptide from the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide. In some embodiments, once expressed, the randomized antigen polypeptide is C-terminal to the  $\beta$ 2-microglobulin on the single polypeptide, and the HLA polypeptide is C-terminal to the randomized antigen polypeptide on the single polypeptide. In some embodiments, the nucleotide sequence

encoding the first flexible polypeptide linker separates the nucleotide sequence encoding the  $\beta$ 2-microglobulin polypeptide from the nucleotide sequence encoding the randomized antigen polypeptide, and the nucleotide sequence encoding the second flexible polypeptide linker separates the nucleotide sequence encoding the randomized antigen polypeptide from the nucleotide sequence encoding the HLA polypeptide.

**[00114]** In other embodiments, components of the antigen screening libraries of a plurality HLA-antigen polypeptide complexes are expressed as more than one polypeptide despite being encoded by a single nucleic acid. In these embodiments, a nucleotide encoded cleavage sequence separates components of the antigen screening libraries of a plurality HLA-antigen polypeptide complexes from one another. For example, once expressed, the randomized peptide antigen is separated from the HLA polypeptide and/or from the Beta-2 ( $\beta$ 2) microglobulin polypeptide by the cleavage sequence. As another example, once expressed, the HLA peptide is separated from the Beta-2 ( $\beta$ 2) microglobulin polypeptide by the nucleotide encoded cleavage sequence. In these embodiments, a portion of the HLA polypeptide is expressed separately from other components of the antigen screening libraries of a plurality HLA-antigen polypeptide complexes and, when expressed separately, pairs naturally with the other components of the HLA-antigen polypeptide complexes inside the cell.

**[00115]** In some embodiments, the randomized antigen polypeptide of the HLA-antigen complex is encoded by a nucleic acid set forth in any one of SEQ ID NOs: 210 to 411. In some embodiments, the HLA polypeptide of the HLA-antigen complex is encoded by a nucleic acid at least 70%, 75%, 80%, 85%, 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% homologous to any one of SEQ ID NOs: 210 to 411. In some embodiments, the randomized antigen polypeptide of the HLA-antigen complex is encoded by a nucleic acid set forth in any one of SEQ ID NOs: 412 to 426. In some embodiments, the HLA polypeptide of the HLA-antigen complex is encoded by a nucleic acid at least 70%, 75%, 80%, 85%, 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% homologous to any one of SEQ ID NOs: 280 to

308. In some embodiments, one or more of the nucleic acids such as one or more of the nucleic acids of SEQ ID NOs: 210 to 411 and 412 to 426 are expressed by a plurality of cells. In some embodiments, each cell of the plurality of cells comprises a nucleic acid encoding a HLA-antigen complex. In some embodiments, the plurality of cells are a plurality of yeast cells. In some embodiments, the plurality of yeast cells are a plurality of cells of the of the EBY100 strain of *Saccharomyces cerevisiae*.

**[00116]** Nucleic acids encoding one or more components of the HLA-antigen polypeptide complexes can be delivered to the plurality of cells with a nucleic acid or a vector, such as an exogenous nucleic acid or exogenous vector. Suitable exogenous nucleic acids and exogenous vectors include plasmids, bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs), transposons, and viral vectors. These exogenous nucleic acids and exogenous vectors can further comprise components that allow for replication of the nucleic acids encoding one or more components of the HLA-antigen polypeptide complexes, permit antibiotic selection to allow for selection of cells or other organisms expressing the nucleic acids encoding one or more components of the HLA-antigen polypeptide complexes, genes that complement yeast autotrophies to select for yeast transformants expressing the nucleic acids encoding one or more components of the HLA-antigen polypeptide complexes, promoters or enhancers for prokaryotic or eukaryotic expression of the HLA-antigen polypeptide complexes, polyadenylation sites, or marker genes that allow for visualization of transformed cells. In some embodiments, the nucleic acids that comprise a nucleic acid encoding the HLA-antigen polypeptide complexes of the current disclosure comprise an inducible promoter.

**[00117]** Methods of using the HLA-antigen polypeptide complexes and nucleic acids encoding such complexes minimally comprise contacting one or more cells, such as a plurality of cells, expressing the HLA antigen polypeptide complexes with a TCR and selecting for one or more cells that interact with the TCR. Selection can be performed, for example, by using the TCR in a “panning step” to capture the one or more cells expressing HLA-antigen polypeptide complexes that interact with the

TCR, and washing away any non-interacting cells, such as one or more cells that do not express the HLA-antigen polypeptide complexes that do not interact with the TCR. Nucleic acids from interacting cells can be harvested and sequenced to elucidate the amino acid sequences of the randomized antigen polypeptide that interacted with the TCR. These nucleic acids can be re-transfected, transformed, or transduced into one or more different cells for another round of selection. This method can be iterated for any number of rounds of selection, such as 1, 2, 3, 4, 5, or more times (e.g., in cycles) to enrich for HLA-antigen polypeptide complexes that strongly interact with the TCR.

**[00118]** Sequencing platforms that can be used in the present disclosure include, but are not limited to: pyrosequencing, sequencing-by-synthesis, single-molecule sequencing, second-generation sequencing, nanopore sequencing, sequencing by ligation, or sequencing by hybridization. Preferred sequencing platforms are those commercially available from Illumina (RNA-Seq) and Helicos (Digital Gene Expression or “DGE”). “Next generation” sequencing methods include, but are not limited to those commercialized by: 1 ) 454/Roche Lifesciences including but not limited to the methods and apparatus described in Margulies 2005 and in US Patent Nos. 7,244,559; 7,335,762; 7,211,390; 7,244,567; 7,264,929; and 7,323,305; 2) Helicos Biosciences Corporation (Cambridge, MA) as described in US Patent Nos. 7,501,245; 7,491,498; and 7,276,720; and in U.S. Patent Publ. Nos. 2006/0024711; 2009/0061439; 2008/0087826; 2006/0286566; 2006/0024711; 2006/0024678; 2008/0213770; and 2008/0103058; 3) Applied Biosystems (e.g. SOLiD sequencing); 4) Dover Systems (e.g., Polonator G.007 sequencing); 5) Illumina as described US Patent Nos. 5,750,341; 6,306,597; and 5,969,119; and 6) Pacific Biosciences as described in US Patent Nos. 7,462,452; 7,476,504; 7,405,281; 7,170,050; 7,462,468; 7,476,503; 7,315,019; 7,302,146; and 7,313,308; and in US Patent Publ. Nos. 2009/0029385; 2009/0068655; 2009/0024331; and 2008/0206764.

## **Methods**

**[00119]** Described herein are methods of using the HLA-antigen polypeptide complexes of the

present disclosure to select or enrich for antigens that bind to a TCR, such as a specific TCR. In some embodiments, the method includes selecting an antigen comprising contacting one or more cells, such as a plurality of cells, expressing at the HLA antigen polypeptide complexes with a TCR using one or more transgenic HLA-antigen polypeptide cell libraries, such as a transgenic HLA-antigen polypeptide yeast cell libraries. The methods described herein include methods for constructing one or more transgenic HLA-antigen polypeptide yeast cell libraries.

**[00120]** After construction of the one or more transgenic HLA-antigen polypeptide yeast cell libraries, the methods further include validating one or more transgenic HLA-antigen polypeptide yeast cell libraries using limiting dilution methods which include limited dilution of one or more cultures of proliferating yeast cells that each express at least one of the HLA-antigen polypeptides with nutrient-deficient yeast media. In some embodiments, the methods further include counting yeast from diluted yeast cultures and estimating HLA-antigen polypeptide yeast cell libraries with diversities of at least about  $10^6$ ,  $10^7$ ,  $10^8$ , or  $10^9$  unique HLA-antigen polypeptide sequences (e.g., clones). In some embodiments, expression of an epitope tag by a yeast cell is measured to determine if any of the  $10^6$ ,  $10^7$ ,  $10^8$ , or  $10^9$  clones are displayed on a yeast cell surface. For example, expression of the epitope tag can be determined as a surrogate value for total HLA-antigen polypeptide expression in the plurality of yeast cells and percent expression can be calculated. In some embodiments, the percent expression is an estimate of a number of yeast cells expressing a certain HLA-antigen polypeptide relative to the HLA-antigen polypeptide sequence library.

**[00121]** Referring to **FIG. 2**, the plurality of cells **201**, such as yeast, can be transformed, transfected, or electroporated with the plurality of nucleic acids encoding the HLA-antigen polypeptide complexes of the present disclosure **202**. The plurality of cells expressing the plurality of nucleic acids encoding the HLA-antigen peptide complexes is referred to as a transgenic HLA-antigen polypeptide cell library **203**. The transgenic HLA-antigen polypeptide cell library **203** is expanded through cell proliferation

and expression of HLA-antigen polypeptide complexes **204** by the plurality of cells is induced by methods known in the art, for example, by galactose, lactose, or isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG). Cells expressing an HLA-antigen polypeptide complex that interact with a TCR are positively selected using that TCR **205**. In some embodiments, the TCR is immobilized on a substrate. In some embodiments, the TCR is expressed by a cell or a plurality of cells. This selection process illustrated in **FIG. 2** can be repeated for any number of rounds of selection, such as 1, 2, 3, 4, 5, or more times to arrive at a single or small number of HLA-antigen polypeptide complexes that interact with the TCR. In some embodiments, the HLA-antigen polypeptide complexes include a polypeptide antigen. In some embodiments, the polypeptide antigen is a non-naturally occurring polypeptide antigen, such as a polypeptide antigen that does not naturally occur in a human. Deep sequencing or next-generation sequencing reactions can be performed on nucleic acids extracted from the selected cells **205** after each round of selection, or after the last round of selection.

**[00122]** In some embodiments, greater than at least  $1 \times 10^4$ , at least  $1 \times 10^5$ , at least  $1 \times 10^6$ , at least  $1 \times 10^7$ , at least  $1 \times 10^8$ , at least  $1 \times 10^9$ , at least  $1 \times 10^{10}$ , at least  $1 \times 10^{11}$ , at least  $1 \times 10^{12}$ , at least  $1 \times 10^{13}$ , at least  $1 \times 10^{14}$ , or at least  $1 \times 10^{15}$  different HLA-antigen polypeptide complexes are screened with methods of the present disclosure, such as those illustrated in **FIG. 2**. In some embodiments, the methods of the present disclosure result in identification of less than  $10^4$ ,  $10^3$ ,  $10^2$ , 10, 9, 8, 7, 6, 5, 4, 3, or 2 different HLA-antigen polypeptide complexes. In some embodiments, greater than 90%, 95%, 97%, 98%, or 99% of the HLA-antigen polypeptide complexes remaining after 1, 2, 3, 4, or 5 rounds of selection comprise less than 10, 9, 8, 7, 6, 5, 4, 3, or 2 different HLA-antigen polypeptide complexes. In some embodiments, greater than 90%, 95%, 97%, 98%, or 99% of the HLA-antigen polypeptide complexes remaining after 1, 2, 3, 4, or 5 rounds of selection comprise less than 10, 9, 8, 7, 6, 5, 4, 3, or 2 different antigenic polypeptide sequences within the HLA-antigen polypeptide complexes. In some embodiments, greater than 90%, 95%, 97%, 98%, or 99% of the HLA-antigen polypeptide

complexes remaining after 1, 2, 3, 4, or 5 rounds of selection comprise a single HLA-antigen polypeptide complex. In some embodiments, greater than 90%, 95%, 97%, 98%, or 99% of the HLA-antigen polypeptide complexes remaining after 1, 2, 3, 4, or 5 rounds of selection comprise a single antigenic polypeptide sequence within a single HLA-antigen polypeptide complex.

**[00123]** Expression of naive yeast libraries, such as the HLA-antigen polypeptide sequence libraries described herein, minimally express at about 15% of total antigen polypeptide sequences in an antigen polypeptide sequence library for a single length 9mer peptide presented by HLA-A1 (Gee 2018b) and less than about 5% of a single length peptide (e.g., 8mer) expression in an antigen polypeptide sequence library having mixed length peptides (e.g., 8mer, 9mer, 10mer, 11mer, 12mer) presented by HLA-A2 (Gee 2018a). Despite less than about 5% single length expression of the antigen polypeptide sequence library having 8mer length peptides, TCRs isolated target 8mer antigens from the antigen polypeptide sequence library that stimulated the TCR in an *in vitro* co-culture assay (Gee 2018a). These antigen polypeptide sequence libraries have been screened and isolate peptides against TCRs of known specificity (Gee 2018a). While a minimum level of expression necessary for a functional library has not yet been determined, data shows that less than 15% expression can result in an antigen polypeptide sequence library useful with the methods described herein.

**[00124]** In some embodiments, methods of the present disclosure further include identifying a polypeptide antigen that interact with a TCR. For example, a method for determining TCR interacting polypeptide antigens can comprise any of the following steps:

1. Generation and HLA-antigen polypeptide complex construct design: In some embodiments, step (1) includes, but is not limited to, generating one or more DNA constructs and/or designs to display one or more HLA polypeptides with a naturally occurring protein sequence, a synthetic protein sequences, or a combination thereof.
2. Test expression of the HLA-antigen polypeptide complex construct via yeast expression:

In some embodiments, step (2) includes, but is not limited to, transforming one or more electro- or chemically competent yeast with a plasmid encoding a single peptide or library of peptides including the HLA of interest, such as the HLA polypeptide. The plasmid is designed for the single peptide construct or library of peptides constructs to display from the N-terminus of Aga2, a yeast protein. In some embodiments, expression confirmation can include antibody staining of an epitope tag (e.g., V5, VSVg, c-Myc, HA) or fluorescent TCR tetramer, dimer, or dextramer staining of yeast displaying a single peptide-HLA construct or library of peptide-HLA constructs.

3. Optional validation step for HLA display: In some embodiments, step (3) includes, but is not limited to, antibody-based staining of the epitope tag or fluorescent TCR tetramer, dimer, or dextramer staining of yeast displaying a single peptide-HLA construct or library of peptide-HLA constructs of step (2). In some embodiments, validation can also include staining a peptide-HLA construct with a TCR of known specificity or for selecting a diverse peptide library presented by the HLA.
4. Optional step to re-engineer the HLA for display: In some embodiments, step (4) includes, but is not limited to, random mutagenesis via an error-prone polymerase followed by electroporation into chemically and/or electro-competent yeast. Yeast cells expressing the library and/or libraries of the present technology are selected with cell separation by magnetic cell sorting (MACS) or fluorescence-activated cell sorting (FACS) based on a TCR of interest. In some embodiments, isolated yeast clones are sequenced or deep-sequenced to identify any functional HLA mutants that properly display antigenic peptides of interest. Step (4) is included in some embodiments if the construct or library is improperly displayed.
5. Generation of the peptide-HLA library: In some embodiments, step (5) includes, but is not

limited to, randomized encoded peptide ligands or explicitly encoded peptide ligands. For example, the randomized encoded peptide ligands or explicitly encoded peptide ligands are uniquely designed for each HLA allele based on a preference for which peptides each HLA allele can present. In some embodiments, step (5) also includes generating genetic material from one or more polymerase-chain reactions.

6. Selection of the peptide-HLA library with a TCR of interest: In some embodiments, step (6) includes, but is not limited to, iterative MACS-based or FACS-based selection. For example, the TCR of interest, or other macromolecule having one or more antigen binding domains, acts as bait and can be multimerized on magnetic beads, streptavidin, dextran, or other substrates suitable for multimerization. In some embodiments, output from one or more selection rounds includes physically isolating one or more yeast cells with the TCR. Following isolation, the yeast are propagated and re-induced for protein expression. These iterative rounds enrich for binding yeast populations.
7. Deep-sequencing and data analysis: This process can involve extracting the genetic information of the yeast library and selection, and sequencing the products to identify the nature of peptides from the selected library. These data can then be analyzed to identify potential targets of TCRs and/or fed into algorithms to make predictions about TCR specificity.

#### **T cell receptors (TCRs)**

**[00125]** The transgenic HLA-antigen polypeptide cell libraries and antigens of the HLA-antigen polypeptide complexes described herein can be used in conjunction with a given TCR. For example, the TCR, or other macromolecule having one or more antigen binding domains, is a positive selector or bait and once bound to an antigen (e.g., HLA-antigen polypeptide complex), identifies its cognate antigen. The TCRs described herein can be native or exogenous (e.g., recombinant) and expressed by

a cell, such as a primary T cell, an immortalized T cell, or a non-T cell. In some embodiments, the TCR is immobilized on a solid support such as a column, a polystyrene plate or well of a multi-well plate, or a bead. In some embodiments, the TCR is multimerized as a plurality of TCRs immobilized on a bead. For example, the TCR can be multimerized on but not limited to magnetic beads, streptavidin, or dextran.

**[00126]** In some embodiments, the TCR is a soluble protein comprising at least one or more binding domains of a TCR of interest, e.g. TCR $\alpha/\beta$ , TCR $\gamma/\delta$ . The soluble protein may be a single chain, or a heterodimer. In some embodiments, the soluble TCR is modified by the addition of a biotin acceptor peptide sequence at the C terminus of one polypeptide. After biotinylation at the acceptor peptide, the TCR can be multimerized or added to substrate by binding to biotin binding partner, e.g. avidin, streptavidin, traptavidin, neutravidin, etc. In some embodiments, the biotin binding partner can comprise a detectable label, e.g. a fluorophore, mass label, etc., or can be bound to a particle, e.g. a paramagnetic particle. Selection of ligands bound to the TCR can be performed by flow cytometry, magnetic selection, and the like as known in the art.

**[00127]** To the extent the foregoing materials and/or any other materials incorporated herein by reference conflict with the present disclosure, the present disclosure controls.

**[00128]** The following examples provide further representative embodiments of the presently disclosed technology.

## EXAMPLES

**[00129]** The following examples are provided to further illustrate embodiments of the present technology and are not to be interpreted as limiting the scope of the present technology. To the extent that certain embodiments or features thereof are mentioned, it is merely for purposes of illustration and, unless otherwise specified, is not intended to limit the present technology. One skilled in the art may develop equivalent means without the exercise of inventive capacity and without departing from

the scope of the present technology. It will be understood that many variations can be made in the procedures herein described while remaining within the bounds of the present technology. Such variations are intended to be included within the scope of the presently disclosed technology. As such, embodiments of the presently disclosed technology are described in the following representative examples.

#### **Example 1: Design of An Antigen Polypeptide Library**

[00130] This example describes design of the antigen libraries of the present disclosure for use with a polypeptide antigen HLA complex. An exemplary algorithm to design and select anchor residues for each HLA allele is as follows, using data of known HLA binding epitopes ligands from a website such as [www.IEDB.org/](http://www.IEDB.org/):

[00131] Step 1: download list of polypeptides that bind to a given allele which may comprise several hundred peptides or several thousand peptides.

[00132] Step 2: construct a frequency matrix of residues per position of the peptide based upon the downloaded known peptides.

[00133] Step 3: select composition of “anchors” for library design by using a cutoff of the top 4 residues at each position.

#### **Example 2: Electroporation of pHLA Library**

[00134] This example describes electroporating yeast cells with nucleic acids encoding an exemplary antigen library of the present disclosure having all HLA allotypes and using peptides of 8-11 amino acids in length (8mer-11mer). In this example, yeast cells were electroporated with nucleic acids encoding the antigen library of HLA-antigen polypeptide complexes (pHLA library).

[00135] The electroporation methods for expression of pHLA on yeast are as follows:

[00136] Day 0:

1. Autoclave three 2.5L baffled flasks and one 250 ml baffled flask for expanding proliferating yeast cultures.
2. Prepare Yeast Peptone Dextrose media (YPD), which includes bacto peptone, glucose, and yeast extract.
3. Prepare two 5ml EBY100 yeast cultures and shake at 30°C overnight.
4. Prepare pYAL\_3T plasmid (10µg) restriction enzyme digested with HindIII, NheI or NheI, and BamHI and insert containing libraries of SEQ ID NO: 210 to 411 (50µg). pYAL\_3T vector (SEQ ID NO: 485) is a derivative of pCT vector (SEQ ID NO: 486; Invitrogen), **Table 8**, and the maps are provided in **FIGS. 3A and 3B**. Features of pYAL\_3T and pCT are included in **Tables 9 and 10**, respectively. pYAL\_3T differs from pCT at least by the orientation of the display protein scaffold (Aga2) being C-terminal of the pHLA library, the addition of human B2M, and connecting linkers. pYAL\_3T has been described (Gee 2018a).

**Table 8: Nucleotide sequences of pYAL\_3T and pCT vectors**

| SEQ ID NO: | Nucleotide Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 485        | acggattagaagccgcccagcggggtgacagccctccgaaggaaagactctctccgtgcgtcctcgtcttcaccgg<br>tcgcgttctgaaacgcagatgtgcctcgcgccgactgctccgaacaataaagattctacaatactagcttttatgg<br>ttatgaagaggaaaaattggcagtaacctggccccacaaaccttcaaatgaacgaatcaaatcaaacataggat<br>gataatgcgattagtttttagccttatttctggggaatataatcagcgaagcgatgattttgatctattaacagatatata<br>aatgcaaaaactgcataaccactttaactaatactttcaacatttccggttgattacttctattcaaatgtaataaaagt<br>atcaacaaaaaattgtaatactctatactttaacgtcaaggagaaaaaacccggatcgactactagcagctg<br>taatacgactcactatagggaatattaagctaattctacttcatacatttcaattaagatgcagttacttcgctgttttca<br>atatcttctgtattgctagcgttttgctggtggaggaggttctggaggtggtgtagtggtggtggtggttccataca<br>aagaactccaaagatccaagttacagtagacatcctgctgaaaacggtaaatctaatttctgaactgttacgtctcc<br>gggttccaccaagtgatagaaagtgactgttgtaaaaatgggtgaaagaatcgaaaaggtgaacattcagattgt<br>cttttctaaggactggtccttctatttgtgtactacacagaattcactccaactgaaaaggatgaatacgttcgaga<br>gtaaatcatgaactgtctcaacataaaatcgtaagtgggatagacatgggtggaggtggaagtggaggtgg<br>cgggtcaggtggtggcgggtcgggtggaggtggatccgaacaaaagcttatctccgaagaagactgggtggtggt<br>ggatctggtggtggtggtggtggtggtggttctcaggaaactgacaactatatgcgagcaaaccctcaccaac<br>ttagaatcgacgccgtactcttgcacgactactatttggccaacgggaaggcaatgaaggagttttgaatatt<br>acaaatcagtaacgtttgtcagtaattgcggttctcaccctcaacaactagcaaaggcagccccataaacacacag<br>tatgtttttgagtttaaacccgctgatctgataacaacagtgtagatgtaacaaaatcgactttgttccactgtacttta<br>gctcgtaaaaatacaatacttttcttctccgtaaaacaacatgtttccatgtaataccttttctattttcgttcggt<br>accaactttacacatactttatagctattcacttctatacactaaaaaactaagacaattttaatttgcctgctgcata<br>ttcaatttgtataaattcctataatttatcctattagtagctaaaaaagatgaatgtgaatcgaatcctaagagaattg<br>ggcaagtcacaaacaatacttaataataactactcagtaataacatttcttagcattttgacgaaattgctatttt<br>gtagagcttttacaccatttgcctccacacctccgcttatcaacaccaataacgccatttaactaagcgcatcac<br>caacatttttggcgtcagtcaccagtaacataaaatgtaagctctcggggctctcttgccttcaacccagtcag<br>aaatcgagttccaatcaaaaagtcacctgtcccctgcttctgaatcaacaagggaataaacgaatgaggttct<br>gtgaagctgcactgagtagtatgttcagctcttttgaataacgagcttttaataactggcaaaccgaggaactctg<br>gtattctgcccagactcatctccgtgcagttggacgatataatgccgtaatcattgaccagagccaaaacatcctc |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p> cttaggttgattacgaacacgccaaccaagtatttcggagtgccctgaactattttatatgctttacaagacttgaat<br/> tttccttgcaataaccgggtcaattgtctctttctattgggcacacataataatccagcaagtcagcatcggaatcta<br/> gagcaccattctgcggtcctctgtgctctgcaagccgcaaactttaccaatggaccagaactacctgtgaattaataa<br/> cagacatactccaagctgcctttgtgtgcttaatcacgtatactcacgtgctcaatagtcaccaatgccctccctctgg<br/> ccctctccttttcttttcgaccgaatttctgaagacgaaaggccctctgatacgcctattttataggtaagtcatg<br/> ataataatggtttcttaggacggatcgcttgcctgtaacttacacgcgcctctgtatctttaatgatggaataattggga<br/> atttactctgtgtttattttttatgtttgtatttgatttagaaagtaaataaagaaggtagaagagttacggaatgaa<br/> gaaaaaaaaataaacaagggttaaaaaattcaacaaaaagcgtactttacatatattttatagacaagaaaagca<br/> gattaaatagatatatacttcgattaacgataagtaaaatgtaaaatcacaggatttctgtgtggtctttacacagac<br/> aagatgaacaattcggcattaatactgagagcaggaagagcaagataaaaggtagtattgttggcgatccccct<br/> agagtcttttacatcttcggaacaaaaactttttctttaatttctttttacttctatttttaatttatattatattataa<br/> aaatttaattataattttttatagcacgtgatgaaaaggaccaggtggcacttttcggggaatgtgcgcggaac<br/> ccctattgtttttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgctcaataattg<br/> aaaagggaagagtatgattcaacatttccgtgtcgccttattccctttttgcggcatttgccttctgttttctc<br/> accagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggggttacatcgaactggatct<br/> caacagcggtgaagatccttgagagttttcgcccgaagaacgtttccaatgatgagcacttttaagtctgctatgt<br/> ggcgcggtattatcccggttgacgccgggcaagagcaactcggcgccgcatactattctcagaatgacttgggt<br/> tgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgagtgctgccataacca<br/> tgagtataacactgcggccaacttactctgacaacgatcggaggaccgaaggagctaaccgctttttgcacaac<br/> atgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccatacacaacgacgagcgtgac<br/> accacgatgcctgtagcaattgcaacaacgttgcgcaaaactttaactggcgaaactacttactctagcttccggca<br/> acaattaatagactggatggaggcggaataaagtgcaggaccacttctgcgctcggccctccggctggctggttta<br/> ttgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggaagccctc<br/> ccgtatcgtagtattctacacgacgggcagtcaggcaactatggatgaacgaaatagacagatcgtgagataggt<br/> gcctcactgattaagcattgtaactgtcagaccaagtttactcatatatactttagattgatttaaaactcatttttaatt<br/> aaaaggatctaggtgaagatccttttgataatctcatgacaaaatccctaacgtgagtttctgctcactgagcgtc<br/> agaccccgtagaaaagatcaaggatcttcttgatcctttttctgcgcgtaatctgctgcttgcaacaaaaaaa<br/> ccaccgctaccagcggtggtttgttgcggatcaagagctaccaactcttttccgaaggaactggcttcagcaga<br/> gcgcatgatacacaatactgtccttctagttagcgttagtagccaccacttcaagaactctgtagcaccgctaca<br/> tacctcgctctgtaactctgttaccagtggctgctgccagtggcgataagtcgtgttaccgggttggaactaaga<br/> cgatagttaccggataaggcgagcggtcgggtgaacggggggtcgtgcacacagcccagcttgagcgaa<br/> cgacctacaccgaactgagatactacagcgtgagcattgagaaagcgccacgcttccgaaggagaaaggc<br/> ggacaggtatccggtgaagcggcagggctcggaacaggagagcgcacgagggagcttccaggggggaacgcct<br/> ggtatctttatagtcctgtcgggttgcgacacctctgacttgagcgtcgattttgtgatgctcgtcaggggggagg<br/> cctatggaaaaacgcagcaacgcggccttttacgggtcttgccttttctggtcctttgtcacatgttcttctgc<br/> gttatccctgattctgtgataaccgtattaccgctttgagttagctgataccgctcgcgcagccgaacgaccg<br/> agcgagcagtgatgagcaggaagcggaagagcgcccaatacgaacgcctctccccgcggttggc<br/> cgattcattaatgagctggcagcagaggttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtga<br/> gttacctcactcattaggaacccaggtttacatttatgcttccggctcctatgttgttgggaattgtgagcggataa </p> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     | caatttcacacaggaacagctatgacctgattacgccaagctcggaaattaacccctactaaagggaacaaaagc<br>tggctagt                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 486 | acggattagaagccgcccagcggtgacagccctccgaaggaaagactctcctcgtgcgtcctcgtcttcaccgg<br>tcgcttctgaaacgcagatgtgcctcgcgccgactgtccgaacaataaagattctacaatactagcttttatgg<br>ttatgaagaggaaaaattggcagtaacctggccccacaaacctcaaatgaacgaatcaaatcaaacatagggat<br>gataatgcgattagtttttagccttatttctggggtaattaatcagcgaagcgatgattttgatctattaacagatatata<br>aatgcaaaaactgcataaccactttaactaatactttcaacatttgcgtttgtattacttctattcaaatgataaaaagt<br>atcaacaaaaaattgtaatactctatactttaacgtcaaggagaaaaaaccccgatcgactactagcagctg<br>taatacgaactactatagggaaatattaagctaattctacttcatacattttcaattaagatgcagttacttcgctgttttca<br>atatcttctgttattgtctcagtttagcacaggaactgacaactatatgcgagcaaatccctcaccactttagaatcg<br>acgccgtactctttgtcaacgactactattttggccaacgggaaggcaatgcaaggagttttgaaattacaaatcag<br>taacgtttgtcagtaattgcggttctacccctcaacaactagcaaggcagccccataaacacacagatgtttttaa<br>gcttctgcaggctagtgggtgggtggttctgggtgggtggttctgggtgggtggttctgtagcatgactgggtgg<br>acagcaaatgggtcgggatctgtacgacgatgacgataagggtaccaggatccagtgtgggtggaattctgcagatat<br>ccagcacagtggcgccgctcgagtctagagggcccttcgaaggtaagcctatccctaaccctctcctcggtctcg<br>attctacggtaccggtcatcatcaccatcaccattgagtttaaacccgctgactgtataacaacgtgtagatgtaac<br>aaaatcgactttgtccactgtacttttagctctgacaaaatacaatacttttcttccgtaaaacaatgttttcc<br>catgtaatactcttttctatttttctggtccgttaccactttacacatactttatagctattcacttctatacactaaaaaact<br>aagacaatttaattttgctgcctgccatatttcaattgttataaattcctataatttctctattagtagctaaaaaagat<br>gaatgtgaatcgaatcctaagagaattgggcaagtgcacaaacaatacttaataaatactactcagtaataacattt<br>tcttagcatttttgacgaaatttgctattttgtagagtctttacaccattgtctccacacctcggttacatcaacacca<br>ataacgccatttaataagcgcacacccaacatttctggcgctcagtcaccagctaacataaaatgtaagctctcgg<br>ggctctctgcttccaaccagtcagaaatcaggttccaatccaaaagttcacctgtccacctgcttctgaatcaaa<br>caagggaataaacgaatgaggtttctgtgaagctgcaactgagtagtatgttgagctttttggaaatacagagtcttta<br>ataactggcaaacgaggaactcttggtatttctgccagactcatctccgtgcagttggacgatcaatgccgtaa<br>tcattgaccagagccaaaacatcctccttaggttattacgaaacacgccaaccaagtatttcggagtgcctgaacta<br>tttttatatgctttacaagacttgaaattttccttgcaataaccgggtcaattgttctcttttattgggcacacataata<br>cccagcaagtcagcatcggaatctagagcacattctcgcgccctctgtgctctgcaagccgcaaactttaccaatgg<br>accagaactacctgtgaaattaataacagacatactccaagctgcctttgtgtgcttaatacaggtatactcacgtgctc<br>aatagtcaccaatgccctccctcttgccctctccttttcttttcgaccgaatttctgaaagacgaaaggccctcgtga<br>tacgcctattttataggttaatgtcatgataataatgggttcttaggacggatcgcttgccgttaacttacacgcgcctc<br>gtatcttttaatgatggaataattgggaatttactctgtgtttattttttatgttttgtatttggaatttagaaagtaataa<br>agaaggtagaagagttacggaatgaagaaaaaaaataaacaagggttaaaaaatttcaacaaaaagcgacttta<br>catatatatttattagacaagaaaagcagattaataagatatattcagtaaacgataagtaaaatgtaaaatcacagg<br>attttcgtgtgtgcttctacacagacaagatgaaacaattcggcattaatacctgagagcaggaagagcaagata<br>aaaggtagtattttgtggcgatccccctagagtcttttacatcttcggaaaacaaaaacttttttcttaatttctttttta<br>cttctatttttaatttatattatattaaaaatttaattataattttttatagcacgtgatgaaaaggaccaggtgg<br>cacttttcggggaaatgtgcgcggaaccctatttgttttttctaaatacattcaaatatgtatccgctcatgagaca<br>ataacctgataaatgcttcaataatattgaaaaaggaaagtatgatttcaacatttccgtgtcgcccttattccctt<br>ttttgcggcattttgccttctgtttttgctcaccagaaacgctgggtgaaagtaaaagatgctgaagatcagttgggtg |

|  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | cacgagtggttacatcgaactggatctcaacagcggtaagatccttgagagtttcgccccgaagaacgtttcca<br>atgatgagcacttttaagttctgctatgtggcgcggtattatcccgtgttgacggggcaagagcaactcggtcgc<br>cgcatatactattctcagaatgacttggtgagtgactcaccagtcacagaaaagcatcttacggatggcatgacagta<br>agagaattatgcagtgtgcccataaccatgagtataactgcggccaacttacttctgacaacgatcggaggac<br>cgaaggagctaaccgctttttgcacaacatgggggatcatgtaactgccttgatcgttgggaaccggagctgaat<br>gaagccatacacaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaactattaaact<br>ggcgaaactacttactctagcttcccggcaacaattaatagactggatggaggcgataaagttgcaggaccacttct<br>gcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattg<br>cagcactggggccagatggttaagccctcccgtatcgtagtattctacacgacgggcagtcaggcaactatggatg<br>aacgaaatagacagatcgtgagataggtgcctcactgattaagcattggtaactgtcagaccaagtftactcatata<br>tacttttagattgatttaaaactcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgacaaaatcc<br>cttaacgtgagtttcttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgatccttttttctgc<br>gcgtaatctgctgcttgcacaacaaaaaccaccgctaccagcgggtggtttgttgcggatcaagagctaccaact<br>cttttccgaaggtaactggcttcagcagagcgcagatacacaatactgtccttctagttagccgtagttagggccac<br>cacttcaagaactctgtagcaccgcctacatacctcgtctgctaactcctgttaccagtggctgctgccagtggcgat<br>aagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctcgggctgaacgggggggt<br>tcgtgcacacagcccagcttgaggcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaag<br>cgccacgcttcccgaaggaggaaaggcggacaggtatccggtgaagcggcaggggtcggaaacaggagagcgcga<br>cgaggagcgttccaggggggaacgcctgtatctttatagtcctgtcgggttccgacctctgacttgagcgtcga<br>ttttgtgatgctcgtcagggggccgagcctatggaaaaacgccagcaacgcggccttttacgggtcctggccttt<br>tgctggccttttgcacatgttcttctcgttaccctgattctgtggataaccgtattaccgcctttgagtgaactg<br>ataccgctcggcgagccgaacgaccgagcgcagcagtcagtgagcagggaagcgggaagagcgcgccaatac<br>gcaaaccgcctctcccgcgcttgggccgattcattaatgcagctggcagcagaggttcccgactggaaagcgg<br>gcagtgagcgcgaacgaattaatgtgagttacctcactcattaggcaccacaggctttacactttatgcttccgctc<br>ctatgttgtgtgaattgtgagcggataacaatttcacacaggaaacagctatgacctgattacccaagctcgga<br>attaaccctcactaaagggaacaaaagctggctagt |
|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Table 9: Features of pYAL\_3T Vector

| Feature       | Nucleotide Position |
|---------------|---------------------|
| GAL1 promoter | 5007-450            |
| T7            | 473-494             |
| Aga2 Leader   | 534-587             |
| Aga2          | 588-794             |
| Linker        | 813-858             |
| CEN-ARS Prs   | 2289-2799           |
| Ampr          | 3129-3788           |
| Laco          | 4908-4930           |

**Table 10: Features of pCT Vector**

| <b>Feature</b>     | <b>Position</b> |
|--------------------|-----------------|
| GAL1 promotor      | 5217-451        |
| T7                 | 475-494         |
| Aga2 Leader        | 534-587         |
| linker             | 588-632         |
| hB2m               | 633-929         |
| linker             | 930-989         |
| epitope tag (cMyc) | 990-1019        |
| linker             | 1020-1064       |
| Aga2               | 1065-1271       |
| CEN-ARS pRS        | 2499-3009       |
| AmpR               | 3339-3998       |
| LacO               | 5118-5140       |

**[00137]** Day 1: Passage the two yeast cultures from Day 0 step 3 by adding 100 µl of each of the two yeast cultures to 5 ml of fresh YPD and shake at 30°C overnight.

**[00138]** Day 2:

1. Measure optical density (OD) of overnight cultures from Day 1.
2. Prepare a new culture with 300 ml of an OD 0.3 yeast culture from step 1 using YPD in a 2.5L baffled flask.
3. Prepare 3 ml of 1M Tris pH 8.0/1 M 1,4-dithiothreitol (DTT)
4. Prepare 15 ml of 2 M lithium acetate (LiAc)/ 10 mM Tris, 1 mM EDTA (TE)
5. Propagate culture to an OD of 1.6-2.0.
6. Add 3 ml of Tris/DTT.
7. Add 15 ml of 2 M LiAc/TE.
8. Propagate culture at 30 °C for 15 minutes while shaking at 225 rounds per minute (rpm).
9. Centrifuge culture at 3000 xg for 3 minutes.
10. Resuspend the pellet in 50 mL cold E-buffer.
11. Centrifuge the suspension from step 10 at 3000xg for 3 min, at 4°C.
12. Repeat steps 10 and 11 twice.

13. Remove residual buffer.
14. Resuspend pellet in 600 $\mu$ L E-buffer.
15. Add the 50  $\mu$ g insert and 10  $\mu$ g digested plasmid from step 4 of Day 0, total volume of buffer, insert, plasmid, and yeast should be about 1mL.
16. Aliquot 150  $\mu$ L of the suspension from step 15 into ice-chilled 2 mm gap electroporation cuvettes.
17. Electroporate each cuvette at 2.5 kV. The time constant should be between 3 and 4 ms<sup>1</sup>.
18. Add three 1 mL volumes of cold YPD, then bring the total volume up to 200 mL with YPD
19. Culture electroporated yeast at 30 °C at 225 rpm, for 1 hour in a 250 ml baffled flask.
20. Centrifuge culture at 3500 xg for 3 minutes to form a yeast cell pellet, decant the supernatant, and re-suspend the yeast cell pellet in 10 mL of SDCAA (dextrose casamino acids, which also includes, yeast nitrogen base without amino acids and with ammonium sulfate, sodium citrate, and citric acid monohydrate, at a pH of 4.5).

**[00139]** Day 2: Determine Titer

1. Add 990  $\mu$ L of SDCAA to each of four Eppendorf tubes.
2. Add 10  $\mu$ L from step 20 above to one tube containing 990  $\mu$ L SDCAA.
3. Pipet 100  $\mu$ L of the 10<sup>4</sup> solution into a tube containing only 990  $\mu$ L of SDCAA.
4. Pipet 100  $\mu$ L of the 10<sup>5</sup> solution into a tube containing only 990  $\mu$ L of SDCAA.
5. Pipet 100  $\mu$ L of the 10<sup>6</sup> solution into a tube containing only 990  $\mu$ L of SDCAA.
6. Spread 100  $\mu$ L of each dilution in steps 2-5 onto a separate SDCAA plates and incubate at 30°C for 3 days. Count the colonies on the plates to determine the titer. From step 2, the colonies counted represent the diversity of the library x 10<sup>4</sup>. From step 3, the colonies counted represent the diversity of the library x 10<sup>5</sup>. From step 4, the colonies counted represent the diversity of the library x 10<sup>6</sup>. From step 5, the colonies counted represent the diversity of the library x 10<sup>7</sup>.
7. Add 490 ml of pH 4.5 SDCAA to the remaining cell suspension from step 20 and culture at 30 °C overnight.

**[00140]** Day 3: Measure the OD of the passage from step 8 after 24 hours. The OD should be at least 5. Passage the culture to an OD of 1 in a total volume of 500 mL SDCAA.

**[00141]** Day 4: Passage cells to an OD of 1 in a total volume of 500 mL SDCAA.

**[00142]** Day 5: 72 hours after step 18 from Day 2 was performed, induce in SGCAA (galactose casamino acids, which also includes, yeast nitrogen base without amino acids and with ammonium sulfate, sodium citrate, and citric acid monohydrate, at a pH of 4.5).

**[00143]** Recipes:

- 1) E-buffer, 500ml
  - 0.6 g Tris base,
  - 91.09 g Sorbitol (1M)
  - 73.50 mg CaCl<sub>2</sub> (1mM; consider making 1M stock solution) in ddH<sub>2</sub>O to a final volume of 500 ml, pH to 7.5. Filter through 0.22 µm membrane.
- 2) 1 M Tris/1 M DTT, 3ml
  - 0.462 g 1,4-dithiothreitol in 3 ml 1 M Tris, pH 8.0 and sterilize by filtration.
- 3) 2 M LiAc/TE solution, 15ml
  - 1.98 g LiAc in 10ml of TE (10mM Tris, 1mM EDTA), sterilize by filtration.

### **Example 3: Characterization of pHLA Expression**

**[00144]** This example describes characterizing expression of HLA-antigen polypeptide complexes on the electroporated yeast cells of Example 2. These expression measurements include FACS analysis to determine the levels of peptide-MHC displayed on the surface of yeast cells and indicate functionality of the random yeast display library. The characterization methods for expression of pHLA on yeast are as follows:

#### **[00145] Materials**

1. Yeast library from Example 2
2. PBSM (1x PBS, 1 g/L bovine serum albumin, EDTA, pH 7.4; filtered)
3. Anti-myc (FITC fluorophore-conjugated) antibody
4. 96-well U-bottom plate
- Optional:
5. Anti-V5 (647 fluorophore-conjugated) antibody
6. Anti-HA (BV421 fluorophore-conjugated) antibody
7. Anti-VSV (PE fluorophore-conjugated) antibody

#### **[00146] Cell Preparation**

1. Measure optical density of yeast culture on NanoDrop. OD<sub>600</sub> readings between 0 to 1 are in the linear range for cultures induced 2 to 3 days at 20°C at a 1:20 culture:SDCAA dilution.
2. Transfer samples of yeast cultures into wells of a 96-well plate. For yeast cultures having an OD<sub>600</sub> of about 10, use 25 µL of the culture. Include single-color and unstained controls.
3. Add PBSM to 200 µl final volume to each well.
4. Centrifuge the 96-well plate at 2500xg for 2 minutes.
5. Remove supernatants.

**[00147] Staining Cells**

1. Re-suspend each cell pellet in 100 µl PBSM
2. Add 1 µl antibody as appropriate.
3. Incubate at 4°C, protected from light (e.g., dark) for 30 minutes.

**[00148] Washing Cells and Determining pHLA Expression**

1. Centrifuge the 96-well plate at 2500 xg for 2 minutes.
2. Remove supernatants.
3. Re-suspend each pellet in 200 µl PBSM.
4. Centrifuge the 96-well plate at 2500 xg for 2 minutes.
5. Remove supernatants.
6. Re-suspend each pellet in 200 µl PBSM.
7. Analyze samples in each well with CytoFlex.

**[00149] Results:** Expression of the HLA-antigen polypeptide complexes (peptides of SEQ ID Nos: 8, 11, 14, 18, 21 + 24, 28, 32, 36, 40-44, 47, 50, 53, 56, 65, 75, 69, 77 + 80, 89, 95, 99, 102 + 106, 108, 111 + 114, 117 + 120, 124, and 125) was determined by flow cytometry and shown in **FIG. 4**. Antibodies targeting epitope tags expressed by HLA-antigen polypeptide complexes were used to stain electroporated yeast cells. FITC-A staining corresponds to HLA-antigen polypeptide complexes expressing a c-Myc tag. Antibody-epitope tag binding was used as a proxy to determine pHLA expression, which, as shown in **FIG. 4**, ranged from 8.99% for SEQ ID NO: 56 to 26.3% expression for SEQ ID NOs: 177+180.

**Example 4: Functional Validation of pHLA Expression**

**[00150]** This example describes functionally validating expression of pHLAs on the electroporated yeast cells of Example 2 with a candidate TCR. Expected target antigens of the pHLAs can be identified from up to 6 libraries when the candidate TCR is allotype-matched. The functional validation methods for expression of pHLA on yeast are as follows:

**[00151]** HLA-antigen polypeptide sequence libraries, such as those disclosed herein, minimally express about 25% of total antigen polypeptide sequences for a single length 9mer peptide presented by HLA-A1 (Gee 2018b) and express less than about 5% of a single length peptide (e.g., 8mer) having

mixed length peptides (e.g., 8mer, 9mer, 10mer, 11mer) presented by HLA-A2 (Gee 2018a). Despite less than about 5% single length peptide expression of the HLA-antigen polypeptide sequence having 8mer length peptides, isolated TCRs of interest target 8mer antigens from the HLA-antigen polypeptide complexes. These isolated TCR of interest were stimulated by one or more HLA-antigen polypeptide complexes in an *in vitro* co-culture assay (Gee 2018a; see Figs. 5C and 7A therein). HLA-antigen polypeptide libraries have been screened and peptides which bind TCRs of known specificity have been isolated (Gee 2018a). While a minimum level of expression that is necessary for a functional HLA-antigen polypeptide library of the present disclosure has not yet been determined, data shows that less than 15% expression can result in an HLA-antigen polypeptide library that is useful with the methods described herein.

**[00152]** While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention.

**[00153]** All publications, patent applications, issued patents, and other documents referred to in this specification are herein incorporated by reference as if each individual publication, patent application, issued patent, or other document was specifically and individually indicated to be incorporated by reference in its entirety. Definitions that are contained in text incorporated by reference are excluded to the extent that they contradict definitions in this disclosure.

## REFERENCES

- Altschul et al. "Basic local alignment search tool." J Mol Biol 215(3):403-410 (1990)
- Altschul & Karlin. "Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes." Proc Natl Acad Sci USA 87(6):2264-2268 (1990)

- Altschul & Karlin. "Applications and statistics for multiple high-scoring segments in molecular sequences." *Proc Natl Acad Sci USA* 90(12):5873-5877 (1993)
- Arden et al. "Human T-cell receptor variable gene segment families." *Immunogenetics* 42(6):455-500 (1995)
- Bowie et al. "A method to identify protein sequences that fold into a known three-dimensional structure." *Science* 253(5016):164-170 (1991)
- Brenner et al. "Population statistics of protein structures: lessons from structural classifications." *Curr Opin Struct Biol* 7(3):369-376 (1997)
- Chou & Fasman. "Prediction of protein conformation." *Biochemistry* 13(2):222-245 (1974a)
- Chou & Fasman. "Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins." *Biochemistry* 113(2):211-222 (1974b)
- Chou & Fasman. "Prediction of the secondary structure of proteins from their amino acid sequence." *Adv Enzymol Relat Areas Mol Biol* 47:45-148 (1978a)
- Chou & Fasman. "Empirical predictions of protein conformation." *Annu Rev Biochem* 47:251-276 (1978b)
- Chou & Fasman. "Prediction of beta-turns." *Biophys J* 26:367-384 (1979)
- Gee et al. "Antigen identification for orphan T cell receptors expressed on tumor-infiltrating lymphocytes." *Cell* 172(3):549-563 (2018a)
- Gee et al. "Facile method for screening clinical T cell receptors for off-target peptide-HLA reactivity." *bioRxiv* 472480 (2018b)
- Gribskov et al. "Profile analysis: detection of distantly related proteins." *Proc Natl Acad Sci USA* 84(13):4355-4358 (1987)
- Gribskov et al. "Profile analysis." *Methods Enzymol* 183:146-159 (1990)

- Holm & Sander. "Protein folds and families: sequence and structure alignments." *Nucleic Acids Res* 27(1):244–247 (1999)
- Jones. "Progress in protein structure prediction." *Curr Opin Struct Biol* 7(3):377–87 (1997)
- Jones et. al. "Engineering and Characterization of a Stabilized  $\alpha 1/\alpha 2$  Module of the Class I Major Histocompatibility Complex Product L." *J. of Biol. Chem.* 281(35):25734-25744 (2006)
- Kotsiou et al. "Properties and Applications of Single-Chain Major Histocompatibility Complex Class I Molecule." *Antioxid. Redox Signal.* 15(3):645-655 (2011)
- Kyte & Doolittle. "A simple method for displaying the hydropathic character of a protein." *J Mol Biol* 157(1):105-132 (1982)
- Mackelprang et al. "Sequence diversity, natural selection and linkage disequilibrium in the human T cell receptor alpha/delta locus." *Hum Genet* 119(3):255-266 (2006)
- MacLennan et al. "Structure-function relationships in the Ca(2+)-binding and translocation domain of SERCA1: physiological correlates in Brody disease." *Acta Physiol Scand Suppl* 643:55-67 (1998)
- Margulies et al. "Genome sequencing in microfabricated high-density picolitre reactors." *Nature* 437(7057):376-380 (2005)
- Mottez et al. "Cells Expressing a Major Histocompatibility Complex Class I Molecule with a Single Covalently Bound Peptide Are Highly Immunogenic." *J. Exp. Med.* 181: 493-502 (1995)
- Moult. "The current state of the art in protein structure prediction." *Curr Opin Biotechnol* 7(4):422–427 (1996)
- Pandey et al. "Current strategies for protein production and purification enabling membrane protein structural biology." *Biochem Cell Biol.* 94(6) 507-527 (2016)
- Rowen et al. "The complete 685-kilobase DNA sequence of the human beta T cell receptor locus." *Science* 272(5269):1755-1762 (1996)

- Sasaki & Sutoh. "Structure-mutation analysis of the ATPase site of Dictyostelium discoideum myosin II." *Adv Biophys* 35:1–24 (1998)
- Sippl & Flockner. "Threading thrills and threats." *Structure* 4(1):15–19 (1996)
- Tafuro et al. "Reconstitution of Antigen Presentation in HLA Class I-Negative Cancer Cells with Peptide- $\beta$ 2m Fusion Molecules." *Eur. J. Immunol.* 31: 440-449 (2001)
- White et al. "Soluble Class I MHC with  $\beta$ 2-Microglobulin Covalently Linked Peptides: Specific Binding to a T Cell Hybridoma." *J. Immunol.* 162: 2671-2676 (1999)

## CLAIMS

1. An antigen screening library comprising a plurality of Human Leukocyte Antigen (HLA)-antigen polypeptide complexes, the HLA-antigen polypeptide complexes comprising:
  - a) an HLA polypeptide, the HLA polypeptide comprising a peptide binding cleft;
  - b) a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209, wherein the randomized antigen polypeptide specifically binds to the peptide binding cleft of the HLA polypeptide; and
  - c) a Beta-2 ( $\beta$ 2) microglobulin polypeptide.
2. The antigen screening library of claim 1, wherein the plurality of HLA-antigen complexes comprises an HLA polypeptide selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, B57, C1, C2, C3, C4, C5, C6, C7, C8, and E.
3. The antigen screening library of claim 1, wherein the plurality of HLA-antigen complexes comprises at least five, ten, fifteen, twenty, or twenty-five different HLA polypeptides selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E.
4. The antigen screening library of claim 1, wherein the plurality of HLA-antigen complexes comprises all of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E HLA polypeptides.
5. The antigen screening library of claim 1, wherein the plurality of HLA-antigen complexes comprises an HLA polypeptide comprising an amino acid sequence at least 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 427 to 455.

6. The antigen screening library of any one of claims 1 to 5, wherein the plurality of the HLA-antigen polypeptide complexes comprises at least about  $10^5$  different HLA-antigen polypeptide complexes comprising at least about  $10^5$  different randomized antigen polypeptides.

7. The antigen screening library of any one of claims 1 to 6, wherein the HLA polypeptide, the randomized antigen polypeptide, and the  $\beta$ 2-microglobulin polypeptide comprise a single polypeptide.

8. The antigen screening library of claim 7, wherein the single polypeptide further comprises a first flexible polypeptide linker and a second flexible polypeptide linker.

9. The antigen screening library of claim 8, wherein the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

10. The antigen screening library of claim 9, wherein the first flexible polypeptide linker separates the HLA polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the HLA polypeptide from the  $\beta$ 2-microglobulin polypeptide.

11. The antigen screening library of claim 8, wherein the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

12. The antigen screening library of claim 11, wherein the first flexible polypeptide linker separates the HLA polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the HLA polypeptide from the  $\beta$ 2-microglobulin polypeptide.

13. The antigen screening library of claim 8, wherein the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

14. The antigen screening library of claim 13, wherein the first flexible polypeptide linker separates the randomized antigen polypeptide from the  $\beta$ 2-microglobulin polypeptide, and a second flexible polypeptide linker separates the  $\beta$ 2-microglobulin polypeptide from the HLA polypeptide.

15. The antigen screening library of claim 8, wherein the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

16. The antigen screening library of claim 15, wherein the first flexible polypeptide linker separates the HLA polypeptide from the  $\beta$ 2-microglobulin polypeptide, and a second flexible polypeptide linker separates the randomized antigen polypeptide from the HLA polypeptide.

17. The antigen screening library of claim 8, wherein the  $\beta$ 2-microglobulin polypeptide is C-terminal to the HLA polypeptide on the single polypeptide, and the HLA polypeptide is N-terminal to the randomized antigen polypeptide on the single polypeptide.

18. The antigen screening library of claim 17, wherein the first flexible polypeptide linker separates the HLA polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the randomized antigen polypeptide from the  $\beta$ 2-microglobulin polypeptide.

19. The antigen screening library of claim 8, wherein the randomized antigen polypeptide is C-terminal to the  $\beta$ 2-microglobulin on the single polypeptide, and the HLA polypeptide is C-terminal to the randomized antigen polypeptide on the single polypeptide.

20. The antigen screening library of claim 19, wherein the first flexible polypeptide linker separates the  $\beta$ 2-microglobulin polypeptide from the randomized antigen polypeptide, and a second flexible polypeptide linker separates the randomized antigen polypeptide from the HLA polypeptide.

21. The antigen screening library of any one of claims 1 to 20, wherein each of the HLA-antigen complexes of the plurality of the HLA-antigen complexes do not comprise an epitope tag.

22. The antigen screening library of any one of claims 1 to 20, wherein at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes comprise an epitope tag.

23. The antigen screening library of any one of claims 1 to 20, wherein at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does not comprise an epitope tag and at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does comprise an epitope tag.

24. The antigen screening library of claims 22 or 23, wherein the epitope tag comprises a FLAG tag, a c-Myc tag, a HIS-tag, a hemagglutinin (HA) tag, a VSVg tag, or a V5 tag.

25. The antigen screening library of any one of claims 1 to 24, wherein the HLA-antigen complexes each comprise a membrane tethering domain.

26. The antigen screening library of claim 25, wherein the membrane tethering domain comprises Aga2.

27. The antigen screening library of any one of claims 1 to 26, wherein the antigen screening library is expressed on a plurality of cells.

28. The antigen screening library of claim 27, wherein the plurality of cells are a plurality of yeast cells.

29. The antigen screening library of claim 28, wherein the plurality of yeast cells are a plurality of yeast cells of the EBY100 strain of *Saccharomyces cerevisiae*.

30. The antigen screening library of any one of claims 27 to 29, wherein each cell of the plurality of cells expresses a specific HLA-antigen complex.

31. An antigen screening library comprising a plurality of Human Leukocyte Antigen (HLA)-antigen polypeptide complexes, the HLA-antigen polypeptide complexes comprising:

- a) an HLA polypeptide, the HLA polypeptide comprising a peptide binding cleft; and

b) a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209, wherein the randomized antigen polypeptide specifically binds to the peptide binding cleft of the HLA polypeptide.

32. The antigen screening library of claim 31, wherein the plurality of HLA-antigen complexes comprises an HLA polypeptide selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, B57, C1, C2, C3, C4, C5, C6, C7, C8, and E.

33. The antigen screening library of claim 31, wherein the plurality of HLA-antigen complexes comprises at least five, ten, fifteen, twenty, or twenty-five different HLA polypeptides selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E.

34. The antigen screening library of claim 31, wherein the plurality of HLA-antigen complexes comprises all of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E HLA polypeptides.

35. The antigen screening library of claim 31, wherein the plurality of HLA-antigen complexes comprises an HLA polypeptide comprising an amino acid sequence at least 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 427 to 455.

36. The antigen screening library of any one of claims 31 to 35, wherein the plurality of the HLA-antigen polypeptide complexes comprises at least about  $10^5$  different HLA-antigen polypeptide complexes comprising at least about  $10^5$  different randomized antigen polypeptides.

37. The antigen screening library of any one of claims 31 to 36, wherein the HLA polypeptide, and the randomized antigen polypeptide comprise a single polypeptide.

38. The antigen screening library of claim 37, wherein the single polypeptide further comprises a first flexible polypeptide linker separating the HLA polypeptide from the randomized antigen polypeptide.

39. The antigen screening library of claim 38, wherein the randomized antigen polypeptide is N-terminal to the HLA polypeptide on the single polypeptide.

40. The antigen screening library of claim 38, wherein the randomized antigen polypeptide is C-terminal to the HLA polypeptide on the single polypeptide.

41. The antigen screening library of any one of claims 31 to 40, wherein each of the HLA-antigen complexes of the plurality of the HLA-antigen complexes do not comprise an epitope tag.

42. The antigen screening library of any one of claims 31 to 40, wherein at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes comprise an epitope tag.

43. The antigen screening library of any one of claims 31 to 40, wherein at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does not comprise an epitope tag and at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does comprise an epitope tag.

44. The antigen screening library of claims 42 or 43, wherein the epitope tag comprises a FLAG tag, a c-Myc tag, a HIS-tag, a hemagglutinin (HA) tag, a VSVg tag, or a V5 tag.

45. The antigen screening library of any one of claims 31 to 44, wherein the HLA-antigen complexes each comprise a membrane tethering domain.

46. The antigen screening library of claim 45, wherein the membrane tethering domain comprises Aga2.

47. The antigen screening library of any one of claims 31 to 46, wherein the antigen screening library is expressed on a plurality of cells.

48. The antigen screening library of claim 47, wherein the plurality of cells are a plurality of yeast cells.

49. The antigen screening library of claim 48, wherein the plurality of yeast cells are a plurality of yeast cells of the EBY100 strain of *Saccharomyces cerevisiae*.

50. The antigen screening library of any one of claims 47 to 49, wherein each cell of the plurality of cells expresses a specific HLA-antigen complex.

51. The antigen screening library of any one of claims 47 to 50, wherein each cell of the plurality of cells expresses a  $\beta$ 2-microglobulin polypeptide.

52. An antigen screening library comprising

a) a plurality of antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes, the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes comprising:

i) a randomized antigen polypeptide comprising an amino acid sequence set forth in any one of SEQ ID NOs: 1 to 209, wherein the randomized antigen polypeptide specifically binds to the peptide binding cleft of the HLA polypeptide; and

ii) a Beta-2 ( $\beta$ 2) microglobulin polypeptide; and

b) a plurality of HLA polypeptides constitutively expressed by one or more yeast cells and comprising a peptide binding cleft.

53. The antigen screening library of claim 52, wherein the plurality of HLA polypeptides is selected from the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, B57, C1, C2, C3, C4, C5, C6, C7, C8, and E.

54. The antigen screening library of claim 52, wherein the plurality of HLA polypeptides comprises at least five, ten, fifteen, twenty, or twenty-five different HLA polypeptides selected from

the list consisting of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E.

55. The antigen screening library of claim 52, wherein the plurality of HLA polypeptides comprises all of A3, A11, A23, A24, A26, A30, A31, A33, A68, B7, B8, B15, B27, B40, B44, B51, B53, C1, C2, C3, C4, C5, C6, C7, C8, and E HLA polypeptides.

56. The antigen screening library of claim 52, wherein the plurality of HLA polypeptides comprises an HLA polypeptide comprising an amino acid sequence at least 87.5%, 90%, 95%, 97%, 98%, 99%, or 100% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 427 to 455.

57. The antigen screening library of any one of claims 52 to 56, wherein the plurality of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes comprises at least about  $10^5$  different antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes comprising at least about  $10^5$  different randomized antigen polypeptides.

58. The antigen screening library of any one of claims 52 to 57, wherein the randomized antigen polypeptide and the  $\beta$ 2-microglobulin polypeptide comprise a single polypeptide.

59. The antigen screening library of claim 58, wherein the single polypeptide further comprises a first flexible polypeptide linker.

60. The antigen screening library of claim 59, wherein the randomized antigen polypeptide is N-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

61. The antigen screening library of claim 60, wherein the randomized antigen polypeptide is C-terminal to the  $\beta$ 2-microglobulin polypeptide on the single polypeptide.

62. The antigen screening library of any one of claims 52 to 61, wherein each of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes of the plurality of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes do not comprise an epitope tag.

63. The antigen screening library of any one of claims 52 to 61, wherein at least one of the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes of the plurality of antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes comprise an epitope tag.

64. The antigen screening library of any one of claims 52 to 61, wherein at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does not comprise an epitope tag and at least one of the HLA-antigen complexes of the plurality of HLA-antigen complexes does comprise an epitope tag.

65. The antigen screening library of claims 63 or 64, wherein the epitope tag comprises a FLAG tag, a c-Myc tag, a HIS-tag, a hemagglutinin (HA) tag, a VSVg tag, or a V5 tag.

66. The antigen screening library of any one of claims 52 to 65, wherein the antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complexes each comprise a membrane tethering domain.

67. The antigen screening library of claim 66, wherein the membrane tethering domain comprises Aga2.

68. The antigen screening library of any one of claims 52 to 67, wherein the antigen screening library is expressed on a plurality of cells.

69. The antigen screening library of claim 68, wherein the plurality of cells are a plurality of yeast cells.

70. The antigen screening library of claim 69, wherein the plurality of yeast cells are a plurality of yeast cells of the EBY100 strain of *Saccharomyces cerevisiae*.

71. The antigen screening library of any one of claims 68 to 70, wherein each cell of the plurality of cells expresses a specific antigen polypeptide-Beta-2 ( $\beta$ 2) microglobulin polypeptide complex.

72. A plurality of nucleic acids encoding the antigen screening library of any one of claims 1 to 71.

73. The plurality of nucleic acids of claim 72, wherein the HLA polypeptide is encoded by a nucleic acid that is at least about 85%, 87.5%, 90%, 95%, 97%, 98%, or 99% homologous to any one of SEQ ID NOs: 456 to 484.

74. The plurality of nucleic acids of claim 73, wherein the randomized antigen polypeptide is encoded by a nucleic acid set forth in any one of SEQ ID NOs: 210 to 426.

75. The plurality of nucleic acids of any one of claims 72 to 74, wherein the plurality of nucleic acids is expressed by a plurality of cells.

76. A plurality of cells expressing the antigen screening library of any one of claims 1 to 75.

77. The plurality of cells of claim 76, wherein the plurality of cells is a plurality of yeast cells.

78. The plurality of cells of claim 77, wherein the plurality of yeast cells is a plurality of cells of the EBY100 strain of *Saccharomyces cerevisiae*.

79. The plurality of cells of any one of claims 76 to 78, wherein each cell of the plurality of cells comprises a nucleic acid of the plurality of nucleic acids encoding a specific of HLA-antigen complex.

80. A method of selecting an antigen comprising contacting the plurality of cells of any one of claims 76 to 79 with a T cell receptor (TCR) or other macromolecule having one or more antigen binding domains.

81. The method of claim 80, wherein the TCR or other macromolecule having one or more antigen binding domains is immobilized on a substrate.

82. The method of claim 80, wherein the TCR or other macromolecule having one or more antigen binding domains is expressed by a cell.

83. The method of any one of claims 76 to 79, wherein the selection is repeated for 2, 3, 4, or 5 cycles.

84. The method of claim 83, wherein the antigen is a polypeptide antigen.

85. The method of claim 84, wherein the antigen is a polypeptide antigen that does not naturally occur.

86. The method of claim 85, wherein the antigen is a polypeptide antigen that does not naturally occur in a human.