

every possible pMHC combination is extremely high, and since there are few if any 'holes' characterized in the TCR repertoire, it has been posited that a large degree of TCR cross-reactivity is a requirement of functional antigen recognition. How the T cell repertoire can simultaneously be MHC restricted, cross-reactive enough to ensure all potential antigenic challenges can be met, yet still specific enough to avoid aberrant autoimmunity, has remained an open and pressing question in immunology.

[0006] There have been a number of strategies used to determine the specificity of orphan TCRs (Birnbaum et al., (2012) *Immunol Rev* 250, 82-101). Mass spectrometry can provide an unbiased method of antigen isolation, but is restricted to experiments requiring large cell numbers, typically 10^7 to 10^9 , and the targets must still be presented by the correct HLA. Traditionally, most studies of T cell antigen specificities have involved testing candidate antigens empirically. For example, studies of anti-tumor T cell specificities have correctly postulated that there are productive T cell responses towards neo-antigens. Such studies involve sequencing of tumors to identify mutations, using epitope prediction algorithms to predict immunogenic mutant peptides, and testing for T cell responses directed at these mutant peptides (Kreiter et al., (2015) *Nature* 520, 692-696; Rajasagi et al., (2014) *Blood* 124, 453-462; Tran et al., (2014) *Science* 344, 641-645). Other strategies query established T cell specificities in patients by using pHLA multimers (Bentzen et al., (2016) *Nat Biotechnol* 34, 1037-1045; Newell et al., (2013) *Nat Biotechnol* 31, 623-629).

[0007] High-throughput and sensitive approaches to determining the specificity of 'orphan' TCRs (i.e. TCRs of unknown antigen specificity) that could help uncover potential targets for cancer immunotherapy, autoimmunity, and infection and provide mechanistic insight into disease pathogenesis are of great interest.

SUMMARY

[0001] In one aspect, the present disclosure refers to a T cell receptor or antibody comprising the amino acid sequence of SEQ ID NO: 258 and the amino acid sequences of SEQ ID NO: 259, SEQ ID NO: 260, or a combination thereof.

[0001a] In one aspect, the present disclosure refers to an immune cell engineered to comprise a T cell receptor as disclosed herein, or the antibody as disclosed herein.

[0001b] In one aspect, the present disclosure refers to a peptide ligand comprising an amino acid sequence of any of SEQ ID NO:1-SEQ ID NO:257.

[0001c] In one aspect, the present disclosure refers to a peptide ligand consisting of an amino acid sequence of any of SEQ ID NO:1-SEQ ID NO:257.

[0001d] In one aspect, the present disclosure refers to a polynucleotide encoding a peptide as disclosed herein.

[0001e] In one aspect, the present disclosure refers to a pharmaceutical composition comprising polynucleotide as disclosed herein, a peptide or combination of peptides as disclosed herein; and a pharmaceutically acceptable excipient.

[0001f] In one aspect, the present disclosure refers to an antigen presenting cell comprising the peptide or the combination of peptides as disclosed herein.

[0001g] Compositions are provided for ligands for a T cell receptor (TCR) of interest in a defined MHC context. The composition may comprise or consist of a defined peptide, or may comprise or consist of a polynucleotide encoding such a peptide. Such peptides may be fragments of naturally occurring antigenic proteins; may be fragments of neoantigenic proteins that are the subject of somatic mutation during tumorigenesis, or may be a synthetically generated mimic of an antigenic protein. The synthetic peptides can act as highly potent agonists of T cell receptors. In some examples a peptide, or encoding sequence, is selected from sequences provided herein, including without limitation any one or a combination of the peptide sequences set forth in SEQ ID NO:1-257. A peptide may be provided as short antigenic sequence active in stimulating T cells; or may be provided in the form of the larger protein, e.g. an intact domain, a soluble protein portion, a complete protein, etc. In some examples, peptide antigens are identified that are shared between patients

and provide a means for broadly applicable therapy. In other examples identification of antigens provides for a personalized medicine approach.

[0002] Identification of T cell receptors and cognate antigens provides targets for immunotherapy, including screening of patient T cells for responsiveness, vaccination with peptides or nucleic acids encoding such peptides, cell-based therapies, protein-based therapies, etc. The peptides and methods disclosed herein are useful in classifying TCRs based on peptide antigen specificities, which allows the identification of clinical candidate TCRs that recognize shared antigens across patients.

[0002a] In one aspect, the present disclosure refers to method of screening for a set of peptide antigens of a TCR of interest, the method comprising:

(a) contacting the TCR of interest with a population of host cells which express a multiplexed library of at least 10^8 different polynucleotides encoding single chain polypeptides on the cell surface, the single chain polypeptides comprising—

binding domains of the MHC protein; and
a peptide ligand;

(b) selecting for host cells expressing a single chain polypeptide that binds to the TCR of interest;

(c) iterating step (b) for at least three rounds;

(d) sequencing a nucleic acid encoding the single chain polynucleotides of the final selected population to determine a dataset of possible amino acids for each position of the peptide ligand;

(e) inputting the dataset to computer readable medium to generate a search algorithm; and

(f) searching a sequence database with the search algorithm to identify the set of peptide antigens that bind to the T cell receptor.

[0002b] In one aspect, the present disclosure refers to a set of peptide antigens for a TCR of interest identified by a method comprising:

(a) contacting the TCR of interest with a population of host cells which express a multiplexed library of at least 10^6 different polynucleotides encoding single chain polypeptides on the cell surface, the single chain polypeptides comprising—

binding domains of an MHC protein; and
a peptide ligand;

(b) selecting for host cells expressing a single chain polypeptide that binds to the TCR of interest;

(c) iterating step (b) for at least three rounds to generate a final selected population of host cells;

(d) sequencing a nucleic acid encoding the single chain polynucleotides of the final selected population to determine a dataset of possible amino acids for each position of the peptide ligand;

(e) inputting the dataset to computer readable medium to generate a search algorithm;

(f) searching a sequence database with the search algorithm to identify the set of peptide antigens for the TCR of interest.

[0003] In some examples, methods are provided for vaccination against cancer, for example colorectal cancer, the method comprising administering an effective dose of a vaccine composition, which composition may comprise a peptide identified herein; a combination of peptides, e.g. 2, 3, 4, 5, 6, 7, 8, 9, 10 or more distinct peptides; a complex of a peptide and at least a portion of an MHC protein; an autologous or allogeneic T cell that has been stimulated to respond to an antigenic peptide identified herein; a nucleic acid encoding an antigenic peptide identified herein; and optionally a pharmaceutically acceptable excipient, which may comprise a vaccine adjuvant. The peptide vaccination strategy may be used to initially prime an immune response, e.g. with a synthetic peptide provided herein, followed by a boost with the corresponding known wildtype antigen or wildtype whole protein.

[0003a] In one aspect, the present disclosure refers to use of an effective dose of the pharmaceutical composition as disclosed herein, or the antigen presenting cell as disclosed herein in the manufacture of a medicament for inducing an immune response to a cancer cell antigen in a subject in need thereof.

[0003b] In one aspect, the present disclosure refers to a method of determining a subject's response to an antigen, the method comprising: analyzing a sample comprising T cells from the subject for T cell stimulation in response to a peptide according to any SEQ ID NO: 1-257, wherein T cell stimulation is indicative that an immune response is inducible in the subject according to the use as disclosed herein.

[0004] The defined peptides are identified by screening peptide-MHC libraries by yeast-display was used to identify the recognition landscape of individual T cell receptors. The screening method may be utilized in a multiplex method to screen a plurality of peptide libraries simultaneously, e.g. screening 2, 3, 4 or more libraries simultaneously. Multiplexing allows improved efficiency of antigen discovery. Each library may comprise a unique epitope tag, e.g. an epitope targetable by an antibody, to allow identification; may comprise DNA barcodes; protein barcodes; etc. Each library utilizing the epitope tags were generated separately and diversities calculated, e.g. based on colony counts from limiting dilution of the initial libraries on growth plates. Pooling T cell receptors for library selection can further multiplex the selection, e.g. multiplexing of peptide sequence, peptide lengths, collections of different MHC or HLA alleles, *etc.* For selections, each barcode, epitope tag, etc. may be monitored via anti-epitope tag staining to detect the level of peptide-specific enrichment, statistical algorithms and machine-learning algorithms may be used for identification.

[0005] In some examples sequences of T cell receptors responsive to cancer antigens are provided. T cell receptor sequences may include, without limitation, the proteins having an alpha chain with sequence set forth in SEQ ID NO:258, optionally combined with a beta chain sequence of SEQ ID NO:259 or SEQ ID NO:260. The binding regions (CDR) sequences of these T cell receptors may be grafted onto an antibody framework to provide a TCR-like antibody. Because T cell receptors are adaptable and often unique from patient-to-

patient, the individual T cell receptor sequences may differ between patients. Despite these differences, different TCR can still recognize the same target. Thus, different T cell receptors may have slight sequence variations from these T cell receptors that can bind the same target. Additionally, T cell receptors may be modified to introduce amino acid substitutions that will allow binding to the same antigen. Such cases include affinity maturation of the T cell receptor for the specific target or receptor modification to improve the specificity of the T cell receptor for its target. The recognition portion of a T cell receptor can be grafted onto other protein scaffolds to be used as a therapeutic reagent. Because T cell receptors are somewhat cross-reactive, the list of synthetic peptides is not exhaustive. Slight modifications to peptide sequences can still result in T cell stimulation.

[0006] In some examples the T cells from which TCR sequences for screening are obtained are isolated from tumor sites, and may include without limitation tumor infiltrating T cells (TILs). In other examples the T cells are obtained from an individual responsive to an infection, e.g. bacterial, viral, protozoan, etc. infection. In other examples the T cells are obtained from a graft recipient, and may be isolated from the site of a graft.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The invention is best understood from the following detailed description when read in conjunction with the accompanying drawings. The patent or application file contains at least one drawing executed in color. It is emphasized that, according to common practice, the various features of the drawings are not to-scale. On the contrary, the dimensions of the various features are arbitrarily expanded or reduced for clarity. Included in the drawings are the following figures.

[0009] Figure 1. Design of the peptide-HLA-A*02:01 yeast-display library. (A) Methodology for selecting a yeast-display library of pHLA. Each yeast display a unique peptide that is genetically encoded. A typical library contains $\sim 10^8$ unique peptides, which is selected by a TCR of interest. Yeast are enriched in an affinity-based selection using bead-multimerized TCR and grown for iterative rounds of selection. Peptides are successively enriched and all yeast DNA is deep-sequenced. These synthetic peptide sequences are used to generate a model to make predictions for TCR ligands derived from the human proteome and/or patient-specific exome. (B) The goal of the study is to use the yeast-display selection to de-orphanize a TCR of unknown antigen specificity. The peptides selected by a TCR from the yeast-display selection generates a recognition landscape for a particular TCR, which is then used to make predictions of antigen specificity for orphan TCRs. Predicted targets can be validated in a T cell stimulation assay. (C) The construct utilizes a single-chain design to display the pHLA-A*02:01 complex tethered to an epitope tag and Aga2p, which binds to the native Aga1 protein on yeast. Each component is connected covalently by a Gly-Ser linker. The epitope

WHAT IS CLAIMED IS:

1. A T cell receptor or antibody comprising the amino acid sequence of SEQ ID NO: 258 and the amino acid sequences of SEQ ID NO: 259, SEQ ID NO: 260, or a combination thereof.
2. The T cell receptor of claim 1, comprising the amino acid sequence of SEQ ID NO: 258 and the sequence of SEQ ID NO: 259 or the sequence of SEQ ID NO: 260.
3. An immune cell engineered to comprise a T cell receptor of claim 1 or claim 2, or the antibody of claim 1.
4. The T cell receptor of any of claims 1 to 3, wherein SEQ ID NO: 258 is a T cell receptor alpha chain of 2A and 3B.
5. The T cell receptor of any of claims 1 to 3, wherein SEQ ID NO: 259 is a T cell receptor beta chain of 2A.
6. The T cell receptor of any of claims 1 to 3, wherein SEQ ID NO: 260 is a T cell receptor beta chain of 3B.
7. A peptide ligand comprising an amino acid sequence of any of SEQ ID NO:1-SEQ ID NO:257.
8. A peptide ligand consisting of an amino acid sequence of any of SEQ ID NO:1-SEQ ID NO:257.
9. A polynucleotide encoding a peptide of claim 7 or claim 8.
10. A pharmaceutical composition comprising polynucleotide of claim 9, a peptide or combination of peptides of claim 7 or claim 8; and a pharmaceutically acceptable excipient.
11. A pharmaceutical composition of claim 10, further comprising a vaccine adjuvant.
12. A pharmaceutical composition of claim 10 or claim 11, wherein the peptide or the combination of peptides is complexed with an MHC antigen.
13. An antigen presenting cell comprising the peptide or the combination of peptides of claim 7 or claim 8.

14. Use of an effective dose of the pharmaceutical composition of any of claims 10-12, or the antigen presenting cell of claim 13 in the manufacture of a medicament for inducing an immune response to a cancer cell antigen in a subject in need thereof.

15. A method of determining a subject's response to an antigen, the method comprising:
analyzing a sample comprising T cells from the subject for T cell stimulation in response to a peptide according to any SEQ ID NO: 1-257, wherein T cell stimulation is indicative that an immune response is inducible in the subject according to the use of claim 14.

16. A set of peptide antigens for a TCR of interest identified by a method comprising:
(a) contacting the TCR of interest with a population of host cells which express a multiplexed library of at least 10^6 different polynucleotides encoding single chain polypeptides on the cell surface, the single chain polypeptides comprising—
binding domains of an MHC protein; and
a peptide ligand;
(b) selecting for host cells expressing a single chain polypeptide that binds to the TCR of interest;
(c) iterating step (b) for at least three rounds to generate a final selected population of host cells;
(d) sequencing a nucleic acid encoding the single chain polynucleotides of the final selected population to determine a dataset of possible amino acids for each position of the peptide ligand;
(e) inputting the dataset to computer readable medium to generate a search algorithm;
(f) searching a sequence database with the search algorithm to identify the set of peptide antigens for the TCR of interest.

17. The set of peptide antigens of claim 16, wherein the peptide ligand is from 8 to 20 amino acids in length.

18. The set of peptide antigens of claim 17, wherein the multiplexed library contains the peptide ligand randomized at multiple positions.

19. The set of peptide antigens of claim 18, wherein the multiplexed library of the peptide ligand has limited diversity at the MHC anchor positions.

20. The set of peptide antigens of any one of claims 16-19, wherein the MHC binding domains comprise the alpha 1 and alpha 2 domains of a class I MHC protein and $\beta 2$ microglobulin.

21. The set of peptide antigens of claim 20, wherein the class I MHC is an allele of HLA-A2.

22. The set of peptide antigens of claim 21, wherein the HLA-A2 allele comprises the amino acid change {Y84A}.

23. A method of screening for a set of peptide antigens of a TCR of interest, the method comprising:

(a) contacting the TCR of interest with a population of host cells which express a multiplexed library of at least 10^8 different polynucleotides encoding single chain polypeptides on the cell surface, the single chain polypeptides comprising—

binding domains of the MHC protein; and

a peptide ligand;

(b) selecting for host cells expressing a single chain polypeptide that binds to the TCR of interest;

(c) iterating step (b) for at least three rounds;

(d) sequencing a nucleic acid encoding the single chain polynucleotides of the final selected population to determine a dataset of possible amino acids for each position of the peptide ligand;

(e) inputting the dataset to computer readable medium to generate a search algorithm; and

(f) searching a sequence database with the search algorithm to identify the set of peptide antigens that bind to the T cell receptor.