METHODS FOR PROFILING THE T-CELL RECEPTOR REPERTOIRE

The present disclosure relates to methods for profiling subject specific and personalized T cell receptor (TCR) repertoires using a single-cell sequencing method. More particularly, disclosed are methods for determining binding of T cell receptors to subject specific neoantigens. In addition, the techniques herein may identify the antigenic targets of T cell receptors in the context of tumor neoantigens. Moreover, the present disclosure enables the discovery of T cell targets in numerous diseases, with implications for understanding the basic mechanisms of the mammalian immune response and for developing antigen-specific diagnostic markers and therapies. Finally, cloned TCRs can be used to formulate personalized immunotherapies for those afflicted with a disease, such as cancer.
METHODS FOR PROFILING THE T-CELL RECEPTOR REPERTOIRE

RELATED APPLICATIONS AND INCORPORATION BY REFERENCE


[0002] The foregoing applications, and all documents cited therein or during their prosecution ("appln cited documents") and all documents cited or referenced in the appln cited documents, and all documents cited or referenced herein ("herein cited documents"), and all documents cited or referenced in herein cited documents, together with any manufacturer's instructions, descriptions, product specifications, and product sheets for any products mentioned herein or in any document incorporated by reference herein, are hereby incorporated herein by reference, and may be employed in the practice of the invention. More specifically, all referenced documents are incorporated by reference to the same extent as if each individual document was specifically and individually indicated to be incorporated by reference.

FIELD OF THE INVENTION

[0003] The present disclosure relates to methods for profiling the T cell receptor repertoire of single subjects in need thereof and preparing subject specific treatments based on the T cell receptors.

FEDERAL FUNDING LEGEND

[0004] This invention was made with government support under Grant No. CA155010 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0005] The T cell receptor (TCR) is a molecule found on the surface of T lymphocytes (i.e. T cells) that is responsible for recognizing antigens bound to major histocompatibility complex (MHC) molecules. The TCR is a heterodimer composed of two different protein chains. In most T cells (about 95%), these two protein chains are termed the alpha (α) and beta (β) chains. However, in a small percentage of T cells (about 5%), these two protein chains are termed the gamma and delta (γ/δ) chains. The ratio of TCRs comprised of α/β chains versus γ/δ chains may change during a diseased state. When the TCR engages with antigenic peptide and MHC
(peptide/MHC), the T lymphocyte is activated through a series of biochemical events mediated by associated enzymes, co-receptors, specialized adaptor molecules, and activated or released transcription factors.

[0006] The genetically programmed variability of TCRs and immunoglobulins (Ig) underlies immune recognition of diverse antigens. The selection of antigen-specific T and B cells under different pressures—such as infections, vaccines, autoimmune diseases, allergy, and tumors—can dramatically alter the repertoire in individuals either transiently or permanently. However, since the active receptor consists of paired chains (e.g., TCRα/TCRβ or IgH/IgL) within single cells, determination of active paired chains requires the sequencing of single cells.

[0007] The immune system is a vital component in preventing and eliminating cancer. Cytotoxic T cells (CTL) and natural killer cells (NK) have potent ability to kill tumor cells and numerous studies show that effector T cells at the tumor site predict favorable outcome across many cancers. Additionally, tumors vary greatly between each individual, requiring subject specific T cells targeting subject specific tumor antigens to kill a tumor.

[0008] Methods for sequencing TCR receptors have been described. Linnemann, et al., describe a method for identifying unpaired TCR sequences using genomic DNA from a large number of samples and assemble a library of TCRs (Nature Medicine 2013 Nov;19(1 l):1534-41). The library was proposed to be used for autologous TCR gene therapy without knowledge of antigen specificity. Dossinger, et al., describe isolation of paired full-length TCR sequences from non-expanded antigen-specific T cells using a PCR-based method (TCR-SCAN) (PLoS One. 2013 Apr 26;8(4):e61384). The method allowed isolation of TCRs of known oncogenes. Seitz, et al., describes a method to identify TCRaP pairs from archival tissue (Proc Natl Acad Sci USA. 2006 Aug 8;103(32):12057-62). However, TCR transfected cells were not reactive to the antigens assayed.

[0009] Personalized treatments based on the matching of neoantigens to T cells has not been previously described. Thus, there remains an unmet need in the art to develop methods to treat cancer patients in need thereof with a personalized immunotherapy.

[0010] Citation or identification of any document in this application is not an admission that such document is available as prior art to the present invention.
SUMMARY OF THE INVENTION

[0011] The present disclosure relates to methods for profiling the T cell repertoire of individual subjects and matching of TCR pairs with subject specific neoepitopes. Specifically, T cell receptors are identified that can recognize a subject specific tumor. Therefore, it is an object of the present invention to treat a subject in need thereof with T cells expressing personalized T cell receptors to provide a subject specific and effective therapy.

[0012] In a first aspect TCRs from individual T cells are identified. In one embodiment the TCR repertoire is quantified to determine the number of cells expressing a subject specific TCR. In one embodiment T cells are obtained from a subject with cancer. In another embodiment T cells are obtained that have infiltrated a subjects tumor. In another embodiment T cells are isolated from blood. In another embodiment T cells are isolated from peripheral blood mononuclear cells (PBMC). In another embodiment T cells are enriched by binding of a ligand to T cell specific markers. In one embodiment, the markers may be CD3, CD4, CD8, CD28, or any combination therewith. In one embodiment the markers are CD3/CD28. In one embodiment the ligands are antibodies. In one embodiment the antibodies are conjugated to beads. In one embodiment the antibodies are fluorescently labeled. In one embodiment the cells are separated by cell sorting.

[0013] In one embodiment single T cells are sequenced. In one embodiment single T cells are diluted such that each well of a plate contains a single cell. In one embodiment the single T cells are expanded in tissue culture. In one embodiment the nucleic acid from the single expanded T cell clones are sequenced. In one embodiment the nucleic acid from the single cells is sequenced without expanding the cells.

[0014] In another embodiment single cells are sequenced using a microfluidic system. Single cells may be selected for and then sequenced using a microfluidic system. In one embodiment single T cells are selected by FACS and then sequenced. Single cells may be encapsulated in droplets. The droplets may include reagents for synthesizing nucleic acids. The droplets may include individual markers to identify that the nucleic acid present in the droplet originated from a single cell. The individual marker may be a barcode. The barcode may be a polynucleotide with a unique sequence. The barcode may be introduced on a bead that is incorporated into the droplet.
In one embodiment, the present disclosure relates to methods for a platform for profiling the T cell receptor repertoire using single-cell bar-coded droplets (SCBD). In one embodiment a high-throughput single cell analysis system that integrates micron-scale engineered emulsions (droplets), novel molecular barcoding, hydrogel materials, and massively parallel sequencing techniques is used. This technology enables, a general method for the capture of single cell genomic information, including the pairing of multiple amplicon sequences per individual cell from an extremely large cell population (>10^6). SCBD may be used to profile the T cell repertoire and to quantify the relative abundance of each T cell clone within a population. In addition, methods are provided to identify the antigenic targets of T cell receptors in the context of tumor neoantigens.

In a second aspect neoantigens are identified for single subjects. In addition to identifying individual TCRs present in a subject with cancer, neoantigens targeted by T cells expressing the TCRs are identified. In one embodiment neoantigens are determined by whole exome sequencing. In one embodiment neoantigens are determined based on individual HLA type of a subject. In one embodiment molecular modeling is used to determine the TCR that binds to a neoantigen.

The invention comprehends performing methods as in U.S. patent application No. 20110293637, incorporated herein by reference, e.g., a method of identifying a plurality of at least 4 subject-specific peptides and preparing a subject-specific immunogenic composition that upon administration presents the plurality of at least 4 subject-specific peptides to the subject’s immune system, wherein the subject has a tumor and the subject-specific peptides are specific to the subject and the subject’s tumor, said method comprising:

(i) identifying, including through
nucleic acid sequencing of a sample of the subject’s tumor and
nucleic acid sequencing of a non-tumor sample of the subject,
a plurality of at least 4 tumor-specific non-silent mutations not present in the non-tumor sample; and

(ii) selecting from the identified non-silent mutations the plurality of at least 4 subject-specific peptides, each having a different tumor neo-epitope that is an epitope specific to the tumor of the subject, from the identified plurality of tumor specific mutations,
wherein each neo-epitope is an expression product of a tumor-specific non-silent mutation not present in the non-tumor sample, each neo-epitope binds to a HLA protein of the subject, and selecting includes
determining binding of the subject-specific peptides to the HLA protein, and

(iii) formulating the subject-specific immunogenic composition for administration to the subject so that upon administration the plurality of at least 4 subject-specific peptides are presented to the subject's immune system,

wherein the selecting or formulating comprises at least one of:

including in the subject-specific immunogenic composition a subject-specific peptide that includes an expression product of an identified neo-ORF, wherein a neo-ORF is a tumor-specific non-silent mutation not present in the non-tumor sample that creates a new open reading frame, and

including in the subject-specific immunogenic composition a subject-specific peptide that includes an expression product of an identified point mutation and has a determined binding to the HLA protein of the subject with an IC50 less than 500 nM, whereby, the plurality of at least 4 subject-specific peptides are identified, and the subject-specific immunogenic composition that upon administration presents the plurality of at least 4 subject-specific peptides to the subject's immune system, wherein the subject-specific peptides are specific to the subject and the subject's tumor, is prepared; or a method of identifying a neoantigen comprising:
a. identifying a tumor specific mutation in an expressed gene of a subject having cancer;
b. wherein when said mutation identified in step (a) is a point mutation:
   i. identifying a mutant peptide having the mutation identified in step (a), wherein said mutant peptide binds to a class I HLA protein with a greater affinity than a wild-type peptide; and has an IC50 less than 500 nm;
c. wherein when said mutation identified in step (a) is a splice-site, frameshift, read-through or gene-fusion mutation:
   i. identifying a mutant polypeptide encoded by the mutation identified in step (a), wherein said mutant polypeptide binds to a class I HLA protein; or a method of inducing a tumor specific immune response in a subject comprising administering one or more peptides or
polypeptides identified and an adjuvant: or a method of vaccinating or treating a subject for cancer comprising:

a. identifying a plurality of tumor specific mutations in an expressed gene of the subject wherein when said mutation identified is a:
   i. point mutation further identifying a mutant peptide having the point mutation; and/or
   ii. splice-site, frameshift, read-through or gene-fusion mutation further identifying a mutant polypeptide encoded by the mutation;

b. selecting one or more mutant peptides or polypeptides identified in step (a) that binds to a class I HLA protein;

c. selecting the one or more mutant peptides or polypeptides identified in step (b) that is capable of activating anti-tumor CDS T cells, and

d. administering to the subject the one or more peptides or polypeptides, autologous dendritic cells or antigen presenting cells pulsed with the one or more peptides or polypeptides selected in step (c); or preparing a pharmaceutical composition comprising one identified peptide(s), and performing method(s) as herein discussed. Thus, the neoplasia vaccine or immunogenic composition herein can be as in U.S. patent application No. 201 10293637.

[0018] In a third aspect the present invention provides for functional analysis of subject specific TCRs. In one embodiment the TCRs are cloned into a vector that allows expression. In one embodiment the TCRs are cloned into a plasmid. In another embodiment the TCRs are cloned into a viral vector. In another embodiment the TCRs are expressed in T cells. In one embodiment the T cells are transformed with a plasmid. In another embodiment the T cells are transduced with a virus. In one embodiment TCRs are expressed in cells that do not express endogenous TCRs. In one embodiment TCRs are transduced into a mouse cell line. In one embodiment TCRs are transduced into a human cell line. In one embodiment the T cell can release IL2. In one embodiment the T cells can express the cellular machinery to function in cytolytic killing of a tumor cell. In one embodiment the cells are BW5147 cells. In one embodiment the cells are peripheral blood lymphocytes. In one embodiment the T cells expressing cloned TCRs are used to assay cytolytic activity against subject specific tumor cells in vitro. In one embodiment T cells expressing cloned TCRs are used to assay binding to tumor cells obtained from a subject. In one embodiment TCRs that bind neoantigens are determined. In one embodiment the TCRs identified in a subject are used to model binding to the neoantigens
present in the subject. In one embodiment T cells expressing cloned TCRs are used to assay binding to subject specific neoantigens. In one embodiment soluble recombinant TCRs are used to bind subject specific neoantigens. In another embodiment T cells expressing subject specific TCRs are incubated with antigen presenting cells that present subject specific neoantigens to the engineered T cells. In one embodiment at least one reporter is used to detect binding of TCRs to antigen. In one embodiment the T cell line includes a polynucleotide sequence encoding a reporter gene. The reporters may be expressed in the T cells that express the cloned TCRs. In another embodiment the expression of cytokines is used to assay TCR binding to antigen.

[0019] In another embodiment the TCRs present in a subject are monitored before, after and/or simultaneously with the administration of an immunogenic composition that includes neoantigens. TCRs may be monitored by PCR using primers specific to each T cell receptor pair.

[0020] In a fourth aspect the present invention provides a personalized treatment for a subject using the identified T cell repertoire. In one embodiment identification of the T cell repertoire is used to determine an immunogenic composition or vaccine to be administered to a subject in need thereof. In one embodiment the immunogenic composition is a neoantigen vaccine. In another embodiment the neoantigen vaccine may comprise subject specific neoantigen peptides. In one embodiment neoantigen peptides to be included in a neoantigen vaccine are selected based on the quantification and identity of subject specific TCRs. In one embodiment the neoantigen peptides are selected based on the binding affinity of the peptide to a TCR. In one embodiment the selecting is based on a combination of both the quantity and the binding affinity. Not being bound by a theory, a TCR that binds strongly to a neoantigen in a functional assay, but that is not highly represented in the TCR repertoire is a good candidate for a neoantigen vaccine because T cells expressing the TCR would be advantageously amplified.

[0021] In one embodiment the subject in need thereof is treated with T cells expressing T cell receptors. In one embodiment the T cell receptors are autologous. In one embodiment T cells are transduced with subject specific TCRs. In one embodiment T cells are obtained from the subject being treated. In another embodiment the T cell receptors target neoantigens. In one embodiment the T cells are administered to a subject. In one embodiment the T cells are administered after immunoablation. Not being bound by a theory this may be an effective treatment because the drug regimen used to treat a patient may have killed the majority of the
tumor cells and an immunotherapy that includes T cells can be effective to remove any remaining cells. In another embodiment the T cells are administered within a standard of care for a cancer.

Accordingly, it is an object of the invention to not encompass within the invention any previously known product, process of making the product, or method of using the product such that Applicants reserve the right and hereby disclose a disclaimer of any previously known product, process, or method. It is further noted that the invention does not intend to encompass within the scope of the invention any product, process, or making of the product or method of using the product, which does not meet the written description and enablement requirements of the USPTO (35 U.S.C. § 112, first paragraph) or the EPO (Article 83 of the EPC), such that Applicants reserve the right and hereby disclose a disclaimer of any previously described product, process of making the product, or method of using the product.

It is noted that in this disclosure and particularly in the claims and/or paragraphs, terms such as "comprises", "comprised", "comprising" and the like can have the meaning attributed to it in U.S. Patent law; e.g., they can mean "includes", "included", "including", and the like; and that terms such as "consisting essentially of" and "consists essentially of" have the meaning ascribed to them in U.S. Patent law, e.g., they allow for elements not explicitly recited, but exclude elements that are found in the prior art or that affect a basic or novel characteristic of the invention.

These and other embodiments are disclosed or are obvious from and encompassed by, the following Detailed Description.

**BRIEF DESCRIPTION OF THE DRAWINGS**

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

The following detailed description, given by way of example, but not intended to limit the invention solely to the specific embodiments described, may best be understood in conjunction with the accompanying drawings.

**Figure 1** is an exemplary schematic showing that recognition of MHC-bound peptide by the combined TCRP and TCRa proteins occurs primarily in the CDR3 regions.
[0028] **Figure 2** illustrates that a T cell receptor (TCR) recognizes a specific peptide presented on HLA.

[0029] **Figure 3** illustrates the amplification of the CDR3 region and addition of sequencing adaptors by RT-PCR and PCR. Sequences P5 and P7 are configured to bind to a flow cell and sequences SBS3 and SBS12 are configured to bind sequencing primers.

[0030] **Figure 4** illustrates TCRaP sequencing data analysis.

[0031] **Figure 5** illustrates generating TCRaP constructs using a pre-assembled library of Va and Vβ chains.

[0032] **Figure 6** illustrates a Golden Gate Assembly of TCRa/β vectors. A pre-made library of Vβ and Va plasmids is linked to a synthesized CDR3.

[0033] **Figure 7** illustrates a reporter assay for detecting a TCR specific for a neoantigen.

[0034] **Figure 8** illustrates a workflow of a proof of concept experiment for obtaining specific TCRs using known TCR sequences. A TCR specific for EBNA3A is shown.

[0035] **Figure 9** illustrates expression of EBNA3A-specific TCRaP in JurkatAαP cells and the specificity of the T cells for EBNA3A. Shown are three methods of detecting TCR specificity.

[0036] **Figure 10** illustrates functional analysis of EBNA3A-specific TCR expressing JurkatAαP cells using antigen presenting cells loaded with no peptide, EBNA3A peptide or BRLF1 peptide.

[0037] **Figure 11** illustrates a workflow of a proof of concept experiment for isolating antigen-specific TCRs against a known pool of antigens (CEF).

[0038] **Figure 12** illustrates the expansion and isolation of CEF-specific T cells.

[0039] **Figure 13** illustrates TCR sequencing of CEF-specific T cells.

[0040] **Figure 14** illustrates TCRap pairs in PBMCs and PMBCs stimulated with CEF peptides. Shown are the CEF peptides and enrichment of TCRaP pairs specific for BMLF1.

[0041] **Figure 15** illustrates IFN-γ elispot experiments with T cells from the same donor stimulated with CEF and then tested against individual antigens to confirm the pipeline of sequencing to antigen screening.

[0042] **Figure 16** illustrates a workflow for isolating T cell receptors with specificity for an antigen panel.
Figure 17 illustrates a workflow for stimulating T cells with a neoantigen peptide pool and isolating T cell receptors with specificity for a neoantigen.

Figure 18 illustrates CLL-specific CD8+ T cell immunity in CLL patients following allo-HSCT/whole tumor cell vaccination. (A) Target cell panel and expected reactivity pattern. (B) Mean tumor- or alloantigen-specific IFNγ spot production of CD8+ T cells isolated from vaccinated, GvHD or control patients. (C) Number of T cell clones specifically recognizing CLL-associated antigens per patient.

Figure 19 illustrates CD8+ naive TCR diversity increases from post-transplant day 30 to 365 in vaccinated patients. The statistical measure 'normalized entropy' characterizes the shape of distribution of TCR-β clonotype frequencies within the TCR-β repertoire. Values close to 1 indicate a relatively even distribution of TCR-β clonotypes.

Figure 20A-D Illustrates that post-HSCT/vaccination T cells of Patient 2 are (A) specifically reactive to mutated- but not wild-type-FNDC3B peptide (ELISPOT), and (B) these neoantigen-reactive T cells are detectable by neoantigen-specific tetramers, and express the cytolytic marker CD107a (C). (D) The kinetics of the mutated-FNDC3B specific T cell response in relation to molecular tumor burden.

Figure 21 illustrates an experimental workflow for functionally analysing the TCR repertoire.

Figure 22 illustrates that BW5 147 cells produce IL2 in response to PMA/ionomycin.

Figure 23 illustrates an experimental workflow for determining identified TCRs that bind to antigen presenting cells expressing neoantigens.

Figure 24 illustrates a therapeutic vaccine strategy based on tumor neoantigens. First, tumor mutations are discovered rapidly through DNA and RNA sequencing of tumor and normal tissue. Second, personalized tumor-specific mutated peptides are identified on the basis of predictive HLA-binding algorithms. Third, peptides based on neoORFs and missense neoantigens are synthesized. Finally, the peptides are delivered to patients with a powerful immune adjuvant and coupled with complementary immunotherapeutics, such as checkpoint-blockade inhibitors. (Hacohen, Cancer Immunology Research. 2013; 1(1)).
DETAILED DESCRIPTION OF THE INVENTION

[0051] Unless specifically stated or obvious from context, as used herein, the term "about" is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. About can be understood as within 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from context, all numerical values provided herein are modified by the term about.

[0052] By "agent" is meant any small molecule chemical compound, antibody, nucleic acid molecule, or polypeptide, or fragments thereof.

[0053] By "alteration" is meant a change (increase or decrease) in the expression levels or activity of a gene or polypeptide as detected by standard art known methods such as those described herein. As used herein, an alteration includes a 10% change in expression levels, preferably a 25% change, more preferably a 40% change, and most preferably a 50% or greater change in expression levels.

[0054] By "analog" is meant a molecule that is not identical, but has analogous functional or structural features. For example, a tumor specific neo-antigen polypeptide analog retains the biological activity of a corresponding naturally-occurring tumor specific neo-antigen polypeptide, while having certain biochemical modifications that enhance the analog's function relative to a naturally-occurring polypeptide. Such biochemical modifications could increase the analog's protease resistance, membrane permeability, or half-life, without altering, for example, ligand binding. An analog may include an unnatural amino acid.

[0055] By "control" is meant a standard or reference condition.

[0056] By "disease" is meant any condition or disorder that damages or interferes with the normal function of a cell, tissue, or organ.

[0057] The terms "droplet library" or "droplet libraries" are also referred to herein as an "emulsion library" or "emulsion libraries." These terms are used interchangeably throughout the specification.

[0058] By "fragment" is meant a portion of a polypeptide or nucleic acid molecule. This portion contains, preferably, at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the entire length of the reference nucleic acid molecule or polypeptide. A fragment may
contain 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more nucleotides or amino acids.

[0059] The terms "frequency" or "frequencies" refer to the rate at which microdroplets of certain species are delivered to a specific location. Moreover, this frequency or rate is a number per unit time, typically several hundred to tens of thousands per second. Furthermore the terms "frequency" or "frequencies" refers to the number of times at which droplets of certain species are delivered to a specific location.

[0060] "Hybridization" means hydrogen bonding, which may be Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary nucleobases. For example, adenine and thymine are complementary nucleobases that pair through the formation of hydrogen bonds.

[0061] The term "hydrogel" as used herein refers to a gel in which water is the dispersion medium. Non-limiting examples of hydrogels include cellulose gels, such as agarose and derivatized agarose (e.g., low melting agarose, monoclonal anti-biotin agarose, and streptavidin derivatized agarose); xanthan gels; synthetic hydrophilic polymers, such as crosslinked polyethylene glycol, polydimethyl acrylamide, polyacrylamide, polyacrylic acid (e.g., cross-linked with dysfunctional monomers or radiation cross-linking), and micellar networks; and combinations thereof.

[0062] By "immune response" is meant any cellular or humoral response against an antigen.

[0063] By "isolated polynucleotide" is meant a nucleic acid (e.g., a DNA) that is free of the genes which, in the naturally-occurring genome of the organism—or in the genomic DNA of a neoplasia/tumor derived from the organism—the nucleic acid molecule of the disclosure is derived. The term therefore includes, for example, a recombinant DNA (e.g., DNA coding for a neoORF, read-through, or InDel derived polypeptide identified in a patient's tumor) that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In addition, the term includes an RNA molecule that is transcribed from a DNA molecule, as well as a recombinant DNA that is part of a hybrid gene encoding additional polypeptide sequence.
By an "isolated polypeptide" is meant a polypeptide of the disclosure that has been separated from components that naturally accompany it. Typically, the polypeptide is isolated when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, a polypeptide of the invention. An isolated polypeptide of the disclosure may be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

The CEF Control Peptides are 8-12 amino acids in length, with sequences derived from the human Cytomegalovirus, Epstein-Barr Virus and Influenza Virus (CEF). These peptides are used in the stimulation of IFNg release from CD8+ T cells in individuals with defined ULA types, they are useful in applications such as ELISPOT, intracellular cytokine and CTL assays.

A "ligand" is to be understood as meaning a molecule which has a structure complementary to that of a receptor and is capable of forming a complex with the receptor. According to the invention, a ligand is to be understood as meaning a peptide or peptide fragment that has a suitable length and suitable binding motifs in its amino acid sequence, so that the peptide or peptide fragment is capable of forming a complex with proteins of MHC class I or MHC class II.

"Mutation" for the purposes of this document means a DNA sequence found in the tumor DNA sample of a patient that is not found in the corresponding normal DNA sample of that same patient. "Mutation" may also refer to patterns in the sequence of RNA from a patient that are not attributable to expected variations based on known information for an individual gene and are reasonably considered to be novel variations in, for example, the splicing pattern of one or more genes that has been specifically altered in the tumor cells of the patient.

"Neo-antigen" or "neo-antigenic" means a class of tumor antigens that arises from a tumor-specific mutation(s) which alters the amino acid sequence of genome encoded proteins.

By "neoplasia" is meant any disease that is caused by or results in inappropriately high levels of cell division, inappropriately low levels of apoptosis, or both. For example, cancer is an example of a neoplasia. Examples of cancers include, without limitation, leukemia (e.g.,
acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, acute myeloblastic leukemia, acute promyelocytic leukemia, acute myelomonocytic leukemia, acute monocytic leukemia, acute erythroleukemia, chronic leukemia, chronic myelocytic leukemia, chronic lymphocytic leukemia), polycythemia vera, lymphoma (e.g., Hodgkin's disease, non-Hodgkin's disease), Waldenstrom's macroglobulinemia, heavy chain disease, and solid tumors such as sarcomas and carcinomas (e.g., fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, nile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor, cervical cancer, uterine cancer, testicular cancer, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodenroglioma, schwannoma, meningioma, melanoma, neuroblastoma, and retinoblastoma). Lymphoproliferative disorders are also considered to be proliferative diseases.

[0070] The term "neoplasia vaccine" is meant to refer to a pooled sample of neoplasia/tumor specific neoantigens, for example at least two, at least three, at least four, at least five, or more neoantigenic peptides. A "vaccine" is to be understood as meaning a composition for generating immunity for the prophylaxis and/or treatment of diseases (e.g., neoplasia/tumor). Accordingly, vaccines are medicaments which comprise antigens and are intended to be used in humans or animals for generating specific defence and protective substance by vaccination. A "neoplasia vaccine composition " can include a pharmaceutically acceptable excipient, carrier or diluent.

[0071] Unless specifically stated or obvious from context, as used herein, the term "or" is understood to be inclusive. Unless specifically stated or obvious from context, as used herein, the terms "a," "an," and "the" are understood to be singular or plural.

[0072] The term "patient" or "subject" refers to an animal which is the object of treatment, observation, or experiment. By way of example only, a subject includes, but is not limited to, a
mammal, including, but not limited to, a human or a non-human mammal, such as a non-human primate, bovine, equine, canine, ovine, or feline.

[0073] "Pharmaceutically acceptable" refers to approved or approvable by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, including humans.

[0074] "Pharmaceutically acceptable excipient, carrier or diluent" refers to an excipient, carrier or diluent that can be administered to a subject, together with an agent, and which does not destroy the pharmacological activity thereof and is nontoxic when administered in doses sufficient to deliver a therapeutic amount of the agent.

[0075] A "pharmaceutically acceptable salt" of pooled tumor specific neo-antigens as recited herein may be an acid or base salt that is generally considered in the art to be suitable for use in contact with the tissues of human beings or animals without excessive toxicity, irritation, allergic response, or other problem or complication. Such salts include mineral and organic acid salts of basic residues such as amines, as well as alkali or organic salts of acidic residues such as carboxylic acids. Specific pharmaceutical salts include, but are not limited to, salts of acids such as hydrochloric, phosphoric, hydrobromic, malic, glycolic, fumaric, sulfuric, sulfamic, sulfanilic, formic, toluenesulfonic, methanesulfonic, benzene sulfonic, ethane disulfonic, 2-hydroxyethylsulfonic, nitric, benzoic, 2-acetoxybenzoic, citric, tartaric, lactic, stearic, salicylic, glutamic, ascorbic, pamoic, succinic, fumaric, maleic, propionic, hydroxymaleic, hydroiodic, phenylacetic, alkanolic such as acetic, HOOC-(CH₂)n-COOH where n is 0-4, and the like. Similarly, pharmaceutically acceptable cations include, but are not limited to sodium, potassium, calcium, aluminum, lithium and ammonium. Those of ordinary skill in the art will recognize further pharmaceutically acceptable salts for the pooled tumor specific neo-antigens provided herein, including those listed by Remington's Pharmaceutical Sciences, 17th ed., Mack Publishing Company, Easton, PA, p. 1418 (1985). In general, a pharmaceutically acceptable acid or base salt can be synthesized from a parent compound that contains a basic or acidic moiety by any conventional chemical method. Briefly, such salts can be prepared by reacting the free acid or base forms of these compounds with a stoichiometric amount of the appropriate base or acid in an appropriate solvent.

[0076] By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification.
As used herein, the terms "prevent," "preventing," "prevention," "prophylactic treatment," and the like, refer to reducing the probability of developing a disease or condition in a subject, who does not have, but is at risk of or susceptible to developing a disease or condition.

"Primer set" means a set of oligonucleotides that may be used, for example, for PCR. A primer set would consist of at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 30, 40, 50, 60, 80, 100, 200, 250, 300, 400, 500, 600, or more primers.

"Proteins or molecules of the major histocompatibility complex (MHC)," "MHC molecules," "MHC proteins" or "HLA proteins" are to be understood as meaning, in particular, proteins capable of binding peptides resulting from the proteolytic cleavage of protein antigens and representing potential T cell epitopes, transporting them to the cell surface and presenting them to specific cells there, in particular naive T cells, cytotoxic T-lymphocytes or T-helper cells. The major histocompatibility complex in the genome comprises the genetic region whose gene products are expressed on the cell surface and are important for binding and presenting endogenous and/or foreign antigens, and thus for regulating immunological processes. The major histocompatibility complex is classified into two gene groups coding for different proteins: molecules of MHC class I and MHC class II. The molecules of the two MHC classes are specialized for different antigen sources. The molecules of MHC class I typically present but are not restricted to endogenously synthesized antigens, for example viral proteins and tumor antigens. The molecules of MHC class II present protein antigens originating from exogenous sources, for example bacterial products. The cellular biology and the expression patterns of the two MHC classes are adapted to these different roles.

MHC molecules of class I consist of a heavy chain and a light chain and are capable of binding a peptide of about 8 to 11 amino acids, but usually 9 or 10 amino acids, if this peptide has suitable binding motifs, and presenting it to naive and cytotoxic T-lymphocytes. The peptide bound by the MHC molecules of class I typically but not exclusively originates from an endogenous protein antigen. The heavy chain of the MHC molecules of class I is preferably an HLA-A, HLA-B or HLA-C monomer, and the light chain is β-2-microglobulin.

MHC molecules of class II consist of an α-chain and a β-chain and are capable of binding a peptide of about 15 to 24 amino acids if this peptide has suitable binding motifs, and presenting it to T-helper cells. The peptide bound by the MHC molecules of class II usually
originates from an extracellular or exogenous protein antigen. The a-chain and the β-chain are in particular HLA-DR, HLA-DQ and HLA-DP monomers.

Ranges provided herein are understood to be shorthand for all of the values within the range. For example, a range of 1 to 50 is understood to include any number, combination of numbers, or sub-range from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50, as well as all intervening decimal values between the aforementioned integers such as, for example, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, and 1.9. With respect to sub-ranges, "nested sub-ranges" that extend from either end point of the range are specifically contemplated. For example, a nested sub-range of an exemplary range of 1 to 50 may comprise 1 to 10, 1 to 20, 1 to 30, and 1 to 40 in one direction, or 50 to 40, 50 to 30, 50 to 20, and 50 to 10 in the other direction.

A "receptor" is to be understood as meaning a biological molecule or a molecule grouping capable of binding a ligand. A receptor may serve, to transmit information in a cell, a cell formation or an organism. The receptor comprises at least one receptor unit and frequently contains two or more receptor units, where each receptor unit may consist of a protein molecule, in particular a glycoprotein molecule. The receptor has a structure that complements the structure of a ligand and may complex the ligand as a binding partner. Signaling information may be transmitted by conformational changes of the receptor following binding with the ligand on the surface of a cell. According to the invention, a receptor may refer to particular proteins of MHC classes I and II capable of forming a receptor/ligand complex with a ligand, in particular a peptide or peptide fragment of suitable length.

A "receptor/ligand complex" is also to be understood as meaning a "receptor/peptide complex" or "receptor/peptide fragment complex," in particular a peptide- or peptide fragment-presenting MHC molecule of class I or of class II.

By "reduces" is meant a negative alteration of at least 10%, 25%, 50%, 75%, or 100%.

By "reference" is meant a standard or control condition.

A "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset of, or the entirety of, a specified sequence; for example, a segment of a full-length cDNA or genomic sequence, or the complete cDNA or
genomic sequence. For polypeptides, the length of the reference polypeptide sequence will generally be at least about 10-2,000 amino acids, 10-1,500, 10-1,000, 10-500, or 10-100. Preferably, the length of the reference polypeptide sequence may be at least about 10-50 amino acids, more preferably at least about 10-40 amino acids, and even more preferably about 10-30 amino acids, about 10-20 amino acids, about 15-25 amino acids, or about 20 amino acids. For nucleic acids, the length of the reference nucleic acid sequence will generally be at least about 50 nucleotides, preferably at least about 60 nucleotides, more preferably at least about 75 nucleotides, and even more preferably about 100 nucleotides or about 300 nucleotides or any integer thereabout or there between.

[0088] By "specifically binds" is meant a compound or antibody that recognizes and binds a polypeptide of the invention, but which does not substantially recognize and bind other molecules in a sample, for example, a biological sample.

[0089] Nucleic acid molecules useful in the methods of the disclosure include any nucleic acid molecule that encodes a polypeptide of the disclosure or a fragment thereof. Such nucleic acid molecules need not be 100% identical with an endogenous nucleic acid sequence, but will typically exhibit substantial identity. Polynucleotides having "substantial identity" to an endogenous sequence are typically capable of hybridizing with at least one strand of a double-stranded nucleic acid molecule.


[0091] For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and more preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and more preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the
inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50% formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

[0092] For most applications, washing steps that follow hybridization will also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include a temperature of at least about 25°C, more preferably of at least about 42°C, and even more preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art. Hybridization techniques are well known to those skilled in the art and are described, for example, in Benton and Davis (Science 196:180, 1977); Grunstein and Hogness (Proc. Natl. Acad. Sci., USA 72:3961, 1975); Ausubel et al. (Current Protocols in Molecular Biology, Wiley Interscience, New York, 2001); Berger and Kimmel (Guide to Molecular Cloning Techniques, 1987, Academic Press, New York); and Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.

[0093] By "subject" is meant a mammal, such as a human patient or an animal (e.g., a rodent, bovine, equine, porcine, ovine, canine, feline, or other domestic mammal).

[0094] By "substantially identical" is meant a polypeptide or nucleic acid molecule exhibiting at least 50% identity to a reference amino acid sequence (for example, any one of the amino acid sequences described herein) or nucleic acid sequence (for example, any one of the
nucleic acid sequences described herein). Preferably, such a sequence is at least 60%, more preferably 80% or 85%, and more preferably 90%, 95% or even 99% identical at the amino acid level or nucleic acid to the sequence used for comparison.

Sequence identity is typically measured using sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, BLAST, BESTFIT, GAP, or PILEUP/PRETTYBOX programs). Such software matches identical or similar sequences by assigning degrees of homology to various substitutions, deletions, and/or other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. In an exemplary approach to determining the degree of identity, a BLAST program may be used, with a probability score between e^-3 and e^-100 indicating a closely related sequence.

A "T cell epitope" is to be understood as meaning a peptide sequence that can be bound by MHC molecules of class I or II in the form of a peptide-presenting MHC molecule or MHC complex and then, in this form, be recognized and bound by naive T cells, cytotoxic T-lymphocytes or T-helper cells.

By "tumor specific T cell" is meant a T cell with special affinity for a tumor antigen and whose function is restricted to a tumor.

The present invention relates to methods for profiling subject specific T cell receptor (TCR) repertoires. More particularly, the present invention relates to methods for determining binding of T cell receptors to subject specific neoantigens and for determining cytolytic activity targeting the individual subject's tumor using a single-cell sequencing method. The techniques described herein enable, for the first time, a transformative and general method to profile the T cell repertoire and to quantify the relative abundance of each T cell clone within a population. In addition, the techniques herein may identify the antigenic targets of T cell receptors in the context of tumor neoantigens. Additionally, the present disclosure enables the discovery of T cell targets in numerous diseases, with implications for understanding the basic mechanisms of the mammalian immune response and for developing antigen-specific diagnostic markers and therapies (whether immunizing or tolerizing). Finally, cloned TCRs can be used to formulate personalized immunotherapies for those inflicted with cancer.
As noted herein, T cells express specific TCR pairs that determine the target of a T cell. The immune system can be classified into two functional subsystems: the innate and the acquired immune system. The innate immune system is the first line of defense against infections, and most potential pathogens are rapidly neutralized by this system before they can cause, for example, a noticeable infection. The acquired immune system reacts to molecular structures, referred to as antigens, of the intruding organism. There are two types of acquired immune reactions, which include the humoral immune reaction and the cell-mediated immune reaction. In the humoral immune reaction, antibodies secreted by B cells into bodily fluids bind to pathogen-derived antigens, leading to the elimination of the pathogen through a variety of mechanisms, e.g. complement-mediated lysis. In the cell-mediated immune reaction, T cells capable of destroying other cells are activated. For example, if proteins associated with a disease are present in a cell, they are fragmented proteolytically to peptides within the cell. Specific cell proteins then attach themselves to the antigen or peptide formed in this manner and transport them to the surface of the cell, where they are presented to the molecular defense mechanisms, in particular T cells, of the body. Cytotoxic T cells recognize these antigens and kill the cells that harbor the antigens.

The molecules that transport and present peptides on the cell surface are referred to as proteins of the major histocompatibility complex (MHC). MHC proteins are classified into two types, referred to as MHC class I and MHC class II. The structures of the proteins of the two MHC classes are very similar; however, they have very different functions. Proteins of MHC class I are present on the surface of almost all cells of the body, including most tumor cells. MHC class I proteins are loaded with antigens that usually originate from endogenous proteins or from pathogens present inside cells, and are then presented to naive or cytotoxic T-lymphocytes (CTLs). MHC class II proteins are present on dendritic cells, B-lymphocytes, macrophages and other antigen-presenting cells. They mainly present peptides, which are processed from external antigen sources, i.e. outside of the cells, to T-helper (Th) cells. Most of the peptides bound by the MHC class I proteins originate from cytoplasmic proteins produced in the healthy host cells of an organism itself, and do not normally stimulate an immune reaction. Accordingly, cytotoxic T-lymphocytes that recognize such self-peptide-presenting MHC molecules of class I are deleted in the thymus (central tolerance) or, after their release from the thymus, are deleted or inactivated, i.e. tolerized (peripheral tolerance). MHC molecules are capable of stimulating an
immune reaction when they present peptides to non-tolerized T-lymphocytes. Cytotoxic T-lymphocytes have both T cell receptors (TCR) and CD8 molecules on their surface. T cell receptors are capable of recognizing and binding peptides complexed with the molecules of MHC class I. Each cytotoxic T-lymphocyte expresses a unique T cell receptor which is capable of binding specific MHC/peptide complexes.

The peptide antigens attach themselves to the molecules of MHC class I by competitive affinity binding within the endoplasmic reticulum, before they are presented on the cell surface. Here, the affinity of an individual peptide antigen is directly linked to its amino acid sequence and the presence of specific binding motifs in defined positions within the amino acid sequence. If the sequence of such a peptide is known, it is possible to manipulate the immune system against diseased cells using, for example, peptide vaccines.

T cells play an important role in numerous diseases, and yet in most cases, the critical TCR heterodimers and their cognate antigens have not been identified. Furthermore, the principles underlying thymic positive and negative selection, generation of regulatory T cells, and induction of peripheral anergy are not well understood. Accordingly, there is a great need to identify the TCRs within single T cells and monitor the dynamics of the TCR repertoire during these processes. Using such an approach, it should be feasible to derive the rules of self vs, non-self antigen recognition and the evolution of the repertoire and to identify the TCRs that drive disease. For example, the ability to systematically discover the functional CD8 T cell TCRs that target tumor antigens may make it possible to develop more rational cancer vaccines. Moreover, T cells expressing TCRs that have potent cytolytic activity can be expanded and used for cancer therapies. Finally, TCRs that have high affinity for a tumor can be expressed in T cells and administered as a therapy.

The highly polymorphic TCR is generated by joining of non-contiguous gene segments (\(v\beta\), D\(\beta\), J\(\beta\) for TCRP and Va, Ja for TCRa) together with deletion/insertion of random sequences at junctions and Recombination Signal Sequences (RSS) to form the highly variable CDR3 regions. The recognition of MHC-bound peptide by the combined TCR\(\beta\) and TCRa proteins occurs primarily by the CDR3 regions (see e.g., FIG. 1: Robins et al., 2010 Sci Transl Med. 2:47ra64; Krogsgaard et al., 2005; Nicholson et al., 2005). Although there is a theoretical possibility of forming as many as \(5 \times 10^{11}\) unique TCRP chains, the actual number of unique TCRP genes found in humans is closer to 0.1% of this estimate (Robins et al., 2010).
Without being bound by theory, this reduction in complexity may be due to thymic education (positive/negative selection) and antigen exposure (e.g. pathogens, tumors, self antigens), processes that select specific T cell clones. The techniques described herein will facilitate understanding of the evolution of the repertoire, and also identify functional TCRs.

Identifying TCR pairs from individual subject specific T cells

[00104] In a first aspect, T cell receptors that are expressed on individual T cells in a subject are identified. In one embodiment a source of T cells is obtained from a subject. The term "subject" is intended to include living organisms in which an immune response can be elicited (e.g., mammals). The present disclosure provides, at least in part, a method for comprehensive analysis of the TCR repertoire. In an exemplary embodiment, the present disclosure provides a platform that would include both TCR and immunoglobulin sequencing and cloning into vectors. Advantageously, the exemplar)' platform disclosed herein is generalizable to any disease setting (human or animal) in which the TCR repertoire and antigen specificity are important to study, including tumors, infections, autoimmunity, transplant and allergy/asthma.

[00105] In one embodiment T cells that have infiltrated a tumor are isolated. T cells may be removed during surgery. T cells may be isolated after removal of tumor tissue by biopsy. T cells may be isolated by any means known in the art. In one embodiment the method may comprise obtaining a bulk population of T cells from a tumor sample by any suitable method known in the art. For example, a bulk population of T cells can be obtained from a tumor sample by dissociating the tumor sample into a cell suspension from which specific cell populations can be selected. Suitable methods of obtaining a bulk population of T cells may include, but are not limited to, any one or more of mechanically dissociating (e.g., mincing) the tumor, enzymatically dissociating (e.g., digesting) the tumor, and aspiration (e.g., as with a needle).

[00106] The bulk population of T cells obtained from a tumor sample may comprise any suitable type of T cell. Preferably, the bulk population of T cells obtained from a tumor sample comprises tumor infiltrating lymphocytes (TILs).

[00107] The tumor sample may be obtained from any mammal. Unless stated otherwise, as used herein, the term "mammal" refers to any mammal including, but not limited to, mammals of the order Logomorpha, such as rabbits; the order Carnivora, including Felines (cats) and Canines (dogs); the order Artiodactyla, including Bovines (cows) and Swines (pigs); or of the order Perssodactyla, including Equines (horses). The mammals may be non-human primates, e.g., of
the order Primates, Ceboids, or Siamoids (monkeys) or of the order Anthropoids (humans and apes). In some embodiments, the mammal may be a mammal of the order Rodentia, such as mice and hamsters. Preferably, the mammal is a non-human primate or a human. An especially preferred mammal is the human.

[00108] T cells can be obtained from a number of sources, including peripheral blood mononuclear cells, bone marrow, lymph node tissue, spleen tissue, and tumors. In certain embodiments of the present invention, T cells can be obtained from a unit of blood collected from a subject using any number of techniques known to the skilled artisan, such as Ficoll separation. In one preferred embodiment, cells from the circulating blood of an individual are obtained by apheresis or leukapheresis. The apheresis product typically contains lymphocytes, including T cells, monocytes, granulocytes, B cells, other nucleated white blood cells, red blood cells, and platelets. In one embodiment, the cells collected by apheresis may be washed to remove the plasma fraction and to place the cells in an appropriate buffer or media for subsequent processing steps. In one embodiment of the invention, the cells are washed with phosphate buffered saline (PBS). In an alternative embodiment, the wash solution lacks calcium and may lack magnesium or may lack many if not all divalent cations. Initial activation steps in the absence of calcium lead to magnified activation. As those of ordinary skill in the art would readily appreciate a washing step may be accomplished by methods known to those in the art, such as by using a semi-automated "flow-through" centrifuge (for example, the Cobe 2991 cell processor) according to the manufacturer's instructions. After washing, the cells may be resuspended in a variety of biocompatible buffers, such as, for example, Ca-free, Mg-free PBS. Alternatively, the undesirable components of the apheresis sample may be removed and the cells directly resuspended in culture media.

[00109] In another embodiment, T cells are isolated from peripheral blood lymphocytes by lysing the red blood cells and depleting the monocytes, for example, by centrifugation through a PERCOLL™ gradient. A specific subpopulation of T cells, such as CD28+, CD4+, CDC, CD45RA+, and CD45RO+ T cells, can be further isolated by positive or negative selection techniques. For example, in one preferred embodiment, T cells are isolated by incubation with anti-CD3/anti-CD28 (i.e., 3*28)-conjugated beads, such as DYNABEADS® M-450 CD3/CD28 T, or XCYTE DYNABEADS™ for a time period sufficient for positive selection of the desired T cells. In one embodiment, the time period is about 30 minutes. In a further embodiment, the
time period ranges from 30 minutes to 36 hours or longer and all integer values there between.
In a further embodiment, the time period is at least 1, 2, 3, 4, 5, or 6 hours. In yet another
preferred embodiment, the time period is 10 to 24 hours. In one preferred embodiment, the
incubation time period is 24 hours. For isolation of T cells from patients with leukemia, use of
longer incubation times, such as 24 hours, can increase cell yield. Longer incubation times may
be used to isolate T cells in any situation where there are few T cells as compared to other cell
types, such in isolating tumor infiltrating lymphocytes (TIL) from tumor tissue or from
immunocompromised individuals. Further, use of longer incubation times can increase the
efficiency of capture of CD8+ T cells.

[00110] Enrichment of a T cell population by negative selection can be accomplished with a
combination of antibodies directed to surface markers unique to the negatively selected cells. A
preferred method is cell sorting and/or selection via negative magnetic immunoadherence or flow
cytometry that uses a cocktail of monoclonal antibodies directed to cell surface markers present
on the cells negatively selected. For example, to enrich for CD4+ cells by negative selection, a
monoclonal antibody cocktail typically includes antibodies to CD14, CD20, CD1ib, CD16,
HLA-DR, and CDS.

[00111] Further, monocyte populations (i.e., CD14+ cells) may be depleted from blood
preparations by a variety of methodologies, including anti-CD14 coated beads or columns, or
utilization of the phagocytotic activity of these cells to facilitate removal. Accordingly, in one
embodiment, the invention uses paramagnetic particles of a size sufficient to be engulfed by
phagocytotic monocytes. In certain embodiments, the paramagnetic particles are commercially
available beads, for example, those produced by Life Technologies under the trade name
Dynabeads™. In one embodiment, other non-specific cells are removed by coating the
paramagnetic particles with "irrelevant" proteins (e.g., serum proteins or antibodies). Irrelevant
proteins and antibodies include those proteins and antibodies or fragments thereof that do not
specifically target the T cells to be isolated. In certain embodiments the irrelevant beads include
beads coated with sheep anti-mouse antibodies, goat anti-mouse antibodies, and human serum
albumin.

[00112] In brief, such depletion of monocytes is performed by preincubating T cells isolated
from whole blood, apheresed peripheral blood, or tumors with one or more varieties of irrelevant
or non-antibody coupled paramagnetic particles at any amount that allows for removal of
monocytes (approximately a 20:1 bead:cell ratio) for about 30 minutes to 2 hours at 22 to 37 degrees C, followed by magnetic removal of cells which have attached to or engulfed the paramagnetic particles. Such separation can be performed using standard methods available in the art. For example, any magnetic separation methodology may be used including a variety of which are commercially available, (e.g., DYNAL® Magnetic Particle Concentrator (DYNAL MPC®)). Assurance of requisite depletion can be monitored by a variety of methodologies known to those of ordinary skill in the art, including flow cytometric analysis of CD14 positive cells, before and after depletion.

For isolation of a desired population of cells by positive or negative selection, the concentration of cells and surface (e.g., particles such as beads) can be varied. In certain embodiments, it may be desirable to significantly decrease the volume in which beads and cells are mixed together (i.e., increase the concentration of cells), to ensure maximum contact of cells and beads. For example, in one embodiment, a concentration of 2 billion cells/ml is used. In one embodiment, a concentration of 1 billion cells/ml is used. In a further embodiment, greater than 100 million cells/ml is used. In a further embodiment, a concentration of cells of 10, 15, 20, 25, 30, 35, 40, 45, or 50 million cells/ml is used. In yet another embodiment, a concentration of cells from 75, 80, 85, 90, 95, or 100 million cells/ml is used. In further embodiments, concentrations of 125 or 150 million cells/ml can be used. Using high concentrations can result in increased cell yield, cell activation, and cell expansion. Further, use of high cell concentrations allows more efficient capture of cells that may weakly express target antigens of interest, such as CD28-negative T cells, or from samples where there are many tumor cells present (i.e., leukemic blood, tumor tissue, etc). Such populations of cells may have therapeutic value and would be desirable to obtain. For example, using high concentration of cells allows more efficient selection of CD8+ T cells that normally have weaker CD28 expression.

In a related embodiment, it may be desirable to use lower concentrations of cells. By significantly diluting the mixture of T cells and surface (e.g., particles such as beads), interactions between the particles and cells is minimized. This selects for cells that express high amounts of desired antigens to be bound to the particles. For example, CD4+ T cells express higher levels of CD28 and are more efficiently captured than CD8+ T cells in dilute concentrations. In one embodiment, the concentration of cells used is 5x10⁶/ml. In other
embodiments, the concentration used can be from about $1 \times 10^{5}/\text{ml}$ to $1 \times 10^{7}/\text{ml}$, and any integer value in between.

[00115] T cells can also be frozen. Wishing not to be bound by theory, the freeze and subsequent thaw step provides a more uniform product by removing granulocytes and to some extent monocytes in the cell population. After a washing step to remove plasma and platelets, the cells may be suspended in a freezing solution. While many freezing solutions and parameters are known in the art and will be useful in this context, one method involves using PBS containing 20% DMSO and 8% human serum albumin, or other suitable cell freezing media, the cells then are frozen to -80$^\circ$ C at a rate of 1$^\circ$ per minute and stored in the vapor phase of a liquid nitrogen storage tank. Other methods of controlled freezing may be used as well as uncontrolled freezing immediately at -20$^\circ$ C. or in liquid nitrogen.

[00116] T cells for use in the present invention may also be antigen-specific T cells. For example, tumor-specific T cells can be used. In certain embodiments, antigen-specific T cells can be isolated from a patient of interest, such as a patient afflicted with a cancer or an infectious disease as described herein. In one embodiment neoepitopes are determined for a subject and T cells specific to these antigens are isolated. In certain embodiments, antigen-specific T cells can be induced by vaccination of a subject with a particular antigen, either alone or in conjunction with an adjuvant or pulsed on dendritic cells. In one embodiment a subject is vaccinated with a neoantigen identified as described herein. Antigen-specific cells for use in expansion may also be generated in vitro using any number of methods known in the art, for example, as described in U.S. Patent Publication No. US 20040224402 entitled, Generation And Isolation of Antigen-Specific T Cells, or in U.S. Pat. Nos. 6,040,177. Antigen-specific cells for use in the present invention may also be generated using any number of methods known in the art, for example, as described in Current Protocols in Immunology, or Current Protocols in Cell Biology, both published by John Wiley & Sons, Inc., Boston, Mass.

[00117] In a related embodiment, it may be desirable to sort or otherwise positively select (e.g. via magnetic selection) the antigen specific cells prior to or following one or two rounds of expansion. Sorting or positively selecting antigen-specific cells can be carried out using peptide-MHC tetramers (Altman, et al., Science. 1996 Oct. 4; 274(5284):94-6). In another embodiment the adaptable tetramer technology approach is used (Andersen et al., 2012 Nat Protoc. 7:891-902). Tetramers are limited by the need to utilize predicted binding peptides based on prior
hypotheses, and the restriction to specific HLAs. Peptide-MHC tetramers can be generated using techniques known in the art and can be made with any MHC molecule of interest and any antigen of interest as described herein. In a preferred embodiment, neoantigens are used. Specific epitopes to be used in this context can be identified using numerous assays known in the art. For example, the ability of a polypeptide to bind to MHC class I may be evaluated indirectly by monitoring the ability to promote incorporation of $^{125}$I labeled P2-microglobulin ($\beta_2\pi$) into MHC class I/p2m peptide heterotrimeric complexes (see Parker et al., J. Immunol. 152:163, 1994).

[00118] In one embodiment cells are directly labeled with an epitope-specific reagent for isolation by flow cytometry followed by characterization of phenotype and TCRs. In one embodiment antigen-specific T cells are isolated by contacting the T cells with antibodies specific for T cell activation markers. Antibodies that can be used with the methods of the present invention include, but are not limited to, anti-CD25, anti-CD54, anti-CD69, anti-CD38, anti-CD45RO, anti-CD49d, anti-CD40L, anti-CD137, anti-IFN-\(\gamma\), IL-2, IL-4, and other activation induced cytokines, and anti-CD134 antibodies. Sorting of antigen-specific T cells, or generally any cells of the present invention, can be carried out using any of a variety of commercially available cell sorters, including, but not limited to, MoFlo sorter (DakoCytomation, Fort Collins, Colo.), FACSArray™, FACSAria™, FACSArray™, FACSVantage™, BD™ LSR II, and FACSCaibur™ (BD Biosciences, San Jose, Calif.).

[00119] In a preferred embodiment, the method comprises selecting cells that also express CD3. The method may comprise specifically selecting the cells in any suitable manner. Preferably, the selecting is carried out using flow cytometry. The flow cytometry may be carried out using any suitable method known in the art. The flow cytometry may employ any suitable antibodies and stains. Preferably, the antibody is chosen such that it specifically recognizes and binds to the particular biomarker being selected. For example, the specific selection of CDS, CDS, TTM-3, LAG-3, 4-iBB, or PD-1 may be carried out using anti-CD3, anti-CD8, aiiti-TIM-3, anti-LAG-3, anti-4-iBB, or anti-PD-1 antibodies, respectively. The antibody or antibodies may be conjugated to a bead (e.g., a magnetic bead) or to a fluorochrome. Preferably, the flow cytometry is fluorescence-activated cell sorting (FACS). TCRs expressed on T cells can be selected based on reactivity to autologous tumors. Additionally, T cells that are reactive to tumors can be selected for based on markers using the methods described in patent publication
Nos. WO2014133567 and WO2014133568, herein incorporated by reference in their entirety. Additionally, activated T cells can be selected for based on surface expression of CD107a.

[00120] In one embodiment of the invention, the method further comprises expanding the numbers of T cells in the enriched cell population. Such methods are described in U.S. Patent No. 8,637,307 and is herein incorporated by reference in its entirety. The numbers of T cells may be increased at least about 3-fold (or 4-, 5-, 6-, 7-, 8-, or 9-fold), more preferably at least about 10-fold (or 20-, 30-, 40-, 50-, 60-, 70-, 80-, or 90-fold), more preferably at least about 100-fold, more preferably at least about 1,000 fold, or most preferably at least about 100,000-fold. The numbers of T cells may be expanded using any suitable method known in the art. Exemplary methods of expanding the numbers of cells are described in patent publication No. WO 2003057171, U.S. Patent No. 8,034,334, and U.S. Patent Application Publication No. 2012/0244133, each of which is incorporated herein by reference.

[00121] In one embodiment, ex vivo T cell expansion can be performed by isolation of T cells and subsequent stimulation or activation followed by further expansion. In one embodiment of the invention, the T cells may be stimulated or activated by a single agent. In another embodiment, T cells are stimulated or activated with two agents, one that induces a primary signal and a second that is a co-stimulatory signal. Ligands useful for stimulating a single signal or stimulating a primary signal and an accessory molecule that stimulates a second signal may be used in soluble form. Ligands may be attached to the surface of a cell, to an Engineered Multivalent Signaling Platform (EMSP), or immobilized on a surface. In a preferred embodiment both primary and secondary agents are co-immobilized on a surface, for example a bead or a cell. In one embodiment, the molecule providing the primary activation signal may be a CD3 ligand, and the co-stimulatory molecule may be a CD28 ligand or 4-IBB ligand.

[00122] Once T cells are isolated the TCRs present on individual cells may be determined. Many approaches have been used to analyze the TCR repertoire previously. In one embodiment the present invention determines the TCRs of single cells by subcloning primary T cells in culture. After subcloning, each clonal population is sequenced. In a preferred embodiment the TCR pairs are amplified by PCR and then sequenced.

[00123] Single cells may be sequenced by any method known in the art. The present invention utilizes single cell sequencing to identify TCR pairs. In one embodiment T cells are sorted into single wells of a plate and each well is sequenced individually. As described herein, specific sets
of primers may be used to amplify TCR pairs for sequencing. The T cells may be sorted by FACS. The T cells may be sorted based on IFNy or any other cell surface marker.

[00124] In another embodiment, single cell analysis is performed by digital polymerase chain reactions (PCR), e.g., Fluidigm C. Digital polymerase chain reaction (digital PCR, DigitalPCR, dPCR, or dePCR) is a refinement of conventional polymerase chain reaction methods that can be used to directly quantify and clonally amplify nucleic acids including DNA, cDNA or RNA. The key difference between dPCR and traditional PCR lies in that PCR carries out one reaction per single sample and dPCR carries out a single reaction within samples separated into a large number of partitions wherein the reactions are carried out in each partition individually. A sample is partitioned so that individual nucleic acid molecules within the sample are localized and concentrated within many separate regions. The capture or isolation of individual nucleic acid molecules may be effected in micro well plates, capillaries, the dispersed phase of an emulsion, and arrays of miniaturized chambers, as well as on nucleic acid binding surfaces.

[00125] In a preferred embodiment single cell sequencing is performed using microfluidics. Microfluidics involves micro-scale devices that handle small volumes of fluids. Because microfluidics may accurately and reproducibly control and dispense small fluid volumes, in particular volumes less than 1 μl, application of microfluidics provides significant cost-savings. The use of microfluidics technology reduces cycle times, shortens time-to-results, and increases throughput. Furthermore, incorporation of microfluidics technology enhances system integration and automation. Microfluidic reactions are generally conducted in microdroplets. The ability to conduct reactions in microdroplets depends on being able to merge different sample fluids and different microdroplets. See, e.g., US Patent Publication No. 20120219947 and PCT publication NO.WO2014085802 Al.

[00126] Droplet microfluidics offers significant advantages for performing high-throughput screens and sensitive assays. Droplets allow sample volumes to be significantly reduced, leading to concomitant reductions in cost. Manipulation and measurement at kilohertz speeds enable up to $10^8$ samples to be screened in a single day. Compartmentalization in droplets increases assay sensitivity by increasing the effective concentration of rare species and decreasing the time required to reach detection thresholds. Droplet microfluidics combines these powerful features to enable currently inaccessible high-throughput screening applications, including single-cell and single-molecule assays. See, e.g., Guo et al., Lab Chip, 2012,12, 2146-2155.

In one embodiment RNA from single cells is used to make cDNA within the droplet. In one embodiment the reagents are delivered to the droplet during droplet formation. In one embodiment the reagents for producing cDNA cause the single cells to be lysed within a droplet. In one embodiment the cDNA from single cells is attached to a barcode. In one embodiment the barcode is attached to a bead. In one embodiment the beads are hydrogel beads. In one embodiment droplets are formed to include a single cell, a single barcoded bead, and reagents for producing cDNA. In one embodiment the reagents include primers specific for all TCR a and b chains. In another embodiment the cDNA is sequenced by any method of sequencing known to one of ordinary skill. In a preferred embodiment massively parallel sequencing or a next generation sequencing platform is used. Not being bound by a theory single cell sequencing allows the ability to tag the desired nucleic acids in each droplet with a unique computationally designed sequencing-compatible barcode, allowing droplets to be subsequently broken and their contents pooled for sequencing. After sequencing, the unique barcodes representing individual cells (droplets) are then re-associated in silico.
[00129] Single T cells of the present invention may be divided into single droplets using a microfluidic device. RNA and/or DNA in single cells in such droplets may be further labeled with a barcode. In this regard reference is made to Macosko et al., 2015, "Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets" Cell 161, 1202-1214 and Klein et al., 2015, "Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells" Cell 161, 1187-1201 all the contents and disclosure of each of which are herein incorporated by reference in their entirety.

[00130] Barcoding may be performed based on any of the compositions or methods disclosed in patent publication WO 2014047561 Al, Compositions and methods for labeling of agents, incorporated herein in its entirety. The term "barcode" as used herein, refers to any unique, non-naturally occurring, nucleic acid sequence that may be used to identify the originating source of a nucleic acid fragment. Such barcodes may be sequences including but not limited to, TTGAGCCT, AGTTGCTT, CCAGTTAG, ACCAACTG, GTATAACA or CAGGAGCC. Although it is not necessary to understand the mechanism of an invention, it is believed that the barcode sequence provides a high-quality individual read of a barcode associated with a viral vector, shRNA, or cDNA such that multiple species can be sequenced together.

[00131] DNA barcoding is a taxonomic method that uses a short genetic marker in an organism's DNA to identify it as belonging to a particular species. It differs from molecular phylogeny in that the main goal is not to determine classification but to identify an unknown sample in terms of a known classification. Kress et al., "Use of DNA barcodes to identify flowering plants" Proc. Natl. Acad. Sci. U.S.A. 102(23):8369-8374 (2005). Barcodes are sometimes used in an effort to identify unknown species or assess whether species should be combined or separated. Koch H., "Combining morphology and DNA barcoding resolves the taxonomy of Western Malagasy Liotrigona Moure, 1961" African Invertebrates 51(2): 413-421 (2010); and Seberg et al., "How many loci does it take to DNA barcode a crocus?" PLoS One 4(2):e4598 (2009). Barcoding has been used, for example, for identifying plant leaves even when flowers or fruit are not available, identifying the diet of an animal based on stomach contents or feces, and/or identifying products in commerce (for example, herbal supplements or wood). Soininen et al., "Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures" Frontiers in Zoology 6:16 (2009).
It has been suggested that a desirable locus for DNA barcoding should be standardized so that large databases of sequences for that locus can be developed. Most of the taxa of interest have loci that are sequencable without species-specific PCR primers. CBOL Plant Working Group, "A DNA barcode for land plants" PNAS 106(31): 12794-12797 (2009). Further, these putative barcode loci are believed short enough to be easily sequenced with current technology. Kress et al., "DNA barcodes: Genes, genomics, and bioinformatics" PNAS 105(8): 2761-2762 (2008). Consequently, these loci would provide a large variation between species in combination with a relatively small amount of variation within a species. Lahaye et al., "DNA barcoding the floras of biodiversity hotspots" Proc Natl Acad Sci USA 105(8): 2923-2928 (2008).

DNA barcoding is based on a relatively simple concept. For example, most eukaryote cells contain mitochondria, and mitochondrial DNA (mtDNA) has a relatively fast mutation rate, which results in significant variation in mtDNA sequences between species and, in principle, a comparatively small variance within species. A 648-bp region of the mitochondrial cytochrome c oxidase subunit 1 (COI) gene was proposed as a potential 'barcode'. As of 2009, databases of COI sequences included at least 620,000 specimens from over 58,000 species of animals, larger than databases available for any other gene. Ausubel, J., "A botanical macroscope" Proceedings of the National Academy of Sciences 106(31): 12569 (2009).

Software for DNA barcoding requires integration of a field information management system (FIMS), laboratory information management system (LIMS), sequence analysis tools, workflow tracking to connect field data and laboratory data, database submission tools and pipeline automation for scaling up to eco-system scale projects. Geneious Pro can be used for the sequence analysis components, and the two plugins made freely available through the Moorea Biocode Project, the Biocode LEVIS and Genbank Submission plugins handle integration with the FEVIS, the LEVIS, workflow tracking and database submission.


Identifying TCR pairs induced by neoantigens
In a second aspect, neoantigens are identified that bind to subject specific TCRs. Neoantigens are a unique class of tumor antigen characterized by their potential to generate high avidity T cells and their exquisite tumor-specificity. Abundant evidence in mouse and man supports the role of neoantigens in cancer control (reviewed in Hacohen et al, CIR 2013; Fritsch et al, CIR 2014). A complementary approach to the identification of and immunization with neoantigens is an understanding of the interacting T cell receptors (TCR). The repertoire of TCRs within tumor infiltrating lymphocytes, prior to or following immunotherapy, is a measure of T cell responsiveness and reflects the repertoire of epitopes against which these T cells react. These epitopes could be derived from predicted neoantigens, from neoantigens missed by the prediction algorithm, or from other tumor-associated antigens active in the tumor. In one embodiment a subject is immunized with at least one antigen before identifying TCRs. In another embodiment the antigen is a neoantigen. In another embodiment the subject is immunized with a neoplasia vaccine that includes at least one neoantigen. In one embodiment tumor specific T cells are enriched by stimulation with autologous tumor ex vivo (Burkhardt et al., J Clin Invest. 2013;123(9):3756-3765). Whole genome/exome sequencing may be used to identify all, or nearly all, mutated neoantigens that are uniquely present in a neoplasia/tumor of an individual patient, and that this collection of mutated neoantigens may be analyzed to identify a specific, optimized subset of neoantigens for use as a personalized cancer vaccine or immunogenic composition for treatment of the patient's neoplasia/tumor. For example, a population of neoplasia/tumor specific neoantigens may be identified by sequencing the neoplasia/tumor and normal DNA of each patient to identify tumor-specific mutations, and the patient's ULA allotype can be identified. The population of neoplasia/tumor specific neoantigens and their cognate native antigens may then be subject to bioinformatics analysis using validated algorithms to predict which tumor-specific mutations create epitopes that could bind to the patient's ULA allotype. Based on this analysis, a plurality of peptides corresponding to a subset of these mutations may be designed and synthesized for each patient, and pooled together for use as a cancer vaccine or immunogenic composition in immunizing the a subject. TCRs may be identified after an initial vaccination or any time after the initial vaccination. TCRs may be identified after a boosting dose. TCRs may be identified after a second boosting dose. TCRs may be identified before vaccination and after vaccination. The TCRs identified before and after vaccination can be compared in order to determine new TCRs present in the T
cell repertoire. Additionally, TCRs may be found to be present before and after vaccination, however, the amount of a TCR pair may be increased after vaccination. In one embodiment new TCRs are determined. In another embodiment TCR pairs are quantified to determine a change in their representation in the T cell repertoire.

In another embodiment TCRs are determined by incubating PBMCs with neoantigens followed by sequencing of the TCR repertoire as described herein.

**Functional Analysis of TCR pairs**

In a third aspect, the identified TCR pairs are used in functional assays. In a first embodiment a nucleic acid encoding the TCR pairs are cloned. A further embodiment provides a nucleic acid (e.g., a polynucleotide) encoding protein of interest (e.g., a TCRa chain, a TCRP chain, a TCR pair, a tumor neoantigen, and the like), which may be used to produce the protein in vitro. The polynucleotide may be, e.g., DNA, cDNA, PNA, CNA, RNA, either single- and/or double-stranded, or native or stabilized forms of polynucleotides, such as e.g. polynucleotides with a phosphorothiate backbone, or combinations thereof and it may or may not contain introns so long as it codes for the peptide. A still further embodiment of the disclosure provides an expression vector capable of expressing a polypeptide according to the invention. Expression vectors for different cell types are well known in the art and can be selected without undue experimentation. Generally, the DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, the DNA may be linked to the appropriate transcriptional and translational regulatory control nucleotide sequences recognized by the desired host (e.g., bacteria), although such controls are generally available in the expression vector. The vector is then introduced into the host bacteria for cloning using standard techniques (see, e.g., Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).

TCRP genes in a T cell, the appropriate segments are mixed with a synthesized CDR3 segment, and assembled into a final vector that expresses both TCRa and TCRp.

[00140] The disclosure further embraces variants and equivalents which are substantially homologous to the identified tumor TCR pairs or neo-antigens described herein. These can contain, for example, conservative substitution mutations, i.e., the substitution of one or more amino acids by similar amino acids. For example, conservative substitution refers to the substitution of an amino acid with another within the same general class such as, for example, one acidic amino acid with another acidic amino acid, one basic amino acid with another basic amino acid, or one neutral amino acid by another neutral amino acid. What is intended by a conservative amino acid substitution is well known in the art.

[00141] The disclosure also includes expression vectors comprising the isolated polynucleotides, as well as host cells containing the expression vectors. It is also contemplated within the scope of the disclosure that the polynucleotides may be provided in the form of RNA or cDNA molecules encoding the desired TCR peptides. The disclosure also provides that the one or more TCR peptides of the disclosure may be encoded by a single expression vector. The disclosure also provides that the one or more TCR peptides of the disclosure may be encoded and expressed in vivo using a viral based system (e.g., an adenovirus, AAV, or retrovirus described in more detail herein).

[00142] The term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequences for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequences. The polynucleotides of the disclosure can be in the form of RNA or in the form of DNA. DNA includes cDNA, genomic DNA, and synthetic DNA; and can be double-stranded or single-stranded, and if single stranded can be the coding strand or non-coding (anti-sense) strand.

[00143] In embodiments, the polynucleotides can comprise the coding sequence for the TCR peptide fused in the same reading frame to a marker sequence that allows, for example, for purification of the encoded polypeptide. For example, the marker sequence can be a hexa-histidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, a molecular barcode, or the marker sequence can be a hemagglutinin (HA) tag derived from the influenza hemagglutinin protein when a mammalian host (e.g., COS-7 cells) is used. Additional tags include, but are not limited to,
Calmodulin tags, FLAG tags, Myc tags, S tags, SBP tags, Softag 1, Softag 3, V5 tag, Xpress tag, Isopeptag, SpyTag, Biotin Carboxyl Carrier Protein (BCCP) tags, GST tags, fluorescent protein tags (e.g., green fluorescent protein tags), maltose binding protein tags, Nus tags, Strep-tag, thioredoxin tag, TC tag, Ty tag, molecular barcodes, and the like.

[00144] In embodiments, the present disclosure provides isolated nucleic acid molecules having a nucleotide sequence at least 60% identical, at least 65% identical, at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, or at least 96%, 97%, 98% or 99% identical to a polynucleotide encoding a TCR peptide of the present invention.

[00145] By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence can include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence.

In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence can be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence can be inserted into the reference sequence. These mutations of the reference sequence can occur at the amino- or carboxy-terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

[00146] As a practical matter, whether any particular nucleic acid molecule is at least 80% identical, at least 85% identical, at least 90% identical, and in some embodiments, at least 95%, 96%, 97%, 98%, or 99% identical to a reference sequence can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). Bestfit uses the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981), to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set such that the percentage of identity is
calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

[00147] The isolated TCR peptides described herein can be produced in vitro (e.g., in the laboratory) by any suitable method known in the art. Such methods range from direct protein synthetic methods to constructing a DNA sequence encoding isolated polypeptide sequences and expressing those sequences in a suitable transformed host. In some embodiments, a DNA sequence is constructed using recombinant technology by isolating or synthesizing a DNA sequence encoding a wild-type protein of interest. Optionally, the sequence can be mutagenized by site-specific mutagenesis to provide functional analogs thereof. See, e.g., Zoeller et al., Proc. Nat'l. Acad. Sci. USA 81:5662-5066 (1984) and U.S. Pat. No. 4,588,585.

[00148] Once assembled (e.g., by synthesis, site-directed mutagenesis, or another method), the polynucleotidide sequences encoding a particular isolated polypeptide of interest will be inserted into an expression vector and optionally operatively linked to an expression control sequence appropriate for expression of the protein in a desired host. Proper assembly can be confirmed by nucleotide sequencing, restriction mapping, and expression of a biologically active polypeptide in a suitable host. As well known in the art, in order to obtain high expression levels of a transfected gene in a host, the gene can be operatively linked to transcriptional and translational expression control sequences that are functional in the chosen expression host.

[00149] Recombinant expression vectors may be used to amplify and express DNA encoding the TCR peptides. Recombinant expression vectors are replicable DNA constructs which have synthetic or cDNA-derived DNA fragments encoding a subject or tumor specific TCR or a bioequivalent analog operatively linked to suitable transcriptional or translational regulatory elements derived from mammalian, microbial, viral or insect genes. A transcriptional unit generally comprises an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, transcriptional promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription and translation initiation and termination sequences, as described in detail herein. Such regulatory elements can include an operator sequence to control transcription. The ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants can additionally be incorporated. DNA regions are operatively linked when they are functionally related to each other. For example, DNA for a
signal peptide (secretory leader) is operatively linked to DNA for a polypeptide if it is expressed as a precursor which participates in the secretion of the polypeptide; a promoter is operatively linked to a coding sequence if it controls the transcription of the sequence; or a ribosome binding site is operatively linked to a coding sequence if it is positioned so as to permit translation. Generally, operatively linked means contiguous, and in the case of secretory leaders, means contiguous and in reading frame. Structural elements intended for use in yeast expression systems include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it can include an N-terminal methionine residue. This residue can optionally be subsequently cleaved from the expressed recombinant protein to provide a final product.

[00150] The choice of expression control sequence and expression vector will depend upon the choice of host. A wide variety of expression host/vector combinations can be employed. Useful expression vectors for eukaryotic hosts, include, for example, vectors comprising expression control sequences from SV40, bovine papilloma virus, adenovirus and cytomegalovirus. Useful expression vectors for bacterial hosts include known bacterial plasmids, such as plasmids from *Escherichia coli*, including pCR 1, pBR322, pMB9 and their derivatives, wider host range plasmids, such as M13 and filamentous single-stranded DNA phages.

[00151] Suitable host cells for expression of a polypeptide include prokaryotes, yeast, insect or higher eukaryotic cells under the control of appropriate promoters. Prokaryotes include gram negative or gram positive organisms, for example *E. coli* or bacilli. Higher eukaryotic cells include established cell lines of mammalian origin. Cell-free translation systems could also be employed. Appropriate cloning and expression vectors for use with bacterial, fungal, yeast, and mammalian cellular hosts are well known in the art (see Pouwels et al., *Cloning Vectors: A Laboratory Manual*, Elsevier, N.Y., 1985).

[00152] Various mammalian or insect cell culture systems are also advantageously employed to express recombinant protein. Expression of recombinant proteins in mammalian cells can be performed because such proteins are generally correctly folded, appropriately modified and completely functional. Examples of suitable mammalian host cell lines include the COS-7 lines of monkey kidney cells, described by Gluzman (*Cell* 23:175, 1981), and other cell lines capable of expressing an appropriate vector including, for example, L cells, C127, 3T3, Chinese hamster
ovary (CHO), HeLa and BHK cell lines. Additionally, T cell lines as described herein may be used in order to determine binding of a T cell to a tumor or antigen. Mammalian expression vectors can comprise nontranscribed elements such as an origin of replication, a suitable promoter and enhancer linked to the gene to be expressed, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' nontranslated sequences, such as necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, and transcriptional termination sequences. Baculovirus systems for production of heterologous proteins in insect cells are reviewed by Luckow and Summers, *Bio/Technology* 6:47 (1988).

The proteins produced by a transformed host can be purified according to any suitable method. Such standard methods include chromatography (e.g., ion exchange, affinity and sizing column chromatography, and the like), centrifugation, differential solubility, or by any other standard technique for protein purification. Affinity tags such as hexahistidine, maltose binding domain, influenza coat sequence, glutathione-S-transferase, and the like can be attached to the protein to allow easy purification by passage over an appropriate affinity column. Isolated proteins can also be physically characterized using such techniques as proteolysis, nuclear magnetic resonance and x-ray crystallography.

For example, supernatants from systems which secrete recombinant protein into culture media can be first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. Following the concentration step, the concentrate can be applied to a suitable purification matrix. Alternatively, an anion exchange resin can be employed, for example, a matrix or substrate having pendant diethylaminoethyl (DEAE) groups. The matrices can be acrylamide, agarose, dextran, cellulose or other types commonly employed in protein purification. Alternatively, a cation exchange step can be employed. Suitable cation exchangers include various insoluble matrices comprising sulfopropyl or carboxymethyl groups. Finally, one or more reversed-phase high performance liquid chromatography (RP-FIPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify a cancer stem cell protein-Fc composition. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a homogeneous recombinant protein.
Recombinant protein produced in bacterial culture can be isolated, for example, by initial extraction from cell pellets, followed by one or more concentration, salting-out, aqueous ion exchange or size exclusion chromatography steps. High performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of a recombinant protein can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

In certain embodiments nucleic acids encoding T cell receptors are introduced into a cell for expression of the T cell receptor. Transfection can be carried out using calcium phosphate, by electroporation, by cell squeezing or by mixing a cationic lipid with the material to produce liposomes, which fuse with the cell membrane and deposit their cargo inside. In one embodiment the TCR is transfected into a cell by using a plasmid. In other embodiments a viral vector is used to transduce a TCR into a cell. In one embodiment the viral vectors include AAV, adenovirus, or a retrovirus. Plasmids that can be used for adeno associated virus (AAV), adenovirus, and lentivirus delivery have been described previously (see e.g., U.S. Patent Nos. 6,955,808 and 6,943,019, and U.S. Patent application No. 20080254008, hereby incorporated by reference).

Among vectors that may be used in the practice of the invention, integration in the host genome of a T cell is possible with retrovirus gene transfer methods, often resulting in long term expression of the inserted transgene. In a preferred embodiment the retrovirus is a lentivirus. Additionally, high transduction efficiencies have been observed in many different cell types and target tissues. The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. A retrovirus can also be engineered to allow for conditional expression of the inserted transgene, such that only certain cell types are infected by the lentivirus. Additionally, cell type specific promoters can be used to target expression in specific cell types. Lentiviral vectors are retroviral vectors (and hence both lentiviral and retroviral vectors may be used in the practice of the invention). Moreover, lentiviral vectors are preferred as they are able to transduce or infect non-dividing cells and typically produce high viral titers. Retroviral vectors are comprised of cis-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum cis-acting LTRs are sufficient for replication and packaging of the vectors, which are then used to integrate the desired nucleic acid into the target cell to provide permanent expression. Separate retroviral

[00158] Also useful in the practice of the invention is a minimal non-primate lentiviral vector, such as a lentiviral vector based on the equine infectious anemia virus (EIAV) (see, e.g., Balagaan, (2006) J Gene Med; 8:275 - 285, Published online 21 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845). The vectors may have cytomegalovirus (CMV) promoter driving expression of the target gene. Accordingly, the invention contemplates amongst vector(s) useful in the practice of the invention: viral vectors, including retroviral vectors and lentiviral vectors.

[00159] Also useful in the practice of the invention is an adenovirus vector. One advantage is the ability of recombinant adenoviruses to efficiently transfer and express recombinant genes in a variety of mammalian cells and tissues in vitro and in vivo, resulting in the high expression of the transferred nucleic acids. Further, the ability to productively infect quiescent cells, expands the utility of recombinant adenoviral vectors. In addition, high expression levels ensure that the products of the nucleic acids will be expressed to sufficient levels to determine function in infected cells (see e.g., U.S. Patent No. 7,029,848, hereby incorporated by reference).

[00160] Also useful in the practice of the invention is an adenovirus associated virus (AAV) vector. AAV is advantageous over other viral vectors due to low toxicity and low probability of causing insertional mutagenesis because it doesn't integrate into the host genome. AAV has a packaging limit of 4.5 or 4.75 Kb. Constructs larger than 4.5 or 4.75 Kb result in significantly reduced virus production. There are many promoters that can be used to drive nucleic acid molecule expression. AAV ITR can serve as a promoter and is advantageous for eliminating the need for an additional promoter element. For ubiquitous expression, the following promoters can be used: CMV, CAG, CBh, PGK, SV40, Ferritin heavy or light chains, etc. Promoters used
to drive RNA can include: Pol III promoters such as U6 or HI. The use of a Pol II promoter and intronic cassettes can also be used to express TCRs.

[00161] In one embodiment purified recombinant TCRs are used in in vitro assays. In one embodiment in vitro binding assays to antigens are performed. TCRs in a soluble form can be made by methods known in the art (Boulter et al. (2003), Protein Engineering 16; 9: 707-711). Additionally binding assays such as Biacore have been previously described (Laugel et al., (2005), The Journal of Biological Chemistry 280; 3: 1882-1892; Gadola et al., (2006), Journal of Experimental Medicine 203: 699-710; Cole et al., (2007), The Journal of Immunology 178: 5727-5734; Sami et al., (2007), Protein Engineering, Design & Selection 20; 8: 397-403; Cole et al., (2008), Molecular Immunology 45: 2700-2709).

[00162] In one embodiment T cell lines are transfected or transduced with the subject specific TCRs of the invention for use in functional assays. In one embodiment T cell lines are deficient for endogenous TCRs. In one embodiment mammalian cells are used. In one embodiment the T cells are murine cells. In one embodiment the T cells are human cells. In one embodiment the T cells respond to antigen stimulation. In one embodiment the T cells have cytolytic activity. In one embodiment the cells are peripheral blood lymphocytes (PBLs). In one embodiment the PBLs are from a healthy subject that has not had cancer previously and does not have a tumor. In another embodiment the PBLs are from a subject with a different HLA type than the subject where the subject specific TCRs were cloned. This would assure that TCRs present on the PBL would not react with antigens that are presented by the subject specific HLA proteins. In one embodiment the cells are BW5 147.3 (ATCC TIB-47; BW) cells. In one embodiment the T cells are the human TCR-deficient Jurkat76 cells (Heemskerk, M. H., et al., (2003) Blood 102, 3530-3540).

[00163] In one embodiment functional assays are performed using T cells that express subject specific TCRs. In one embodiment neoantigens are presented by antigen presenting cells in functional assays. In one embodiment neoantigens are presented that were used to immunize the subject. In one embodiment sequence analysis of the TCRs determines the antigens to be presented. In one embodiment autologous tumor cells are presented to T cells. In one embodiment cytolytic activity is determined. In another embodiment TCR binding to antigen is determined.
Assays for detection of CTLs have relied on direct determination of cell lysis as measured by the classical assay for CTL activity namely the chromium release assay (Walker et al., (1987) Nature: 328:345-348; Scheibenbogen et al., (2000) J Immunol Methods: 244(1-2):81-89.). Effector Cytotoxic T Lymphocytes (CTL) bind targets bearing antigenic peptide on Class I MHC and signal the targets to undergo apoptosis. If the targets are labeled with $^{51}$Chromium before the CTL are added, the amount of $^{51}$Cr released into the supernatant is proportional to the number of targets killed. Antigen-specific lysis is calculated by comparing lysis of target cells expressing disease or control antigens in the presence or absence of patient effector cells, and is usually expressed as the % -specific lysis. Percent specific cytotoxicity is calculated by (specific release - spontaneous release) ÷ (maximum release - spontaneous release) and may be 20%-85% for a positive assay. Percent specific cytotoxicity is usually determined at several ratios of effector (CTL) to target cells (E:T). Additionally, the standard lytic assay is qualitative and must rely on a limiting dilution analysis (LDA) for quantitative results, and the LDA frequently underestimates the true level of CTL response. Although CTL can each kill many targets in vivo, in vitro this assay requires numbers of CTL equal to or greater than the number of targets for detectable killing. In one embodiment CTL responses are measured by the chromium release assay, monitoring the ability of T cells (Effector cells) to lyse radiolabelled HLA matched "target cells" that express the appropriate antigen-MHC complex.

Another method of measuring cytotoxicity, is the ELISPOT assay where the CD8+ CTL response, which can be assessed by measuring IFN-γ production by antigen-specific effector cells, is quantitated by measuring the number of Spot Forming Units (SFU) under a stereomicroscope (Rininsland et al., (2000) J Immunol Methods: 240(1-2):143-155.). In this assay, antigen-presenting cells (APC) are immobilized on the plastic surface of a micro titer well, and effector T cells are added at various effectortarget ratios. Antigen presenting cells are preferably B cells or dendritic cells. More preferably the B cells or dendritic cells are from the subject wherein the TCRs expressed on the effector T cells were identified. The binding of APC's by antigen-specific effector cells triggers the production of cytokines including IFN-γ by the effector cells (Murali-Krishna et al., (1998) Adv Exp Med Biol.: 452:123-142). In one embodiment subject specific T cells are used in the ELISPOT assay.

Another method for quantifying the number of circulating antigen-specific CD8+ T cells is the tetramer assay that is used to measure CTL activity. In this assay, a specific epitope
is bound to synthetic tetrameric forms of fluorescent labeled MHC Class I molecules. Since CD8+ T cells recognize antigen in the form of short peptides bound to Class I molecules, cells with the appropriate T cell receptor will bind to the labeled tetramers and can be quantified by flow cytometry. Although this method is less time-consuming than the ELISpot assay, the tetramer assay measures only binding, not function. Not all cells that bind a particular antigen necessarily become activated.

[00167] In another embodiment cytolytic activity is determined by FACS. In one embodiment target cells are incubated with T cells. The target cells may be incubated with agents that stain activated apoptotic proteins such as caspases. The target cells may also be incubated with stains that indicate cell death, such as 7-AAD. Analysis of the cells by FACS indicates the level of cytolytic activity. The present invention can use any commercially available assay to detect cytolytic activity. FACS based assays for cytolytic activity are well known in the art and have been described previously (Lee-MacAry et al., (2001). J. Immunology. Met 252, 83-92; Gogoy-Ramirez et al., (2000). Journal of Immunology. Met 239, 35-44; Goldberget al., (1999). Journal of Immunology. Methods 224, 1; Hatam et al., (1994). Cytometry 16,59; De Clerck et al., (1994) J. Immunol. Meth. 172, 115; Bronner-Fraser, J. Cell Biol. 101, 610 (1985); Rabinovitch et al., (1986) J. Immunol. 136, 2769 (1986); Su, X,J. (1996) Immunol. 156, 156, 4198).

[00168] In another embodiment TCRs will be expressed on any of the T cells described herein and further incorporating a reporter gene. The reporter gene may be activated upon binding of the TCR to antigen. The reporter gene may express a fluorescent protein. Transcription of the reporter gene may be activated upon activation of a signalling cascade initiated by TCR binding to antigen. In a preferred embodiment the reporter gene is controlled by the nuclear factor of activated T cells (NFAT) (Szymczak et al., Nat Biotechnol. 2004;22:589-594; Jones et al., Hum Gene Ther. 2009;20:630-640.). In one embodiment fluorescence indicates TCR binding to an antigen. In another embodiment T cells are analysed by FACS after activation of the reporter.

Treatment of patients with personalized TCR therapy

[00169] In a fourth aspect a subject in need thereof is treated based on the TCR repertoire of the subject. In one embodiment a neoantigen vaccine is selected based on the TCRs. In another embodiment a subject is treated with T cells expressing subject specific TCRs. The ability to effectively profile the TCR repertoire and to link individual T cells containing specific TCRs to
an epitope thereby provides an essential approach to the identification of therapy-critical T cell targets. Once identified, such TCRs provide molecular reagents to prove the functionality of epitope-specific T cells against tumor targets and to follow highly specific T cells longitudinally in a patient and also enable adoptive therapy with T cells engineered to contain these epitope-specific TCRs.

[00170] In one embodiment a neoantigen immunogenic composition or vaccine is selected based on the TCRs identified. In one embodiment identification of the T cell repertoire and testing in functional assays as described herein is used to determine an immunogenic composition or vaccine to be administered to a subject in need thereof. In one embodiment the immunogenic composition is a neoantigen vaccine. In another embodiment the neoantigen vaccine may comprise subject specific neoantigen peptides. In one embodiment neoantigen peptides to be included in a neoantigen vaccine are selected based on the quantification of subject specific TCRs that bind to the neoantigens. In one embodiment the neoantigen peptides are selected based on the binding affinity of the peptide to a TCR. In one embodiment the selecting is based on a combination of both the quantity and the binding affinity. Not being bound by a theory, a TCR that binds strongly to a neoantigen in a functional assay, but that is not highly represented in the TCR repertoire is a good candidate for a neoantigen vaccine because T cells expressing the TCR would be advantageously amplified.

[00171] Adoptive cell therapy (ACT) can refer to the transfer of cells, most commonly immune-derived cells, back into the same patient or into a new recipient host with the goal of transferring the immunologic functionality and characteristics into the new host. If possible, use of autologous cells helps the recipient by minimizing GVHD issues. The adoptive transfer of autologous tumor infiltrating lymphocytes (TIL) (Besser et al., (2010) Clin. Cancer Res 16 (9) 2646-55; Dudley et al., (2002) Science 298 (5594): 850-4; and Dudley et al., (2005) Journal of Clinical Oncology 23 (10): 2346-57.) or genetically re-directed peripheral blood mononuclear cells (Johnson et al., (2009) Blood 114 (3): 353-46; and Morgan et al., (2006) Science 314(5796) 126-9) has been used to successfully treat patients with advanced solid tumors, including melanoma and colorectal carcinoma, as well as patients with CD19-expressing hematologic malignancies (Kalos et al., (2011) Science Translational Medicine 3 (95): 95ra73). In one embodiment TCRs are selected for administering to a subject based on binding to neoantigens as described herein. In one embodiment T cells are expanded using the methods described herein.
Expanded T cells that express tumor specific TCRs may be administered back to a subject. In another embodiment PBMCs are transduced or transfected with polynucleotides for expression of TCRs and administered to a subject. T cells expressing TCRs specific to neoantigens are expanded and administered back to a subject. In one embodiment T cells that express TCRs that result in cytolytic activity when incubated with autologous tumor tissue are expanded and administered to a subject. In one embodiment T cells that express TCRs that when used in the functional assays described herein result in binding to neoantigens are expanded and administered to a subject. In another embodiment TCRs that have been determined to bind to subject specific neoantigens are expressed in T cells and administered to a subject.


[00173] As an alternative to, or addition to, TCR modifications, chimeric antigen receptors (CARs) may be used in order to generate immunoresponsive cells, such as T cells, specific for selected targets, such as malignant cells, with a wide variety of receptor chimera constructs having been described (see U.S. Patent Nos. 5,843,728; 5,851,828; 5,912,170; 6,004,811; 6,284,240; 6,392,013; 6,410,014; 6,753,162; 8,211,422; and, PCT Publication WO9215322). Alternative CAR constructs may be characterized as belonging to successive generations. First-generation CARs typically consist of a single-chain variable fragment of an antibody specific for an antigen, for example comprising a V<sub>L</sub> linked to a V<sub>H</sub> of a specific antibody, linked by a flexible linker, for example by a CD8α hinge domain and a CD8α transmembrane domain, to the
transmembrane and intracellular signaling domains of either CD3ζ or FcRy (scFv-CD3ζ or scFv-FcRy; see U.S. Patent No. 7,741,465; U.S. Patent No. 5,912,172; U.S. Patent No. 5,906,936). Second-generation CARs incorporate the intracellular domains of one or more costimulatory molecules, such as CD28, OX40 (CD134), or 4-1BB (CD137) within the endodomain (for example scFv-CD28/OX40/4-1BB-CD3Q see U.S. Patent Nos. 8,911,993; 8,916,381; 8,975,071; 9,101,584; 9,102,760; 9,102,761). Third-generation CARs include a combination of costimulatory endodomains, such a CD3C-chain, CD97, GDI la-CD18, CD2, ICOS, CD27, CD154, CDS, OX40, 4-1BB, or CD28 signaling domains (for example scFv-CD28-4-1BB-CD3C or scFv-CD28-OX40-CD3Q see U.S. Patent No. 8,906,682; U.S. Patent No. 8,399,645; U.S. Patent No. 5,686,281; PCT Publication No. WO2014134165; PCT Publication No. WO2012079000). Alternatively, costimulation may be orchestrated by expressing CARs in antigen-specific T cells, chosen so as to be activated and expanded following engagement of their native αβ TCR, for example by antigen on professional antigen-presenting cells, with attendant costimulation. In addition, additional engineered receptors may be provided on the immunoresponsive cells, for example to improve targeting of a T-cell attack and/or minimize side effects.

[00174] Alternative techniques may be used to transform target immunoresponsive cells, such as protoplast fusion, lipofection, transfection or electroporation. A wide variety of vectors may be used, such as retroviral vectors, lentiviral vectors, adenoviral vectors, adeno-associated viral vectors, plasmids or transposons, such as a Sleeping Beauty transposon (see U.S. Patent Nos. 6,489,458; 7,148,203; 7,160,682; 7,985,739; 8,227,432), may be used to introduce CARs, for example using 2nd generation antigen-specific CARs signaling through CD3ζ and either CD28 or CD137. Viral vectors may for example include vectors based on HIV, SV40, EBV, HSV or BPV.

[00175] Cells that are targeted for transformation may for example include T cells, Natural Killer (NK) cells, cytotoxic T lymphocytes (CTL), regulatory T cells, human embryonic stem cells, tumor-infiltrating lymphocytes (TIL) or a pluripotent stem cell from which lymphoid cells may be differentiated. T cells expressing a desired CAR may for example be selected through coculture with γ-irradiated activating and propagating cells (AaPC), which co-express the cancer antigen and co-stimulatory molecules. The engineered CAR T-cells may be expanded, for example by co-culture on AaPC in presence of soluble factors, such as IL-2 and IL-21. This expansion may for example be carried out so as to provide memory CAR+ T cells (which may
for example be assayed by non-enzymatic digital array and/or multi-panel flow cytometry). In this way, CAR T cells may be provided that have specific cytotoxic activity against antigen-bearing tumors (optionally in conjunction with production of desired chemokines such as interferon-γ). CAR T cells of this kind may for example be used in animal models, for example to threat tumor xenografts.

[00176] Approaches such as the foregoing may be adapted to provide methods of treating and/or increasing survival of a subject having a disease, such as a neoplasia, for example by administering an effective amount of an immunoresponsive cell comprising an antigen recognizing receptor that binds a selected antigen, wherein the binding activates the immunoreponsive cell, thereby treating or preventing the disease (such as a neoplasia, a pathogen infection, an autoimmune disorder, or an allogeneic transplant reaction). Dosing in CAR T cell therapies may for example involve administration of from 10⁶ to 10⁹ cells/kg, with or without a course of lymphodepletion, for example with cyclophosphamide.

[00177] To guard against possible adverse reactions, engineered immunoresponsive cells may be equipped with a transgenic safety switch, in the form of a transgene that renders the cells vulnerable to exposure to a specific signal. For example, the herpes simplex viral thymidine kinase (TK) gene may be used in this way, for example by introduction into allogeneic T lymphocytes used as donor lymphocyte infusions following stem cell transplantation. In such cells, administration of a nucleoside prodrug such as ganciclovir or acyclovir causes cell death. Alternative safety switch constructs include inducible caspase 9, for example triggered by administration of a small-molecule dimerizer that brings together two nonfunctional icasp9 molecules to form the active enzyme. A wide variety of alternative approaches to implementing cellular proliferation controls have been described (see U.S. Patent Publication No. 20130071414; PCT Patent Publication WO2011146862; PCT Patent Publication WO2014011987; PCT Patent Publication WO2013040371; Zhou et al. BLOOD, 2014, 123/25:3895 - 3905; Di Stasi et al., The New England Journal of Medicine 2011; 365:1673-1683; Sadelain M, The New England Journal of Medicine 2011; 365:1735-173; Ramos et al., Stem Cells 28(6): 1107-15 (2010)). In a further refinement of adoptive therapies, genome editing may be used to tailor immunoresponsive cells to alternative implementations, for example providing edited CAR T cells (see Poirot et al., 2015, Multiplex genome edited T-cell
manufacturing platform for "off-the-shelf adoptive T-cell immunotherapies, Cancer Res 75 (18): 3853).

[00178]  Cell therapy methods often involve the ex-vivo activation and expansion of T-cells. In one embodiment T cells are activated before administering them to a subject in need thereof. Activation or stimulation methods have been described herein and is preferably required before T cells are administered to a subject in need thereof. Examples of these type of treatments include the use tumor infiltrating lymphocyte (TIL) cells (see U.S. Patent No. 5,126,132), cytotoxic T-cells (see U.S. Patent No. 6,255,073; and U.S. Patent No. 5,846,827), expanded tumor draining lymph node cells (see U.S. Patent No. 6,251,385), and various other lymphocyte preparations (see U.S. Patent No. 6,194,207; U.S. Patent No. 5,443,983; U.S. Patent No 6,040,177; and U.S. Patent No. 5,766,920). These patents are herein incorporated by reference in their entirety.

[00179]  For maximum effectiveness of T-cells in cell therapy protocols, the ex vivo activated T-cell population should be in a state that can maximally orchestrate an immune response to cancer, infectious diseases, or other disease states. For an effective T-cell response, the T-cells first must be activated. For activation, at least two signals are required to be delivered to the T-cells. The first signal is normally delivered through the T-cell receptor (TCR) on the T-cell surface. The TCR first signal is normally triggered upon interaction of the TCR with peptide antigens expressed in conjunction with an MHC complex on the surface of an antigen-presenting cell (APC). The second signal is normally delivered through co-stimulatory receptors on the surface of T-cells. Co-stimulatory receptors are generally triggered by corresponding ligands or cytokines expressed on the surface of APCs.

[00180]  Due to the difficulty in maintaining large numbers of natural APC in cultures of T-cells being prepared for use in cell therapy protocols, alternative methods have been sought for ex-vivo activation of T-cells. One method is to by-pass the need for the peptide-MHC complex on natural APCs by instead stimulating the TCR (first signal) with polyclonal activators, such as immobilized or cross-linked anti-CD3 or anti-CD2 monoclonal antibodies (mAbs) or superantigens. The most investigated co-stimulatory agent (second signal) used in conjunction with anti-CD3 or anti-CD2 mAbs has been the use of immobilized or soluble anti-CD28 mAbs. The combination of anti-CD3 mAb (first signal) and anti-CD28 mAb (second signal) immobilized on a solid support such as paramagnetic beads (see US Pat No 6,352,694, herein

[00181] It is contemplated that the T cells obtained by the inventive methods can be used in methods of treating or preventing cancer. In this regard, the invention provides a method of treating or preventing cancer in a subject, comprising administering to the subject the pharmaceutical compositions or cell populations obtained by any of the inventive methods described herein in an amount effective to treat or prevent cancer in the subject. Another embodiment of the invention provides a method of treating or preventing cancer in a subject, comprising administering a cell population enriched for tumor-reactive T cells to a subject by any of the inventive methods described herein in an amount effective to treat or prevent cancer in the mammal.

[00182] For purposes of the inventive methods, wherein populations of cells are administered, the cells can be cells that are allogeneic or autologous to the subject. In one embodiment the T cells are autologous and the TCRs are allogeneic. In one embodiment the TCRs are autologous and the T cells are allogeneic. In one embodiment the TCRs are autologous and the T cells are autologous. Preferably, the cells are autologous to the subject.

[00183] The terms "treat," and "prevent" as well as words stemming therefrom, as used herein, do not necessarily imply 100% or complete treatment or prevention. Rather, there are varying degrees of treatment or prevention of which one of ordinary skill in the art recognizes as having a potential benefit or therapeutic effect. In this respect, the inventive methods can provide any amount or any level of treatment or prevention of cancer in a mammal.

[00184] Furthermore, the treatment or prevention provided by the inventive method can include treatment or prevention of one or more conditions or symptoms of the disease, e.g., cancer, being treated or prevented. Also, for purposes herein, "prevention" can encompass delaying the onset of the disease, or a symptom or condition thereof.

[00185] With respect to the inventive methods, the cancer can be any cancer, including any of sarcomas (e.g., synovial sarcoma, osteogenic sarcoma, leiomyosarcoma uteri, and alveolar rhabdomyosarcoma), lymphomas (e.g., Hodgkin lymphoma and non-Hodgkin lymphoma), hepatocellular carcinoma, glioma, head-neck cancer, acute lymphocytic cancer, acute myeloid
leukemia, bone cancer, brain cancer, breast cancer, cancer of the anus, anal canal, or anorectum, cancer of the eye, cancer of the intrahepatic bile duct, cancer of the joints, cancer of the neck, gallbladder, or pleura, cancer of the nose, nasal cavity, or middle ear, cancer of the oral cavity, cancer of the vulva, chronic lymphocytic leukemia, chronic myeloid cancer, colon cancer (e.g., colon carcinoma), esophageal cancer, cervical cancer, gastrointestinal cancer (e.g., gastrointestinal carcinoid tumor), hypopharynx cancer, larynx cancer, liver cancer, lung cancer, malignant mesothelioma, melanoma, multiple myeloma, nasopharynx cancer, ovarian cancer, pancreatic cancer, peritoneum, omentum, and mesentery cancer, pharynx cancer, prostate cancer, rectal cancer, renal cancer, small intestine cancer, soft tissue cancer, stomach cancer, testicular cancer, thyroid cancer, uterine cancer, and urinary bladder cancer.

The method may comprise combining the cell population of tumor-reactive T cells expressing subject specific TCRs with a pharmaceutically acceptable carrier to obtain a pharmaceutical composition comprising a personalized cell population of tumor-reactive T cells. Preferably, the carrier is a pharmaceutically acceptable carrier. With respect to pharmaceutical compositions, the carrier can be any of those conventionally used for the administration of cells. Such pharmaceutically acceptable carriers are well-known to those skilled in the art and are readily available to the public. It is preferred that the pharmaceutically acceptable carrier be one which has no detrimental side effects or toxicity under the conditions of use. A suitable pharmaceutically acceptable carrier for the cells for injection may include any isotonic carrier such as, for example, normal saline (about 0.90% w/v of NaCl in water, about 300 mOsm/L NaCl in water, or about 9.0 g NaCl per liter of water), NORMOSOL R electrolyte solution (Abbott, Chicago, IL), PLASMA-LYTE A (Baxter, Deerfield, IL), about 5% dextrose in water, or Ringer's lactate. In an embodiment, the pharmaceutically acceptable carrier is supplemented with human serum albumen.

The T cells can be administered by any suitable route as known in the art. Preferably, the T cells are administered as an intra-arterial or intravenous infusion, which preferably lasts approximately 30-60 min. Other examples of routes of administration include intraperitoneal, intrathecal and intralymphatic. T cells may also be administered by injection. T cells may be introduced at the site of the tumor.

For purposes of the invention, the dose, e.g., number of cells in the inventive cell population expressing subject specific TCRs, administered should be sufficient to effect, e.g., a
therapeutic or prophylactic response, in the subject over a reasonable time frame. For example, the number of cells should be sufficient to bind to a cancer antigen, or detect, treat or prevent cancer in a period of from about 2 hours or longer, e.g., 12 to 24 or more hours, from the time of administration. In certain embodiments, the time period could be even longer. The number of cells will be determined by, e.g., the efficacy of the particular cells and the condition of the subject (e.g., human), as well as the body weight of the subject (e.g., human) to be treated.

[00189] Many assays for determining an administered number of cells from the inventive cell population expressing subject specific TCRs are known in the art. For purposes of the invention, an assay, which comprises comparing the extent to which target cells are lysed or one or more cytokines such as, e.g., IFN-γ and IL-2 are secreted upon administration of a given number of such cells to a subject, could be used to determine a starting number to be administered to a mammal. The extent to which target cells are lysed, or cytokines such as, e.g., IFN-γ and IL-2 are secreted, upon administration of a certain number of cells, can be assayed by methods known in the art. Secretion of cytokines such as, e.g., IL-2, may also provide an indication of the quality (e.g., phenotype and/or effectiveness) of a cell preparation.

[00190] The number of the cells administered from the inventive cell population expressing subject specific TCRs may also be determined by the existence, nature and extent of any adverse side effects that might accompany the administration of a particular cell population.

[00191] Typically, the attending physician will decide the number of the cells with which to treat each individual patient, taking into consideration a variety of factors, such as age, body weight, general health, diet, sex, route of administration, and the severity of the condition being treated. By way of example and not intending to limit the invention, the number of cells can be about $10 \times 10^6$ to about $10 \times 10^9$ cells per infusion, about $10 \times 10^6$ cells to about $10 \times 10^{10}$ cells per infusion, or $10 \times 10^7$ to about $10 \times 10^9$ cells per infusion. The cell populations obtained by the inventive methods may, advantageously, make it possible to effectively treat or prevent cancer. Likewise, any suitable dose of T cells can be administered. Preferably, from about $2.3 \times 10^9$ T cells to about $13.7 \times 10^{10}$ T cells are administered, with an average of around $7.8 \times 10^{10}$ T cells, particularly if the cancer is melanoma. With respect to the alternative method, preferably, from about $1.2 \times 10^9$ to about $4.3 \times 10^{10}$ T cells are administered.

[00192] An embodiment of the invention further comprises lymphodepleting the subject prior to administering any of the T cells obtained by any of the inventive methods described herein.
Examples of lymphodepletion include, but may not be limited to, nonmyeloablative lymphodepleting chemotherapy, myeloablative lymphodepleting chemotherapy, total body irradiation, etc.

[00193] In another embodiment, the T cell therapy described herein provides selecting the appropriate point to administer the therapy in relation to and within the standard of care for the cancer being treated for a subject in need thereof. The therapy described herein can be effectively administered even within the standard of care that includes surgery, radiation, or chemotherapy. The standards of care for the most common cancers can be found on the website of National Cancer Institute (http://www.cancer.gov/cancertopics). The standard of care is the current treatment that is accepted by medical experts as a proper treatment for a certain type of disease and that is widely used by healthcare professionals. Standard of care is also called best practice, standard medical care, and standard therapy. Standards of Care for cancer generally include surgery, lymph node removal, radiation, chemotherapy, targeted therapies, antibodies targeting the tumor, and immunotherapy. Immunotherapy can include checkpoint blockers (CBP), chimeric antigen receptors (CARs), and adoptive T cell therapy. The therapy described herein can be incorporated within the standard of care. The therapy described herein may also be administered where the standard of care has changed due to advances in medicine.

[00194] Incorporation of the T cell therapy described herein may depend on a treatment step in the standard of care that causes the immune system to be suppressed. Such treatment steps may include irradiation, high doses of alkylating agents and/or methotrexate, steroids such as glucocorticoids, surgery, such as to remove the lymph nodes, imatinib mesylate, high doses of TNF, and taxanes (Zitvogel et al., 2008). The therapy may be administered before such steps or may be administered after.

[00195] In one embodiment the T cell therapy may be administered after bone marrow transplants and peripheral blood stem cell transplantation. Bone marrow transplantation and peripheral blood stem cell transplantation are procedures that restore stem cells that were destroyed by high doses of chemotherapy and/or radiation therapy. After being treated with high-dose anticancer drugs and/or radiation, the patient receives harvested stem cells, which travel to the bone marrow and begin to produce new blood cells. A "mini-transplant" uses lower, less toxic doses of chemotherapy and/or radiation to prepare the patient for transplant. A "tandem transplant" involves two sequential courses of high-dose chemotherapy and stem cell
transplant. In autologous transplants, patients receive their own stem cells. In syngeneic transplants, patients receive stem cells from their identical twin. In allogeneic transplants, patients receive stem cells from their brother, sister, or parent. A person who is not related to the patient (an unrelated donor) also may be used. In some types of leukemia, the graft-versus-tumor (GVT) effect that occurs after allogeneic BMT and PBSCT is crucial to the effectiveness of the treatment. GVT occurs when white blood cells from the donor (the graft) identify the cancer cells that remain in the patient's body after the chemotherapy and/or radiation therapy (the tumor) as foreign and attack them. Immunotherapy with the T cell therapy described herein can take advantage of this by introducing T cells expressing subject tumor specific TCRs after a transplant.

[00196] In another embodiment T cells expressing subject specific TCRs are administered to a subject in need thereof that has not received a treatment resulting in immunoablation. In one embodiment T cells expressing subject specific TCRs are administered after surgery to remove a tumor. Not being bound by a theory, for tumors that are localized and where the standard of care is surgical removal followed by adjuvant therapy to remove any tumor cells that are present as micrometastases, introducing T cells expressing subject specific TCRs can facilitate removal of any remaining tumor cells.

[00197] In another embodiment TCR expressing cells that were transplanted into a subject are monitored to quantify each T cell clone at various time points following transplant, whole tumor cell vaccination, or immunization with a neoantigen immunogenic composition. Monitoring may be by PCR of nucleic acids isolated from subjects during treatment. The PCR may use TCR specific primers. The TCR primers may hybridize to constant regions of each TCR chain and a unique region specific to a particular TCR chain.

[00198] In another embodiment identified subject specific TCRs are administered by a gene therapy method. TCRs may be cloned and expressed using a vector. The vector can be introduced into a subject. The vector may be a viral vector. The vector preferably targets T cells. The T cells preferably express the TCR. More preferably the T cell is activated. In a preferred embodiment the vector allows expression of the TCR and activation of the target T cell.

[00199] In one embodiment a lentivirus is used. Preferably the lentivirus is administered with about 10 µl of recombinant lentivirus having a titer of 1 x 10^9 transducing units (TU)/ml by an
intrathecal catheter. These sort of dosages can be adapted or extrapolated to use of a retroviral or lentiviral vector in the present invention.

[00200] In an embodiment herein the delivery is via an adenovirus, which may be at a single dose containing at least 1 x 10^5 particles (also referred to as particle units, pu) of adenoviral vector. In an embodiment herein, the dose preferably is at least about 1 x 10^6 particles (for example, about 1 x 10^6-1 x 10^{12} particles), more preferably at least about 1 x 10^7 particles, more preferably at least about 1 x 10^8 particles (e.g., about 1 x 10^8-1 x 10^{11} particles or about 1 x 10^{8}-1 x 10^{12} particles), and most preferably at least about 1 x 10^9 particles (e.g., about 1 x 10^{9}-1 x 10^{10} particles or about 1 x 10^{9}-1 x 10^{12} particles), or even at least about 1 x 10^{10} particles (e.g., about 1 x 10^{10}-1 x 10^{12} particles) of the adenoviral vector. Alternatively, the dose comprises no more than about 1 x 10^{14} particles, preferably no more than about 1 x 10^{13} particles, even more preferably no more than about 1 x 10^{12} particles, even more preferably no more than about 1 x 10^{11} particles, and most preferably no more than about 1 x 10^{10} particles (e.g., no more than about 1 x 10^{9} articles). Thus, the dose may contain a single dose of adenoviral vector with, for example, about 1 x 10^6 particle units (pu), about 2 x 10^6 pu, about 4 x 10^6 pu, about 1 x 10^7 pu, about 2 x 10^7 pu, about 4 x 10^7 pu, about 1 x 10^8 pu, about 2 x 10^8 pu, about 4 x 10^8 pu, about 1 x 10^9 pu, about 2 x 10^9 pu, about 4 x 10^9 pu, about 1 x 10^{10} pu, about 2 x 10^{10} pu, about 4 x 10^{10} pu, about 1 x 10^{11} pu, about 2 x 10^{11} pu, about 4 x 10^{11} pu, about 1 x 10^{12} pu, about 2 x 10^{12} pu, or about 4 x 10^{12} pu of adenoviral vector. See, for example, the adenoviral vectors in U.S. Patent No. 8,454,972 B2 to Nabel, et. al., granted on June 4, 2013; incorporated by reference herein, and the dosages at col 29, lines 36-58 thereof. In an embodiment herein, the adenovirus is delivered via multiple doses.

[00201] In an embodiment herein, the delivery is via an AAV. A therapeutically effective dosage for in vivo delivery of the AAV to a human is believed to be in the range of from about 20 to about 50 ml of saline solution containing from about 1 x 10^{10} to about 1 x 10^{50} functional AAV/ml solution. The dosage may be adjusted to balance the therapeutic benefit against any side effects. In an embodiment herein, the AAV dose is generally in the range of concentrations of from about 1 x 10^5 to 1 x 10^{50} genomes AAV, from about 1 x 10^8 to 1 x 10^{20} genomes AAV, from about 1 x 10^{10} to about 1 x 10^{16} genomes, or about 1 x 10^{11} to about 1 x 10^{16} genomes AAV. A human dosage may be about 1 x 10^{13} genomes AAV. Such concentrations may be delivered in from about 0.001 ml to about 100 ml, about 0.05 to about 50 ml, or about 10 to
about 25 ml of a carrier solution. In a preferred embodiment, AAV is used with a titer of about 2 x 10^{13} viral genomes/milliliter, and each of the striatal hemispheres of a mouse receives one 500 nanoliter injection. Other effective dosages can be readily established by one of ordinary skill in the art through routine trials establishing dose response curves. See, for example, U.S. Patent No. 8,404,658 B2 to Hajjar, et al., granted on March 26, 2013, at col. 27, lines 45-60.


The practice of the present invention employs, unless otherwise indicated, conventional techniques for generation of genetically modified mice. See Marten H. Hofker and Jan van Deursen, TRANSGENIC MOUSE METHODS AND PROTOCOLS, 2nd edition (2011).

The main advantage of the methods described herein are that TCRs that are subject specific and tumor specific are identified and these TCRs are matched to subject specific neoantigens. The identification allows the formulation of a subject specific immunotherapy that can be coordinated with an immunotherapy that includes a neoantigen immunogenic composition or can help guide the proper neoantigen immunogenic composition.

Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined in the appended claims.

The present invention will be further illustrated in the following Examples which are given for illustration purposes only and are not intended to limit the invention in any way.

Examples

Example 1

Combined TCRα/TCRβ chains within single cells are responsible for T cell specificity.
The highly polymorphic TCR is generated by joining of non-contiguous gene segments (\( \nu \beta, D \beta \beta \) for TCR\( \beta \) and Va, Ja for TCRa) together with deletion/insertion of random sequences to form the CDR3 regions. Although there is a theoretical possibility of forming as many as \( 5 \times 10^{11} \) unique TCR\( \beta \) chains, the actual number of unique TCR\( \beta \) genes found in humans is closer to 0.1% of this estimate (Brahmer et al., N Engl J Med. 2012;366:2455-2465). The recognition of MHC-bound peptide by the combined TCR\( \beta \) and TCRa proteins occurs primarily by the CDR3 regions (Fig. 1 and 2; Topalian et al., N Engl J Med. 2012;366:2443-2454; and Wolchok et al., N Engl J Med. 2013;369:122-133). During thymic education (positive/negative selection) and antigen exposure (e.g. pathogens), specific T cell clones are selected and alter the actual repertoire. Since antigen-specific T cells with high-avidity TCRs (e.g. against tumor neoantigens) are rare relative to the large repertoire of T cells, it remains critical to develop a method to comprehensively characterize the TCR repertoire and thus enable discovery of TCRs and their target tumors antigens.

Example 2

Isolation and sequencing of T cell receptors from T cells associated with tumor

Applicants generate RNA-sequencing libraries from single T cells from a patient sample. Applicants can dissociate cells and generate single cell RNA-sequencing libraries from all T cells in a biopsy. While the synthesis of genome-wide RNA-seq libraries uses oligo-dT primers, this may not always amplify the TCR genes. Since Applicants need to guarantee amplification and sequencing of the TCR genes, and are not concerned with quantifying TCR transcript numbers, Applicants add to the reaction mixture a pool of 38 Va and 36 \( \nu \beta \) primers to amplify the 45 and 48 TCRa and TCR\( \beta \) variable regions, respectively, together with primers in the constant region of TCRa and TCRP (Fig. 3). Amplification productions are sequenced in parallel and deconvoluted based on barcodes that tag transcripts from each cell. Applicants analyze the resulting sequences to identify the Va and \( \nu \beta \) regions, including the CDR3, of each cell. If Applicants sequence -1000 T cells per biopsy, Applicants expect to identify clones with frequencies of 0.3% at 90% power (and lower when isolating and sequencing higher numbers of T cells).

Single cells are sequenced by any method described herein. The TCR single cell sequencing data is aligned to a database to get dominant clonotypes per cell (Fig. 4).
Applicants use an approach for amplifying and sequencing paired CDR3-Vα and -β chain sequences from single T cells using a ‘multiprimer approach’: a modified version of a recently described method, in which 36 T cell receptor TCRα V primers and 36 TCRβ V primers were used to perform multiplexed amplification of TCR gene sequences (Fig. 3; Han et al., Nature Biotechnology, Vol 32:7, 2014). Primers of reaction 1 and 2 for TRAV and TRBV are shown (Table 1 and Table 2).

Table 1

<table>
<thead>
<tr>
<th>Oligo Name</th>
<th>Reaction 1</th>
<th>Reaction 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAV1</td>
<td>CTCGACGTACCAGACAGACATCCTGGGTTCAG</td>
<td>AGGATGTGTCTTCTCATCCTTACGTCAC</td>
</tr>
<tr>
<td>TRAV2</td>
<td>GGTCTAAAAGACCTCCTCAGCGGTG</td>
<td>AGCTAACAATGACCTATGAACGG</td>
</tr>
<tr>
<td>TRAV3</td>
<td>GGATAACCTGGTTAAAGCACGCTA</td>
<td>CTTTGAGCTGATTTAACAAGAGCC</td>
</tr>
<tr>
<td>TRAV4</td>
<td>GGTACAAAGACAGAAAAGTACACACGAG</td>
<td>CCTCCCTTTTCATCCCTGCAGAC</td>
</tr>
<tr>
<td>Ti1V5</td>
<td>GCTGACGTATATTCTTCAATATGGA</td>
<td>AAACAGACACAAAGACTCAGTCAGAC</td>
</tr>
<tr>
<td>Ti3V6</td>
<td>GGATACCCAGCCTTCTCTGAGAACGT</td>
<td>GATCACAAGGGTTTGGCAGATTAA</td>
</tr>
<tr>
<td>TRAV6</td>
<td>GGAAGAGGCCCTGGTTTCTTGCT</td>
<td>AAGACTGAAAGTCCCTTGTACAC</td>
</tr>
<tr>
<td>TRAV7</td>
<td>GGTGGATATGAGAAGCAAGAAGGGA</td>
<td>ACTAAATGCTCATTACATGAAAGTC</td>
</tr>
<tr>
<td>TRAV8</td>
<td>AGGACTCCTCACGCTTCTCTGAGAACGT</td>
<td>GATCACAAGGGTTTGGCAGATTAA</td>
</tr>
<tr>
<td>Ti4V7</td>
<td>GATGTCTGATATTCTTCAATATGGA</td>
<td>AAACAGACACAAAGACTCAGTCAGAC</td>
</tr>
<tr>
<td>TRAV9</td>
<td>CAGTGAGAACACAAAGTCGAACGG</td>
<td>TACAGCAACTCAGTTGACAGACAC</td>
</tr>
<tr>
<td>TRAV12-1</td>
<td>CCTAAGTGCTGCTGATCCTGATAC</td>
<td>GAAGATGGAAGGTTTACAGCACA</td>
</tr>
<tr>
<td>TRAV12-2</td>
<td>GGGAAAACCCCTCAGTTGATATGT</td>
<td>GAAGATGGAAGGTTTACAGCACA</td>
</tr>
<tr>
<td>TRAV12-3</td>
<td>GGTGTACGATACATACACTCAGTGG</td>
<td>GAAGATGGAAGGTTTACAGCACA</td>
</tr>
<tr>
<td>TRAV13-1</td>
<td>CTTCTGCTATGAGAAGCAAGGGA</td>
<td>ACTACATCAGTCATTAAAGGGA</td>
</tr>
<tr>
<td>TRAV13-2</td>
<td>CTTCAATTCTAAAATAGACATTCTGC</td>
<td>GGCAAGGCAAAGAGTCACGCT</td>
</tr>
<tr>
<td>TRAV14</td>
<td>GCAATAATGCAAACAGAGTGCCCTA</td>
<td>TCCAGAAAGCGAAAATCCGCA</td>
</tr>
<tr>
<td>TRAV16</td>
<td>TAGAGAGAGCTCAGCTAGTGG</td>
<td>GTCGACCTTCTAACAAGGCGAGACA</td>
</tr>
<tr>
<td>TRAV17</td>
<td>GTCTAAAATGTTAAAGAGAAGTGGA</td>
<td>TAAAGAGCAGCTCTGACACTCTCA</td>
</tr>
<tr>
<td>TRAV18</td>
<td>CTTCAATTCTCAAATGAAAGGGAACAG</td>
<td>TTAAGAGCAGCTCTGACACTCTCA</td>
</tr>
<tr>
<td>TRAV19</td>
<td>GGTGCGGTATTTCTTGGAACACTCAG</td>
<td>GCCACAGCCAAGAGAGTCACGCT</td>
</tr>
<tr>
<td>TRAV20</td>
<td>GCTGCCCCGTTTACTTACTTTACAG</td>
<td>GCCACATTAACAAAAGAG6AAAGC</td>
</tr>
<tr>
<td>TRAV21</td>
<td>GTCAGAGAGAGCTAAAGAGGGAAG</td>
<td>GCTGAGCTGATATAATCAGA</td>
</tr>
<tr>
<td>TRAV22</td>
<td>GACAAACAGAAATGAAAGTGGA</td>
<td>ACTACAGTCATGCAGACAGC</td>
</tr>
<tr>
<td>TRAV23</td>
<td>CCAGATGGATGAAAGAAAAGGAAG</td>
<td>CACATCTCCTCACAAGGATGC</td>
</tr>
<tr>
<td>TRAV24</td>
<td>GACTTAGAAATGAGAGTAAAGAGAAGAAG</td>
<td>AGAATAATGQCCACTCTATAC</td>
</tr>
<tr>
<td>TRAV25</td>
<td>GGAGAAGTGAAGAAGCAGAGAAGAG</td>
<td>GTTGGAGAAGCAAAAGAAGCAGCT</td>
</tr>
<tr>
<td>TRAV26-1</td>
<td>CCAATGGAATGGCTCTCTCTGAC</td>
<td>CAGAAGACAGAAAGGTGGAGAC</td>
</tr>
<tr>
<td>TRAV26-2</td>
<td>GCAATGGAACACAAAGATGCGCT</td>
<td>ATGCTGAGACAGAAAGTGGAGAC</td>
</tr>
<tr>
<td>TRAV27</td>
<td>GGTGGAGAAGTGAAGAAGCTGAAG</td>
<td>ACTACAGCCAAGAGAGTGGAGAC</td>
</tr>
<tr>
<td>oligo name</td>
<td>reaction $^1$</td>
<td>reaction $^2$</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>TRBV2</td>
<td>CTGAAATTTCACTGATGATCAATTCTCAC</td>
<td>GCTGTAGTGATCAAATCTCTCG</td>
</tr>
<tr>
<td>TRBV3-1</td>
<td>TCATTATAAGGAAAGATGCCAATTCTGAG</td>
<td>TCTCAGCTAAATCTCCAGGAACAGCT</td>
</tr>
<tr>
<td>TRBV4</td>
<td>AGTGGCCAAGTCGCTCTCTAC</td>
<td>CCTGAAATGCCCCACAGCTCTC</td>
</tr>
<tr>
<td>TRBV5-1</td>
<td>GAGACACAGAGAAAAAAAGGAAACTTAC</td>
<td>CGATTCTCAGGCGCTAGTTCTC</td>
</tr>
<tr>
<td>TRBV5-4</td>
<td>CAGAGAACACTCCCTCTCTAGATTT</td>
<td>CTGTGAGCTGAAATGTGAAAGC</td>
</tr>
<tr>
<td>TRBV5-8</td>
<td>CAGGGAACACTCCTCTCTAGATT</td>
<td>CTTGTAGTGCATAATGTCTCCAGA</td>
</tr>
<tr>
<td>TRBV6-1</td>
<td>GGTACCTCAGACAAAGGAGAAGCTTCC</td>
<td>TGCTCAATGCTCTCCAGATTAAACAA</td>
</tr>
<tr>
<td>TRBV6-2</td>
<td>GAGGATCAACTGCCAAGAGGAGGAT</td>
<td>GCTGCCAATGACAAACACAGAGGAT</td>
</tr>
<tr>
<td>TRBV6-3</td>
<td>GAGGATCAACTGCCAAGAGGAGGAT</td>
<td>GCTGCCAATGACAAACACAGAGGAT</td>
</tr>
<tr>
<td>TRBV6-4</td>
<td>GGAAGAGGAAAGGAGAAGCTGATG</td>
<td>GGTCTCCCAGAACAACAGTATT</td>
</tr>
<tr>
<td>TRBV6-5</td>
<td>AAGGGAAGAAGTTCCAAATGCTACAA</td>
<td>GCTCCAGAGTAAACACAGAGGAG</td>
</tr>
<tr>
<td>TRBV6-6</td>
<td>AAGGGAAGAAGTTCCAAATGCTACAA</td>
<td>GCTCCAGAGTAAACACAGAGGAG</td>
</tr>
<tr>
<td>TRBV6-8</td>
<td>CTAAGAAAGAAAGTCCCAAACTGCTAC</td>
<td>GCTCTAGATTTAAAAACACAGAGGATTC</td>
</tr>
<tr>
<td>TRBV6-9</td>
<td>CACTGCAAGAAAGAAGATCCCAACAGG</td>
<td>GCTCAGAAATGTTCAGAATCAGAA</td>
</tr>
<tr>
<td>TRBV7-2</td>
<td>AGACAAATACAGGGCCTGCAAGGACTG</td>
<td>TGCTCTTCTGAGAGAGGACTG</td>
</tr>
<tr>
<td>TRBV7-3</td>
<td>GACTCGACGGCTGCTAGGAGG</td>
<td>CGGTCTTTGTGACTGAGCCAGTC</td>
</tr>
<tr>
<td>TRBV7-4</td>
<td>GTTTCTGCTAGAGAGAGGCTAGG</td>
<td>TCTCCAGCTGAMGATCCAGGCA</td>
</tr>
<tr>
<td>TRBV7-6</td>
<td>GTTTCTGCTAGAGAGAGGCTAGG</td>
<td>TCTCCAGCTGAMGATCCAGGCA</td>
</tr>
<tr>
<td>TRBV7-7</td>
<td>GCAGAGCTAGGTGTCTTCGTGCTC</td>
<td>GCAGAGAGGCTGACAGATCAT</td>
</tr>
<tr>
<td>TRBV7-8</td>
<td>CCAGAATGAAGCTCAACTAGACAA</td>
<td>CCAGTGATCGCTTCTTTGCAAAA</td>
</tr>
<tr>
<td>TRBV7-9</td>
<td>GCATTCTCCAGAATAAGCTTACACT</td>
<td>CTGCAGAGAGGCTAAGGATCT</td>
</tr>
<tr>
<td>TRBV9</td>
<td>GAGCAGAAGAAACATTTCTGAAAGGATT</td>
<td>CTCCGCAACACAGTTCCCTGAGCT</td>
</tr>
<tr>
<td>TRBV10-1</td>
<td>GGCAGTATCATTCACTATAAAGTGGTGTG</td>
<td>CAGATGGCTGAATGCTCTAGGATCAAA</td>
</tr>
<tr>
<td>TRBV10-2</td>
<td>GATAAAGAAGGAGATCCCAAGTGGCTGCTC</td>
<td>GGTGCTCCAGATCAAAGACAGAAG</td>
</tr>
<tr>
<td>TRBV10-3</td>
<td>GGCAGTATCATTCACTATAAAGTGGTGTG</td>
<td>CAGATGGCTGAATGCTCTAGGATCAAA</td>
</tr>
<tr>
<td>TRBV11</td>
<td>GATCCACAGTGGCTTAAAGGATGCTGAT</td>
<td>GCAGAGAGGCTAAAGGAGTAGCT</td>
</tr>
</tbody>
</table>
Example 3

Cloning of full TRCa and TCRβ into expression vector

In order to probe the functional activity of TCR sequences obtained from single cell sequencing, Applicants clone and express any pair of TCRs for functional analysis. Because TCRs are generated through recombination of pre-fixed modules, Applicants have created a fixed plasmid library in which each plasmid contains one of -48 Vβ or -45 Va regions. In addition, each plasmid also contains the constant region of the other TCR, a leader sequence for surface expression, and BsaI restriction sites for seamless assembly into a complete TCR by Golden Gate assembly methods. In the assembled constructs a 2A-peptide sequence separates the TCRa and TCRβ chains, and is cleaved following translation to allow simultaneous expression of both genes at equal levels. Finally, the CDR3 sequences are synthesized as oligos for both TCRs and co-assembled (Fig. 5,6). As describe herein, Applicants have generated a synthetic library of all V, D, and J components, so that any TCR plasmid construct can be readily generated through multi-component DNA assembly (e.g., Golden Gate assembly) to include discovered CDR3 Va and -Vβ sequences. The sequences for segments in the pre-made library are shown in the sequence listing. The pre-made library consists of the TRAV sequences inserted in "master pUC57-Kan Cbl-F2A" and "master pUC57-Kan Cb2-F2A", and the TRBV sequences...
inserted in "master pUC57-Kan Ca short" or "master pUC57-Kan Ca." TRBV sequences in "master pUC57-Kan Ca" are: TRBV5-6, 5-8, 6-8, 6-9, 7-4, 7-8, 10-2, 11-1, 11-2, 11-3, 16, 19, 29-1. The rest are ALL in "master pUC57-Kan Ca short"

**Example 4**

**Expression of TCR in TCR-deleted T cells**

[00212] The plasmids containing both TCR sequences are transfected or infected (by producing lentiviruses) into a Jurkat Δαβ cell line (Fig. 7). Expression is verified by antibody surface staining for TCR. The Jurkat cells also express CD8 and Thy1.1 under control of an NFAT promoter to respond to TCR stimulation. If the patient tumor expresses MHC Class II, Applicants also use Jurkat cells expressing CD4.

**Example 5**

**A library of candidate antigens to screen for TCR reactivity**

[00213] Candidate sets of neoantigens are selected from 2 groups: (i) predicted neoantigens from each patient's tumor using our MHC Class I prediction pipeline and also MHC class II binding prediction if a tumor expresses MHC class II; (ii) neoantigen peptides identified on tumors by mass spectrometry, some of which may not be predicted by predictive algorithms. In addition, Applicants test native tumor antigen peptides observed by mass spectrometry, and published tumor antigen peptides corresponding to patient-specific HLAs. Applicants then test hundreds of neoantigens and other antigens for identifying cognate TCRs that bind these neoantigens bound to autologous HLA proteins. Applicants create a library of antigens as an expression library in lentiviral vectors, and infect the library into immortalized B cells generated from the blood of the same patient (for candidate class II antigens, Applicants also insert the antigens into a separate lentiviral vector for antigen fusion with ATG8/LC3 to enhancer antigen trafficking to the MHC Class II compartment Schmid, et al., Immunity. 2007 Jan;26(1):79-92. Epub 2006 Dec 21).

**Example 6**

**Screening TCRs for reactivity against neoantigens**
[00214] Applicants screen for TCR recognition of tumor antigens using the assembled library of TCR-expressing T cells and a library of antigen-expressing B cells. First, Applicants incubate all transgenic T cells with B cells expressing each of the neoantigens in separate wells and sort out reactive T cells from positive wells using the TCR-induced molecules, such as CD83 or CD69 on the surface of cells, or by a surface of fluorescent reporter downstream of NFTA binding sites (that is activated by TCR stimulation). Applicants focus on CD8 T cells, but add CD4 T cells if MHC Class II is expressed on tumor cells. Second, Applicants then incubate reactive T cells with tumor cells to determine if T cells recognize antigen endogenously presented by tumor cells.

Example 7

Tracking tumor-specific TCRs in the blood

[00215] Using the subset of TCRs that detect tumors, Applicants use bulk TCR sequencing of longitudinal blood samples to monitor anti-tumor immunity. This allows for a correlation of changes in tumor-reactive clones over the short and long term with tumor growth and spread.

Example 8

Proof of concept with known TCR sequences

[00216] As proof of principle, Applicants have cloned the EBNA3A-specific TCR pair by Golden Gate assembly into a lentiviral vector (Fig. 8, 9 and 10). Expression of EBNA3A-specific TCRaP on JurkatAaP is shown by binding of TCR v β5-1 antibody, EBNA3A tetramer binding and IL-2 release. Additionally, the reporter T cells may be stimulated with antigen presenting cells loaded with EBNA3A peptide or control peptide.

Example 9

Proof of concept for isolation of TCR sequences specific for peptides from CMV, EBV and influenza (CEF)

[00217] Applicants have cloned a library of TCR pairs in T cells from patients expanded with peptides from CMV, EBV and influenza (CEF). A workflow is described in Figure 11. Expansion and isolation of CEF-specific T cells is performed by stimulation of healthy donor PBMCs with the CEF peptide pool, culturing for 10 days to expand CEF-specific T cells, performing an IFN-γ catch assay to isolate CEF-specific T cells, and FACS sorting IFN-γ+CD8+
T cells into 384 well plates (Fig. 12). T cells stimulated with CEF peptides have more dominant TCRα and TCRβ chains (Fig. 13). Comparison of the TCRα and TCRβ pairs between CEF treated and PBMCs shows less diversity and dominant TCR pairs (Fig. 14). Applicants also found a published TCR pair specific for the BMLF1 peptide that is included in the CEF peptide pool. Applicants use IFN-γ elispot to confirm the screening pipeline (Fig. 15). Stimulation with individual CEF peptides shows signal above background in T cells re-stimulated with BMLF1. Thus, Applicants have developed a method to isolate TCR sequences specific for antigens.

Example 10

Screening patient T cells

Applicants have developed a pipeline to screen patient T cells for T cell pairs with specificity against an antigen panel, such as a neoantigen panel (Fig. 16). T cell pairs are sequenced, cloned and assayed against the panel. The TCR pairs may then be used in a therapeutic such as adoptive transfer of transgenic T cells. Additionally, neoantigen peptides reactive to T cells may be used in a neoantigen immunogenic composition. Applicants further developed another pipeline to isolate TCR pairs reactive to a neoantigen peptide pool (Fig. 17). Applicants stimulate patient PBMCs with a neoantigen peptide pool (or transfection with plasmids or infection with viruses to express neoantigens), followed by IFN-γ catch (or other method to capture activated T cells) to isolate neoantigen-specific T cells. The paired TCRαPs are sequenced in single cells, and screened against the individual neoantigens in a reporter cell line. The isolated TCR pairs that are shown to recognize patient tumor neoantigens may then be used as a therapeutic such as adoptive transfer of transgenic T cells.

Example 11

Multi-epitope personal vaccines for CLL is an effective therapeutic strategy

The curative basis of allogeneic hematopoietic stem cell transplantation (HSCT) relies on the immunologic recognition and elimination of malignant cells by normal donor cells that have reconstituted the hematopoietic system of the host (called the graft-versus-leukemia effect [GvL]) (Smith et al., Nat Protoc. 2009;4:372-384; Wu et al., Nat Methods. 2014;1 1:41-46). Strong evidence in support of the potency of GvL comes from (1) durable remissions observed following donor lymphocyte infusion (DLI) by which lymphocytes from the original donor are infused in the absence of further chemotherapy or radiotherapy, and in which
remissions have been observed in 80% of patients with chronic myeloid leukemia (CML) and 15-40% of patients with CLL (Wu et al., Nat Methods. 2014;1 1:41-46); and, (2) long-term remissions following reduced-intensity HSCT, where it is acknowledged that the intensity of the conditioning regimen alone is insufficient to generate durable leukemia control, and hence GvL is presumed to underly the clinically evident responses (DeKosky et al., Nat Biotechnol. 2013;31:166-169; Mazutis et al., Nat Protoc. 2013;8:870-891; Abate-Daga et al., PLoS One. 2014;9:e93321). Based on these concepts, Applicants devised and tested a strategy to enhance GvL by vaccinating 18 patients with advanced CLL with irradiated autologous whole-tumor cells, administered early following reduced-intensity HSCT (between days 30-45) (Burkhardt et al., J Clin Invest. 2013;123(9):3756-3765; Horowitz et al., Blood. 1990;75:555-562). In addition to promising clinical activity (an estimated 2-year progression-free and overall survival rate of 82% and 88% respectively), Applicants also observed CD8+ T cells from vaccinated patients, but not non-vaccinated patients, to consistently react against autologous tumor (with increased secretion of the effector cytokine IFNγ) but not to autologous alloantigen-bearing cells (Fig. 18A,B). Applicants further confirmed that approximately 15-30% of CD8+ T cell clones isolated from 4 vaccinated patients by limiting dilution solely reacted against CLL-associated antigens (Fig. 18C). These studies support the existence of CLL-specific antigens and that targeting of these antigens is associated with cancer control.

In ongoing studies, Applicants have analyzed whether perturbation by vaccination in the early post-transplant period impacts immune reconstitution. Applicants isolated naive and memory CD4+ and CD8+ T cells from peripheral blood mononuclear cells (PBMC) of 14 patients with advanced CLL who underwent a reduced-intensity allo-HSCT. Time points included post-transplant day 30, day 120 (a time point informative for thymic-independent T cell immune recovery) and day 365 (a time point informative for thymic-dependent immune recovery). From these T cell subpopulations, genomic DNA was extracted and a template library for sequencing on an Illumina GA2 system was generated through PCR amplification of the TCR-β CDR3 region using an established panel of 45 vβ- and 13 Jβ-specific primers (Adaptive Biotechnologies, Seattle, WA). Applicants obtained a median of 394,872 (range 0-26,426,784) productive reads across 168 samples. As a comparison group, Applicants further studied repertoire data from naive and memory CD4+ and CD8+ T cells collected from 9 healthy adult volunteers. Analyses revealed that CD8+ naive T cells exhibited greater TCR diversity, as
defined by a clonality measure, at post-transplant day 365 in vaccinated compared to non-vaccinated/non-graft-versus-host disease (GvHD) individuals (Fig. 19; Burkhardt et al. CMT presentation). These results suggest that immunologic intervention early following hematopoietic stem cell infusion can indeed alter the T cell repertoire diversity of patients.

Example 12

**Development of a tumor neoantigen discovery pipeline.**

Neoantigens are a highly valuable class of tumor-expressed antigens generated as a result of somatic mutation. To systematically define CLL neoantigens, Applicants developed a pipeline that incorporates existing Broad Institute bioinformatic algorithms that have been used and validated in large-scale cancer genome projects such as the TCGA to precisely identify the tens to thousands of protein-coding changes in the DNA of each tumor (Wu et al., Adv Immunol. 2006;90:133-173; Brown et al., Leukemia. 2013;27:362-369; and NetMHCpan, one of the top rated prediction algorithms for HLA binding). Using recently reported results of large-scale whole-exome sequencing (WES) of CLL (Wang et al, NEJM 2011; Landau et al, Cell 2013:29,30) Applicants used NetMHCpan to predict candidate CLL-specific peptides generated from missense mutations with the potential to bind personal class-I HLA proteins. Applicants predicted an average of -22 binding epitopes (IC50<500nM) per CLL patient and evaluated the binding affinity of more than 100 predicted peptides using a competitive class I binding assay (Burkhardt et al., J Clin Invest. 2013;123:3756-3765) to understand the boundaries of accurate predictivity.

Example 13

**Identification of the CLL neoantigen FNDC3B**

Applicants further established that this approach could identify neoantigens which were immunogenic in vivo. In one CLL patient who achieved long-term remission following HSCT/whole tumor cell vaccination, Applicants found CTLs that were reactive to a predicted neoantigen peptide (from mutated but not wild-type FNDC3B). These mutFNDC3B-specific T cells could be detected by neoantigen-specific tetramers (Fig. 20B) and were found to be cytolytic to autologous tumor (based on CD107a staining, Fig. 20C), were long-lived (>32 months) and Vβ1 restricted (Fig. 20D).
Example 14

Developing the pipeline and prioritization criteria for selecting neoepitopes to include in a personalized neoantigen-based vaccination study

[00223] Leveraging publicly-available TCGA WES, Applicants used the pipeline to predict missense mutation-generated neoepitopes across 13 cancer types, including CLL. Applicants predicted 10’s-1000’s of neoantigens per tumor, suggesting that neoantigens are frequent.

[00224] Applicants retrospectively applied NetMHCpan to 40 neoepitopes identified as spontaneously occurring targets of T cells isolated from cancer patients, most of whom had regressing or long-term stable disease. This analysis demonstrated that all neoepitopes naturally observed in cancer patients would have been predicted by NetMHCpan and thus establishing criteria for epitope prioritization based on bona fide human T cell responses (Fritsch et al, Cancer Immunol Res 2014, Jun; 2(6):522-9).

Example 15

Exemplary workflow to directly identify the paired a and β chains of the TCR complex in individual T cells following epitope recognition

[00225] An experimental workflow is shown to directly identify the paired a and β chains of the TCR complex in individual T cells following epitope recognition (Fig. 21). TCRα/TCRP pairs of individual tumor-reactive T cells and their reactivity against predicted personal neoepitopes is described. The experiments described in panel A shows the methodology to identify the known and novel TCRα/β pairs for two distinct target antigens (differing significantly in TCRs). Additionally, a T cell reporter cell line is generally applicable to the analysis of all isolated TCRα/β pairs to demonstrate antigen specificity. In panel B the tools in panel A are applied to identify the neoantigens associated with the tumor-specific T cell responses observed in the recently reported study of whole CLL cell vaccination following allotransplant (Rajasagi et al., Blood 2014, Jul 17; 124(3):453-62; Horowitz et al., Blood. 1990;75:555-562).

[00226] Well-characterized primers specific for the 3’ constant region and for each of the 5' Va and Vβ gene segments has been described (Robins et al., Sci Transl Med. 2010;2:47ra64). Validation of the primer set includes amplifying CDR3 using a primer in the constant region and specific primers from each Va and Vβ segment. The primers generate amplicons corresponding
to all of the Va and v β families. The products are sequenced on an Illumina MiSeq with 250nt paired end reads. This strategy can be compared to an alternative 5'-RACE approach that does not require 5' V primers but rather uses a template switching oligonucleotide for second strand synthesis; while the downside of this latter approach is reduced efficiency, it requires only a single primer for RT to capture each TCR.

P00227  PCR primers are validated by confirming the ability to correctly identify known TCRα/β pairs by performing multiplex amplification and sequencing of single T cells with known paired TCRs (i.e. Jurkat and HuT78 cell lines).

P00228  Panel A shows a schematic representation of single cell sequencing being used to sequence the TCRs from T cells that recognize a known viral antigen (influenza M1) and a single tumor neoantigen (mutated FNDC3B from CLL) (Fig. 21A).

P00229  To generate M1-specific T cells, PBMCs are collected from HLA-A2+ healthy adult volunteers after generating M1-reactive T cells by standard stimulation with the HLA-A2-restricted M1 (GILGFVFTL) peptide (as performed in Naito et al, Cancer Immunol Immunother 2013, Feb; 62(2):347-57; Sidney et al., Curr Protoc Immunol. 2013;Chapter 18:Unit 18 13). About 10,000 CD8+ M1-tetramer + cells may be isolated by flow sorting and analyzed by TCRα and TCRP amplification and sequencing. In addition, M1-responsive CD8+ T cells may be isolated based on IFNγ secretion (IFNy Catch Assay, Miltenyi) to compare the TCR repertoire with that identified by tetramer sorting.

P00230  To generate tumor mutated FND3CB neoepitope specific T cells, PBMCs from a well-characterized CLL patient are stimulated with the mutFNDC3B peptide. About 10,000 tetramer+ CD8+ T cells may be isolated (NIH Tetramer Core Facility at Emory; see Fig 4B) for TCRα and TCRP amplification and sequencing.

P00231  Reads are aligned with the IMGT database of germline TCR genes to reveal the identity and abundance of paired TCRα/β chains. One example of how the Illumina MiSeq paired-end reads are analyzed are:

(i) Each read is queried against the FMGT-curated library of V, D, and J gene sequences using IgBlast (a well-established algorithm to align TCR). Paired-end reads are kept if at least one read aligns to the V gene with high confidence (p<10^-6). This read end is assumed to map upstream in the pair.
(ii) Each (upstream) read is then superimposed upon its assigned germline V gene segment. In cases of mismatch, the sequenced base is preferred if the quality score is high (Phred>30). In addition, algorithms may be developed to utilize unique molecular identifiers to determine sequence accuracy.

(iii) The reconstructed V segments are tested for overlap with their corresponding paired end read. If the overlap (>10nts) fails, 1 base will be deleted from the 3’ end of the reconstructed V segment and retried (to account for chewback) up to 5 times before giving up. When successful overlap occurs, the reconstructed V segment and the paired read are fused to form a fully reconstructed sequence.

(iv) Each fully reconstructed sequence is then queried using IgBlast. Sequences with V-gene and J-gene assignments both scoring p<10^-6 are retained, and their IgBlast-identified CDR3 sequences are recorded as long as all CDR3 bases have a Phred score > 30.

(v) The most frequent CDR3 reconstruction is designated as a motif, and all reconstructions within a Hamming distance of 1 are assigned to that motif. This process is repeated on the remaining reconstructions until fewer than 2 sequences can be assigned to a motif.

(vi) Sequence and abundance of unique TCR sequences are reported based on CDR3 assignments.

Panel A also describes a TCR-deficient reporter cell for expression of identified antigen-specific TCRα/TCRβ pairs for functional analysis. The reporter is activated upon binding to antigen. The reactivity of the TCRs to cognate antigen are tested by introducing them into a TCR-deficient reporter cell that can be used to test for antigen responsiveness.

The identified TCRα/β paired chains are cloned and expressed in the reporter T cell line. An expression vector containing the distinct paired specific TCRα and β chains linked by a 2A-peptide encoding sequence (-60 bp); the 2A peptide is cleaved following translation and allows reliable simultaneous expression of both genes at equal levels (Boria et al., BMC Immunol. 2008;9:50). The -1.8 kb insert consists of fixed sequences (Co. and C β - the invariant portions of each TCR chain) as well as the linking region. Only the -400 bp variable regions require synthesis or PCR-based insertion to construct the complete vector. The expression cassette is cloned as a TCRα/β cassette into the lentiviral vector.
pRRLSIN.cPPT.PGK-GFP.WPRE (Addgene, Cambridge, MA; GFP can be exchanged with YFP in this vector), which has been shown to effectively express TCR genes (Naito et al., Cancer Immunol Immunother. 2013;62:347-357).

To study the identified TCRs, the paired TCR α and β chains are expressed in a TCR-deficient reporter cell line. The reporter cell line may be constructed by transducing the murine TCR-deficient thymoma line BW5 147.3 (ATCC TIB-47; BW) with multi-cistronic murine stem cell virus (MSCV)-based lentiviral constructs encoding the human CD3 complex and the human CD8 α and β chains. Both constructs contain an IRES and different antibiotic resistance markers. In addition, the reporter cell line is transduced with CFP controlled by the nuclear factor of activated T cells (NFAT-CFP reporter) similar to studies from others (Szymczak et al., Nat Biotechnol. 2004;22:589-594; Jones et al., Hum Gene Ther. 2009;20:630-640.). An advantage of using a murine cell line is that human TCRs positively selected for personal HLAs do not recognize murine MHC-molecules as allogeneic antigens. Furthermore, BW cells are functionally competent and secrete IL-2 upon antigen stimulation (Fig. 22). The lentiviral TCR constructs are introduced to create engineered BW T cells.

The expression of CD3 and CD8 is verified by flow cytometry with specific antibodies. The functionality of the NFAT-CFP reporter is confirmed by stimulating the parent reporter cells with PMA/ionomycin and monitoring CFP expression (Fig. 22).

Described is a method to screen the reactivity of engineered T cells against cognate antigens. To determine whether the TCRs isolated from tetramer+ T cells are reactive to M1 and mutFNDC3B peptides, BW cells transduced with specific TCRs are studied in two ways. First, flow cytometry is used to test for expression of antigen-specific TCR using fluorescent M1- or mutFNDC3B-specific tetramers. Second, CFP expression via NFAT signaling is monitored by flow cytometry after exposure of cells to cognate antigen on HLA-A2+ expressing antigen presenting cells (i.e. K562-A2+ cells, which have been previously used, Naito et al., 2013) (Rajasagi et al., Blood 2014, Jul 17; 124(3):453-62; Sidney et al., Curr Protoc Immunol. 2013;Chapter 18:Unit 18 13). The antigen presenting cells are pulsed with M1 (GILGFVFTL) or mutFNDC3B (VVMSWAPPV) peptides or control peptides (i.e. the HTLV-1-derived Tax peptide LLFGYPVYV), or transfected with an expression plasmid encoding minigenes for M1 or mutFNDC3B. Several TCR sequences recognizing each antigen are expected to be identified, and further studies are performed to address which TCRs are able to induce cytolysis on
influenza-infected target cells (for M1-specific T cells) or tumor cells (mutFNDC3B-specific T cells).

Results with mutFNDC3B-tetramer+ T cells from a CLL patient showed that approximately half of neoantigen-reactive T cells appeared to be cytolytic to autologous tumor on the basis of surface expression of the standard marker CD107a, while the other half were negative for CD107a (Fig. 20C). To determine whether the TCRs with these different phenotypes are similar or different, CD107a-positive and -negative cells are sorted, and the sequences of paired TCR chains in each population are compared. These TCRs can then be directly tested to determine whether TCRs found in CD107a+ T cells are cytolytic for mutated FNDC3B-pulsed HLA-A2+ targets. While the murine thymoma line BW5 147.3 is well-suited for initial TCR screening experiments, it lacks the cellular machinery for cytotoxicity studies. Primary peripheral blood lymphocytes (PBLs) are used for transduction with the paired TCRα/β chains (Cibulskis et al., Nat Biotechnol. 2013;31:213-219) and are assessed for cytotoxic activity of TCR-transduced PBLs on pulsed targets with a standard chromium release assay.

The described methods are used to deconvolute complex T cell populations that target tumors following whole tumor cell vaccination. Described in Fig. 21B are methods to determine if CTL reactivity following tumor cell vaccination is directed against CLL neoantigens. The experimental framework in Fig. 21A may be used to identify TCRs and their target neoantigens in patients who have received autologous tumor cell vaccination.

Isolating tumor-reactive T cells for which the identity of the target antigens are unknown can be performed. Such patient tumor-specific T cells are enriched following stimulation with autologous tumor ex vivo (see Burkhardt et al, JCI 2013; Horowitz et al., Blood. 1990;75:555-562). Based on limiting dilution experiments and patterns of reactivity of these T cell clones against a panel of target cells (Fig. 18C), it is estimated that 15-30% of clones are tumor-restricted. Thus, to further enrich for tumor-specific rather than alloantigen-specific T cells (resulting from transplantation), Applicants re-stimulate against three sets of targets: autologous tumor, recipient PHA blasts or recipient skin fibroblasts. For each stimulus, IFNγ+ cells can be isolated and processed as single cells for TCR profiling. Based on the sequences of paired TCRα/β chains in each group of T cells, TCRs that are reactive to tumor but not to non-malignant recipient cells (PHA blasts or skin cells) are identified.
Droplet sequencing and cloning of tumor-reactive TCRs is performed. Based on bulk IFN\(\gamma\) ELISpot data, if 1 million CD8+ T cells are processed, it can be expected that an average of 4000 IFN\(\gamma\) secreting cells are isolated. If 20\% of these cells are tumor-reactive, then sequencing is expected to reveal fewer than 800 unique paired TCRa/\(\beta\) chains. 96 of the most abundant and unique paired TCRa/\(\beta\) chain sequences are cloned into a lentiviral expression vector for stable expression in BW T cells, using the same clone strategy described herein.


Tumor-reactive TCRs can be screened against candidate neoantigens. To screen tumor-reactive TCRs for neoantigen specificity, the array of 96 distinct TCR-expressing BW cells are stimulated with irradiated autologous EBV-transformed B cells (APCs) pulsed with a mixture of 8-10mer predicted personal neopeptides. Neoantigen-reactive TCR-transduced BW cells (CFP+) are selected for further testing against individual neopeptides to match TCRs to neoeptitopes (deconvolution).

The described methods can be used to test if the TCRs recognizing personal tumor neoeptitopes lead to cytolysis of autologous CLL cells. To test for cytotoxicity against the tumor PBLs can be transduced with paired TCRa/\(\beta\) chains and the specificity of the cytotoxic activity can be assessed against: i) neoantigen- or unmutated native-antigen-pulsed HLA-matched targets; ii) autologous CLL cells; iii) HLA-matched CLL cells pulsed with native variants of neoantigens.

The described methods can be used to determine whether the tumor-specific TCR repertoire is similar in the marrow and periphery. Prior evidence leads to the hypothesis that tumor infiltrating T cells have higher avidity TCRs that are more reflective of an effective vaccination than peripheral blood T cells (Siewert et al., Nat Med. 2012;18:824-828). Since the bone marrow is a common site of persistent CLL disease as well as a reservoir for memory T cells (Ohashi et al., Nature. 1985;316:606-609; Melenhorst et al., Blood. 2009;1 13:2238-2244), the TCR repertoire can be monitored in matched peripheral blood and marrow samples.
The described methods can be used to track the kinetics of the tumor-specific TCR repertoire. Based on CDR3 sequence of neoantigen-reactive cytolytic TCRs, quantitative real-time PCR assays can be designed with a probe to the highly diverse V-D-J junction region of each TCRβ CDR3 sequence (as done previously Zhang & Choi, Clin Cancer Res 2010, May 15; 16(10):2729-39; Feuerer et al., Nat Med. 2003;9:1 151-1 157). These qPCR assays are used to quantify each T cell clone at various time points following transplant and whole tumor cell vaccination, as shown in the example of the v β1 1-restricted clone recognizing mutated FNDC3B (Fig. 20D). This allows a correlation of tumor regression with expansion of neoantigen-specific TCRs to be made.

Though the methods have focused on cytolytic CD8+ T cell responses (given the higher predictivity of MHC class I prediction algorithms and the more common cytolytic capabilities of CD8+ compared to CD4+ cells), there may be relevant CD4+ cells (cytolytic and helper). Therefore, a similar pipeline can be used to discover TCRs of CD4+ T cells and their target antigens.

Example 16

Figure 23 outlines a similar functional assay workflow. T cells are selected by FACS using tetramer analysis and IFNγ expression. Selected single cells are sequenced individually either by sub-cloning, sequencing cells sorted into separate wells of a plate or single cell microfluidic based methods. TCRs are cloned and inserted into lentiviral vectors. TCR deficient reporter T cells are transduced with lentiviruses expressing cloned TCRα/TCRP pairs. The reporter cells are then incubated with antigen presenting cells (B cells) that were pulsed with neoantigens.

Protocol for Cloning TCRs

Based on the sequences of the TCRα/TCRP pairs, the corresponding Va an Vb plasmids are selected from a pre-made library and assembled with the synthesized CDR3 sequence. The detailed protocol is shown below.

Materials

phusion high fidelity PCR master (NEB M053 1L)
NEB 5- alpha competent E coli (C2987H)
NEB golden gate assembly mix (E1600S)
Agel-HF (R3552L) 20,000U/ml
Sall-HF (R3138L) 20,000U/ml
T4 DNA ligase (M0202L) 400,000U/ml

**Annealing single strand ultramer oligos ordered from IDT**

Make STE buffer (10 mM Tris, 50 mM NaCl, 1 mM EDTA) by adding 2M NaCl (1.2g NaCl into 10ml dH20, filter with 0.45um, x40) into 1xTE buffer from Qiagen kit (1ml TE buffer+25ul 2M NaCl)

Spin down the oligo tubes

Dissolve the oligo in 40ul STE buffer to get lOuM for all 4 oligos

Vortex and spin

Mix two strands in equal molarities (lul of top +lul of bottom + 8ul STE) in 1.5ml eppendorf tube

94C, 3min on heat block

Cool gradually at RT

**Golden Gate Assembly Protocol using NEB Golden Gate Assembly Mix**

1. Set up assembly reactions as follows:

<table>
<thead>
<tr>
<th>REAGENT</th>
<th>NEGATIVE CONTROL</th>
<th>ASSEMBLY REACTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEB Golden Gate Buffer (10X)</td>
<td>2 µl</td>
<td>2 µl</td>
</tr>
<tr>
<td>Destination Plasmid* (user provided)</td>
<td>75 ng</td>
<td>75 ng</td>
</tr>
<tr>
<td>Inserts (user provided):</td>
<td>-</td>
<td>75–100 ng each plasmid</td>
</tr>
<tr>
<td>- if precloned**</td>
<td></td>
<td>2.1 molar ratio</td>
</tr>
<tr>
<td>- if in amplicon form***</td>
<td></td>
<td>(insert:vector backbone of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>destination plasmid)</td>
</tr>
</tbody>
</table>
2. *Destination plasmids must contain two Bsal restriction sites to define the sequence functioning as the vector backbone.*  
**Precloned inserts must possess Bsal restriction sites at both ends of the insert sequence.**  
***Amplicon inserts must possess 5' flanking bases and Bsal restriction sites at both ends of the amplicon.*

3. Choose the appropriate assembly protocol.

<table>
<thead>
<tr>
<th>INSERT NUMBER</th>
<th>SUGGESTED ASSEMBLY PROTOCOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>For 1-4 Inserts</td>
<td>37°C, 1 hr → 55°C, 5 min</td>
</tr>
</tbody>
</table>

### BMLF1

<table>
<thead>
<tr>
<th></th>
<th>Concentration (ng/µL)</th>
<th>Dilution first</th>
<th>Mass (ng)</th>
<th>Volume (µL)*</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAV5</td>
<td>2787</td>
<td>x100</td>
<td>75</td>
<td>2.7</td>
<td>937</td>
</tr>
<tr>
<td>TRBV29-1</td>
<td>2278</td>
<td>x100</td>
<td>75</td>
<td>3.3</td>
<td>3418</td>
</tr>
<tr>
<td>CDR3a</td>
<td>619.2</td>
<td>x1000</td>
<td>3.3</td>
<td>5.3</td>
<td>75</td>
</tr>
<tr>
<td>CDR3b</td>
<td>719.6</td>
<td>x1000</td>
<td>3.6</td>
<td>5</td>
<td>81</td>
</tr>
</tbody>
</table>

### EBNA3A

<table>
<thead>
<tr>
<th></th>
<th>Concentration (ng/µL)</th>
<th>Dilution first</th>
<th>Mass (ng)</th>
<th>Volume (µL)*</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAV8-1</td>
<td>197.8</td>
<td>x10</td>
<td>75</td>
<td>3.8</td>
<td>966</td>
</tr>
<tr>
<td>TRBV5-1</td>
<td>349.2</td>
<td>x10</td>
<td>75</td>
<td>2.1</td>
<td>4228</td>
</tr>
<tr>
<td>CDR3a</td>
<td>715</td>
<td>x1000</td>
<td>2.9</td>
<td>4.1</td>
<td>82</td>
</tr>
<tr>
<td>CDR3b</td>
<td>579.5</td>
<td>x1000</td>
<td>2.7</td>
<td>4.7</td>
<td>75</td>
</tr>
</tbody>
</table>

Destination plasmid: Vb (75ng/reaction)  
Precloned inserts: Va (75ng/reaction)  
Amplicon form: CDR3a, CDR3b (2:1 molar ratio, insert:vector of destination plasmid)  

ex. 2x75ng destination plasmid/3418bp= Xng CDR3a/75bp, X=3.3ng
Volumes refer to diluted solutions

Dilutions

<table>
<thead>
<tr>
<th>Volume</th>
<th>Solution 1</th>
<th>Solution 2</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAV5</td>
<td>luL TRAV5</td>
<td>+ 99uL dH20</td>
<td></td>
</tr>
<tr>
<td>TRBV29-1</td>
<td>luL TRBV29-1</td>
<td>+ 99uL dH20</td>
<td></td>
</tr>
<tr>
<td>CDR3a</td>
<td>luL CDR3a+</td>
<td>9uL dH20</td>
<td>Take luL + 99uL dH20</td>
</tr>
<tr>
<td>CDR3b</td>
<td>luL CDR3b+</td>
<td>9uL dH20</td>
<td>Take luL + 99uL dH20</td>
</tr>
</tbody>
</table>

Assembly Reactions for 4 (complete BMLFl and negative control: no CDR3b, EBNA3A as positive control)

Mix 1: 5uL Assembly Mix + 10uL NEB Golden Gate Buffer

Mix 2: 7uL Mix 1 + 54uL TRAV5 + 66uL TRBV29-1 + 106uL BMLFl CDR3a
   split Mix 2 into 2 tubes, 14.8 ul/tube
   BMLFl complete: 14.8 uL Mix + 5 uL CDR3b + 5.2 uL dH20 (final 25ul)
   BMLFl negative control: 14.8 uL Mix + 5 uL dH20 + 5.2 uL dH20

Mix 3: 7uL Mix 1 + 7.6 ul TRAV + 4.2 ul TRBV+ 8.2 ul EBNA3A CDR3a
   split Mix 3 into 2 tubes, 13.5 ul/tube
   EBNA3A complete: 13.5 uL Mix + 4.7 uL CDR3b + 6.8 uL dH20
   EBNA3A negative control: 13.5 uL Mix + 4.7 uL dH20 + 6.8 uL dH20

Assembly Protocol

Incubate at 37C for 1 hour, then 55C for 5 minutes on PCR machine

Transformation Protocol

Thaw 50ul of NEB 5-alpha competent E coli on ice for 10min

Add 2ul of Assembly reaction, gently mix by flicking the tube 4-5 times

Incubate on ice for 30min
Heat shock at 42C for 30 sec
Place on ice for 5min
Add 950ul of RT SOC
37C, 30min, shake vigorously by using a rotation device

**Plating Protocol**

**Warm LB agar plates containing kanamycin (50ng/ml) at RT**
Mix the cells thoroughly by flicking the tube and inverting
Add 75ul of the 1ml outgrowth onto each plate
Add 8 beads and shake,
Remove beads
Incubate the plate 16h, 37C (lid facing down), Mayer577

**Pick colonies and colony PCR**
Confirm success of golden gate by PCR.

The cloned TCR sequence specific for an antigen is encoded in a lentiviral vector, and transfected in 293T cells with second-generation viral packaging plasmids to generate lentivirus. This is used to transduce the Jurkat Δαβ cell line and express the TCR. Expression is verified by antibody surface staining for TCR and tetramer staining by flow cytometry. In addition, as CD3 expression is down regulated in TCR-deficient cell lines, successful expression of the TCR construct can be measured by CD3-rescue.

**Sequence Listing**

"master pUC57-Kan Cbl-F2A" = following sequence in pUC57-Kan backbone
ggtctctACCTGAACAAGGTGTTCCCACCCGAGGTCGCTGTGTTTGAAGCCATCAAGAGC
AGAGATCTCCACACCCAAAGGCCACACTGGGTGGGTGCTGGCCACAGGCTTCTTCC
CCGACCACGCTGGAGCTGGCTGGGTGGATGGAATGGGAAGGAGGTGACAGTCAGGGG
CAGCAGGGACCAGCCAGCCCTCAAGGAGCAGCCGCCCCCTCAATGACTCCAGATACT
"master pUC57-Kan Cb2-F2A" = following sequence in pUC57-Kan backbone

Full "master pUC57-Kan Cbl-F2A"
Full "master pUC57-Kan Cb2-F2A"

tcgccgctttcgggtatgagcgtggaacagcttggaaaacccctctgacacatgcagctcccggagacggtcacagctgtctgtaagcggatgccgggagca
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gcgagcggtatcagctcactcaaaggcggtaatacggtctatttacctggccgtctctcaatcctgactctcgtctgagcagttaagttcaagcag
tccagtttccagtcagcagctgttataatcagagctgaatcatactcataattgtttcctgtgtgaaattgttatccgctcacaattccacacaacacatcagcggcag
gacaagccgcatagcggcgcggcgtgtggttgccggtggctgggttctgctctgctctgactctgctgcgctcgggtcgttcggctgc
gttccataggatggcaagatcctggtctgcgattccgactcgtccaacatcaatacaacctattaatttcccctcgtcaaaaaataggttatcaaatgagaatcaccatgagtgcagctactgcgataatggtgagaatggcaaaaagtttatgcatttctttccagactgttcaacaggccagcgtacatcacaataacctgcgtcgcgagacgaaatacgcgatcgctgttaaaaggacaattacaacaggaataaatgcaacggcgcaggaacactgccagcagcaaaatatttttcaccaaatgagatattttctataacgctcgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgcacctttgccatgtttcagaaacaactctggcgcacgggcttcccatacaagcgatagattgtcgcacctgattgccgacattatcgcgagcccatttatacccatataaatcagcatccatgttggaatttaatcgcggcctcgacgtttcccgttgaatatggctcataacaccccttgattactcattttatgaaatgctgctgtttttcccgggagttcgcagctggtgagtaaccatgcatcatcaggagtacggataaaatgcttgatggtcggaagagcagggagtgcagctggtgaccagcgcgagcgcagagctgcgcagagctgcgttccttcctagtttatttttatcttgtgcaatgtaacatcagagattttgagacacgggccagagctgca

TRAV (46 total)

TRAV1

atgtggggagctttcctctctatgtttccatgaagatgggaggcactgcaGGACAAGCCTTGAGCAGCCTCTTG
AAGTGACAGCTGTGGGAGGAGGCATTGTCCAGATAAAACTGCACGTACCCAGACATCT
GGGTITTATGGGCTCTCTGTCGCCAGCAACATGAGTGCGGAGCAGCACCACATTTCTT
TCTTACAATGCTCTCGGATGTTTGGAGGAGACAGGTCGGTTTTTCTTCATTCCTTATG
CGCTCTGATAGTTATGGTTCCTCCTTCTACAGGAGCTCAGAGCCAGAATCTCAGC
TCTTACtgagacc

TRAV1-2

atgtggggaggttctccctctcttcatgtaagagatgggaggcactacaGGACAAAAACATTGGACCAGGCCACTGA
GATGACAGCTACGGAAGGTGCCCTTGAGTCCAGATCAACTGCACGTACCAGACATCT
GGTTCACCGGGCAGTGTCTCGGATGTTTGGAGGAGACAGGTCGGTTTTTCTTCATTCCTTAG
CGCTCTAAAGGGTGACAGCTTACCTCCTCCTTGAAGGGAGCTCCAGATGAAAGACTCAGC
TCTTACtgagacc

TRAV2

atgtgctgcagagcactgagggcgctgttggtgctctcctcaactctctctgagaaggtgagaacgAAGGACCAGGTGTTTACGCCCTCCAGACGTCTTACGAGGGAGGAGCTTGTTGGAATCTTCTTG
AATCAGCTCTGTGCTCAGATGTTCTCTCTCTTCTGACTCCTCCTCCCCGAGATGT
GCACCAAGACTCCTGTATTAAAGGCACCTCAAGCTGAGGACGATACAACATGACCTATGAACGGTTCTCTTCATCGCTGCTCATCCTCCAGGTGCGGGAGGCAGATGCTGCTGTTTACtgagacc

**TRAV3**

atggcccttgaccacactctccagatgtcagctcttacatgaggtggtgagaGCTCAGTCACTGGCTGACCCGGGATAGGTCAAGCTGAGGATGCTGACTTGAAATGCACCTTATTCA
GTCTCTGGAAACCCTATCTTTTTTTGATGTTGAATACCCCCACCCGCCAGCCGTCCAGTTCCTTCTGAAATACATCACAGGGGATAACCTGGTTAAAGGCAGCTATGGCTTTGAAAGCTGAATTTAACAAGAGCCAAACCTCCTCTCCACCTGAAAGAACCATCTGCCCCTTG
TGAGCGACTCCGCTTTTGATACtgagacc

**TRAV4**

atgaggcaagtggccagaggtcagctcttgacctgaccttttgtgacCTTGCTAAGACCACCAACAGCCCATCTCCATGGACTCATATGAAGGACAAGAAGTGAACATAACCTGTAGCCACAACAATCT
GCTACAAATGATTATATACAGCTGTTACAAACAGTTTCTCCACCTGACCTGATTTATCC
CTGCCGACAGAAAGCTCCAGCACTCTGTAGCCCTGCCCCGGGTTTCCCTCTGAGCGACACTGCTGTGTACtgagacc

**TRAV5**

ATGAAGACATTTGCTGGATTTTCGTTCCTGTTTTTGTGGCTGCAGCTGGACTGTATGAGTAGAGGAGAGGATGTGGAGCAGAGTCTTTTCCTGGGTCCGAGAGGGAGACAGCTCCGTTATAAACCAGCAACTGCTCTCCACCTCCTCCACCTACTCATGTTATAAGCAAGGAAAGCTGGAGCCGGGaCTCCAGTGGCTGACGTATATTCTTTTCAATAATGGGACA
TGAAACAAGCAAGACACAGCTCTGTTTCTATTTGAATAAAAAGGATGAAACATCTGTCTCTCGCGCATTCAGACACCAGACTGGGGACTCAGCTATCTACtgagacc

**TRAV6**

atggagctattctgcggaggtggttttgctgattttgtggcttcaagtggactgggtgaagAGCCAAAAGATAGAACAGAATTCCGAGGCCCTGAACATTCAGGAGGGTAAAACGGCCACCCTGACCTGCAACTATA
GCACCAAGACTCCTGTATTAAAGGCACCTCAAGCTGAGGACGATACAACATGACCTATGAACGGTTCTCTTCATCGCTGCTCATCCTCCAGGTGCGGGAGGCAGATGCTGCTGTTTACtgagacc

**TRAV3**

atggcccttgaccacactctccagatgtcagctcttacatgaggtggtgagaGCTCAGTCACTGGCTGACCCGGGATAGGTCAAGCTGAGGATGCTGACTTGAAATGCACCTTATTCA
GTCTCTGGAAACCCTATCTTTTTTTGATGTTGAATACCCCCACCCGCCAGCCGTCCAGTTCCTTCTGAAATACATCACAGGGGATAACCTGGTTAAAGGCAGCTATGGCTTTGAAAGCTGAATTTAACAAGAGCCAAACCTCCTCTCCACCTGAAAGAACCATCTGCCCCTTG
TGAGCGACTCCGCTTTTGATACtgagacc

**TRAV4**

atgaggcaagtggccagaggtcagctcttgacctgaccttttgtgacCTTGCTAAGACCACCAACAGCCCATCTCCATGGACTCATATGAAGGACAAGAAGTGAACATAACCTGTAGCCACAACAATCT
GCTACAAATGATTATATACAGCTGTTACAAACAGTTTCTCCACCTGACCTGATTTATCC
CTGCCGACAGAAAGCTCCAGCACTCTGTAGCCCTGCCCCGGGTTTCCCTCTGAGCGACACTGCTGTGTACtgagacc

**TRAV5**

ATGAAGACATTTGCTGGATTTTCGTTCCTGTTTTTGTGGCTGCAGCTGGACTGTATGAGTAGAGGAGAGGATGTGGAGCAGAGTCTTTTCCTGGGTCCGAGAGGGAGACAGCTCCGTTATAAACCAGCAACTGCTCTCCACCTCCTCCACCTACTCATGTTATAAGCAAGGAAAGCTGGAGCCGGGaCTCCAGTGGCTGACGTATATTCTTTTCAATAATGGGACA
TGAAACAAGCAAGACACAGCTCTGTTTCTATTTGAATAAAAAGGATGAAACATCTGTCTCTCGCGCATTCAGACACCAGACTGGGGACTCAGCTATCTACtgagacc

**TRAV6**

atggagctattctgcggaggtggttttgctgattttgtggcttcaagtggactgggtgaagAGCCAAAAGATAGAACAGAATTCCGAGGCCCTGAACATTCAGGAGGGTAAAACGGCCACCCTGACCTGCAACTATA
CAAACTATTCCCCAGCAT ACTTACAGTGGTACCGACAAGATCCAGGAAGAGGCCCT
GTTTTCTTGCTACTCATACGTGAAAATGAGAAAGAAAAAAGGAAAGAAAGACTGA
AGGTCACCTTTTGATACCACCCTTAAACAGAGTTTGTTTCATATCACAGCCTCCCAGC
CTGCAGACTCAGCTACCTACtgagacc

TRAV7
atggagaagatgcggaggcctgtcctaattatattttgtctatgtcttggctgggcaaatggaGAAAACCAGGTGGAGGCAC
AGCCCTCATTTTCTGGGACCCCCAGCAGGGAGACGTGCTCATGAGCTGCACGTA
CTCTGTGCACTGGTTTTTAAATTTTGCAATGTTACAGGCAAAATACAGGGATGGGTC
CTAAACACCTATTATCCATGTATTACGTGGGATATGGAAGCAGAAAGGAAGAC
AGATTCAGCCACCTATA Tgagacc

TRAV8-1
atgctcctgttgetctataaccagtagctgggagatgttttgcctgagagatgcgcagaGCCCAGTCTGTGAGCCACGATA
ACCACCAGTAATTCTCTCTGAAAGCGACCTCACTGAGTTGGGATGCAACTATTCCT
ATGAGGAACCTGTGTTAATTCTCTGTGATGTGACGTACCCTGGGATACACCTTCAGC
TTCTCCTCAAAGTACTTTTTCAGGGGATCCACTGGTTAAAGGCATCAAGGGCTTTGAGG
CTGAATTTATAAAGAGTAATTCTCTCTTTATCTGAGGAAAACTCTCTGTGAGG
GTCAGACACGCTGAGTACtgagacc

TRAV8-2
atgctcctgtctcgctccctcgtggagagttttacttgtggagacagagaagcagagaGCCCACTGCTGAGCCACGCTTG
ACAGCCACGTCTGTCTCTGAAAGGAACCCCGGTGCTGAGGAGGTGCAACTACTCA
TCTCTCTTTTCCATCTCTCTGTTGATGTGCAACACCACCCAAAGGACTCCAG
CTTCACAGGTACACATCAGCGCCACCTGGTTAAAGGGCATCAACCGGTTTGA
GGCTGAATTATAAGAGATGGAACCTCTCTCCACCTGAGCAAACCCCTCAGCCCATGA
TGAGCGACGCGGCTGAGTA Ctgagacc

TRAV8-3
atgctctggagctatcccaactgcgaggatatattgactcgagacaactgcagaACCCAGTCAGTGACCCAGCTG
ACATCCACATCACTGTCCTGTAAGGACCTCCTACTGGAAGTGGAGATGTAACTATTCT
ATGGAAGCAACACCTTATCCTCTGTTATGTCCAGTCCCCCGGCAGGCCTCCAGC
TGCTCCTGAAGTAGTACTTTTCAGGAAGACACTCTCTGGTTCAAGGCAATTAAGGCTTTTAG
GCTGAATTTAAAGAGGAGTCAATCTCCTCTCAATCTGAGGAAACCTCTGTGCAATGG
AGTGATGCTGCTGAGTACtgagacc

TRAV8-4
atgcctctgctgtgtccagctgtcggggtatattttatcctcggaggaaccagaGCCCCAGTCGGTGACCCAGCTT
GGCACAGCCACGTCTCTGCTCTGAGAGCCCTGTTCTGAGGTGGAAGCTACTTCT
ATCGTCTGTGGTCCATATCTCTTCTGGTAGATGTGCAATACCCCAACCCAGACTCCA
GCTTCTCCTGAAGTAGCACACACGGCAGCCAACCTCTGGTTAAGGCAATCAACGGTTTTTG
AGGCTGAATTTAAAGAGGAGTCAATCTCCTCTCAATCTGAGGAAACCTCTGTGCAATGG
AGTGATGCTGCTGAGTACtgagacc

TRAV8-6
atgcctctgctgtgtccagctgtcggggtatattttatcctcggaggaaccagaGCCCCAGTCGGTGACCCAGCTT
GGCACAGCCACGTCTCTGCTCTGAGAGCCCTGTTCTGAGGTGGAAGCTACTTCT
ATCGTCTGTGGTCCATATCTCTTCTGGTAGATGTGCAATACCCCAACCCAGACTCCA
GCTTCTCCTGAAGTAGCACACACGGCAGCCAACCTCTGGTTAAGGCAATCAACGGTTTTTG
AGGCTGAATTTAAAGAGGAGTCAATCTCCTCTCAATCTGAGGAAACCTCTGTGCAATGG
AGTGATGCTGCTGAGTACtgagacc

TRAV8-7
atgcctctagtgctctgtctggagatgttacctcagaaaccagaACCCAGTCGGTGACCCAGCTT
GGCACAGCCACGTCTCTGCTCTGAGAGCCCTGTTCTGAGGTGGAAGCTACTTCT
ATCGTCTGTGGTCCATATCTCTTCTGGTAGATGTGCAATACCCCAACCCAGACTCCA
GCTTCTCCTGAAGTAGCACACACGGCAGCCAACCTCTGGTTAAGGCAATCAACGGTTTTTG
AGGCTGAATTTAAAGAGGAGTCAATCTCCTCTCAATCTGAGGAAACCTCTGTGCAATGG
AGTGATGCTGCTGAGTACtgagacc
TRAV9-1
atgaattctctccaggaccacgcattcttaatgtttgggggaatcaatGGAGATTCAGTGGTCCAGACAG
AAGGCCAAGTGCTCCCCTCTGAAGGGGATTCCCTGATTGTGAACTGCTCCTATGAA
ACCACACAGTACCTTCCCTCTTTTTTGGTATGTCCAATATCCTGGAGAAGGTCCACAG
CTCCTACCTGAAAGCCATGAAGGCCAATGACAAGGGAAGGAACAAAGGTTTTGAAG
CCATGTACCGTAAGAAAACCACTTCTTTCCACTTGGAGAAAGACTCAGTTCAAGAG
TCAGACTCTCGCTGTGTAACgagacc

TRAV9-2
atgaactattctccaggcttagtatctctgatactcttactgcttggaagaacccgtGGAAATTCAGTGACCCAGATGG
AAGGGGCAGTGACTCTCTCGAGAAGGCTTCCCTCCTGACTATAAAACTGCACGTACACA
GCCACAGGATACCTTCCCTCTTTTTGTTCTGTCCAATATCCTGGAGAAGGTCTACAG
CTCCTCTGAAAGCCAGAAGGCTGATGACAAGGGAAGCAACAAAGGTTTTGAAG
CCACATACCGTAAGAAAACCACTTCTTTCCACTTGGAGAAAGACTCAGTTCAAGTG
TCAGACTCTCGGTGTAACgagacc

TRAV10
atgaaaaagcatctgacgaccttcttggtgattttgtggctttatttttatagggggaatggcAAAAACCAAGTGGAGCAGA
GTCCTCAGTCCCTGATCATCCTGGAGGGAAAGAACTGCACTCTTCAATGCAATTAT
ACAGTGAGGCCCTTCCAGAATCTTAAAGGTGTTATAAGCAAGATCTGGGAGAGGTCCC
TGTTTCTGACAATCATGACTTCTCTCAGTGAAGACCAACAAAAGTGAACAGGAAAGATATA
CAGCAACTCTTGAGTCGACAAACAAAACAAGCGCTCTCTCAGAATACAGCAGGCTCCAG
CTCAGCGATTCAGCCTCTACgagacc

TRAV12-1
atgatatctctgaggattttactggtgatcctgtggcttcagttaagctgggtttggagccaaCGGAAGGAGGTGGAGCAG
GATCCTGGACCCTCCCTGACGCTCTGGAGGGAAAGAACTGCACTCTCCTCTCGATCTTAC
AGCAACAGTGCTTTCTCAGTCTTTCTTCTGTGAATACAGCAGGATTCAGGGAAAGAACC
TAAGTGGCTGATGCTCTCTCCAGTGGTATAGGAAAGATGGGAAAAGTTACAGCAG
AGCTCAATAGAGGCCAGCAGAGTATATTTCCCTGTGCTCAGAGACCTCAGAACGCTCAGT
GATTCAGCCACCTACgagacc
TRAV12-2
atgaaatcctgagatgttactgatcctgtggactgtttggagccaaCAGAAGGAGGTGGAGCAG
AATTCTGGACCCCTCAGTGTTCCAGAGGGAGCCATTGCTCTCTCAACTGCACTCAGT
AGTGACCGAGGTTCCAGTCCTTCTTCTGCTACAGACAAATATTTCTGGGAAAAGGCC
TGAGTGGTATAATGGACTATATCACTCCCAATGGTGACAGGATAAGGTGATTTACAG
CACAGCTCAATAAAGCCAGCCAGCATGATTTCTCTGCTCATCAGAGACTCCCAGCCC
AGTGATTCAGCCACCTACTgagacc

TRAV12-3
atgatgaaatccttgagagttttacctggtgatcctgtggcttcagttagctgggtttggagccaaCAGAAGGAGGTGGAGCAG
AGGATCCTGGACCACTCAGTGTTCCAGAGGGAGCCATTGTTTCTCTCAACTGCACTT
ACAGAAGCTGCTTTTCCAATACTCTTCTGTTGCTACAGACGGGTATTTCCAGAAAAGGC
CCTGAGTTGTGTATGAACACATACACTCCAGTGGAAACAAAGAGATGGGGAAGGTACAG
AGCAGAGTCCGATAAAATCCAGCAAGTATATATCTCCTTGTTCATCAGAGACTACAGGAC
CCAGTGGATTCCACACCTACTgagacc

TRAV13-1
atgacatccattcgagctgtatttatatctggtggctgagttaactgtgggctgggctggagGGAGAGAATGTGGAGCAGCATC
CTTCAACCCCTGAGTGGTCCAGAGGGAGACAGCGCTTTGATCAAGTGTACTTATTCA
GAGACGTGCCTAAAACACTTCTTCTGTTATAAGCAAGAATTTGGAAAAGGACCTCA
GCTTTATATAGACATCATTGCTTTCAAAATGTGGGCGAAAGAAAAGAGACCAACAGATTCG
TTCATTGGAACAAAGACAGCCAAAAACATTCTCCCTGCACATCAGAGACaCAACCT
GAAAGACTCGGCTGTCTACgagacc

TRAV13-2
atggcaggccattcactggtgattatattactgtgctgactgtaacgtgggagacagGGAGAGATGGTGCGCTG
ATCTTCTACCTGAGTGGTCCAGAGGGTGCAACTCTATTATCAACTCTGTTATT
CAACACGGCCTCAGACTACTTCTTGTACAGCAAGAATCTGGGAAAAGGTCCT
CAATTCCATTAGACATTGGCTTTCAAAATATGGACAAAGGCAAAGGCCAAAGAGTCAC
CGTTTTATTGAATAAGACAGTGAAACATCTCTCTCTGCAAATTGCAGCTACTCAACC
TGGAGACTCAGCTGTCTACtgagacc

**TRAV14**
atgtcactttctagcctgtaaggtggtcagctcagtgtgctaggacctgccatGCCCAGAAGATAACTCAACACCAACCAGGAATGTTCTGAGGAAAAAGGAGGCTGTGACTCTGGACTGACATATGACACCAGTGATCAGAGGTTATGGTCTATTCTGGTACAGCCAGCAGCAGTGGGAAAATGATTGTTCTTTATTTATCAGGGGTCTTTATGACCAGCATAATGCAACAGAGGGTCGC
TACTCATTGAATTTCCAGAAGGCAAGAAAATCCGCAACCCTTGTCTCATTCCCGCTTCA
CAACTGGGGGAAGCTCAGAATATGTGATCAGACCagacc

**TRAV16**
atgaagccaccctactctcagcttgatgtgataatatttatactcagaggaacaagaGCCCAGAGAGTGACTCAGCCGAGAAGCTCCTCTCTGTCTTTAAAGGGGCCCCAGTGGAGCTGAAGTGCAACTATTCC
TATTCTGGGGAGTCTCAGACTCTTCTTCTGTTATGGTACAGCTCAGAACAAGCCTCCAG
TTACTCTTTGAGACACATCTCTCTAGAGAGAGCATCAAAGGCGACACCCCTCGTAAGGGACTGACCTTGACCTTTCA
CAAAAGGCGAGACATCTTTCACCTCAAGAAAAACCATTGGCTCAAGAGGAAGACTCAGCCATGTATagacc

**TRAV17**
atggaaactctctctgggagttcttttctgttctatggcttcactggctagggaacAGTCAACAGGGAGAAGAGG
ATCCCTCAGGCTTGAGCACATCCAGGAGGTTGAAAATGCCACCATGAGCTGACAGTTAC
AAAACTAGTATAAAAAAATCTACGTGGTATAGCAAAAAATTCTAGGTAGGAGGCCCTTGT
CCACCCTATTAAAAATCTGTTCACAGGAGAAAACAGATGGAGAGAGAGTTAGAAGATTAAGAGTCACGCTTGACACTTCCAGGAGCTGACCTTTGAGACACCTTGAACGAGCAGTGGTCGTTCAAGGAGGAGACTCAGAGCAGACACTGCTTCTTACagacc

**TRAV18**
atgctgtctgtctgtctgctcagcttgtgtgcatctgtgtatgtcagacagtGGAGACTCGGTTACCCAGACAGAAGGCCCAGTTACCCTCCCTGAGAGGGCAGCTCTGACATTAAACTGCACTTATCAG
TCCAGCTATTCAACCTTTCTATTCTGGTATGTCAGTATCATCTAATACACAAAGAGACGCTGAG
CTCCTCCTGAAAAAGTTCCAGAAAAACCAGGAGACGGACAGCAGAGGTTTTTCAGGCCAGTCCTATCAAGAGTGACAGTTCCTTCCACCTGGAGAAGCCCTCGGTGCAGCTGTCGGACTCTGCACTCCAGACCTGATTCTCCAGGAGGGAGCCAATTCCACGCTGCGGTGCAATTTTCTCCTCCTGAAAAGTTCAGAAAACCAGGAGACGGACAGCAGAGGTTTTTCAGGCCAGTCCTATCAAGAGTGACAGTTCCTTCCACCTGGAGAAGCCCTCGGTGCAGCTGTCGGACTCTGCCGTGTACtgagacc

TRAV19
atgtgactgccagcttgaggcagctcatagctccatctgttgttgatccagcttgGCTCAGAAGGTAACTCAAGCGCAGACTGAAATTTCTGTGGTGGAGAAGGAGGATGTGACCTTGGACTGTGTGTATGAAACCCGTGATACTACTATTACTATTATTCTGGTACAGCAAGCAACCACAAAGTGGGAAGTTGTTTTCTTTATTCGTCGGAACTCTTTTGATGAGCAAAATGAAATAAGTGGTCGGTATTCTTGGAACTTCCAGAAATCCACCAGTTCTCCTCAACTTCACCATCACAGCCTCAAGTCGTGGACTCAGCAGTATACtgagacc

TRAV20
atggagaaaatgtggaggtgtgcattcatagctttggtcgcttggttggttggtgtaAGAAGACCAGGTGACGCAGAGTCCCGAGGCCCTGAGACTCCAGGAGGGAGAGAGTAGCAGTCTTAACTGCAGTTACACAGTCAGCGGTTTAAGAGGGCTGTTCTGGTATAGGCAAGATCCTGGGAAAGGCCCTGAATTCCTCTTCACCCTGTATTCTCAGCTGGGGAAGAAAAGGAGAAAGAAAGCCTAAAGCCACATTAACAAAGAAGGAAAGCTTTCTGCACATCACAGCCCCTAAACCTGAAAGACTCAGCCACTTATgagacc

TRAV21
atggagacactcttgggctgttatctcttggcttcagctgcaatatgggtgagcagcCCTGCAGCTCTGAGTATCCAGAAGGAGAAAACCTGGTTCTCAACTGCAGCTTACATGATAGCGCTATTTACAACCTCCAGTGGTTAGGCAGGACCCTGGGAAAAAGGaCTCACTATCTGCTTCTTCTTACAGTCAAGTGAAGACTTAATGCCTCGCTGGATAAATCATCAGGACGTAGTACTTTATACATTGCAGCTTCTCAGCCTGGTGACTCAGCCACCTACTagacc

TRAV22
atgaagaggtatttgggagctctgctggggctcttgagtgcccaggtttgctgtgtgagaGGAATACAAGTGGAGCAGAGTCCTCCAGACCTGATTCTCCAGGAGGGAGCCAATTCCACGCTGCGGTGCAATTTTCTCCTCCTGAAAAGTTCAGAAAACCAGGAGACGGACAGCAGAGGTTTTTCAGGCCAGTCCTATCAAGAGTGACAGTTCCTTCCACCTGGAGAAGCCCTCGGTGCAGCTGTCGGACTCTGCCGTGTACtgagacc

TRAV22
atgaagaggtatttgggagctctgctggggctcttgagtgcccaggtttgctgtgtgagaGGAATACAAGTGGAGCAGAGTCCTCCAGACCTGATTCTCCAGGAGGGAGCCAATTCCACGCTGCGGTGCAATTTTCTCCTCCTGAAAAGTTCAGAAAACCAGGAGACGGACAGCAGAGGTTTTTCAGGCCAGTCCTATCAAGAGTGACAGTTCCTTCCACCTGGAGAAGCCCTCGGTGCAGCTGTCGGACTCTGCCGTGTACtgagacc
CTGACTCTGTGAACAATTTGCAGTGGTTTCATCAAAACCCTTGGGGACAGCTCATCA
ACCTGTTTTACATTTCCCTCAGGGACAAAACAGAATGGAAGATTAAGCGCCACGACT
GTCGCTACGGAACGCTACAGCTTATTGTACATTTTCCCTTCCAGACCACAGACTCA
GGCGTTTATtgagacc

TRAV23
atggacaagatctttagagcatcattatatttttattcttttaggtgatggtgagtggcaacaagagaaatgacCAGCA
GCAAGGTGAAACAAAACTCCTCAATCTCTTGGATAGTGAGAGAGAGTGAGAGATGGTTCAATTA
TAAACTGTTGCTTTATGAGAACCACACTGCGTTTGAACACTTTTCCATGTTACCAACAATTTCC
CTGGGAAAGGGCCCTGACATTATTGATAGCGCATAGCTCCAGATGTTGAGTGAADAGAAAAG
GAAGGAAGATTCAACAATCTCCTTCAATAAAGAGTGCCAAGCAGTTCTCATTGATAT
CATGGATTCCCAGCCTGGAGACTCGACAGCCACACATtgagacc

TRAV24
atggagaagaatcctttggcagccccattactaactatctctttggtcttgactgcgtgagcatATACTGAACGTGGAAAC
AAAGTCCTCAGTCAGTGCGATGTTCCAGGAGACAGCAACAAATTTCACCTGCAGC
TTCCCTTCCAGCAATTAGTGTCAAACACTGGTACAGATGGGAACACTGCAAAAAAGC
CCCGAGGCCTTGGTGTGAATGACTTTAATAATGAGGATAGAAGAAAGAGACGAAT
AAGTGCCACTTATAATCCAAGGAGGGTTACAGCTATTGTACATCAAAAGGATCCC
AGCCTGAGAAGCTCAGCCACACATtgagacc

TRAV25
atgcctactacatcaatgtgtggcttataggtgcaattgtcagaggtgaatGGACAACAGGTAATGCAAATTCCT
CAGTACCAGCATGTAACAAGAAGGAGAGGACTTCCACCGACTCGAATTTCCCTCAAC
TACCTTTAAGCAATATACAGTGGTTAAAGCAAAAGGCTTGGTAGACATCCCGTTTTT
GATACAGTTAGTGAAAGATGGGAAGATGGAAGAAAGCAGAAAGAGACTGCAATTTCCAG
TTGGAGAAAGCAAAAAGAAAAGACAGCTCCCTGACATCAAGCACCACACAGACTACAG
ATGTAAGGAACCTACtgagacc

TRAV26-1
atgaggctgggtgcaaagagaactgtgttctgacctttggaactataattGATGCTAAGACCACCCAGCCCCCCTC
CATGGATTGCGCTGAAGGAAGAGCTGCAAACCTGCCTTGTAATCACTCTCATCAC
ATCATTCATGGTCTAAAAAAACAATGAAACCAATGAAATGGCCTTACTGCAGACTG
CTGTGTACtgagacc

TRAV26-2
atgaagtgtgtgacaagcatcactgtactctctttggtattatgggtGATGCTAAGACCACACAGCCAAATTC
AATGGGAGAGTACCAAGAAGAGCGCTTCTACTTGCTTGTAAACACTCCACAAATCA
GTGGAAACTGATTACATACAAGGGTGATACAGCTCTCCCTCCAGGGTCAGATAC
GTGATTCATGGTCTTACAAGCAATGTGAACAACAGAATGGCCTCTCTGGCAATCGC
TGAAGACAGAAGTCCAGTACCTTGATCCTGCCCCACGCTACGAGAGATGCTG
CTGTGTACtgagacc

TRAV27
atgcctgaaattctccgtgtcattctctttggtacctgtggtccagctggtggctcgtcagccgatgggtgagcACCCAGCTGCTGGAGCAGAGCCCT
CAGTTTCTAAGCAATCCAGAGGGAGAAAATCTCAGTTACTGCAACTCCTCAAG
TGTTTTTCCAGCTTACAATGGGAGGACAGGAGCCTGGGGAAGGTTGTCAGCTCCTCCT
GGTGCACAGTAGGTTACGGGTGGAGAAGTGAGAAGGCTGAAGAGACTTACATCAG
TGAAGGTTGATGCAAGAAGAGGAGACAGTTCTCTCCACACATCAGCGCCGCTGGTGAC
TACAGGCCTCTACtgagacc

TRAV29
atggccatgcctctgggcatcagtctgattctgtggctctcagccgacactgggtaaacgtcaacagaagaatgatGACCAGCA
AGTGAAGCAAAAATCCAACATCCCTGCTGCGTCCAGGAAGGAAGAATTCTATGCA
ACTGGAAGCATAAAACGATGTTGATTATTTCTATGTGAAAATACCTG
CTGAAAGGTCTACATTTCTGATATCTAAGGTCCATTAGGAATAAATAGGAGAT
GGAAGATCTGTTCTTCTTAAACAAAGTGCAAGCGACCTCTCTCTGCAACATTG
CCCTCCAGCGCAGACTCTGCTCGAGTACtgagacc
TRAV30
atggagactctctgaagtgctttcaggcaccttgtgtgctagttgacctgggtgagaagcCAACAACCAGTGCAGAG
TCCTCAAGGCGTGATCTCCTCCGAGAAGGGGAAGATGCTGCATCAACTCAGGTTCCT
CCAAGGCTTTATATTTCTGATACTGGTACAGGCAGATGGAAGCAACCCGTC
TTCCCTGATAATTACTGGAAGGGTGAGAAGAAAGAGGTCATGAAAATAATCTGC
TTCATTTAATGAAAAAGCGACAAAGCTCCCTGTACCTCTTACGGCTCCACAGCCT
GTTAATTGATCAGACCTACtgagacc

TRAV34
atggagacttgtctgaagtgctttcaggcaccttgtgtgctagttgacctgggtgagaagcAGCCAAGAATGGGAGCAGA
GTCCTCAGTCTTGTATGCTCAGAAAGGGGAAGATCCTACCATAAAGCTGCACGTCA
TCAAGAGCGTTATATGGGTTATAGGTATAAGCAGAAAAGGAGAAGGTACTTATATATATGATGATGCTACAGAAAGGTGGGAAGAGAAAATGCATGGCAGAAACT
GCCAAGTTGGGATGAGAAGAAAGCGACAAAGCTCCCTGATACATACAGCCACCCAGCC
cagcacagggcatctacticgtgagacc

TRAV35
atgctctttgaacatttatttattctgttttcggatcttatgcatagctgctgagcCTGAGCAACAGCTCAGAGTCCTC
AATCTATGTGATCCAGGAGGAGAAGATGCTCCTCATTGAACGTCACTCTCCTCAGGCT
ATATTTAACACCTTGCTATGCAAGGCAGGACCTTGGGAAGGGGCTCTCTCTCTCTT
GAGCCCTTATATAGATGCTGAAATTCAGCCTCATAATGGGAAGACTGACTGCTAGT
TTGGATACCAAGGAGGACAGCTTTCTCGATATATCTCAGCATCCTACCCTTAGTGAT
GTGAGGCACTACgagacc

TRAV36
atgatgaaggtcagtttagttcacttgtatcgtctttgacctggctagtggattgggtgaccc acctatatagctgttttggctcagcagcgctgttggagagtGAAAGCAAGGTGGTACAA
AGGCTCTACTCTCTGGTGCTGCAGGAGACACCCTACTGAACTCATTGCAAGTTAT
GAAGTGACTAACCTTTCAGCAGCTACTATCTGTACAGCAGAGAAAGAAAAGCTCCCAC
ATTTCTATTTATGCTAATCTAAGTGCTAAGGGAAAGTCAGGAAGACTAGTA
GCATATTAGATAAGAAGAACGTTTTCCAGCATCCATAGTCAGGGCCACCCAGACC
GGAGACTCGGCACATCATCtgagacc
TRAV38-1
atgacacgagttagctgtggccatctgtccacctgtcttgataattccggcatgGCCCAGACAGTC
TCAACCAGAGATGCTGCTGAGGAGAGCAGAGACTGTGACCCCTGAGTTGCACTATAG
ACACCAGTGAGAATAATTATTATTATTGTCTGGTACAAGAGCAGCCTCCCAGCAGGCAG
ATGATTCTCTGTTATTCGCAAAGAAGCTTTATAAGCAACAGAATGCAACCCAGAGAAC
TTCTCTGTGAACTCAGAGAATATTATTTGTTCTGGTACAAGCAGCCTCCCAGCAGGCAG
TTCTCTGTGACTTCCAGAAAGCAGACCAATACTTTACTCACTGCTCAAGATCTCAGACTC
ACAGCTGGGGGACACTGCGATGTATgagacc

TRAV38-2
atggcatgccctgcctggtgctggtctgctctgccctttactgtccctgtcttgaaatttacatgGCTCAGACAGTC
CTCAGACTCTAGCTCAGTCTCAGTCT
CAACCAGAGATGCTGCTGAGGAGAGCAGAGACGACGTGACCCCTGAGTCTGACACATATG
ACACCAGTGAGAAGTATTTTTTTATTCTGGTACAAGAGCAGCCTCCCAGCAGGCAGA
TGATTCTCTGTTATTCGCAAAGAAGCTTTATAAGCAACAGAATGCAACCCAGAGAAC
TTCTCTGTGACTTCCAGAAAGCAGACCAATACTTTACTCACTGCTCAAGATCTGAGACTC
CAGCTGGGGGATGCCGCGATGTATgagacc

TRAV39
atgaaaagctactagcaatgattctgtggcttcaactagaccgcttaagtggaGAGCTGAAAGTGGAACAAAACCC
TCTGTTCCTGAGTGAGGAGGAGGAAAACACTTATACATCAGCAATATTCAAA
CCACTCCAGACAGACTGTATTTGTCGATACGAGAGGATGCTGCTGGGAAAAGTCTGGAATCT
CTGTTTGTGTTGTCTATCAATAGGAGCAGTGAAGCAGGAGGGACGATTAATGGCCCTC
ACTTGATACAAAGGAGCCCTGCTCAAGCACCCCTCCACATCAGCTGCAGTGCATGACCC
TCTCTGCCACCTACTgagacc

TRAV40
atgaactctctctggactttctaattctgatcttaatgtttggaggaaccagcAGCAATTCAGTCAGCAGACGGGC
CAAATAACCGTCAGGAGGAGCAGCTGACTATGAAGACTGCACATACATCAGCACC
GGGTACCCCTACCTTTTCTGGTATAGGAATACCAGCAAAACCTCAGCTGAGCTTTCTC
TCAGAGAGACAAATGGAAAACAGCAAAAACTTCGGAGGCGGAAATATTAAAGAC

92
AAAACCTCCCCCATTTGGAATAATACGTCAGTCACACCTCAGGCTGTAAC
cacc

TRAV41

atggtgaagatccggaattttgtggctattttgtgcttcagctaa...AATCC
AGTGACAAGTCTGTCTGCCTATTCACCGATTTTGATTCTCAAACAAATGTGTCACAA
AGTAAGGATTCTGATGTGTATATCACAGACAAAACTGTGCTAGAC
ATGAGGTCTATGGACTTCAAGAGCAACAGTGCTGTGGCCTGGAGCAACAAATCTG
CTTTGCATGTGCAAACGCCTTCAACAACAGCATTATTCCAGAAGACCTTCTTCCC
CAGCCCAGAAAGTTCTGTCGAAGCGAACAAAACCTGCTGAGAGCTCTCTTCCTCCTCCTC
CTGAAAG
TGGCCCGGGTTTAAATCTGCTCATGACGCTGCGGCTGCTGTTTCAGCTAAAgcggccgcgtcga
caatca

"master pUC57-Kan Ca short" = following sequence in pUC57-Kan backbone
gggtgggtgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggttgctggt
GGACTTCAAGAGCAACAGTGCTGTGGCCTGGAGCAACAAATCTGACTTTGCATGTG
CAAACGCCTTCAACAACAGCATTATTCCAGAAGAcACCTTCTCCTCCCAGCCCCAGAA
AGTTTCTGTGATGTCAGCTGGTCAGAAAAGCTTTGGAACAGATACGAAACCTAAA
CTTCCAACCTGTCAGTGGATTTGGITTCCGAATCTCCTCCTCCTGAAAGTGGCCCGGGTT
TAATCTGCTCATGACGCTGCGGCTTGGTGTAGCTCACGCTAAAgcggccgcgtcgacaaatcaactacgeaatcc
gggccgccgctgactgcagaggtcatgacaaactggttgaatt GTAgtgactgatgttctctgtgtgtgtaaatftgttatccgctcaaatftccac
acaacatacgacgccgaagctcaataagttgctcctgtgcaaatatgcagctaactacctaatgtatcccaagagatcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggtcagaggt
"master pUC57-Kan Ca" = following sequence in pUC57-Kan backbone
agcgtggcgctttctcatagctcagctagtttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgtt
cagcccgacgccgtccacctatcgcggtaactatctcctggtagttctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgtt
cagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtattttgatatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttt
gtttgcaacgcagcacagattacgcgcagaaaaaagaggtactcaacagagatccttgagatcttttctacgggtacgccgctagtggaacgaaactcagtttaaggagttgtctgcatgagaffactcaacagagttctctactagtttaaaattttaaaaaataagatgttttaaatcaagcacaattctg aataatgttacacaaaccaatctctagttagaaaaactctacgcagcagcactttacagataagagctcagttgcaagtcttggttggccgtagcgcgaagccagtgcagcagggaaagttgcagtttctgggctcttttataataatgaaatctcagagtaagtttttgaatcaccatatttggaagaactcagttgctttcttgatatatttttctcttgtgcaatgtaacatctgtttttttttttaaaaattttctctgttggaagttttttttttttttttttttaaaaattttttttaaaaattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
AAATTCTGAAGATAATTGTTTAGCTACAATAATAAAGGAGCTCATTTAATAATGAAAAC
AGTTCCAAATCGTTCTCACCTAAATCTCCAGACAAAGCTCATTAAATCTTCACAT
CAATTCCCTGGAGCTTGTGGACTCTGCTGTGTATgagacctaa

TRBV4-1

gggatccacgggtgcaccatgggtgccttcctgggagcgctgctgtctgggtactgacctaaGACACTGA
AGTTACCAGACACACAAACACTGGTCATGGGAACGAAATGACAAATGAAAAGCTTTTG
AATGTAACCAACATGCTGGGATACTGGACTATGGTATGGTACAAGCAGGAAACCTCTCTTAAATGAAA
GAAGCCACCCGAGCTCATGTCTGTCTACAGCTATGAGAAACTCTCTCTCTCTTAAACTCCACC
TACACGCCCTCGACAGGAGACTGACGCGCCCTGTATgagacctaa

TRBV4-2

gggatccacgggtgcaccatgggtgccttcctgggagcgctgctgtctgggtactgacctaaGAAACCGG
GAGTTACGCAGACACACCTGGTCATGGGAATGACAAATAAGAAGTCTTTG
AAATGTAACCAACATGCTGGGATACTGGACTATGGTATGGTACAAGCAGGAAACCTCTCTTAAATGAAA
GAAGCCACCCGAGCTCATGTCTGTCTACAGCTATGAGAAACTCTCTCTCTCTTAAACTCCACC
TACACACCCCTGACAGGAGACTGACGCGCCCTGTATgagacctaa

TRBV4-3

gggatccacgggtgcaccatgggtgccttcctgggagcgctgctgtctgggtactgacctaaGAAACCGG
GAGTTACGCAGACACACCTGGTCATGGGAATGACAAATAAGAAGTCTTTG
AAATGTAACCAACATGCTGGGATACTGGACTATGGTATGGTACAAGCAGGAAACCTCTCTTAAATGAAA
GAAGCCACCCGAGCTCATGTCTGTCTACAGCTATGAGAAACTCTCTCTCTCTTAAACTCCACC
TACACACCCCTGACAGGAGACTGACGCGCCCTGTATgagacctaa

TRBV5-1

gggatccacgggtgcaccatgggtgccttcctgggagcgctgctgtctgggtactgacctaaAAGGCTG
GAGTCACCTCAAAACTCAGATATCTGATCAAACGAGGACGACAGGACAGCTGACT
TRBV5-3

GGGATCCACCGTCCACGAGGCAGGAGCGTACGCCCTTGGGATGACAGCACAGGCGGAGGTGCAGGTGGCAGGCTGGGAGCTGGGAAGGAGGTGTGACCCAAAGTCCCACACCTGATCAAAACGAGAGGACAGCAAGTGACTCTGAGATGCTCTCCTATCTCTGGGCACAAGAGTGTGTCCTGGTACCAACAGGTCCTGGGAGCTTCACAGGGGCCCCAGTTTATCTTTCAGTATTATGAGAAAGAAGAGAGGAAGAGGAACTTCCCTGATCGATTCTCAGCTCGCCAGTTCCCTAACTATAGCTCTGAGCTGAA

TRBV5-4

GGGATCCACCGTCCACGAGGCAGGAGCGTACGCCCTTGGGATGACAGCACAGGCGGAGGTGCAGGTGGCAGGCTGGGAGCTGGGAAGGAGGTGTGACCCAAAGTCCCACACCTGATCAAAACGAGAGGACAGCAAGTGACTCTGAGATGCTCTTCTCAGTCTGGGCACAACACTGTGTCCTGGTACCAACAGGTCCTGGGAGCTTCACAGGGGCCCCAGTTTATCTTTCAGTATTATGAGAAAGAAGAGAGGAAGAGGAACTTCCCTGATCGATTCTCAGCTCGCCAGTTCCCTAACTATAGCTCTGAGCTGAA

TRBV5-5

GGGATCCACCGTCCACGAGGCAGGAGCGTACGCCCTTGGGATGACAGCACAGGCGGAGGTGCAGGTGGCAGGCTGGGAGCTGGGAAGGAGGTGTGACCCAAAGTCCCACACCTGATCAAAACGAGAGGACAGCAAGTGACTCTGAGATGCTCTCCTATCTCTGGGCACAAGAGTGTGTCCTGGTACCAACAGGTCCTGGGAGCTTCACAGGGGCCCCAGTTTATCTTTCAGTATTATGAGAAAGAAGAGAGGAAGAGGAACTTCCCTGATCGATTCTCAGCTCGCCAGTTCCCTAACTATAGCTCTGAGCTGAA

TRBV5-6

GGGATCCACCGTCCACGAGGCAGGAGCGTACGCCCTTGGGATGACAGCACAGGCGGAGGTGCAGGTGGCAGGCTGGGAGCTGGGAAGGAGGTGTGACCCAAAGTCCCACACCTGATCAAAACGAGAGGACAGCAAGTGACTCTGAGATGCTCTCCTATCTCTGGGCACAAGAGTGTGTCCTGGTACCAACAGGTCCTGGGAGCTTCACAGGGGCCCCAGTTTATCTTTCAGTATTATGAGAAAGAAGAGAGGAAGAGGAACTTCCCTGATCGATTCTCAGCTCGCCAGTTCCCTAACTATAGCTCTGAGCTGAA
TRBV5-7

TRBV5-8

TRBV6-1
TRBV6-2

```
TRBV6-2

TRBV6-3

TRBV6-4

TRBV6-5
```
TRBV6-6

```plaintext
TRBV6-6

ggggatccacgggtgcaccatgagcctcctctgctgctcggtgaggctccagtgAATGCTGG
TGTCACGACCACCCAAAATTCCCGCTCTGAAGATTAGAGACAGACAGACACTGACCTGC
AGTGTACCCAGGATAGTGAACCCATAACTACATGTACTGTGTATCGACACAGACCCAGGC
ATGGGGCTGAAGCTTGATTTATTTATTTGTTGCTGCTGATCACCTGATAAGAGAGA
AGTCCCCAATGGAACATACGGTCCTCAGATCAACACAGAGGATTCCCGCTCAAGGC
TGGAGGTGGCTGCCCTCCTCCAGACATCTCTGTGTACgagacctaa
```

TRBV6-7

```plaintext
TRBV6-7

ggggatccacgggtgcaccatgagcctcctctgctgctgctgcggtcctctcctctgtgggcaggtcccttgAATGCTGG
TGTCACGACCACCCAAAATTCCACGTCCTGAAGACAGGACAGCAGACACTGACTTCGC
TGTTGTGCCAGGATAGTGAACCATAGAATTACATGTACTCGGTATCGACACAGACCCAGGC
AAAGGGCTGAGGCTATTTACTACTCAGTGGCTGCTGCTCTCAGTCCAAGAGAGA
AGTCCCCAATGGCTAATTGCTCCAGATCAAACACAGAGGATTCCCGCTCAAGGC
TGGAGGTCAATCTGCCTCCTCCAGACATCTCTGTGTACgagacctaa
```

TRBV6-8

```plaintext
TRBV6-8

ggggatccacgggtgcaccatgagcctcctctgctgctgctgcggtcctctcctctgtgggcaggtcccttgAATGCTGG
TGTCACGACCACCCAAAATTCCACATCCTGAAGACAGGACAGCAGACACTGACTTCGC
AGTGTGCAGGATAGTGAACCATAGATACTTGTCCTGGTATCGACACAGACCCAGGC
ATGGGGCTGAGGCTAATTACATGTACTCGGTATCGACACAGACCCAGGC
CCCAATGGCTAATTGCTCTAGATTAAACACAGAGGATTCCCGCTCAAGGC
TGGAGGGCTGCTCCCTCCAGACATCTGTGTACgagacctaa
```

TRBV6-9

```plaintext
TRBV6-9

ggggatccacgggtgcaccatgagcctcctctgctgctgctgcggtcctctcctctgtgggcaggtcccttgAATGCTGG
TGTCACGACCACCCAAAATTCCACATCCTGAAGACAGGACAGCAGACACTGACCTGC
AGTGTGCAGGATAGTGAACCATAGAACTTGTCCTGGTATCGACACAGACCCAGGC
ATGGGGCTGAGGCGATACTTCATCGATTAAACACAGAGGATTCCCGCTCAAGGC
TGGAGGCTGCTCCCTCCAGACATCTGTGTACgagacctaa
```
AGTCCCCGATGGCTACAATGTATCCAGATCAAACACAGAGGATTTCCTCCGCTCACGGC
TGAGTGTCAGCTGCTCCCTCCACACTGATACgagacctaa

TRBV7-1
ggggatccacggtcgcaccatgggacacaggtctctgtggtgcagcatgtctctgtggtgcagatcacaGGTGCTGGAGTCTCCCAGTCCCTGAGACACAAGGTGACAAAGGAAGGAGATGTAGCTCT
CAGATATGATCCAATTTCCAGGTCTAATAAGCCCTTATTAGGTACCAGAGACCCTGG
GGCAGGGGCTGGAGTTTCTTTAACATTTACTTTCAAGAGCAAGATCAGCAACAAATCG
GGGCTTCCCGTGATCTGGTTCTCTCTTGACAGAGGTCTGAGGTGCTTGGGGATACAGATCCACGAGGACTCGGGCTG
TRBV7-2
ggggatccacggtcgcaccatgggacacaggtctctgtggtgcagcatgtctctgtggtgcagatcacaGGAGCTGGAGTCTCCCAGTCCCCCAGTAACAAGGTCAACAGAGAAGGAAAGGATGTAGAGCT
CAGGTGTGATCCAATTTCCAGGTCTAATAAGCCCTTATTAGGTACCAGAGACCCTGG
GGCAGGGGCTGGAGTTTCTTTAACATTTACTTTCAAGAGCAAGATCAGCAACAAATCG
GGGCTGCCAGTGCACTGGTTCTCTCTGCACAGAGGTCTGAGGTGCTTGGGGATACAGATCCACGAGGACTCGGGCTG
TRBV7-3
ggggatccacggtcgcaccatgggacacaggtctctgtggtgcagcatgtctctgtggtgcagatcacaGGTGCTGGAGTCTCCCAGAAGGCTAACAAGGTCAACAGAGAAGGAAATATGTAGAGC
TCAGGTGTGATCCAATTTCCAGGTCTAATAAGCCCTTATTAGGTACCAGACAAAGGACCTGG
GGCAGGGGCAAGGACGTTATTTTTCATCTCTCTCAGCCAGAAGGAGACGGGACTCGGGCTG
TRBV7-4
ggggatccacggtcgcaccatgggacacaggtctctgtggtgcagcatgtctctgtggtgcagatcacaGGTGCTGGAGTCTCCCAGAAGGCTAACAAGGTCAACAGAGAAGGAAATATGTAGAGC
TCAGGTGTGATCCAATTTCCAGGTCTAATAAGCCCTTATTAGGTACCAGACAAAGGACCTGG
GGCAGGGGCAAGGACGTTATTTTTCATCTCTCTCAGCCAGAAGGAGACGGGACTCGGGCTG
CAGGGCTCAGAGGTTCTGACTTACTCCCCAGAGGTGATGCTCAACGAGACAAATCAGG
GCGGCCCAGTGGTCGGTTCTCTGCAGAGAGGCCTGAGAGATCCGTCTCCACTCTGA
AGATCCAGCGCACAGAGCACAGGGGACTCAGCTGTGTATgagacctaa

TRBV7-6

ggggatcccaacggtgcaccatatgggaccagatctctatgtgcctggtctctctgtcttagggacagatcacacaGGTGCTG
GAGTCTCAGTCTCAGGTTACAGTCCAAAAGTACAAGAGGGGACAGGATGTAAGTCTC
AGGTGAGATCAATTTCGCGATGTCATGTAACCTCTTATGTTAGTACCGACAGCCTCGGG
CAGGGCCCAGAGGTTTCTGACTTACTTCAATTATGAAGCCCAACAAGACAAATCAGG
GCTGCCCAATGATCGGTTCTCTGCAGAGAGGCCTGAGGGATCCATCTCCACTCTGA
CGATCCACCGCAGACAGACAGCGGGGACTCAGCCCATGTGTGTgagacctaa

TRBV7-7

ggggatcccaacggtgcaccatatgggaccagatctctatgtgcctggtctctctgtcttagggacagatcacacaGGTGCTG
GAGTCTCAGTCTCAGGTTACAGTCCAAAAGTACAAGAGGGGACAGGATGTAAGTCTC
AGGTGAGATCAATTTCGCGATGTCATGTAACCTCTTATGTTAGTACCGACAGCCTCGGG
CAGGGCCCAGAGGTTTCTGACTTACTTCAATTATGAAGCCCAACAAGACAAATCAGG
GCTGCCCAATGATCGGTTCTCTGCAGAGAGGCCTGAGGGATCCATCTCCACTCTGA
ACGATTCACCGCAGACAGACAGCGGGGACTCAGCCCATGTGTGTgagacctaa

TRBV7-8

ggggatcccaacggtgcaccatatgggaccagatctctatgtgcctggtctctctgtcttagggacagatcacacaGGTGCTG
GAGTCTCAGTCTCAGGTTACAGTCCAAAAGTACAAGAGGGGACAGGATGTAAGTCTC
AGGTGAGATCAATTTCGCGATGTCATGTAACCTCTTATGTTAGTACCGACAGCCTCGGG
CAGGGCCCAGAGGTTTCTGACTTACTTCAATTATGAAGCCCAACAAGACAAATCAGG
GCTGCCCAATGATCGGTTCTCTGCAGAGAGGCCTGAGGGATCCATCTCCACTCTGA
AGATCCACCGCAGACAGACAGCGGGGACTCAGCCCATGTGTGTgagacctaa

TRBV7-9

ggggatcccaacggtgcaccatatgggaccagatctctatgtgcctggtctctctgtcttagggacagatcacacaGGTGCTG
GAGTCTCAGTCTCAGGTTACAGTCCAAAAGTACAAGAGGGGACAGGATGTAAGTCTC
AGGTGAGATCAATTTCGCGATGTCATGTAACCTCTTATGTTAGTACCGACAGCCTCGGG
CAGGGCCCAGAGGTTTCTGACTTACTTCAATTATGAAGCCCAACAAGACAAATCAGG
GCTGCCCAATGATCGGTTCTCTGCAGAGAGGCCTGAGGGATCCATCTCCACTCTGA
AGATCCACCGCAGACAGACAGCGGGGACTCAGCCCATGTGTGTgagacctaa

103
CAGGTGTGATCCAAATTTCTGAACACAACCGCCTTTATTTGGTACCGACAGACCCTGG
GGCAGGGCCCAGAGTTTCTGACTTTACTTTCCAGAATGAAGCTCAACTAGAAAAATCA
AGGCTGCTCAGTGTACGGTTCTCTCTGACAGAGACCCTAAGGGGATCTTTTCTCCACCTTG
GAGATCCAGCGCACAGAGCAGGGGGACTCGGCCATGTATgagacctaa

TRBV9
gggatccacccgtgcaccatgggtctcctctctgctgttgctttttttctctgtgaggcagccagtGATTCTGG
AGTCAACAAAAACCCAAAGCACCCTGATCACAGCAACTGG GCAGAGG TACGCTTG
AGATGCTCCCCTAGGTCTGGAGATCTCTCTGTGTACTGGTACCAAACAGACCTGGAC
CAGGGCCTCCAGTTCCTCTCATTGAGATTATAAATGGGAGAAGAGAGAGCAAAAGGAA
ACATTCTTGAAAGGATTTGCTCCACACAGATTTCCCCTGACTTGCACTCTGAACTAAACC
TGAGCTCTCTGGAGCTGGGGACTCGGCAGCTTTGTATgagacctaa

TRBV10-1
gggatccacccgtgcaccatgggtctcctctctgctgttgctttttttctctgtgaggcagccagtGATGCTGA
AATCACCAGAGCCCAAGACAAAGTCACAGACAGAGAAGGCAGGTGACCTTG
GCGTGTCACCAGACTTGGAACCACAATAATGTTTCGTAACAGACCTGGG
ACATGCGGTGAGCTGATCCATTACCAGTCTCAACGAGATATTTAAGGAG
AAGTCTCAGACGTTCAGTCTCTAGATCAACACAGAGAAGCCTCTCCCTCCTACT
CTGAGTGCCTGCTGCCCTCTCCACGACATCTGTATgagacctaa

TRBV10-2
gggatccacccgtgcaccatgggtctcctctctgctgttgctttttttctctgtgaggcagccagtGATGCTGG
AATCACCAGAGCCCAAGATACAAAGTGACAGACAGAGAAGGCAGGTGACCTTG
ATGTGTCACCAGACTTGGAGCCACAGCTATATGTTTCGTAACAGACCTGGG
ACATGCGGTGAGCTGATCCATTACCAGTCTCAACGAGATATTTAAGGAG
AAGTCCCAAGATGGCTATTTGCTCTCCAGATCCAGACAGAGAATTTCCCCTCCTACT
TGGAGTGCCTGCTGCCCTCTCCACGACATCTGTATgagacctaa

TRBV10-3
```plaintext
TRBV11-1

`ggggatccaccgcctgcaccacatgggcacaaggttgttcttctatgtggccctttgtctcctgtggacaggacacatgGATGCTGG
AATCACCCAGGCCAGGACAGACAGGTTCACAGAGACAGGACATTACTATGTGATGGTACATTGGACACTCTTCGAGATCTTGG
TGGTGTGACTCAGAGATGTGGTACTGAGACTGACCAAGACAGAGGGGATGTCAGTGCCTGACCAGCACCAGCTGACGACAGACG`

TRBV11-2

`ggggatccaccgcctgcaccacatgggcacaaggttgttcttctatgtggccctttgtctcctgtggacaggacacatgGATGCTGG
AATCACCCAGGCCAGGACAGACAGGTTCACAGAGACAGGACATTACTATGTGATGGTACATTGGACACTCTTCGAGATCTTGG
TGGTGTGACTCAGAGATGTGGTACTGAGACTGACCAAGACAGAGGGGATGTCAGTGCCTGACCAGCACCAGCTGACGACAGACG`

TRBV11-3

`ggggatccaccgcctgcaccacatgggcacaaggttgttcttctatgtggccctttgtctcctgtggacaggacacatgGATGCTGG
AATCACCCAGGCCAGGACAGACAGGTTCACAGAGACAGGACATTACTATGTGATGGTACATTGGACACTCTTCGAGATCTTGG
TGGTGTGACTCAGAGATGTGGTACTGAGACTGACCAAGACAGAGGGGATGTCAGTGCCTGACCAGCACCAGCTGACGACAGACG`
```
TRBV12-3

```
gggatccacccggtgccccacaggccaggccactggacctggaccttctgcgtgtgtgcctctttgcacattcggtagcaagcatacaGATGCTGG
AGTTATCCAGTCACCCCCGCACGAGGTGACAGAGATGGGACAAGAAGTGACTCTGA
GATGGATCAATAAACGACACCTCCTTTCTGGTGACAGACAGACCACTGATG
GCGGGGACCTGAGGGGTGTTCTCTTCAGCTCATTTTCAAAACACGTTCCGATAGATGTTTCAG
GATGCCGAAGGATCGATTTCTCAGCTAAAGGCTCAAATGCTATTTCATCTCCACTCCTGAA
GATCCAGCCCTCAGAACCAGGACTCAGCTGCTGTGATCGAGCAGACTTAA
```

TRBV12-4

```
GGGATCCACCCGACGTACCAGTCCGCCACCGACAGGTGACAGAGATGGGACAAGAAGTGACTCTGA
GATGGATCAATAAACGACACCTCCTTTCTGGTGACAGACAGACCACTGATG
GCGGGGACCTGAGGGGTGTTCTCTTCAGCTCATTTTCAAAACACGTTCCGATAGATGTTTCAG
GATGCCGAAGGATCGATTTCTCAGCTAAAGGCTCAAATGCTATTTCATCTCCACTCCTGAA
GATCCAGCCCTCAGAACCAGGACTCAGCTGCTGTGATCGAGCAGACTTAA
```

TRBV12-5

```
GGGATCCACCCGACGTACCAGTCCGCCACCGACAGGTGACAGAGATGGGACAAGAAGTGACTCTGA
GATGGATCAATAAACGACACCTCCTTTCTGGTGACAGACAGACCACTGATG
GCGGGGACCTGAGGGGTGTTCTCTTCAGCTCATTTTCAAAACACGTTCCGATAGATGTTTCAG
GATGCCGAAGGATCGATTTCTCAGCTAAAGGCTCAAATGCTATTTCATCTCCACTCCTGAA
GATCCAGCCCTCAGAACCAGGACTCAGCTGCTGTGATCGAGCAGACTTAA
```

TRBV13

```
GGGATCCACCCGACGTACCAGTCCGCCACCGACAGGTGACAGAGATGGGACAAGAAGTGACTCTGA
GATGGATCAATAAACGACACCTCCTTTCTGGTGACAGACAGACCACTGATG
GCGGGGACCTGAGGGGTGTTCTCTTCAGCTCATTTTCAAAACACGTTCCGATAGATGTTTCAG
GATGCCGAAGGATCGATTTCTCAGCTAAAGGCTCAAATGCTATTTCATCTCCACTCCTGAA
GATCCAGCCCTCAGAACCAGGACTCAGCTGCTGTGATCGAGCAGACTTAA
```

106
ATCATTCTGAAGTGAACATGAGCTCTGTTGGAGCTGGGGGACTCAGCCCTGTACgagaccta

TRBV14

TRBV15

TRBV16

TRBV17
CAAGAAATGAGTTGCTGATTTCCCTCCATGATCACCACACATGAGTTGATTCCAGG
GATGCCAAGGAACGATTCACAGCAGTAAAGACCTAACGGGAAGTCCTTCCACGCTGA
AGATCCATCCCGAGGAGGAGGAGCTCAGCGCCTGATTagacctaa

**TRBV19**

```
AATCACTCAGTCCCCAAAGTACCTGTTCAGAAAGGAAGGACAGAATGTGACCCTGA
```

**TRBV18**

```
GCGTCATGCAGAACCCAAAGACATCTGTTCAGGAGAGGACAGAGGACAGGACTCTCAG
```

**TRBV20-1**

```
AAATCCGAGGAGGAGATCGGCAAGAGTTTGGTCAAAGGAAAAGGACAGAAAACAAAGATG
```

**TRBV23-1**

```
AGTCACTCCAAGACTCCAGGACATTTGGTCAAAGGAAAAGGACAGAAAACAAAGATG
```
GATTGTACCACCCCGAAAAAGGACATCACTTTTGTATTGATATCAACAGAATCAGAA
TAAAGAGTTTAATGCTTTTGATTTCCTTTCAGAATGAACAAGTTCTTCAAGAAACGGA
GATGCACAAGAAGCGATTCTCATCTCAATGCCCAAGAAGCGACACCTGTGAGGCTTGG
CAATCCTGGTCCTCAAGAACCAGGAGACACGACACTGTGATgagacctaa

TRBV24-1

TRBV25-1

TRBV27

TRBV28
Having thus described in detail preferred embodiments of the present invention, it is to be understood that the invention defined by the above paragraphs is not to be limited to particular details set forth in the above description as many apparent variations thereof are possible without departing from the spirit or scope of the present invention.
WHAT IS CLAIMED IS:

1. A method of identifying at least one subject specific T cell receptor pair expressed on a single T cell and preparing a subject-specific immunogenic composition that includes the at least one T cell receptor, wherein the subject has a tumor and the at least one subject-specific T cell receptors are specific to the subject and the subject's tumor, said method comprising:
   (a) identifying TCR pairs from single tumor specific T cells from the subject; and
   (b) formulating the subject-specific immunogenic composition, wherein the immunogenic composition comprises one or more T cells each expressing a single tumor-specific TCR pair identified from the subject.

2. A method of identifying at least one subject specific T cell receptor pair expressed on a single T cell and preparing a subject-specific immunogenic composition that includes the at least one T cell receptor, wherein the subject has a tumor and the at least one subject-specific T cell receptors are specific to the subject and the subject's tumor, said method comprising:
   (a) identifying TCRs from single T cells obtained from the subject, wherein the T cells are specific to subject-specific neoantigens;
   (b) formulating the subject-specific immunogenic composition, wherein the immunogenic composition comprises one or more T cells each expressing a single tumor-specific TCR pair identified from the subject.

3. A method of identifying at least one subject specific T cell receptor pair expressed on a single T cell and preparing a subject-specific immunogenic composition that includes the at least one T cell receptor, wherein the subject has a tumor and the at least one subject-specific T cell receptors are specific to the subject and the subject's tumor, said method comprising:
   (a) immunizing a subject with a vaccine that includes at least one subject specific neoantigen;
   (b) identifying TCRs from single T cells obtained from the subject, wherein the T cells are specific to a subject-specific neoantigen included in the vaccine; and
(c) formulating the subject-specific immunogenic composition, wherein the immunogenic composition comprises one or more T cells each expressing a single tumor-specific TCR pair identified from the subject.

4. A method of identifying at least one subject specific T cell receptor pair expressed on a single T cell and preparing a subject-specific immunogenic composition that includes the at least one T cell receptor, wherein the subject has a tumor and the at least one subject-specific T cell receptors are specific to the subject and the subject's tumor, said method comprising:
   (a) stimulating PBMCs isolated from the subject with at least one subject-specific neoantigen;
   (b) identifying TCRs from T cells present within the PMBCs, wherein the T cells are specific to the at least one subject-specific neoantigen; and
   (c) formulating the subject-specific immunogenic composition, wherein the immunogenic composition comprises one or more T cells each expressing a single tumor-specific TCR pair identified from the subject.

5. The method of claim 1, wherein the T cells used for identifying TCR pairs are selected by detecting stimulation by autologous tumor.

6. The method of any of claims 1-4, wherein the T cells used for identifying TCR pairs are selected by detecting stimulation by the subject-specific neoantigens bound to patient-specific HLA molecules.

7. The method of any of claims 1-6, wherein the identifying TCRs comprises single cell sequencing.

8. The method of any of claims 1-7, wherein the T cell of the immunogenic composition is activated.

9. The method of any of claims 1-8, wherein the formulating includes cloning the T cell receptor.
10. The method of claim 9, wherein the formulating comprises expressing the subject specific TCR in the T cell.

11. The method of claim 10, wherein the expressing comprises introduction of a nucleic acid into the T cell.

12. The method of claim 11, wherein the nucleic acid is a vector.

13. The method of claim 12, wherein the vector is a plasmid.

14. The method of claim 13, wherein the vector is a viral vector.

15. The method of claim 14, wherein the viral vector is selected from the group consisting of a lentivirus, adenovirus, and adeno-associated virus.

16. A method of identifying at least one subject specific TCR expressed on a single T cell, wherein the subject has a tumor and the at least one subject-specific TCRs are specific to the subject and the subject's tumor, said method comprising:
   (a) identifying from single cells, subject and tumor specific TCRs;
   (b) selecting the TCRs that target autologous tumor cells; and
   (c) formulating a subject specific immunotherapy that includes the selected TCRs.

17. The method of claim 16, wherein selecting TCRs comprises:
   (a) cloning the TCRs;
   (b) expressing the TCRs in T cells, wherein the T cells are activated; and
   (c) incubating tumor cells from the subject with the T cells expressing the TCRs.

18. The method of claim 16, wherein selecting TCRs comprises:
19. The method of claim 16, wherein selecting TCR pairs comprises:
(a) cloning the TCR pairs;
(b) generating soluble recombinant TCR pairs; and
(c) incubating at least one subject specific neoepitope bound to HLA with the soluble recombinant TCR pairs, wherein binding can be monitored.

20. A method of identifying a plurality of at least 4 subject-specific peptides and preparing a subject-specific immunogenic composition that upon administration presents the plurality of at least 4 subject-specific peptides to the subject's immune system, wherein the subject has a tumor and the subject-specific peptides are specific to the subject and the subject's tumor, said method comprising:
(i) identifying, including through
nucleic acid sequencing of a sample of the subject's tumor and
nucleic acid sequencing of a non-tumor sample of the subject,
a plurality of at least 4 tumor-specific non-silent mutations not present in the non-tumor sample;
(ii) identifying, including through
single cell sequencing of a subject's TCR repertoire,
the quantity and identity of TCR pairs expressed in a subject; and
(iii) selecting from the identified non-silent mutations the plurality of at least 4 subject-specific peptides, each having a different tumor neo-epitope that is an epitope specific to the tumor of the subject, from the identified plurality of tumor specific mutations,

wherein each neo-epitope is an expression product of a tumor-specific non-silent mutation not present in the non-tumor sample, each neo-epitope binds to a HLA protein of the subject, each neo-epitope binds to a subject specific TCR pair and selecting includes
determining binding of the subject-specific peptides to the HLA protein, and
determining neoantigen binding to an identified TCR pair,
and

(iv) formulating the subject-specific immunogenic composition for administration to
the subject so that upon administration the plurality of at least 4 subject-specific peptides are
presented to the subject's immune system,

wherein the selecting or formulating comprises at least one of:

including in the subject-specific immunogenic composition a subject-specific
peptide that includes an expression product of an identified neo-ORF, wherein a neo-ORF is a
tumor-specific non-silent mutation not present in the non-tumor sample that creates a new open
reading frame, and

including in the subject-specific immunogenic composition a subject-specific
peptide that includes an expression product of an identified point mutation and has a determined
binding to the HLA protein of the subject with an IC50 less than 500 nM,

whereby, the plurality of at least 4 subject-specific peptides are identified, and the subject-
specific immunogenic composition that upon administration presents the plurality of at least 4
subject-specific peptides to the subject's immune system, wherein the subject-specific peptides
are specific to the subject and the subject's tumor, is prepared.

21. A method of treating a subject in need thereof comprising:
   (a) identifying the TCR pairs of single cells from the subject, wherein the
       TCR pairs bind to subject specific neoantigens;
   (b) isolating single cells that express TCR pairs that bind subject specific
       neoantigens;
   (c) expanding the isolated single cells; and
   (d) administering the expanded cells to the subject.

22. A method of treating a subject in need thereof comprising:
   (a) identifying the TCR pairs of single cells from the subject, wherein the
       TCR pairs bind to subject specific neoantigens;
   (b) cloning the TCR pairs; and
   (c) administering activated T cells expressing TCR pairs that bind to subject
       specific neoantigens to the subject.
23. The method of claim 7, wherein single cell sequencing comprises:
   (a) emulsifying (i) single cells of the population of cells and (ii) deformable beads in a population of approximately uniformly-sized aqueous droplets at respective frequencies of less than about 0.1 cells/droplet on average and at least about 0.5 beads/droplet or greater on average, wherein the deformable beads comprise sequencing-compatible barcodes;
   (b) lysing a single cell within a droplet;
   (c) performing RT-PCR upon the droplet such that the RT-PCR amplified cDNA of the single cell is bound to a sequencing-compatible barcode;
   (d) obtaining sequence for at least two cDNAs and the bound sequencing-compatible barcode, wherein the bound sequencing-compatible barcode identifies the at least two cDNA sequences as of a single cell, thereby obtaining nucleic acid sequence information for two or more transcripts of a single cell within a population of cells.

24. The method of claim 23, wherein the two or more transcripts comprise TCRα and TCRρ.

25. The method of claim 23, wherein the population of approximately uniformly-sized aqueous droplets comprises cells at less than about 0.1 cells/droplet on average.

26. The method of claim 23, wherein the population of approximately uniformly-sized aqueous droplets comprises at least about 0.7 beads/droplet or greater on average.

27. The method of claim 23, wherein the population of approximately uniformly-sized aqueous droplets comprises an amount of cells selected from the group consisting of at least 20,000 cells, at least 50,000 cells, at least 100,000 cells, at least 500,000 cells and at least 1,000,000 cells.

28. The method of claim 23, wherein the deformable beads are hydrogel beads.
29. The method of claim 23, wherein the step of obtaining sequence comprises massively parallel sequencing.

30. The method of claim 23, wherein the amplified cDNAs of more than one droplet are combined before the step of obtaining sequence.

31. The method of claim 23, wherein sequence of two or more amplified cDNAs independently bound to different barcodes is obtained.

32. The method of claim 23, wherein lysing the single cell within a droplet comprises contacting the single cell with detergent.

33. The method of claim 23, wherein each droplet consists of a 30-50 pL volume.

34. The method of claim 7, wherein single cell sequencing comprises sorting T cells into single cells and sequencing the single cells.

35. A pharmaceutical composition comprising a non-naturally occurring T cell expressing a subject specific TCR pair, wherein the subject specific TCR pair was identified by sequencing of single cells from the subject.

36. The composition of claim 35, wherein the subject was immunized with at least one subject specific neoantigen.

37. The composition of claim 35, wherein the subject specific TCR binds to autologous tumors.

38. The composition of claim 35, wherein the subject specific TCR binds to a subject specific neoantigen.
Antigen presenting cell

Peptide

HLA

TCR

T cell

α

β

FIG. 2
TCR single-cell sequencing data

Align TCR sequence to IMGT database

Get dominant clonotypes per cell

FIG. 4
### FIG. 5

<table>
<thead>
<tr>
<th>TRAV</th>
<th>TRAJ</th>
<th>CDR3α</th>
<th>TRBV</th>
<th>TRBD</th>
<th>TRJB</th>
<th>CDR3β</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>37</td>
<td>AGRLVDQGGKLIF</td>
<td>29-1</td>
<td>1</td>
<td>1-4</td>
<td>ASSIGLAGYEQYF</td>
</tr>
</tbody>
</table>
Reporter assay

1. Transduce TCR sequence into parental reporter cell line

2. Stimulate with APC pulsed to present neoantigen

3. Detect mCherry expression if TCR is specific for neoantigen

FIG. 7
Paired TCRαβ sequencing

TCR cloning and expression in TCR-deficient reporter cell line

Screen engineered T cell against antigen

EBNA3A-specific TCR

<table>
<thead>
<tr>
<th>TRAV</th>
<th>TRAJ</th>
<th>CDR3α</th>
<th>TRBV</th>
<th>TRBD</th>
<th>TRJB</th>
<th>CDR3β</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>23</td>
<td>AGRLVDQGGKLIF</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>ASSIGLAGEQYF</td>
</tr>
</tbody>
</table>

FIG. 8
FIG. 9
Donor PBMC

Stimulate with CEF peptide pool and IFN-γ to isolate CEF-specific T cells

Paired TCRαβ sequencing

TCR cloning and expression in TCR-deficient reporter cell line

Screen engineered T cell against antigen

Identify CEF antigen that elicit T cell response

FIG. 11
Fig. 12
FIG. 13
Pipeline

Patient T cells

Paired TCRαβ sequencing

TCRαβ cloning

TCRαβ expression in TCRαβ-deficient reporter cell line

Screen engineered T cells against antigen panel

FIG. 16
Patient PBMC

Stimulate with neoantigen peptide pool and IFN-γ catch to isolate neoantigen-specific T cells

Paired TCRαβ sequencing

TCR cloning and expression in TCR-deficient reporter cell line

Screen engineered T cell against antigen

Identify neoantigen that elicit T cell response

FIG. 17
FIG. 18A

<table>
<thead>
<tr>
<th>Autologous target cells:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLL-assoc'd Ag tumor</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PHA blasts</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fibroblasts</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

FIG. 18B

Vaccinated patients

+GVHD patients

Controls (no VAX/GvHD)

Mean TPN p/10^6 CD3+ cells

Days after transplantation

FIG. 18C

<table>
<thead>
<tr>
<th>Patient #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days after HCT</td>
<td>130</td>
<td>148</td>
<td>66</td>
<td>172</td>
</tr>
</tbody>
</table>

Number of T cell clones (%)

- CLL-associated antigen
- Hematopoietically-restricted antigen
- Broadly expressed antigen
- No reactivity

Total

<table>
<thead>
<tr>
<th>Number</th>
<th>Patient #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 (17)</td>
<td>Patient # 1</td>
<td>130</td>
<td>148</td>
<td>66</td>
<td>172</td>
</tr>
<tr>
<td>15 (33)</td>
<td>Patient # 2</td>
<td>130</td>
<td>148</td>
<td>66</td>
<td>172</td>
</tr>
<tr>
<td>37 (15)</td>
<td>Patient # 3</td>
<td>130</td>
<td>148</td>
<td>66</td>
<td>172</td>
</tr>
<tr>
<td>4 (18)</td>
<td>Patient # 4</td>
<td>130</td>
<td>148</td>
<td>66</td>
<td>172</td>
</tr>
</tbody>
</table>

- CLL cells
- PHA blasts
- Fibroblasts

SUBSTITUTE SHEET (RULE 26)
FIG. 19
A Proof of concept: TCR profiling against single, known epitopes

B Deconvolution of complex multiple, unknown epitopes from patient samples

Targets:
- A auto-CLL
- B auto-PHA blasts
- C auto-fibroblasts

Subtracted list of TCRs: [A-(B+C)]

Transduction of a unique TCR vector

TCR-expressing BW reporter cells

Test for:
- Cytolytic activity
- Compare repertoires: PBMC vs. BM
- Track TCR kinetics

FIG. 21
FIG. 22
FIG. 23
FIG. 24
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/10 A61K35/17 A61K39/00

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, CHEM ABS Data, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>CARSTEN LINNEMANN ET AL: &quot;High-throughput identification of anti gen-specific TCRs by TCR gene capture&quot;, NATURE MEDICINE, vol. 19, no. 11, 13 October 2013 (2013-10-13), pages 1534-1541, XP055190765, ISSN: 1078-8956, DOI: 10.1038/nm.3359 page 1538, right-hand column, paragraph 2 page 1539, left-hand column, last paragraph - right-hand column, paragraph 1 page 1540, left-hand column, paragraph 2</td>
<td>1-38</td>
</tr>
</tbody>
</table>

Date of actual completion of the international search

21 April 2016

Date of mailing of the international search report

28/04/2016

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk
Tel. (+31-70) 340.2040, Fax: (+31-70) 340.3016

Authorized officer

Mata Viciante, Teresa

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered to be obvious; invention cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to person skilled in the art

"Z" document member of the same patent family