

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2013327423 B2**

(54) Title
Immunocompetence assessment by adaptive immune receptor diversity and clonality characterization

(51) International Patent Classification(s)
C12Q 1/68 (2006.01)

(21) Application No: **2013327423** (22) Date of Filing: **2013.10.01**

(87) WIPO No: **WO14/055561**

(30) Priority Data

(31) Number	(32) Date	(33) Country
61/708,534	2012.10.01	US

(43) Publication Date: **2014.04.10**

(44) Accepted Journal Date: **2017.06.22**

(71) Applicant(s)
Adaptive Biotechnologies Corporation;Memorial Sloan-Kettering Cancer Center

(72) Inventor(s)
Robins, Harlan;Rubinstein, Julie;Emerson, Ryan;Yuan, Jianda

(74) Agent / Attorney
Pizzeys Patent and Trade Mark Attorneys Pty Ltd, PO Box 291, WODEN, ACT, 2606, AU

(56) Related Art
US 20110207134 A1



- (51) International Patent Classification:
C12Q 1/68 (2006.01)
- (21) International Application Number:
PCT/US2013/062925
- (22) International Filing Date:
1 October 2013 (01.10.2013)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/708,534 1 October 2012 (01.10.2012) US
- (71) Applicants: **ADAPTIVE BIOTECHNOLOGIES CORPORATION** [US/US]; 1551 Eastlake Avenue East, Suite 200, Seattle, WA 98102 (US). **MEMORIAL SLOAN-KETTERING CANCER CENTER** [US/US]; 1275 York Avenue, New York, NY 10065 (US).
- (72) Inventors: **ROBINS, Harlan**; c/o Adaptive Biotechnologies Corporation, 1551 Eastlake Avenue East, Suite 200, Seattle, WA 98102 (US). **RUBINSTEIN, Julie**; c/o Adaptive Biotechnologies Corporation, 1551 Eastlake Avenue East, Suite 200, Seattle, WA 98102 (US). **EMERSON, Ryan**; c/o Adaptive Biotechnologies Corporation, 1551 Eastlake Avenue East, Suite 200, Seattle, WA 98102 (US). **YUAN, Jianda**; c/o Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065 (US).
- (74) Agents: **LEE, Michelle** et al.; Fenwick & West LLP, Silicon Valley Center, 801 California Street, Mountain View, CA 94041 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

— with international search report (Art. 21(3))

(54) Title: IMMUNOCOMPETENCE ASSESSMENT BY ADAPTIVE IMMUNE RECEPTOR DIVERSITY AND CLONALITY CHARACTERIZATION

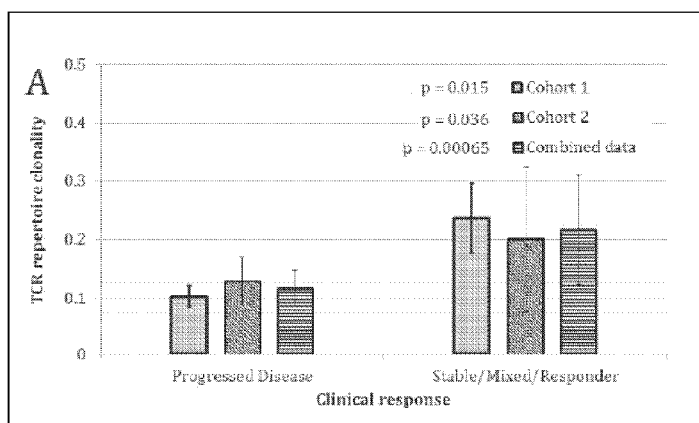


FIGURE 11A

(57) Abstract: Disclosed are methods for determining the immunological status of the adaptive immune system of a subject by identifying and quantifying rearranged DNA (and/or subsequently transcribed RNA) sequences encoding T cell receptor (TCR) and/or immunoglobulin (IG) polypeptides, in a lymphoid DNA-containing sample from the subject. TCR and/or IG sequence diversity and sequence distribution permit immunocompetence and immune repertoire assessment and reflect the degree of T cell or B cell clonality and clonal expansion in the sample. Methods for stratifying patient populations on the basis of immunocompetence including likelihood of responding to immunotherapy are also described.

WO 2014/055561 A1

IMMUNOCOMPETENCE ASSESSMENT BY ADAPTIVE IMMUNE RECEPTOR DIVERSITY AND CLONALITY CHARACTERIZATION

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 61/708,534, filed October 1, 2012, the entire disclosure of which is hereby incorporated by reference in its entirety for all purposes.

STATEMENT OF GOVERNMENT INTEREST

[0002] This invention was made with government support under Grant No. RHL106868A awarded by the National Heart, Blood and Lung Institute of the National Institutes of Health. The government has certain rights in this invention.

BACKGROUND

Technical Field

[0003] The present disclosure relates generally to assessment of immunocompetence of a subject's adaptive immune system by highly sensitive, high throughput DNA sequence-based quantification of the diversity and frequency of occurrence (*e.g.*, clonal expansion) of adaptive immune cells having a particular rearranged T cell receptor (TCR) or immunoglobulin (IG or Ig) encoding gene sequence. Information about the immunological status of a subject or a population of subjects can be used, for example, to characterize an individual or to stratify a patient population with respect to ability to mount an immune response or likelihood of responding to immunotherapy or the likelihood of developing an immune-mediated side effect in response to said therapy, or to otherwise inform a course of clinical immunotherapy management.

Description of the Related Art

[0004] The adaptive immune system protects higher organisms against infections and other pathological events that can be attributable to foreign substances, using adaptive immune receptors, the antigen-specific recognition proteins that are expressed by hematopoietic cells of the lymphoid lineage and that are capable of distinguishing self from non-self molecules in the host. These lymphocytes can be found in the circulation and tissues of a host, and their recirculation between blood and the lymphatics has been described, including their extravasation via lymph node high endothelial venules, as well as at sites of infection, inflammation, tissue injury and other clinical insults. (*See, e.g., Stein et al., 2005 Immunol. 116:1-12; DeNucci et al., 2009 Crit. Rev. Immunol. 29:87-109; Marelli-Berg et al., 2010 Immunol. 130:158; Ward et al., 2009 Biochem. J. 418:13; Gonzalez et al., 2011 Ann. Rev.*

Immunol. 29:215; Kehrl *et al.*, 2009 *Curr. Top. Microb. Immunol.* 334:107; Steinmetz *et al.*, 2009 *Front. Biosci. (Schol. Ed.)* 1:13.)

[0005] Accordingly, the dynamic nature of movement by lymphocytes throughout a host organism is reflected in changes in the qualitative (*e.g.*, antigen-specificity of the clonally expressed adaptive immune receptor (immunoglobulin or T cell receptor), T cell versus B cell, T helper (T_h) cell versus T regulatory (T_{reg}) cell, effector T cell versus memory T cell, etc.) and quantitative distribution of lymphocytes among tissues, as a function of changes in host immune status.

[0006] The adaptive immune system employs several strategies to generate a repertoire of T- and B-cell antigen receptors with sufficient diversity to recognize the universe of potential pathogens. B lymphocytes mature to express antibodies (immunoglobulins, Igs) that occur as heterodimers of a heavy (H) and light (L) chain polypeptide, while T lymphocytes express heterodimeric T cell receptors (TCR). The ability of T cells to recognize the universe of antigens associated with various cancers or infectious organisms is conferred by its T cell antigen receptor (TCR), which is a heterodimer comprising an α (alpha) chain and a β (beta) chain, or a γ (gamma) and a δ (delta) chain. The proteins that make up these chains are encoded by DNA that employs a unique mechanism for generating the tremendous diversity of the TCR. This multi-subunit immune recognition receptor associates with the CD3 complex and binds to peptides presented by the major histocompatibility complex (MHC) class I and II proteins on the surface of antigen-presenting cells (APCs). Binding of TCR to the antigenic peptide on the APC is a central event in T cell activation, which occurs at an immunological synapse at the point of contact between the T cell and the APC.

[0007] Each TCR peptide contains variable complementarity determining regions (CDRs), as well as framework regions (FRs) and a constant region. The sequence diversity of $\alpha\beta$ T cells is largely determined by the amino acid sequence of the third complementarity-determining region (CDR3) loops of the α and β chain variable domains, which diversity is a result of recombination between variable (V_β), diversity (D_β), and joining (J_β) gene segments in the β chain locus, and between analogous V_α and J_α gene segments in the α chain locus, respectively. The existence of multiple such gene segments in the TCR α and β chain loci allows for a large number of distinct CDR3 sequences to be encoded. CDR3 sequence diversity is further increased by deletion and template-independent addition of nucleotides at the V_β - D_β , D_β - J_β , and V_α - J_α junctions during the process of TCR gene rearrangement. In this respect, immunocompetence is reflected in the diversity of TCRs.

[0008] The $\gamma\delta$ TCR is distinctive from the $\alpha\beta$ TCR in that it encodes a receptor that interacts closely with the innate immune system. TCR $\gamma\delta$, is expressed early in development, has specialized anatomical distribution, has unique pathogen and small-molecule specificities, and has a broad spectrum of innate and adaptive cellular interactions. A biased pattern of TCR γ V and J segment expression is established early in ontogeny as the restricted subsets of TCR $\gamma\delta$ cells populate the mouth, skin, gut, vagina, and lungs prenatally. Consequently, the diverse TCR γ repertoire in adult tissues is the result of extensive peripheral expansion following stimulation by environmental exposure to pathogens and toxic molecules.

[0009] Igs expressed by B cells are proteins consisting of four polypeptide chains, two heavy chains (H chains) and two light chains (L chains), forming an H₂L₂ structure. Each pair of H and L chains contains a hypervariable domain, consisting of a V_L and a V_H region, and a constant domain. The H chains of Igs are of several types, μ , δ , γ , α , and β . The diversity of Igs within an individual is mainly determined by the hypervariable domain. Similar to the TCR, the V domain of H chains is created by the combinatorial joining of the V_H, D_H, and J_H gene segments. Hypervariable domain sequence diversity is further increased by deletion and template-independent addition of nucleotides at the V_H-D_H, D_H-J_H, and V_H-J_H junctions during the process of Ig gene rearrangement. In this respect, immunocompetence is reflected in the diversity of Igs.

[0010] Quantitative characterization of adaptive immune cells based on the presence in such cells of functionally rearranged Ig and TCR encoding genes that direct productive expression of adaptive immune receptors has been achieved using biological samples from which adaptive immune cells can be readily isolated in significant numbers, such as blood, lymph or other biological fluids. In these samples, adaptive immune cells occur as particles in fluid suspension. *See, e.g., US 2010/0330571; see also, e.g., Murphy, Janeway's Immunobiology (8th Ed.), 2011 Garland Science, NY, Appendix I, pp. 717-762.*

[0011] The adaptive immune system has long been implicated as having a role in the recognition of cancer cells, and in the ensuing generation of an immune response to eliminate tumors (*e.g., Murphy, Janeway's Immunobiology (8th Ed.), 2011 Garland Science, NY, pp. 682-697; Pandolfi et al., 2011 Clin. Dev. Immunol. Article ID894704; Draghiciu et al., 2011 Clin. Dev. Immunol. Article ID439053*). Variability among the antigenic profiles presented by different cancer cells, and variability in the robustness of different individuals' anti-tumor immune responses, have led to a number of widely different clinical approaches that are generally referred to as cancer immunotherapy, by which efforts are made to induce, recruit,

enhance or otherwise potentiate the adaptive immune response, in this context, by encouraging anti-tumor immunity.

[0012] Such immunotherapeutic approaches represent preferable alternatives to conventional cancer therapies, that can be non-cancer cell-specific and can involve harsh cytotoxic regimens such as radiation and chemotherapy. Cancer immunotherapy is sometimes administered to patients who also receive chemotherapy and/or radiation therapy, but because chemotherapy and radiation are particularly cytotoxic toward dividing cells, whilst immunocyte proliferation can be a cardinal feature of many immune responses, such approaches can counterproductively compromise the adaptive immune system and thus can be accompanied by difficulties in arriving at effective therapeutic regimens.

[0013] Accordingly, it would be desirable to have a detailed understanding of a patient's immunological status in order to design an effective immunotherapy regimen, in the context of cancer and also, for example, in cases of hematopoietic cell transplant or solid organ transplant (*e.g.*, to detect and appropriately treat immunological rejection of a graft), in the treatment of microbial infections, and to optimize responses to vaccines. The presently-described embodiments address these needs and provide other related advantages.

SUMMARY OF INVENTION

[0014] According to certain embodiments of the present invention, a method is provided for determining an immunological status of a test subject, comprising obtaining nucleic acid sequence information generated from one or more samples comprising nucleic acids from lymphoid cells of said test subject, wherein said nucleic acid sequence information comprising sequences for a plurality of unique rearranged nucleic acid sequences, each of said plurality of unique rearranged nucleic acid sequences encoding an AIR polypeptide, said one or more samples obtained from said test subject at one or more time points for said one or more samples, using said nucleic acid sequence information, determining a total number of observed rearranged sequences in said sample; determining a total number of unique rearranged DNA sequences in said sample; quantifying an AIR sequence diversity score for said one or more samples based on said total number of unique rearranged DNA sequences; quantifying an AIR sequence distribution score for said one or more samples by calculating a frequency of occurrence of each unique rearranged DNA sequence as a percentage of said total number of observed rearranged sequences in said one or more samples; and determining a test subject rating score for determining said immunological status of said test subject based on said AIR sequence diversity score and said AIR sequence distribution score of said test subject.

[0015] In some embodiments, the method includes comparing said test subject rating scores for said one or more samples to a second set of control subject rating scores obtained from samples from a control subject and determining said immunological status of said test subject at said one or more time points, wherein said test subject is determined to have a immunological status at said one or more time points that is different from an immunological status of said control subject, if a difference between said test subject rating score and said control subject rating score is statistically significant and wherein said test subject is determined to have the same immunological status of said control subject if there is no statistically significant difference between said test subject rating score and said control subject rating score.

[0016] In certain embodiments, the nucleic acids comprise genomic DNA. In other embodiments, the nucleic acids comprise cDNA. In some embodiments, the nucleic acids comprise messenger RNA.

[0017] In some embodiments, the methods of the invention also include steps for quantifying an AIR sequence distribution score for said subject comprising determining a number of unique rearranged AIR sequences that have a combined frequency of occurrence of up to 50% of the total number of observed rearranged sequences in said sample; and characterizing a AIR sequence distribution score as a low score if the number of unique rearranged AIR sequences that have a combined frequency of occurrence of up to 50% of the total number of observed rearranged sequences in said sample is less than or equal to a predetermined threshold.

[0018] In one embodiment, the method includes quantifying an AIR sequence distribution score for said subject comprising determining a number of unique rearranged AIR sequences that have a combined frequency of occurrence of up to 40% of the total number of observed rearranged sequences in said sample; and characterizing a AIR sequence distribution score as a low score if the number of unique rearranged AIR sequences that have a combined frequency of occurrence of up to 40% of the total number of observed rearranged sequences in said sample is less than or equal to a predetermined threshold.

[0019] In another embodiment, the method includes quantifying an AIR sequence distribution score for said subject comprising determining a number of unique rearranged AIR sequences that have a combined frequency of occurrence of up to 30% of the total number of observed rearranged sequences in said sample; and characterizing a AIR sequence distribution score as a low score if the number of unique rearranged AIR sequences that have

a combined frequency of occurrence of up to 30% of the total number of observed rearranged sequences in said sample is less than or equal to a predetermined threshold.

[0020] In yet another embodiment, the method includes quantifying an AIR sequence distribution score for said subject comprising determining a number of unique rearranged AIR sequences that have a combined frequency of occurrence of up to 20% of the total number of observed rearranged sequences in said sample; and characterizing a AIR sequence distribution score as a low score if the number of unique rearranged AIR sequences that have a combined frequency of occurrence of up to 20% of the total number of observed rearranged sequences in said sample is less than or equal to a predetermined threshold.

[0021] In other embodiments, the method comprises quantifying an AIR sequence distribution score comprising selecting at least one unique rearranged sequence having the highest frequency of occurrence at each time point compared with the frequency of occurrences for each of the remaining unique rearranged sequences in said sample and determining a profile of AIR sequence distribution for said at least one unique rearranged sequence over time in said test subject.

[0022] In another embodiment, the method comprises selecting a plurality of the most abundant unique AIR rearranged sequence having a frequency of occurrence at each time point that is statistically significantly higher than an average frequency of occurrence for the total number of unique rearranged DNA sequences and determining a profile of AIR sequence distribution for each of said most abundant unique rearranged sequences over time in said test subject.

[0023] In yet another embodiment, the method includes quantifying said AIR sequence diversity score comprises determining a total number of unique clones in said sample.

[0024] In certain embodiments, the control subject has a known immunological status. In one embodiment, the control subject is a healthy subject and has an uncompromised immunological status. In another embodiment, the control subject has a compromised immunological status.

[0025] In some embodiments, the control subject has a known outcome of response to immunotherapy. In one embodiment, the response is a positive response to immunotherapy. In an embodiment, the response is a poor response to immunotherapy.

[0026] In another embodiment, the test subject is predicted to have the same outcome of response to immunotherapy as compared to the control subject. In yet another embodiment, test subject is predicted to have a different outcome of response to immunotherapy as compared to the control subject.

[0027] In certain embodiments, the control subject has a known outcome of response to a stem cell transplant. In one embodiment, the response can be a positive response to the stem cell transplant. In another embodiment, the response is a poor response to the stem cell transplant. In yet another embodiment, the test subject is predicted to have the same outcome of response to the stem cell transplant as compared to the control subject. In other embodiments, the test subject is predicted to have a different outcome of response to the stem cell transplant as compared to the control subject.

[0028] In certain aspects, the control subject has a known outcome of response to a treatment. In some aspects, the treatment comprises an immunotherapeutic antibody, a cytokine, a hematopoietic cell transplant, an immunosuppressive agent, or a vaccine.

[0029] In other aspects, the one or more samples comprise solid tissue samples obtained from the test subject.

[0030] In one aspect, the one or more samples comprise blood samples obtained from the test subject. In certain embodiments, where the one or more samples comprise blood samples, a low AIR sequence diversity score and a low AIR sequence distribution score are characterized as a low test subject rating score and are indicative of a high TCR clonality in said test subject. In other embodiments, a low test subject rating score is predictive of a poor response to immunotherapy in said test subject. In another embodiment, a high AIR sequence diversity score and a high AIR sequence distribution score are characterized as a high test subject rating score and are indicative of a low TCR clonality. In other embodiments, a high test subject rating score is predictive of a positive response to immunotherapy in said test subject.

[0031] In other embodiments, the test subject has been treated with immunotherapy. In some embodiments, the immunotherapy comprises administration of an inhibitor of a negative regulator of the immune system. In one embodiment, the negative regulator is selected from a group consisting of CTLA-4 and PD-1. In another embodiment, the negative regulator is CTLA-4. In yet another embodiment, the negative regulator is PD-1. In certain aspects, the inhibitor is an anti-CTLA-4 antibody. In another aspect, the inhibitor is an anti-PD-1 antibody.

[0032] In yet other aspects, the one or more samples comprise solid tumor samples obtained from the test subject. In some embodiments, in solid tumor samples, a high AIR sequence diversity score and a high AIR sequence distribution score are characterized as a low test subject rating score and are indicative of a low TCR clonality in said test subject.

[0033] In one embodiment, in solid tumor samples, a low test subject rating score is predictive of a poor response to immunotherapy. In another embodiment, a low AIR sequence diversity score and a low AIR sequence distribution score are characterized as a high test subject rating score and are indicative of a high TCR clonality. In one embodiment, a high test subject rating score is predictive of a positive response to immunotherapy in said subject.

[0034] In another embodiment, the test subject has been treated with immunotherapy. In one embodiment, the immunotherapy comprises administration of an inhibitor of a negative regulator of the immune system. In some aspects, the negative regulator is selected from a group consisting of CTLA-4 and PD-1. In one embodiment, the negative regulator can be CTLA-4. In another embodiment, the negative regulator can be PD-1. In other embodiments, the inhibitor is an anti-CTLA-4 antibody. In yet other embodiments, the inhibitor is an anti-PD-1 antibody.

[0035] The method also includes determining a side effect of an immunotherapy treatment for said test subject indicated by a clonal expansion of at least one clone that has a frequency of occurrence that is statistically significantly different from a mean frequency of occurrence of a set of remaining clones in a sample obtained after said immunotherapy treatment. In one embodiment, the set of remaining clones comprise clones each having a frequency of occurrence that is in the top 50% of the total clones in said sample. In another embodiment, the set of remaining clones comprise clones each having a frequency of occurrence that is in the top 40% of the total clones in said sample. In yet another embodiment, the set of remaining clones comprise clones each having a frequency of occurrence that is in the top 30% of the total clones in said sample. In other embodiments, the set of remaining clones comprise clones each having a frequency of occurrence that is in the top 20% of the total clones in said sample. In one embodiment, the set of remaining clones comprise clones each having a frequency of occurrence that is in the top 10% of the total clones in said sample. In one aspect, the at least one clone has a frequency of occurrence that is statistically significantly different from clones each having a frequency of occurrence that is in the top quartile of frequency of occurrences in said sample. In other aspects, the clonal expansion of said at least one clone is indicative of a poor response of said test subject to said immunotherapy treatment.

[0036] In certain aspects, the method also includes amplifying nucleic acid sequences obtained from at least one of said samples comprising lymphoid cells of a test subject in a multiplexed polymerase chain reaction (PCR) assay using (1) a plurality of AIR V-segment

oligonucleotide primers and (2) either a plurality of AIR J-segment oligonucleotide primers or a plurality of AIR C-segment oligonucleotide primers.

[0037] In some embodiments, the plurality of AIR V-segment oligonucleotide primers are each independently capable of specifically hybridizing to at least one polynucleotide encoding a mammalian AIR V-region polypeptide, wherein each AIR V-segment oligonucleotide primer comprises a nucleotide sequence of at least 15 contiguous nucleotides that is complementary to at least one functional AIR-encoding gene segment, wherein said plurality of AIR V-segment oligonucleotide primers specifically hybridize to substantially all functional AIR V-encoding gene segments that are present in said sample. In one embodiment, the plurality of J-segment oligonucleotide primers are each independently capable of specifically hybridizing to at least one polynucleotide encoding a mammalian AIR J-region polypeptide, wherein each J-segment primer comprises a nucleotide sequence of at least 15 contiguous nucleotides that is complementary to at least one functional AIR J-encoding gene segment, wherein said plurality of J-segment primers specifically hybridize to substantially all functional AIR J-encoding gene segments that are present in the sample. In another embodiment, the plurality of C-segment oligonucleotide primers are each independently capable of specifically hybridizing to at least one polynucleotide encoding a mammalian AIR C-region polypeptide, wherein each C-segment primer comprises a nucleotide sequence of at least 15 contiguous nucleotides that is complementary to at least one functional AIR C-encoding gene segment, wherein the plurality of C-segment primers specifically hybridize to substantially all functional AIR C-encoding or gene segments that are present in the sample.

[0038] In certain aspects, the plurality of AIR V-segment oligonucleotide primers, and (2) either said plurality of AIR J-segment oligonucleotide primers and said plurality of AIR C-segment oligonucleotide primers are capable of promoting amplification in said multiplex PCR of substantially all rearranged AIR CDR3-encoding regions in said sample to produce a plurality of amplified rearranged DNA molecules from a population of adaptive immune cells in said sample, said plurality of amplified rearranged DNA molecules being sufficient to quantify the full diversity of said AIR CDR3-encoding region in said at least one sample.

[0039] In some embodiments, each functional AIR V-encoding gene segment comprises a V gene recombination signal sequence (RSS) and each functional AIR J-encoding gene segment comprises a J gene RSS, wherein each amplified rearranged DNA molecule comprises (i) at least 10, 20, 30 or 40 contiguous nucleotides of a sense strand of said AIR V-encoding gene segment, wherein said at least 10, 20, 30 or 40 contiguous nucleotides are

situated 5' to said V gene RSS and (ii) at least 10, 20 or 30 contiguous nucleotides of a sense strand of said AIR J-encoding gene segment, wherein said at least 10, 20 or 30 contiguous nucleotides are situated 3' to said J gene RSS. In one embodiment, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 1500 nucleotides in length. In another embodiment, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 1000 nucleotides in length. In yet another embodiment, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 600 nucleotides in length. In other embodiments, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 500 nucleotides in length. In one aspect, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 400 nucleotides in length. In another aspect, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 300 nucleotides in length. In yet another aspect, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 200 nucleotides in length. In some embodiments, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is less than 100 nucleotides in length. In a preferred embodiment, each amplified rearranged DNA molecule in said plurality of amplified rearranged DNA molecules is between 50-600 nucleotides in length.

[0040] In some aspects, the method includes selecting a set of unique rearranged sequences in one of said samples having a frequency of occurrence that is statistically significantly higher compared with other unique rearranged sequences in said sample. In certain aspects, the high frequency of occurrence is determined by a pre-determined threshold percentage. In one aspect, the selected number of unique rearranged sequences in said set is determined by a pre-determined number. In other embodiments, the method includes determining from said set whether one of said high frequency unique rearranged sequences is persistent or transient, wherein a persistent unique rearranged sequence is present across two or more samples obtained from said test subject over subsequent periods of time, and wherein a transient unique rearranged sequence is present in only one sample obtained at one timepoint from said subject.

[0041] The method also includes determining a course of immunotherapy for said subject based on the presence of one or more persistent unique rearranged sequences in said two or more samples of said test subject, wherein the presence of persistent unique rearranged sequences indicates an increased likelihood that said subject has a healthy immune status. In

some embodiments, the presence of said one or more persistent unique rearranged sequences in said subject is predictive of a positive response to immunotherapy treatment by said subject. The method of the invention includes determining a course of immunotherapy for said subject based on a presence of one or more transient unique rearranged sequences in said one or more samples of said test subject, wherein said presence of said one or more transient unique rearranged sequences indicates an increased likelihood that said subject has a compromised immune status. In another embodiment, the presence of said one or more transient unique rearranged sequences in said subject is predictive of a poor response to immunotherapy treatment by said subject.

[0042] In some embodiments, the method of the invention provides steps for categorizing a test subject having a low test subject rating score in said one or more samples as having a lower relative likelihood of responding to immunotherapy in comparison to a second subject having a higher rating score; and stratifying a patient population of test subjects according to relative likelihood of responding to immunotherapy. The method includes determining said test subject rating score comprises extrapolating based on a mathematical model a total AIR repertoire diversity of said test subject by sequencing said nucleic acid sequences from one of said samples and determining a test subject rating score from said total AIR repertoire diversity. In one embodiment, the mathematical model is an unseen species model.

[0043] In another embodiment, determining said test subject rating score comprises calculating a Shannon entropy score and a clonality score and determining a test subject rating score based on said Shannon entropy score and said clonality score. In one embodiment, the clonality score is a transform of the Shannon entropy score.

[0044] In other aspects of the invention, the adaptive immune receptor (AIR) polypeptide is a mammalian AIR polypeptide and is selected from a T cell receptor-gamma (TCRG) polypeptide, a T cell receptor-beta (TCRB) polypeptide, a T cell receptor-alpha (TCRA) polypeptide, a T cell receptor-delta (TCRD) polypeptide, an immunoglobulin heavy-chain (IGH) polypeptide, and an immunoglobulin light-chain (IGL) polypeptide. In some embodiments, the IGH polypeptide is selected from an IgM, an IgA polypeptide, an IgG polypeptide, an IgD polypeptide and an IgE polypeptide. The IGL polypeptide can be selected from an IGL-lambda polypeptide and an IGL-kappa polypeptide. In one embodiment, the mammalian AIR polypeptide is a human AIR polypeptide. In another embodiment, the mammalian AIR polypeptide is selected from a non-human primate AIR polypeptide, a rodent AIR polypeptide, a canine AIR polypeptide, a feline AIR polypeptide and an ungulate AIR polypeptide.

[0045] In certain embodiments, the test subject is selected from: a subject having or suspected of having a malignant condition, a subject who has received a hematopoietic cell transplant, a subject who has received a solid organ transplant, and subject having a microbial infection. In some embodiments, the malignant condition is selected from a hematologic malignancy, a melanoma, a sarcoma and a carcinoma. The malignant condition can be selected from malignant melanoma, small cell lung cancer, non-small cell lung cancer, renal cell carcinoma, pancreatic cancer, breast cancer, ovarian cancer and prostate cancer.

[0046] In other embodiments, the hematopoietic cell transplant is selected from a cord blood transplant, an autologous hematopoietic cell transplant, an allogeneic hematopoietic cell transplant, and a bone marrow transplant. In one embodiment, the hematopoietic cell transplant comprises an autologous T cell transplant.

[0047] In other aspects, the plurality of time points comprise timepoints during or after immunotherapy. In another aspect, the plurality of time points comprise timepoints prior to immunotherapy.

[0048] In other embodiments, the method includes steps for managing a treatment of said test subject who is undergoing immunotherapy based on a determination of said immunological status of said test subject. In one embodiment, the immunotherapy comprises a treatment with an immunotherapy agent that is selected from an immunotherapeutic antibody, a cytokine, a hematopoietic cell transplant, an immunosuppressive agent, and a vaccine.

[0049] In another embodiment, the immunotherapy comprises a treatment with an inhibitor of a negative regulator of an immune response. In some aspects, the negative regulator of an immune response is selected from CTLA4/CD152, LAG3/CD223, and PD-1/CD279. In an embodiment, the negative regulator of an immune response can be CTLA-4/CD152 and said inhibitor of said negative regulator of an immune response can be an anti-CTLA-4 antibody. In another embodiment, the anti-CTLA-4 antibody is selected from ipilimumab and tremelimumab. In some embodiments, the negative regulator of an immune response is PD-1/CD279 and said inhibitor of the negative regulator of an immune response is an anti-PD-1 antibody. In another embodiment, the immunotherapy comprises a treatment with an agent that targets a potentiator of an immune response. In yet another embodiment, the potentiator of an immune response is selected from 41BB/CD137, OX40/CD134 and CD40.

[0050] In other embodiments, the immunotherapy comprises a treatment of an inflammatory condition or an autoimmune disease with an inhibitor of an inflammatory pathway. In certain embodiments, the inflammatory condition or said autoimmune disease is

selected from rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis, Crohn's disease and juvenile idiopathic arthritis. In one embodiment, the inflammatory pathway comprises at least one of tumor necrosis factor-alpha (TNF α), interferon-gamma (IFN γ), interleukin-1 (IL-1), interleukin-6 (IL-6), interleukin-8 (IL-8). In other embodiments, the inflammatory pathway comprises TNF α and said inhibitor of the inflammatory pathway is an agent that specifically binds to TNF α . In another embodiment, the agent that specifically binds to TNF α is selected from an anti-TNF α antibody and an artificial soluble TNF α receptor. In one embodiment, the anti-TNF α antibody is selected from adalimumab and infliximab and said artificial soluble TNF α receptor is etanercept.

[0051] In other embodiments, a computer-implemented method is provided for determining an immunological status of a test subject, comprising: storing data for a control subject obtained from a plurality of samples at various timepoints, said data comprising for each sample, nucleic acid sequence information for a plurality of unique rearranged nucleic acid sequences in said sample, an AIR sequence diversity score for said sample, a frequency of occurrence of each unique rearranged nucleic acid sequence in said sample, and a determined immunological status for said subject; determining rules by a processor for assessing an immunological status of a test subject based on said data of said control subject; inputting data for a test subject for a plurality of samples obtained at various timepoints before and after immunotherapy, said data comprising for each sample, nucleic acid sequence information for a plurality of unique rearranged nucleic acid sequences in said sample, an AIR sequence diversity score for said sample, and a frequency of occurrence of each unique rearranged nucleic acid sequence in said sample; and receiving a determination of an immunological status of said test subject. In some embodiments, the method includes determining a predicted response to immunotherapy of said test subject. In one embodiment, the data for said control subject comprises nucleic acid sequence information obtained from said control subject at a timepoint prior to immunotherapy treatment. In another embodiment, the data for said control subject comprises nucleic acid sequence information obtained from said control subject at a timepoint after immunotherapy treatment.

[0052] These and other aspects of the herein described invention embodiments will be evident upon reference to the following detailed description and attached drawings. All of the U.S. patents, U.S. patent application publications, U.S. patent applications, foreign patents, foreign patent applications and non-patent publications referred to in this specification and/or listed in the Application Data Sheet are incorporated herein by reference in their entirety, as if each was incorporated individually. Aspects and embodiments of the

invention can be modified, if necessary, to employ concepts of the various patents, applications and publications to provide yet further embodiments.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0053] Figure 1 illustrates TCR clones that remained over time after myeloablative treatment. Shown is the range of values for the proportion of each patient's repertoire represented by clones that were held over after stem cell transplant. The bottom quartile ranged to zero. The proportion of holdover clones was calculated as the proportion of total TCR sequencing reads corresponding to clones observed (at any level) before transplant. Values indicated some persistence of pre-transplant clones in these patients' TCR repertoires.

[0054] Figure 2 shows the number of transient TCR clones observed in patients during early immune reconstitution, at 28, 56, 100 and 180 days post-transplant, as compared with healthy controls. For each sample, each of the top 10 TCR clones by frequency was classified as either persistent (observed again in the same patient at a later time point) or transient (not observed again at any level in subsequent samples from the same patient). The number of transient clones was highly variable among patients, ranging from 0 to 9, but the median number of transient clones decreased with time. Four healthy controls were also analyzed, and the number of transient TCR clones ranged from 0 to 2 with a median of 0.

[0055] Figure 3 illustrates TCR repertoire reconstitution after stem cell transplant, shown in TCR repertoire size across all patients following hematopoietic stem cell transplant. Samples were taken before transplant and five times after transplant (at days 28, 56, 100, 180 and 365). TCR repertoire size for each patient was estimated using high-throughput sequencing of TCR rearrangements, and the geometric mean of estimated TCR repertoire size is shown. After transplant, patients had a vastly reduced TCR repertoire that reached its minimum 56 days post-transplant, before beginning a slow recovery.

[0056] Figure 4 shows an estimated TCR repertoire size comparison based on high-throughput sequencing of TCR β rearrangements for all patients with and without eventual non-relapse mortality (NRM). TCR repertoire size values are shown as quartiles for both populations. Significance was assessed using a one-tailed Mann-Whitney U test. Patients who went on to suffer from non-relapse mortality had significantly lower estimated repertoire sizes at 56 and 100 days post-transplant.

[0057] Figure 5 shows a comparison of CD3⁺ counts and estimated TCR repertoire size. CD3⁺ counts (number of cells/mL) were compared to estimated TCR repertoire size for samples from all surviving patients from days 28, 56 and 100 for which both metrics were available. There was a weak correlation between CD3⁺ counts and repertoire size ($r = 0.06$),

indicating that an estimate on the lower bound of TCR β diversity obtained through sequencing revealed information independent of the total density of circulating T cells.

[0058] Figure 6 shows TCR repertoire clonality in blood samples. In blood samples, low TCR repertoire clonality was a predictor of immunotherapy (treatment with ipilimumab (an anti-CTLA-4 mAb)) responder status and high TCR repertoire clonality was a predictor of immunotherapy non-responder status.

[0059] Figure 7 shows results from quantitative sequencing of TCR encoding DNA from tumor tissue samples. The results show an increase in TCR clonality in lymphocytes present in solid tumor tissue samples obtained after administration of immunotherapy (treatment with ipilimumab (an anti-CTLA-4 mAb)) (DT) relative to the level of TCR clonality detected in tumor samples obtained prior to immunotherapy (AT).

[0060] Figure 8 shows dynamics in the relative representations of individual TCR clonal populations over time in blood samples and in solid tumor samples obtained prior to immunotherapy (treatment with ipilimumab (an anti-CTLA-4 mAb)) and post immunotherapy. Timepoints A, B, and C are timepoints taken from blood samples. Timepoint A is before immunotherapy, and timepoints B and C are two timepoints after starting the immunotherapy regimen. Timepoints AT, BT, CT are paired tumor samples (AT is before immunotherapy, and timepoints BT and CT are two timepoints after starting the immunotherapy regimen). The arrow indicates a single clone that has increased in preponderance post-therapy to account for 10% of the repertoire at timepoint CT. Two clones (X1 and X2) that each accounted for 7-9% of TCR sequences in blood at timepoints A and B subsequently declined significantly in relative abundance, while several T cell receptor sequences that initially had very low frequencies in blood increased significantly by timepoint C. The three most numerous clones in tumor samples at timepoint A decreased significantly in their subsequent relative representation, as determined at later timepoints.

[0061] Figure 9 shows TCR sequence diversity and distribution entropies determined in peripheral blood samples obtained prior to (timepoint A) and after (timepoints B and C) initiation of immunotherapy (treatment with ipilimumab (an anti-CTLA-4 mAb)) shows dynamics of individual TCR clonal representations over time, pre-therapy (timepoint A) and post-therapy (timepoints B and C). Arrow indicates a single clone that has increased in preponderance post-therapy to account for greater than 10% of the repertoire at timepoint C.

[0062] Figure 10A shows measurements of TCR repertoire clonality from tissue samples from melanoma lesions obtained from late-stage metastatic melanoma patients before, during and after immunotherapy with anti-PD-1 antibody. The mean and standard deviation of TCR

repertoire clonality (a modified metric based on TCR sequence distribution entropy normalized to the range (0 – 1) by accounting for the number of unique TCR sequences present in each sample) is shown according to response to immunotherapy. Of the 12 patients studied, 8 responded to the treatment (stable disease or partial response), while 4 did not respond (disease progression). TCR repertoire clonality was higher in the 8 patients who responded compared to the 4 patients who did not respond ($p = 0.015$ by two-tailed unpaired t-test).

[0063] Figure 10B shows the mean and standard deviation of T cell infiltration (measured as T cell receptor rearrangements per diploid genome) according to response to immunotherapy obtained from tissue samples of melanoma lesions obtained from a cohort of 12 late-stage metastatic melanoma patients. Measured by a two-tailed unpaired t-test, the levels of T lymphocyte infiltration were higher in the 8 patients who responded compared to the 4 patients who did not respond ($p = 0.056$ by two-tailed unpaired t-test). Immunotherapy was treatment with an anti-PD-1 antibody.

[0064] Figure 11A shows measurements of TCR repertoire clonality and response to immunotherapy for cohort 1 (12 patients), cohort 2 (13 patients), and a combined cohort (obtained from tissue samples of melanoma lesions). Immunotherapy was treatment with an anti-PD-1 antibody. The mean and standard deviation of TCR repertoire clonality (a modified metric based on TCR sequence distribution entropy normalized to the range (0 – 1) by accounting for the number of unique TCR sequences present in each sample) is shown according to response to immunotherapy in cohort 1 (plain square), cohort 2 (slant striped square) and in the combined data (straight striped square). TCR sequence distribution clonality was higher in the patients who responded compared to the patients who did not respond ($p = 0.00065$ in the combined data by a two-tailed unpaired t-test).

[0065] Figure 11B shows the mean and standard deviation of T cell infiltration (measured as T cell receptor rearrangements per diploid genome) according to response to immunotherapy in cohort 1 (plain square), cohort 2 (slanted stripe square) and in the combined data (straight stripe square). T cells were obtained from tissue samples of melanoma lesions. Immunotherapy was treatment with an anti-PD-1 antibody. Levels of T lymphocyte infiltration are higher in the patients who responded compared to the patients who did not respond ($p = 0.0015$ in the combined data by a two-tailed unpaired t-test).

[0066] Figure 11C shows a comparison of each patient's rank (in descending rank, out of 25 patients total) for level of T cell infiltration and TCR repertoire clonality. T cells were obtained from tissue samples of melanoma lesions. Immunotherapy was treatment with an

anti-PD-1 antibody. Compared to responders (diamond), non-responders (circle) simultaneously tend toward low TCR repertoire clonality and low levels of infiltrating T lymphocytes.

[0067] Figure 12 is a high-level block diagram illustrating an example of a computer, according to one embodiment of the invention.

DETAILED DESCRIPTION

I. OVERVIEW

[0068] The present invention provides, in certain embodiments and as described herein, unexpectedly advantageous methods for determining the immunological status of a subject or of a plurality of subjects, including by qualitatively (*e.g.*, by T cell receptor or immunoglobulin sequence diversity) and quantitatively (*e.g.*, by TCR or IG sequence distribution) characterizing adaptive immune cell (*e.g.*, T cell or B cell) clonality, from which immunocompetence of an individual's adaptive immune system can be assessed. The present embodiments thus provide novel methods for assessing the immunocompetence of an individual and for stratifying a population according to immune system status, where determination of both the sequence diversity of TCR and/or IG expressed by lymphoid cells in an individual, and the relative degree of T cell and/or B cell clonality in the individual, are of relevance to prognosis, diagnosis, and outcome, including likelihood of developing immune-related side effects, in a variety of clinical contexts.

[0069] The present embodiments for the first time permit high resolution, large-scale, high throughput assessment of immunocompetence by characterization at the DNA sequence level of (i) TCR and IG repertoire diversity, and (ii) TCR and IG repertoire distribution. The invention includes compositions and methods for quantitative detection of sequences of substantially all possible TCR and IG gene rearrangements that can be present in a sample containing lymphoid cell DNA.

[0070] In certain embodiments, a sample containing lymphoid cell DNA (genomic DNA, cDNA or alternatively, messenger RNA) from a subject is used as a template for multiplexed PCR amplification using a primer set that is specifically designed to be capable of amplifying substantially all possible DNA rearrangements encoding a particular TCR or IG chain. The multiplex PCR amplification products are amenable to rapid, high throughput, high quality quantitative DNA sequencing. Structural TCR or IG repertoire diversity in the sample is determined by identifying a plurality of unique rearranged DNA sequences from the DNA

sequence information, and therefrom determining the total number of unique sequences in the sample.

[0071] Where desired, known estimation or extrapolation methods can be used to determine from the sequence information a repertoire diversity in the subject's entire adaptive immune system. To quantify the relative distribution of each unique sequence, quantitative sequencing methodologies described herein and practiced by those of skill in the art also permit determination of the frequency of occurrence of each particular uniquely rearranged DNA sequence amongst the total number of unique sequences. In certain embodiments, a blood sample can be obtained as the source of lymphoid cells from which lymphoid cell DNA and/or RNA can be extracted to provide PCR templates.

[0072] These and related methods will find a variety of uses as described herein. For example, the methods described herein are used to quantify the diversity and distribution of the adaptive immune receptor (AIR) repertoire within each individual subject's adaptive immune system. The methods described herein are also used to stratify a patient population according to the patient's immunocompetence status or the relative likelihood of individuals to respond to an immunotherapy or develop immune-related side effects. Quantification of AIR sequence diversity (*e.g.*, the number of different unique AIR encoding sequences, identified by obtaining distinctive nucleotide sequence information for all rearranged DNA encoding a particular AIR polypeptide in a sample) and of AIR sequence distribution (*e.g.*, frequency of occurrence of each unique rearranged AIR encoding DNA sequence) advantageously permits correlation of T or B cell clonality, defined with unprecedented precision, with clinically useful information.

[0073] By way of non-limiting theory, this sequence distribution can represent the degree of T cell or B cell clonality in each sample from a patient (*e.g.*, quantitative degree of representation, or relative abundance). Any of a number of known computational tools for processing this distribution parameter can be used to generate distribution values (*e.g.*, the frequency of occurrence of each unique sequence) and diversity values (*e.g.*, the total number of different unique sequences). The distribution and diversity values can be used in a rating step to rate individual samples and compare them to a control sample and/or to one another.

[0074] As described herein, in patients receiving a cord blood transplant to treat hematologic malignancies, a relatively low degree of TCR repertoire diversity in patients following the cord blood transplant was shown to be a predictor of the relative likelihood of susceptibility to infection and of the immunological inability to clear the infection (*e.g.*, poor response). On the other hand, a relatively high degree of TCR repertoire diversity in human

patients following cord blood transplant to treat hematologic malignancies was shown to be a predictor of the relative likelihood of resistance to infection and of immunocompetence, *i.e.*, the immunological ability to clear the infection.

[0075] In blood samples obtained from patients before and after immunotherapy with an inhibitor of a negative regulator of immune response (*e.g.*, anti-CTLA4 antibody), a high TCR sequence diversity and low clonality in the blood of the patient correlated with positive clinical outcomes. On the other hand, a TCR repertoire that was characterized by a low TCR sequence diversity (high clonality) and a lower entropy of TCR sequence distribution was associated with poorer clinical outcomes that were attributable to compromised adaptive immune capability.

[0076] In solid tumor samples obtained from patients before and after immunotherapy with an inhibitor of a negative regulator of immune response (*e.g.*, anti-PD-1 antibody), a high level of infiltrating T cell presence and high clonality (*i.e.*, evidence of T cell migration to the tumor and clonal proliferation within the tumor) were associated with a positive response to immunotherapy. In contrast, a minimal infiltrating T cell repertoire and low clonality in solid tumors (*i.e.*, evidence of a restricted and non-specific T cell response within the tumor) were associated with failure to respond to treatment.

[0077] The presently-disclosed embodiments will find a wide range of uses by profiling a subject's immunocompetence at a given point in time, for example, as a prognostic or diagnostic or to inform a therapeutic strategy, and for other purposes.

II. DEFINITIONS

[0078] Terms used in the claims and specification are defined as set forth below unless otherwise specified.

[0079] As used herein, adaptive immune receptor (AIR) refers to an immune cell receptor, such as a T cell receptor (TCR) or an Immunoglobulin (Ig) receptor found in mammalian cells.

[0080] The term "primer," as used herein, refers to an oligonucleotide capable of acting as a point of initiation of DNA synthesis under suitable conditions. Such conditions include those in which synthesis of a primer extension product complementary to a nucleic acid strand is induced in the presence of four different nucleoside triphosphates and an agent for extension (*e.g.*, a DNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature.

[0081] The term percent "identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that have a specified

percentage of nucleotides or amino acid residues that are the same, when compared and aligned for maximum correspondence, as measured using one of the sequence comparison algorithms described below (e.g., BLASTP and BLASTN or other algorithms available to persons of skill) or by visual inspection. Depending on the application, the percent “identity” can exist over a region of the sequence being compared, *e.g.*, over a functional domain, or, alternatively, exist over the full length of the two sequences to be compared.

[0082] For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

[0083] Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by visual inspection (*see Ausubel et al.*, *infra*).

[0084] One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990). Software for performing BLAST analyses is publicly-available through the National Center for Biotechnology Information website (www.ncbi.nlm.nih.gov).

[0085] The term “sufficient amount” means an amount sufficient to produce a desired effect, *e.g.*, an amount sufficient to modulate immune response in a cell.

[0086] The term “therapeutically effective amount” is an amount that is effective to ameliorate a symptom of a disease. A therapeutically effective amount can be a “prophylactically effective amount” as prophylaxis can be considered therapy.

[0087] Unless specific definitions are provided, the nomenclature utilized in connection with, and the laboratory procedures and techniques of, molecular biology, analytical chemistry, synthetic organic chemistry, and medicinal and pharmaceutical chemistry described herein are those well known and commonly used in the art. Standard techniques can be used for recombinant technology, molecular biological, microbiological, chemical

syntheses, chemical analyses, pharmaceutical preparation, formulation, and delivery, and treatment of patients.

[0088] Unless the context requires otherwise, throughout the present specification and claims, the word “comprise” and variations thereof, such as, “comprises” and “comprising” are to be construed in an open, inclusive sense, that is, as “including, but not limited to.” By “consisting of” is meant including, and typically limited to, whatever follows the phrase “consisting of.” By “consisting essentially of” is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase “consisting essentially of” indicates that the listed elements are required or mandatory, but that no other elements are required and can or cannot be present depending upon whether or not they affect the activity or action of the listed elements.

[0089] In this specification and the appended claims, the singular forms “a,” “an” and “the” include plural references unless the content clearly dictates otherwise. As used herein, in particular embodiments, the terms “about” or “approximately” when preceding a numerical value indicates the value plus or minus a range of 5%, 6%, 7%, 8% or 9%, *etc.* In other embodiments, the terms “about” or “approximately” when preceding a numerical value indicates the value plus or minus a range of 10%, 11%, 12%, 13% or 14%, *etc.* In yet other embodiments, the terms “about” or “approximately” when preceding a numerical value indicates the value plus or minus a range of 15%, 16%, 17%, 18%, 19% or 20%, *etc.*

[0090] Reference throughout this specification to “one embodiment” or “an embodiment” or “an aspect” means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, the appearances of the phrases “in one embodiment” or “in an embodiment” in various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics can be combined in any suitable manner in one or more embodiments.

III. QUANTIFICATION METHODS

[0091] Various methods can be used to quantify and assess the immunocompetence of the subject. In some embodiments, the immunocompetence is assessed by measuring the subject’s adaptive immune receptor (AIR) sequence diversity and AIR sequence distribution.

A. Adaptive Immune Receptor (AIR) Sequence Diversity

[0092] Diversity of unique rearranged TCR or IG encoding DNA sequences in lymphoid cells in a sample reflects the number of different T or B cell clones in a sample from a

subject. Sequence diversity can be determined as the number of clones in a sample of a particular size, such as by direct counting or weighted counting in a sample. A sample can be a blood sample or a tissue sample (solid tumor sample), for example. Alternatively, the number of different clones in a subject can be estimated based on the number of clones in a subsample. In another embodiment, an arbitrary cutoff value can be assigned to estimate the number of different “effective” clones, such as counting toward diversity only those clones that account for greater than 0.01% of all T or all B cells in the sample. Other models for weighted or extrapolated diversity determinations are contemplated for use in certain related embodiments, such as entropy models, the “unseen species model” (*see, e.g., Efron et al., 1976 Biometrika 63:435; Fisher et al., 1943 J. Anim. Ecol. 12:42*) or other suitable models as will be known to those familiar with the art.

[0093] In some embodiments, AIR diversity can be measured by quantitative sequencing of the total AIR observed sequences in a particular sample. Compositions and methods for quantitative sequencing of rearranged adaptive immune receptor gene sequences and for adaptive immune receptor clonotype determination are described, for example, in Robins *et al.*, 2009 *Blood* 114, 4099; Robins *et al.*, 2010 *Sci. Translat. Med.* 2:47ra64; Robins *et al.*, 2011 *J. Immunol. Meth.* doi:10.1016/j.jim.2011.09.001; Sherwood *et al.* 2011 *Sci. Translat. Med.* 3:90ra61; U.S.A.N. 13/217,126, U.S.A.N. 12/794,507, WO/2010/151416, WO/2011/106738 (PCT/US2011/026373), WO2012/027503 (PCT/US2011/049012), U.S.A.N. 61/550,311, and U.S.A.N. 61/569,118, herein incorporated by reference. Therein can also be found details regarding sequences of PCR amplification oligonucleotide primers and sequencing primers, sequencing of PCR amplification products, processing sequencing data, and uses of measurements of adaptive immune receptor diversity, all of which can be employed for use according to the methods described herein.

[0094] In some embodiments, a sequencing program such as Raw HiSeq™ can be used to preprocess sequence data to remove errors in the primary sequence of each read, and to compress the sequence data. A nearest neighbor algorithm can be used to collapse the data into unique sequences by merging closely related sequences, to remove both PCR and sequencing errors.

[0095] Quantitative sequencing of TCR or IG as described herein permits assignment of a diversity score or rating to a sample. In some embodiments, the diversity score or rating can be determined to be low when there are a small number of unique rearranged AIR sequences in the repertoire as compared to the total number of observed rearranged AIR sequences in a sample. The diversity score or rating can be higher when there is a large number of unique

rearranged AIR sequences in the repertoire as compared to the total number of observed rearranged AIR sequences in a sample. The determination of a low or high diversity score or rating can be based on pre-determined thresholds or calculations of statistical significance, as can be determined by one of skill in the art. For example, a predetermined threshold for classifying a diversity score or rating as “low” can be, in some embodiments, a score that is not higher (with statistical significance) than that obtained from blood samples of a subject population, wherein the population can be a population determined to experience a poor outcome in response to an immunotherapeutic intervention. In other embodiments, the predetermined threshold is determined based on calculation of the top or highest 50%, 25%, 10% or 5% of diversity or rating scores determined from rearranged AIR sequences from the sample.

[0096] As a relative scale, the rating system can be varied or adjusted in view of a number of factors, including but not limited to, the sample size, method of diversity quantification (*e.g.*, whether by direct sequencing, or by extrapolation, “hidden species,” *etc.*), clinical signs and symptoms of the patient population from whom samples are obtained, *etc.* For instance, in certain non-limiting examples, members of a patient population can be categorized on the basis of relative diversity and/or distribution ratings, and in certain embodiments, arbitrary segmentation of the population can be practiced. In certain embodiments, the patient population can be stratified according to (i) the degree of sequence diversity or distribution by quartile, quintile, decile, *etc.*, or (ii) by rating relative AIR sequence diversity and distribution entropy in 50, 40, 30, 20 or 10 percent of the total number of sequences as a correlate of clonality, or (iii) by selecting the 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 most abundant unique AIR sequences at each of a succession of timepoints. The categorization provides a set of parameters by which immunocompetence can be assessed.

[0097] As also noted above, using these sequence diversity calculations, it was determined that in blood samples obtained from patients having undergone hematopoietic stem cell transplantation, higher AIR sequence diversity scores correlated with higher likelihoods of favorable clinical outcomes, such as the ability to clear infections. In solid tumor samples obtained from patients undergoing immunotherapy with anti-PD-1 antibody, high levels of T cell infiltration and a low infiltrating T cell repertoire diversity were associated with a positive response to immunotherapy.

[0098] Other known methods for calculating AIR sequence diversity can be used as known to those of skill in the art. For example, the following works, which are incorporated by

reference in their entireties, summarize the current theory and practice of estimating diversity indices from species abundance data, while giving detailed examples of several common embodiments of diversity index measurement. See Anne E. Magurran and Brian J. McGill. 2011. Biological Diversity: Frontiers in Measurement and Assessment. New York: Oxford University Press. Other examples of methods for genetic diversity estimation that can be applied to calculate a diversity score rating can be found in James F. Crow and Motoo Kimura. 2009. An Introduction to Population Genetics Theory. Blackburn Press.

B. Adaptive Immune Receptor (AIR) Sequence Distribution

[0099] In some embodiments, the AIR sequence distribution can be used to determine and assess a subject's immunological status (*e.g.*, immunocompetence). AIR sequence distribution, such as TCR or IG sequence distribution, refers to the variation among the number of different T cell or B cell clones in a sample, *e.g.*, the number of cells that express an identical TCR or IG. For example, AIR sequence distribution can be determined by quantifying the frequency of occurrence of each unique rearranged AIR encoding DNA sequence, as a percentage of the total number of observed rearranged AIR encoding DNA sequences. The quantified distribution of AIR sequences can be used, optionally along with AIR sequence diversity, to rate or rank the immunocompetence of a subject, according to certain presently-contemplated embodiments for determining immunological status.

[00100] In some embodiments, an AIR sequence distribution can be determined by, but not limited to, the following methods: (i) identifying and quantifying at least 1-20 of the most abundant unique rearranged (clonal) AIR sequences in a subject over a time interval, or (ii) by identifying and quantifying the number of unique rearranged (clonal) AIR sequences that are needed to account for a given percentage (*e.g.*, up to 10, 20, 30, 40 or 50%) of the total number of observed rearranged sequences in a sample from a subject. Other calculations can additionally or alternatively be employed to determine AIR sequence distribution of a sample from a subject and to assign a sequence distribution value to a particular sample for purposes of rating the sample in comparison to a control or another sample with a known immunological status. These can include, for example, determining entropy (*i.e.*, Shannon entropy as typically defined in information theory, which can be normalized to the range [0 – 1] by dividing by the logarithm of the number of elements in the sample set) or using other known methods to determine one or more modes of distribution (*e.g.*, mean, skewness, kurtosis, etc.). The present methods permit determination of sequence distribution and clonality with a degree of precision not previously possible and permit a variety of prognostic, diagnostic, prescriptive and other capabilities.

C. Determining Immunological Status

[00101] According to certain embodiments, there is provided a method for determining immunological status of a test subject. In some embodiments, the method includes steps for identifying, quantifying, rating, comparing and categorizing the immunological status of the test subject.

[00102] In some embodiments, identifying DNA sequence information for each of a plurality of unique rearranged DNA sequences that encode an adaptive immune receptor (AIR) polypeptide in one or more samples containing lymphoid cell DNA obtained from a test subject at each of one or a plurality of timepoints, and determining a total number of unique rearranged AIR polypeptide encoding DNA sequences in the test subject at each of the one or a plurality of timepoints to quantify AIR sequence diversity in the subject can be performed as described above and in Robins *et al.*, 2009 *Blood* 114, 4099; Robins *et al.*, 2010 *Sci. Translat. Med.* 2:47ra64; Robins *et al.*, 2011 *J. Immunol. Meth.* doi:10.1016/j.jim.2011.09.001; Sherwood *et al.* 2011 *Sci. Translat. Med.* 3:90ra61; U.S.A.N. 13/217,126, U.S.A.N. 12/794,507, WO/2010/151416, WO/2011/106738 (PCT/US2011/026373), WO2012/027503 (PCT/US2011/049012), U.S.A.N. 61/550,311, and U.S.A.N. 61/569,118, herein incorporated by reference. These references provide description regarding sequences of PCR amplification oligonucleotide primers and sequencing primers, sequencing of PCR amplification products, processing sequencing data, and uses of measurements of adaptive immune receptor diversity.

[00103] After determining the total number of unique rearranged AIR polypeptide encoding DNA sequences in the test subject at each of said one or a plurality of timepoints, the frequency of occurrence of each unique rearranged DNA sequence can be quantified as a percentage of the total number of observed rearranged AIR polypeptide encoding DNA sequences. For example, if an AIR sequence diversity value is determined from a count of actual sequence data, that value can be used to determine AIR sequence distribution. In another example, if AIR sequence diversity data are estimated, such as by extrapolation of a subsample to the subject's full adaptive immune system, or using the "unseen species model," or by any other estimation method, then any of widely known method for capturing properties of a distribution can be employed.

1. Rating the Immunological Status of a Sample from a Subject

[00104] AIR sequence diversity and AIR sequence distribution values for each sample can be used to rate the immunological status of samples. Any of a wide variety of simple, weighted and/or sophisticated rating systems can be employed, as can depend on the diversity

and distribution estimation methods that are used. Preferably and in certain non-limiting embodiments, a low rating is assigned to a test subject's sample in which a small number of unique rearranged sequences in reference to a predetermined threshold have a combined frequency of occurrence of no more than 50 percent of the total number of observed rearranged sequences. A higher test subject rating is assigned to a sample in which a higher number of unique rearranged sequences in reference to a predetermined threshold have a combined frequency of occurrence of no more than 50 percent of the total number of observed rearranged sequences. In other words, the rating is lower where a smaller number of different clones accounts for 50 percent of the total number of observed rearranged sequences, as would be the case where one or a few dominant clones or oligoclonality are present. In similar but related embodiments, a lower test subject rating is assigned to a sample in which a lower number of unique rearranged sequences have a combined frequency of occurrence of no more than 40, 30, 20 or 10 percent of the total number of observed rearranged sequences, and a higher test subject rating is assigned to a sample in which a higher number of unique rearranged sequences have, respectively, a combined frequency of occurrence of no more than 40, 30, 20 or 10 percent of the total number of observed rearranged sequences in the sample.

[00105] In some embodiments, the assigned ratings that can then be compared to control subject ratings generated from control lymphoid cell DNA samples obtained from a second subject with a known immunological status. In certain embodiments, the second subject has a known, compromised immunological status, as defined by one of skill in the art. In other embodiments, the second subject can be a healthy control individual with a known, uncompromised immunological status according to art-established criteria (*e.g.*, Rich *et al.*, *Clinical Immunology: Principles and Practice*, 3rd Ed., Mosby, St. Louis). In some embodiments, the test subject can be categorized as having a compromised immunological status at each of said timepoints at which the test subject rating is lower, in a statistically significant manner, than the control subject rating, such that the immunological status of the test subject is thereby determined. As referred to herein, a "control subject" can refer to a population of control subjects each sharing a relevant clinical phenotype.

[00106] For example, and in certain preferred embodiments, a test subject can be categorized as having a compromised immunological status and/or an unhealthy immune status when a TCR or IG sequence diversity score for a sample from the test subject is, with statistical significance, two standard deviations below that of a sample from a control subject, wherein said control subject is known to have an uncompromised immunological status or a

healthy immune status. Similarly, in certain preferred embodiments a test subject can be regarded as having a compromised immunological status and/or an unhealthy immune status when a TCR or IG sequence distribution (entropy) score for a sample from the test subject is, with statistical significance, two standard deviations below that of a sample from a control subject, wherein said control subject is known to have an uncompromised immunological status or a healthy immune status. Status categorization can then inform diagnosis, prognosis and/or treatment strategies.

[00107] For instance and by way of non-limiting example, age-related decline in adaptive immune system capabilities can be detected according to the herein described methods, such that elderly patients can be immunologically profiled for purposes of predicting whether or not they would be likely to respond immunologically to a vaccine. As another non-limiting example, hematopoietic cell transplant recipients can be tested periodically post-transplant to determine whether or when adaptive immunity has been reconstituted by transplanted cells, so that prophylactic anti-infective (*e.g.*, antibiotic, anti-viral, etc.) and/or immunosuppressive therapies (*e.g.*, to treat graft-versus-host disease (GVHD)) can be adjusted on the basis of each patient's adaptive immune system status instead of on the basis of a fixed regimen. As another non-limiting example, the immune repertoire and immunocompetence of solid organ transplant recipients (*e.g.*, a recipient of all or a portion of a transplanted liver, lung, kidney, pancreas, intestine, heart, or skin) can be tested periodically to determine whether and to what extent the host adaptive immune system can be involved in graft rejection. From such test results, the clinician can adjust immunosuppressive therapies as needed, for example, to palliate rejection or to reduce or avoid potentially deleterious side-effects of excessive immunosuppressive therapy.

[00108] In yet another non-limiting example, immunocompetence can be assessed as described herein in candidate immunotherapy recipients such as oncology patients, in order to predict which patients can be likely to respond positively to immunotherapy and which are unlikely to do so.

[00109] As also noted elsewhere herein, using the presently-disclosed methods, it has been determined that in tumor tissue samples obtained from patients in whom the TCR (TCRB) repertoire exhibited relatively higher TCR sequence diversity and high clonality of TCR sequence distribution (*i.e.*, evidence of T cell migration to the tumor and clonal proliferation within the tumor) were more likely to benefit from immunotherapy designed to inhibit negative regulators of adaptive immune mechanisms (*e.g.*, anti-PD-1 antibodies). Such patients responded beneficially to immunotherapy and had better clinical outcomes that were

attributable to robust immune response within the tumor tissue as detected by the present methods (Figure 10; Figure 11). Thus, in the PD-1 study, it was determined that responders have many more infiltrating T cells than non-responders (*i.e.*, there are more total infiltrating T cells present), but that those T cells are distributed quite unevenly (*i.e.*, high clonality).

[00110] In a different study, it was further observed that in a subset of non-responder subjects, there was a decline in both TCRB sequence diversity and sequence distribution entropy shortly after immunotherapy (with anti-CTLA-4 antibody) was first administered, indicating differential responses to immunotherapy that can be measured by the present methods (Figure 8; Figure 9).

2. Amplification and Primers

[00111] According to these and related embodiments of the herein described methods, the method includes amplifying DNA extracted from or generated from the sample in a multiplexed PCR using (1) a plurality of AIR V-segment oligonucleotide primers and (2) either a plurality of AIR J-segment oligonucleotide primers or a plurality of AIR C-segment oligonucleotide primers. These primers are also described in detail in Robins *et al.*, 2009 *Blood* 114, 4099; Robins *et al.*, 2010 *Sci. Translat. Med.* 2:47ra64; Robins *et al.*, 2011 *J. Immunol. Meth.* doi:10.1016/j.jim.2011.09.001; Sherwood *et al.* 2011 *Sci. Translat. Med.* 3:90ra61; U.S.A.N. 13/217,126, U.S.A.N. 12/794,507, WO/2010/151416, WO/2011/106738 (PCT/US2011/026373), WO2012/027503 (PCT/US2011/049012), U.S.A.N. 61/550,311, and U.S.A.N. 61/569,118. The plurality of V-segment oligonucleotide primers are each independently capable of specifically hybridizing to at least one polynucleotide encoding a mammalian AIR V-region polypeptide, wherein each V-segment primer comprises a nucleotide sequence of at least 15 contiguous nucleotides that is complementary to at least one functional AIR-encoding gene segment and wherein the plurality of V-segment primers specifically hybridize to substantially all functional AIR V-encoding gene segments that are present in the sample,. The plurality of J-segment oligonucleotide primers are each independently capable of specifically hybridizing to at least one polynucleotide encoding a mammalian AIR J-region polypeptide, wherein each J-segment primer comprises a nucleotide sequence of at least 15 contiguous nucleotides that is complementary to at least one functional AIR J-encoding gene segment and wherein the plurality of J-segment primers specifically hybridize to substantially all functional AIR J-encoding or gene segments that are present in the sample. Moreover, the plurality of C-segment oligonucleotide primers are each independently capable of specifically hybridizing to at least one polynucleotide encoding a mammalian AIR C-region polypeptide, wherein each C-segment primer comprises a

nucleotide sequence of at least 15 contiguous nucleotides that is complementary to at least one functional AIR C-encoding gene segment and wherein the plurality of C-segment primers specifically hybridize to substantially all functional AIR C-encoding or gene segments that are present in the sample. In some embodiments, the V-segment and J- or C-segment primers are capable of promoting amplification in said multiplex polymerase chain reaction (PCR) of substantially all rearranged AIR CDR3-encoding regions in the sample to produce said plurality of amplified rearranged DNA molecules from a population of adaptive immune cells in the sample, said plurality of amplified rearranged DNA molecules being sufficient to quantify diversity of the AIR CDR3-encoding region in the population of T cells.

Alternatively, the method can simply involve sequence analysis of the aforementioned amplified DNA sequence data sufficient to characterize the sample with respect to the absolute and/or relative number of distinct clones present in the sample. As used herein, a functional AIR-encoding gene segment refers to a TCR or IG encoding gene segment that has undergone rearrangement in the DNA of a lymphoid cell and that is productively expressed, for instance, such that in preferred embodiments rearrangements that involve pseudogenes are not included, nor are rearrangements that result in an out-of-frame or prematurely terminated AIR polypeptide.

[00112] Methods of amplification, sequencing and primers are discussed in further detail herein.

3. Stratification of Patient Populations

[00113] According to certain related embodiments disclosed herein, there is provided a method for stratifying a patient population according to relative likelihood of responding to immunotherapy, comprising the following steps:

[00114] (a) identifying, in at least one sample (*e.g.*, a solid tumor sample) containing lymphoid cell nucleic acid from each of a plurality of patients who are candidate immunotherapy recipients, nucleic sequence information for each of a plurality of unique rearranged nucleic acid sequences that encode an adaptive immune receptor (AIR) polypeptide, and therefrom determining a total number of unique rearranged AIR polypeptide encoding nucleic acid sequences in each patient to quantify AIR sequence diversity in the subject;

[00115] (b) quantifying, in each of the plurality of patients, a frequency of occurrence of each unique rearranged nucleic sequence identified in (a) as a percentage of the total number of observed rearranged AIR polypeptide encoding nucleic sequences to determine AIR sequence distribution in the subject;

[00116] (c) rating each of the plurality of patients according to the total number of unique rearranged sequences in the patient from (a) and according to the frequency of occurrence of each unique rearranged sequence in the patient from (b), giving lower clonality ratings to patients in whom the Shannon entropy (calculated from the distribution of the frequency of each unique rearranged AIR and normalized to the range [0-1] by dividing by the logarithm of the number of unique rearranged AIRs) is high (*i.e.*, indicating an AIR repertoire with little specific clonal expansion) and giving higher clonality ratings to patients in whom the Shannon entropy (calculated from the distribution of the frequency of each unique rearranged AIR and normalized to the range [0-1] by dividing by the logarithm of the number of unique rearranged AIRs) is low (*i.e.*, indicating an AIR repertoire with extensive specific clonal expansion); and

[00117] (d) categorizing a patient having a lower clonality rating (as measured from a solid tumor sample) as having a lower relative likelihood of responding to immunotherapy than does a patient having a higher clonality rating, and thereby stratifying the patient population according to relative likelihood of responding to immunotherapy. In some embodiments, the at least one sample is a solid tumor sample.

[00118] Practicing these method steps employs compositions and methodologies similar to those described elsewhere herein.

4. Determining Immunological Status for Managing Treatment

[00119] According to certain other related embodiments, there is provided a method for determining immunological status to manage treatment of a test subject undergoing immunotherapy, comprising the following steps:

[00120] (a) identifying, in one or more samples containing lymphoid cell DNA obtained from a test subject at each of one or a plurality of timepoints prior to immunotherapy and at each of one or a plurality of timepoints during or after immunotherapy, nucleic acid sequence information for each of a plurality of unique rearranged nucleic acid sequences that encode an adaptive immune receptor (AIR) polypeptide, and therefrom determining a total number of unique rearranged AIR polypeptide encoding nucleic acid sequences in the test subject at each of said one or a plurality of timepoints to quantify AIR sequence diversity in the subject;

[00121] (b) quantifying, in each of the one or more samples, a frequency of occurrence of each unique rearranged nucleic acid sequence identified in (a) as a percentage of the total number of unique rearranged AIR polypeptide encoding nucleic acid sequences in the test subject at each of said one or a plurality of timepoints to determine AIR sequence distribution in the subject at each of said one or a plurality of timepoints;

[00122] (c) rating each of the one or more samples according to the total number of unique rearranged sequences determined in (a) and according to the frequency of occurrence of each unique rearranged sequence quantified in (b), and using the AIR frequency distribution to extrapolate the number of total unique AIR sequences in the subject at each of said one or a plurality of timepoints, to obtain a profile of AIR sequence diversity over time in the test subject, and (d) assigning an altered course of immunotherapy to the subject based on extrapolated total AIR sequence diversity, wherein high extrapolated total AIR sequence diversity indicates increased likelihood the subject has a healthy immune status and will be able to successfully clear infections and low extrapolated total AIR sequence diversity indicates increased likelihood the subject has an unhealthy immune status and will be unable to successfully clear infections.

[00123] Practicing these method steps employs compositions and methodologies similar to those described elsewhere herein.

IV. IMMUNOTHERAPY AND IMMUNOCOMPETENCE

[00124] Immunocompetence can be usefully understood to include the capacity or potential of an individual's adaptive immune system to mount an effective immune response, such as an immune response that is directed to a particular tumor or to a pathogen (*e.g.*, an infective bacteria, virus, fungus or other microbial or disease-causing agent) such that the tumor or pathogen is eradicated or neutralized. According to certain embodiments of the present disclosure, there are described methods for assessing immunocompetence, which methods can be predictive of an individual's likelihood of responding in a clinically beneficial manner to immunotherapy.

[00125] Hence and as also described elsewhere herein, an immunocompetent adaptive immune system, such as that of a clinically healthy, normal individual, or population of individuals, known by clinical criteria to be free of any risk or presence of disease or immunological disorder, will be characterized by a relatively high degree of AIR sequence diversity and high entropy of AIR sequence distribution in samples obtained from the subject's blood. Conversely, an immunoincompetent adaptive immune system (*e.g.*, relatively poor capacity of an adaptive immune system to mount an immune response) is shown herein to be characterized by relatively low AIR sequence diversity and low entropy of AIR sequence distribution in samples obtained from the subject's blood. AIR sequence diversity and entropy of AIR sequence distribution are herein shown to be dynamic over time, and can tend to decline over time as a correlate of increasing age, increasing

susceptibility to disease, decreasing likelihood of responding robustly to vaccines or to other immunotherapies, and/or other clinically relevant criteria.

[00126] In some embodiments, in solid tumor samples obtained from patients before and after immunotherapy with an inhibitor of a negative regulator of immune response, a high level of infiltrating T cells and high clonality were associated with a positive response to immunotherapy. In contrast, a low level of infiltrating T cells and low clonality in solid tumors were associated with failure to respond to treatment.

[00127] Immunotherapy can include any of a variety of interventions by which the activity levels of one or more cells of the adaptive immune system are altered (*e.g.*, up- or down-regulated in a statistically significant manner). For example, the intervention can induce, recruit, enhance or otherwise potentiate an adaptive immune response, which in preferred embodiments will be an antigen-specific immune response. In certain embodiments, immunotherapy can comprise administration of one or more specific antibodies that recognize adaptive immune system cells to alter the immunological activity of such cells. Other immunotherapeutic approaches include the use of cytokines that similarly can directly or indirectly alter immunocyte activity; vaccines that elicit adaptive immune responses such as antigen-specific responses to tumor-associated antigens; hematopoietic cell transplants which include bone marrow transplants, cord blood transplants and autologous hematopoietic cell transplants including autologous T cell transplants (*e.g.*, Blume and Thomas, 2000 *Biol. Blood Marrow Transpl.* 6(1):1-12); inhibitors of negative regulators of adaptive immune responses such as inhibitors of CTLA4/CD152 (*e.g.*, ipilimumab, tremelimumab; Callahan *et al.*, 2010 *Sem. Oncol.* 37:473), inhibitors of LAG3/CD223 (Huard *et al.*, 1996 *Eur. J. Immunol.* 26:1180; Baixeras *et al.*, 1992 *J. Exp. Med.* 176:327; Hannier *et al.*, 1998 *J. Immunol.* 161:4058; Huard *et al.*, 1994 *Eur. J. Immunol.* 24:3216); and other immunotherapeutic agents including in some cases immunosuppressive agents (*e.g.*, *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, (12th Ed., Brunton *et al.*, Eds., McGraw Hill, NY, 2011, pages 909-1099; 1891-1990; Murphy, *Janeway's Immunobiology* (8th Ed.), 2011 Garland Science, NY, pp. 669-716).

[00128] Accordingly, in certain embodiments, immunotherapy can comprise treatment with an immunotherapy agent, such as an immunotherapeutic antibody, a cytokine, a hematopoietic cell transplant, an immunosuppressive agent, or a vaccine. In certain embodiments, immunotherapy comprises treatment with an inhibitor of a negative regulator of an immune response. The negative regulator of an immune response can be one or more of CTLA4/CD152, LAG3/CD223, and PD-1/CD279. For example, the negative regulator of

an immune response can be CTLA-4/CD152 and the inhibitor of the negative regulator of an immune response is an anti-CTLA-4 antibody, such as ipilimumab (*e.g.*, Lyseng-Williamson *et al.*, 2012 *Am. J. Clin. Dermatol.* 13:349; Jeter *et al.*, 2012 *Clin. Med. Insights Oncol.* 6:275; Waitz *et al.*, 2012 *Canc. Res.* 72:430) or tremelimumab (*e.g.*, Callahan *et al.*, 2010 *Sem. Oncol.* 37:473; Ascieto *et al.* 2011 *J. Transl. Med.* 9:196; Calabro *et al.*, 2010 *Sem. Oncol.* 37:460; Ribas, 2010 *Sem. Oncol.* 37:450). In certain embodiments, the negative regulator of an immune response can be PD-1/CD279, and the inhibitor of the negative regulator of an immune response is an anti-PD-1 antibody. In certain embodiments, immunotherapy can comprise treatment with an agent that targets a potentiator of an immune response. The potentiator of an immune response can be 41BB/CD137 (Kwon *et al.*, 1989 *Proc. Nat. Acad. Sci. USA* 86:1963), OX40/CD134 (GenBank Acc. No. AJ277151) or CD40 (Banchereau *et al.*, 1994 *Ann. Rev. Immunol.* 12:881).

[00129] In certain other embodiments, immunotherapy can comprise treatment of an inflammatory condition or an autoimmune disease with an inhibitor of an inflammatory pathway. Contemplated inflammatory conditions or autoimmune diseases include rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis, Crohn's disease and juvenile idiopathic arthritis. Inflammatory mechanisms have been extensively characterized (*e.g.*, Goodman & Gilman's *The Pharmacological Basis of Therapeutics*, (12th Ed., Brunton *et al.*, Eds., McGraw Hill, NY, 2011, pages 909-1099; 1891-1990; Murphy, *Janeway's Immunobiology* (8th Ed.), 2011 Garland Science, NY), such that in these and related embodiments the inflammatory pathway comprises at least one of tumor necrosis factor-alpha (TNF α), interferon-gamma (IFN γ), interleukin-1 (IL-1), interleukin-6 (IL-6), interleukin-8 (IL-8). For instance, for inflammatory pathways that include TNF α , there are known inhibitors of the inflammatory pathway that specifically bind to TNF α , such as anti-TNF α antibodies (*e.g.*, adalimumab, infliximab) and artificial soluble TNF α receptors (*e.g.*, etanercept).

A. Exemplary Targets for Immunotherapy

[00130] The ability to quantify the immunocompetency of a patient's adaptive immune system (as defined as either diversity in the blood or clonality in a tumor tissue, in the case of solid tumors) prior to treatment with an immunomodulatory drug or treatment regimen is predictive of response to treatment and correlative to overall survival.

[00131] There are many targets in the immune cascade that can be targeted as potential treatments to cancer and other therapeutic areas. Some are expressed on the surface of T cells and are negative regulators of the immune response, and some are expressed on the

surface of antigen presenting cells and are thought to upregulate the immune cascade. Below are exemplary targets in immunotherapy that are or have been used in the clinic (anti-CTLA, ipilimumab) or in active clinical trials, and are currently in use by multiple pharmaceutical manufacturers after reporting successful early data in patients. These targets and inhibitors or regulators thereof can be used in immunotherapy or treatment measures, in accordance with methods of the invention described herein.

[00132] Anti-CTLA. Two proteins on the surface of T cells--CD28 and cytotoxic T-lymphocyte antigen 4 (CTLA-4)--play important roles in the regulation of immune activation and tolerance. CD28 provides positive modulatory signals in the early stages of an immune response, while CTLA-4 signaling inhibits T-cell activation, particularly during strong T-cell responses. CTLA-4 blockade using anti-CTLA-4 monoclonal antibody therapy has great appeal because suppression of inhibitory signals results in the generation of an antitumor T-cell response. Both clinical and preclinical data indicate that CTLA-4 blockade results in direct activation of CD4+ and CD8+ effector cells, and anti-CTLA-4 monoclonal antibody therapy has shown promise in a number of cancers, particularly melanoma. *Oncologist*. 2008;13 Suppl 4:2-9. doi: 10.1634/theoncologist.13-S4-2. There are two anti-CTLA4 compounds: (i) Yervoy, or ipilimumab (Bristol Myers Squibb (BMS)) and (ii) tremelimumab (Medimmune).

[00133] PD-1. Programmed death 1 (PD-1) and its ligands, PD-L1 and PD-L2, deliver inhibitory signals that regulate the balance between T cell activation, tolerance, and immunopathology. Immune responses to foreign and self-antigens require specific and balanced responses to clear pathogens and tumors and yet maintain tolerance to self-antigens. Induction and maintenance of T cell tolerance requires PD-1, and its ligand PD-L1 on nonhematopoietic cells can limit effector T cell responses and protect tissues from immune-mediated tissue damage. The PD-1:PD-L pathway also has been usurped by microorganisms and tumors to attenuate antimicrobial or tumor immunity and facilitate chronic infection and tumor survival. The identification of B7-1 as an additional binding partner for PD-L1, together with the discovery of an inhibitory bidirectional interaction between PD-L1 and B7-1, reveals new ways the B7:CD28 family regulates T cell activation and tolerance. *Annu Rev Immunol*. 2008;26:677-704. doi: 10.1146/annurev.immunol.26.021607.090331. There are at least 5 known PD-1 compounds in development: Merck (MK-3475), or lambrolizumab; BMS (MBS-936558), or nivolumab; Medimmune (MEDI4736); Glaxo (AMP-224); Genentech (MPDL3280A).

[00134] **4-1BB**. 4-1BB (CD137), a member of the TNF receptor superfamily, is an activation-induced T-cell costimulatory molecule. Signaling via 4-1BB upregulates survival genes, enhances cell division, induces cytokine production, and prevents activation-induced cell death in T cells. The importance of the 4-1BB pathway has been underscored in a number of diseases, including cancer. Growing evidence indicates that anti-4-1BB monoclonal antibodies possess strong antitumor properties, which in turn are the result of their powerful CD8⁺ T-cell activating, IFN- γ producing, and cytolytic marker-inducing capabilities. In addition, combination therapy of anti-4-1BB with other anticancer agents, such as radiation, has robust tumor-regressing abilities against nonimmunogenic or poorly immunogenic tumors. *Mol Cancer Ther*; 11(5); 1062-70, 2012 AACR. Two examples of 4-1BB compounds are being developed by Pfizer (PF-05082566) and BMS (BMS-663513).

[00135] **CD40**. CD40 (CD154) is a costimulatory protein found on antigen presenting cells and is required for their activation. The binding of CD154 (CD40L) on T_H cells to CD40 activates antigen presenting cells and induces a variety of downstream effects. The protein receptor encoded by this gene is a member of the TNF-receptor superfamily. This receptor has been found to be essential in mediating a broad variety of immune and inflammatory responses including T cell-dependent immunoglobulin class switching, memory B cell development, and germinal center formation. Entrez Gene: CD40 molecule, TNF receptor superfamily member 5; [En.wikipedia.org/wiki/CD40_\(protein\)](http://En.wikipedia.org/wiki/CD40_(protein)). Exemplary CD40 compounds include, but are not limited to, the following developed by Seattle Genetics/Genentech (dacetuzumab) and Novartis (lucatumumab).

[00136] **LAG-3**. LAG-3 (CD223) is a cell surface molecule expressed on activated T cells (Huard *et al.* *Immunogenetics* 39:213-217, 1994), NK cells (Triebel *et al.* *J Exp Med* 171:1393-1405, 1990), B cells (Kisielow *et al.* *Eur J Immunol* 35:2081-2088, 2005), and plasmacytoid dendritic cells (Workman *et al.* *J Immunol* 182:1885-1891, 2009) that plays an important but incompletely understood role in the function of these lymphocyte subsets. In addition, the interaction between LAG-3 and its major ligand, Class II MHC, is thought to play a role in modulating dendritic cell function (Andreae *et al.* *J Immunol* 168:3874-3880, 2002). Recent preclinical studies have documented a role for LAG-3 in CD8 T cell exhaustion (Blackburn *et al.* *Nat Immunol* 10:29-37, 2009), and blockade of the LAG-3/Class II interaction using a LAG-3 Ig fusion protein is being evaluated in a number of clinical trials in cancer patients. *Curr Top Microbiol Immunol.* 2011;344:269-78. doi: 10.1007/82_2010_114. LAG-3 is being developed as a target, by companies such as BMS.

[00137] Breadth of Targets Along the Immune Cascade. Immune modulation can also be categorized by compound family (versus specific target) into either a member of the immunoglobulin family or the TNF family. *See* Nature Reviews Drug Discovery 12, 130-146 (February 2013) (doi:10.1038/nrd3877). This categorization is useful to highlight the breadth of therapeutic categories outside of cancer that these targets can hit, and for which a measure of immunocompetence can be equally as relevant.

Table 1. Example List of Targets for Immunotherapy

Name	Companies	Type of biologic	Pathways	Roles	Indications	Trial phase
<i>immunoglobulin family</i>						
Tremelimumab	MedImmune/ AstraZeneca	CTLA4-specific human IgG2	CTLA4-B7.1, CTLA4-B7.2, B7H2	T cell priming and activation	Solid tumours	II
Galiximab	Cancer and Leukemia Group B (CALGB)/ Biogen Idec	B7.1-specific chimeric IgG1	B7.1	B cell proliferation	Lymphoma	II
BMS-936558	Bristol-Myers Squibb/ Medarex	PD1-specific human IgG4	PD1-B7H1, PD1-B7DC	T cell activation and tolerance	Multiple cancers; HCV	III
CT-011	CureTech	PD1-specific humanized IgG1	PD1-B7H1, PD1-B7DC	T cell activation and tolerance	Advanced solid tumours; HCV	II
MK-3475	Merck/Schering-Plough	PD1-specific IgG4	PD1-B7H1, PD1-B7DC	T cell activation and tolerance	Advanced or metastatic solid tumours	I
AMP224	Amplimmune/ GlaxoSmithKline	B7DC and human IgG1 fusion protein	PD1-B7H1, PD1-B7DC	T cell activation and tolerance	Multiple cancers	I
BMS-936559	Bristol-Myers Squibb	B7H1-specific human IgG4	PD1-B7H1	T cell activation and tolerance	Advanced or recurrent solid tumours	I
MPDL3280A	Genentech/Roche	B7H1-specific engineered human IgG1	PD1-B7H1	T cell activation and tolerance	Solid tumours	I
MEDI4736	MedImmune/ AstraZeneca	B7H1-specific engineered human IgG1	PD1-B7H1	T cell activation and tolerance	Solid tumours	I
MEDI-570	MedImmune/ AstraZeneca	ICOS-specific human IgG	ICOS-B7H2	T cell-dependent B cell response	SLE	I
AMG 557	Amgen	B7H2-specific human IgG	ICOS, CD28, CTLA4	T cell-dependent B cell response	SLE, psoriasis	I
MGA271	Macrogenics	B7H3-specific, ADCC-enhanced humanized IgG1	B7H3	T cell activation and tolerance	Solid tumours	I
IMP121	Immutep	LAG3 and human IgG1 fusion protein	LAG3-MHCII	DC maturation and T cell activation	Multiple cancers	I/II
<i>TNF family</i>						
BMS-885533	Bristol-Myers Squibb	CD137-specific human IgG4	CD137	T cell activation	Solid tumours	I/II
PF-05082566	Pfizer	CD137-specific human IgG	CD137	T cell activation	Lymphoma	I
CDX-1127	Celidex	CD27-specific human IgG1	CD27	T cell activation	Multiple cancers	I
Anti-OX40	Providence Health & Services	OX40-specific mouse IgG	OX40	CD4 T cell activation	Prostate cancer	II
huMAb OX40L	Genentech/Roche	OX40L-specific human IgG1	OX40-OX40L	CD4 T cell activation	Asthma	II
TRX518	GTR Inc.	GITR-specific humanized IgG1	GITR-GITRL	T cell activation	Solid tumours	I
Atacicept	ZymoGenetics/ EMD Serono	TACI and human IgG1 fusion protein	TACI, BCMA and BAFFR	B cell activation and antibody production	SLE, rheumatoid arthritis, multiple sclerosis and optic neuritis	II/III
CP-870,893	Pfizer	CD40-specific human IgG1	CD40	APC activation and B cell maturation	Multiple cancers	I
Lucatumumab	Novartis	CD40-specific human IgG1	CD40	APC activation and B cell maturation	Lymphoma and leukaemia	I/II
Dacetuzumab	Seattle Genetics	CD40-specific humanized IgG1	CD40	APC activation and B cell maturation	Lymphoma and multiple myeloma	II

ADCC, antibody dependent cell-mediated cytotoxicity; APC, antigen-presenting cell; B7H1, B7 homolog 1; BAFFR, B cell activation factor receptor; BCMA, B cell maturation antigen; CTLA4, cytotoxic T lymphocyte antigen 4; DC, dendritic cell; GITR, glucocorticoid-induced TNFR-related protein; GITRL, GITR ligand; HCV, hepatitis C virus; ICOS, inducible co-stimulator; IgG1, immunoglobulin G1; LAG3, lymphocyte activation gene 3; MHCII, major histocompatibility complex class II; OX40L, OX40 ligand; PD1, programmed cell death protein 1; SLE, systemic lupus erythematosus; TACI, transmembrane activator and CAML interactor; TNF, tumour necrosis factor.

[00138] Nature Reviews Drug Discovery 12, 130-146 (February 2013)

(doi:10.1038/nrd3877).

[00139] Samples and Subjects. The subject or biological source, from which a test biological sample can be obtained, can be a human or non-human animal, or a transgenic or

cloned or tissue-engineered (including through the use of stem cells) organism. In certain preferred embodiments of the invention, the subject or biological source can be known to have, or can be suspected of having or being at risk for having, cancer or another malignant condition, or an autoimmune disease, or an inflammatory condition, or a bacterial, viral, fungal or other microbial infection, or the subject or biological source can be a solid organ transplant recipient (*e.g.*, recipient of all or a portion of a transplanted liver, lung, kidney, pancreas, intestine, heart, or skin). In some embodiments, or the subject or biological source can be a hematopoietic cell transplant recipient (*e.g.*, recipient of a bone marrow transplant, cord blood transplant, autologous T cell transplant, *etc.*). In certain embodiments of the invention, the subject or biological source can be known to be free of a risk or presence of such disease. The test biological sample can be obtained from the subject or biological source at one or a plurality of timepoints, for example, at one or a plurality of timepoints prior to administration of treatment or therapy (*e.g.*, immunotherapy) to the subject or biological source, and also at one or a plurality of timepoints during or after administration of treatment or therapy (*e.g.*, immunotherapy) to the subject or biological source.

[00140] Certain preferred embodiments contemplate a subject or biological source that is a human subject such as a patient that has been diagnosed as having or being at risk for developing or acquiring cancer according to art-accepted clinical diagnostic criteria, such as those of the U.S. National Cancer Institute (Bethesda, MD, USA) or as described in *DeVita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology* (2008, Lippincott, Williams and Wilkins, Philadelphia/ Ovid, New York); Pizzo and Poplack, *Principles and Practice of Pediatric Oncology* (Fourth edition, 2001, Lippincott, Williams and Wilkins, Philadelphia/ Ovid, New York); Vogelstein and Kinzler, *The Genetic Basis of Human Cancer* (Second edition, 2002, McGraw Hill Professional, New York); Dancey *et al.* (2009 *Semin. Oncol.* 36 Suppl.3:S46). Certain embodiments contemplate a human subject that is known to be free of a risk for having, developing or acquiring cancer by such criteria. Examples of malignant conditions that are contemplated according to certain present embodiments can include solid tumors such as melanoma, sarcoma, and carcinoma. Others can also include, for example, malignant melanoma, small cell lung cancer, non-small cell lung cancer, renal cell carcinoma, pancreatic cancer, breast cancer, ovarian cancer and prostate cancer.

[00141] Certain other embodiments contemplate a non-human subject or biological source, for example a non-human primate such as a macaque, chimpanzee, gorilla, vervet, orangutan, baboon or other non-human primate, including such non-human subjects that can be known to the art as preclinical models, including preclinical models for solid tumors and/or other

cancers. Certain other embodiments contemplate a non-human subject that is a mammal, for example, a mouse, rat, rabbit, pig, sheep, horse, bovine, goat, gerbil, hamster, guinea pig or other mammal; many such mammals can be subjects that are known to the art as preclinical models for certain diseases or disorders, including lymphoid hematopoietic malignancies and/or other cancers (*e.g.*, Li *et al.*, 2011 *Dis. Model. Mech.* 4:311; von Euler *et al.*, 2011 *Vet. Comp. Oncol.* 9:1; Goldstein *et al.*, 2010 *Expert Rev. Hematol.* 3:301; Diamond *et al.*, 2009 *J. Bone Min. Res.* 24:1150; Macor *et al.*, 2008 *Curr. Pharm. Des.* 14:2023; Talmadge *et al.*, 2007 *Am. J. Pathol.* 170:793; Kerbel, 2003 *Canc. Biol. Therap.* 2(4 Suppl 1):S134; Man *et al.*, 2007 *Canc. Met. Rev.* 26:737; Cespedes *et al.*, 2006 *Clin. Transl. Oncol.* 8:318). The range of embodiments is not intended to be so limited, however, such that there are also contemplated other embodiments in which the subject or biological source can be a non-mammalian vertebrate, for example, another higher vertebrate, or an avian, amphibian or reptilian species, or another subject or biological source.

[00142] As also noted elsewhere herein, art-accepted clinical diagnostic criteria have been established for these and other cancer types, such as those promulgated by the U.S. National Cancer Institute (Bethesda, MD, USA) or as described in *DeVita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology* (2008, Lippincott, Williams and Wilkins, Philadelphia/ Ovid, New York); Pizzo and Poplack, *Principles and Practice of Pediatric Oncology* (Fourth edition, 2001, Lippincott, Williams and Wilkins, Philadelphia/ Ovid, New York); and Vogelstein and Kinzler, *The Genetic Basis of Human Cancer* (Second edition, 2002, McGraw Hill Professional, New York). Other non-limiting examples of typing and characterization of particular cancers are described, *e.g.*, in Ignatiadis *et al.* (2008 *Pathobiol.* 75:104); Kunz (2008 *Curr. Drug Discov. Technol.* 5:9); and Auman *et al.* (2008 *Drug Metab. Rev.* 40:303).

[00143] Biological samples can be provided by obtaining a blood sample, biopsy specimen, excised tumor specimen such as a solid tumor specimen, tissue explant, organ culture, biological fluid or any other tissue or cell preparation from a subject or a biological source. B cells and T cells can thus be obtained from a biological sample, such as from a variety of tissue and biological fluid samples including bone marrow, thymus, lymph glands, lymph nodes, peripheral tissues and blood, and also from tumor tissues (*e.g.*, tumor-infiltrating lymphocytes), but peripheral blood is most easily accessed. Any peripheral tissue can be sampled for the presence of B and T cells and is therefore contemplated for use in the methods described herein. Tissues and biological fluids from which adaptive immune cells can be obtained include, but are not limited to skin, epithelial tissues, colon, spleen, a

mucosal secretion, oral mucosa, intestinal mucosa, vaginal mucosa or a vaginal secretion, cervical tissue, ganglia, saliva, cerebrospinal fluid (CSF), bone marrow, cord blood, serum, serosal fluid, plasma, lymph, urine, ascites fluid, pleural fluid, pericardial fluid, peritoneal fluid, abdominal fluid, culture medium, conditioned culture medium or lavage fluid. In certain embodiments, adaptive immune cells (*e.g.*, hematopoietic cells of lymphoid lineage such as T cells and B cells) can be isolated from an apheresis sample. Peripheral blood samples can be obtained by phlebotomy from subjects. Peripheral blood mononuclear cells (PBMC) are isolated by techniques known to those of skill in the art, *e.g.*, by Ficoll-Hypaque[®] density gradient separation. In certain embodiments, whole PBMCs are used for analysis.

[00144] In certain related embodiments, preparations that comprise predominantly lymphocytes (*e.g.*, T and B cells) or that comprise predominantly T cells or predominantly B cells, can be prepared for use as a biological sample as provided herein, according to established, art-accepted methodologies. In other related embodiments, specific subpopulations of T or B cells can be isolated prior to analysis using the methods described herein. Various methods and commercially available kits for isolating different subpopulations of T and B cells are known in the art and include, but are not limited to, subset selection immunomagnetic bead separation or flow immunocytometric cell sorting using antibodies specific for one or more of any of a variety of known T and B cell surface markers. Illustrative markers include, but are not limited to, one or a combination of CD2, CD3, CD4, CD8, CD14, CD19, CD20, CD25, CD28, CD45RO, CD45RA, CD54, CD62, CD62L, CDw137 (41BB), CD154, GITR, FoxP3, CD54, and CD28. For example, and as is known to the skilled person, cell surface markers, such as CD2, CD3, CD4, CD8, CD14, CD19, CD20, CD45RA, and CD45RO can be used to determine T, B, and monocyte lineages and subpopulations in flow cytometry. Similarly, forward light-scatter, side-scatter, and/or cell surface markers such as CD25, CD62L, CD54, CD137, CD154 can be used to determine activation state and functional properties of cells.

[00145] Illustrative combinations useful in certain of the methods described herein can include CD8⁺CD45RO⁺ (memory cytotoxic T cells), CD4⁺CD45RO⁺ (memory T helper), CD8⁺CD45RO⁻ (CD8⁺CD62L⁺CD45RA⁺ (naïve-like cytotoxic T cells); CD4⁺CD25⁺CD62L^{hi}GITR⁺FoxP3⁺ (regulatory T cells). Illustrative antibodies for use in immunomagnetic cell separations or flow immunocytometric cell sorting include fluorescently labeled anti-human antibodies, *e.g.*, CD4 FITC (clone M-T466, Miltenyi Biotec), CD8 PE (clone RPA-T8, BD Biosciences), CD45RO ECD (clone UCHL-1,

Beckman Coulter), and CD45RO APC (clone UCHL-1, BD Biosciences). Staining of total PBMCs can be done with the appropriate combination of antibodies, followed by washing cells before analysis. Lymphocyte subsets can be isolated by fluorescence activated cell sorting (FACS), *e.g.*, by a BD FACSAria™ cell-sorting system (BD Biosciences) and by analyzing results with FlowJo™ software (Treestar Inc.), and also by conceptually similar methods involving specific antibodies immobilized to surfaces or beads.

[00146] For nucleic acid extraction, total genomic DNA can be extracted from cells using methods known in the art and/or commercially available kits, *e.g.*, by using the QIAamp® DNA blood Mini Kit (QIAGEN®). The approximate mass of a single haploid genome is 3 pg. Preferably, at least 25,000 to 250,000 cells, for example, at least 50,000 to 125,000 cells, or at least 75,000 to 150,000 cells, or at least 100,000 to 200,000 cells, are used for analysis, *i.e.*, about 0.15 to 1.5 µg, or for instance, 0.6 to 1.2 µg DNA from diploid T or B cells. The number of T or B cells present in a sample can vary considerably when the sample is obtained from a patient having a lymphoid hematological malignancy such as acute T-cell lymphoblastic leukemia (T-ALL). Using peripheral blood mononuclear cells (PBMCs) from a normal healthy adult human as a source, the number of T cells can vary and can be estimated to be about 30% of total cells; the number of B cells can vary and can be estimated to be about 5-15% of total cells in a PBMC preparation.

V. ADAPTIVE IMMUNE RECEPTORS (AIR)

[00147] The native TCR is a heterodimeric cell surface protein of the immunoglobulin superfamily which is associated with invariant proteins of the CD3 complex involved in mediating signal transduction. TCRs exist in $\alpha\beta$ and $\gamma\delta$ forms, which are structurally similar but have quite distinct anatomical locations and probably functions. The MHC class I and class II ligands, which bind to the TCR, are also immunoglobulin superfamily proteins but are specialized for antigen presentation, with a highly polymorphic peptide binding site which enables them to present a diverse array of short peptide fragments at the APC cell surface.

[00148] The extracellular portions of native heterodimeric $\alpha\beta$ and $\gamma\delta$ TCRs consist of two polypeptides each of which has a membrane-proximal constant domain, and a membrane-distal variable domain. Each of the constant and variable domains includes an intra-chain disulfide bond. The variable domains contain the highly polymorphic loops analogous to the complementarity determining regions (CDRs) of antibodies. CDR3 of $\alpha\beta$ TCRs interact with the peptide presented by MHC, and CDRs 1 and 2 of $\alpha\beta$ TCRs interact with the peptide and

the MHC. The diversity of TCR sequences is generated via somatic rearrangement of linked variable (V), diversity (D), joining (J), and constant genes.

[00149] The Ig and TCR gene loci contain many different variable (V), diversity (D), and joining (J) gene segments, which are subjected to rearrangement processes during early lymphoid differentiation. Ig and TCR V, D and J gene segment sequences are known in the art and are available in public databases such as GENBANK.

[00150] The V-D-J rearrangements are mediated via a recombinase enzyme complex in which the RAG1 and RAG2 proteins play a key role by recognizing and cutting the DNA at the recombination signal sequences (RSS), which are located downstream of the V gene segments, at both sides of the D gene segments, and upstream of the J gene segments. Inappropriate RSS reduce or even completely prevent rearrangement. The recombination signal sequence (RSS) consists of two conserved sequences (heptamer, 5'-CACAGTG-3', and nonamer, 5'-ACAAAACCC-3'), separated by a spacer of either 12 +/- 1 bp ("12-signal") or 23 +/- 1 bp ("23-signal"). A number of nucleotide positions have been identified as important for recombination including the CA dinucleotide at position one and two of the heptamer, and a C at heptamer position three has also been shown to be strongly preferred as well as an A nucleotide at positions 5, 6, 7 of the nonamer. (Ramsden *et al.* 1994 *Nucl. Ac. Res.* 22:1785; Akamatsu *et al.* 1994 *J. Immunol.* 153:4520; Hesse *et al.* 1989 *Genes Dev.* 3:1053). Mutations of other nucleotides have minimal or inconsistent effects. The spacer, although more variable, also has an impact on recombination, and single-nucleotide replacements have been shown to significantly impact recombination efficiency (Fanning *et al.* 1996 *Cell. Immunol. Immunopath.* 79:1, Larijani *et al.* 1999 *Nucl. Ac. Res.* 27:2304; Nadel *et al.* 1998 *J. Immunol.* 161:6068; Nadel *et al.*, 1998 *J. Exp. Med.* 187:1495). Criteria have been described for identifying RSS polynucleotide sequences having significantly different recombination efficiencies (Ramsden *et al.* 1994 *Nucl. Ac. Res.* 22:1785; Akamatsu *et al.* 1994 *J. Immunol.* 153:4520; Hesse *et al.* 1989 *Genes Dev.* 3:1053, and Lee *et al.*, 2003 *PLoS* 1(1):E1).

[00151] The rearrangement process generally starts with a D to J rearrangement followed by a V to D-J rearrangement in the case of Ig heavy chain (IgH), TCR beta (TCRB), and TCR delta (TCRD) genes or concerns direct V to J rearrangements in case of Ig kappa (IgK), Ig lambda (IgL), TCR alpha (TCRA), and TCR gamma (TCRG) genes. The sequences between rearranging gene segments are generally deleted in the form of a circular excision product, also called TCR excision circle (TREC) or B cell receptor excision circle (BREC).

[00152] The many different combinations of V, D, and J gene segments represent the so-called combinatorial repertoire, which is estimated to be $\sim 2 \times 10^6$ for Ig molecules, $\sim 3 \times 10^6$ for TCR $\alpha\beta$ and $\sim 5 \times 10^3$ for TCR $\gamma\delta$ molecules. At the junction sites of the V, D, and J gene segments, deletion and random insertion of nucleotides occurs during the rearrangement process, resulting in highly diverse junctional regions, which significantly contribute to the total repertoire of Ig and TCR molecules, estimated to be $> 10^{12}$.

[00153] Mature B-lymphocytes further extend their Ig repertoire upon antigen recognition in follicle centers via somatic hypermutation, a process, leading to affinity maturation of the Ig molecules. The somatic hypermutation process focuses on the V- (D-) J exon of IgH and Ig light chain genes and concerns single nucleotide mutations and sometimes also insertions or deletions of nucleotides. Somatic-mutated Ig genes are also found in mature B-cell malignancies of follicular or post-follicular origin.

VI. AMPLIFICATION PRIMERS AND MULTIPLEX PCR

[00154] In certain preferred embodiments described herein, V-segment and J-segment primers can be employed in a PCR reaction to amplify rearranged TCR or Ig CDR3-encoding DNA regions in a test biological sample, wherein each functional TCR or Ig V-encoding gene segment comprises a V gene recombination signal sequence (RSS) and each functional TCR or Ig J-encoding gene segment comprises a J gene RSS. In these and related embodiments, each amplified rearranged DNA molecule can comprise (i) at least about 10, 20, 30 or 40 contiguous nucleotides of a sense strand of the TCR or Ig V-encoding gene segment, with the at least about 10, 20, 30 or 40 contiguous nucleotides being situated 5' to the V gene RSS and/or each amplified rearranged DNA molecule can comprise (ii) at least about 10, 20 or 30 contiguous nucleotides of a sense strand of the TCR or Ig J-encoding gene segment, with the at least about 10, 20 or 30 contiguous nucleotides being situated 3' to the J gene RSS. In certain preferred embodiments, each amplified TCR or Ig CDR3-encoding region is present in an amplified rearranged DNA molecule that is less than 600 nucleotides in length. Without wishing to be bound by theory, these design features for amplifying CDR3-encoding V-J junctional regions permit V-segment primer hybridization to substantially all functional TCR or Ig V-encoding gene segments, and also permit J-segment primer hybridization to substantially all functional TCR or Ig J-encoding segments, and also permit amplification of CDR3-encoding regions that are amenable to sequencing by the herein described high-throughput sequencing (HTS) platforms while including adequate sequence information to identify all possible V-D-J and V-J combinations.

VII. MULTIPLEX QUANTITATIVE PCR

[00155] As described herein and in view of Robins *et al.*, 2009 *Blood* 114, 4099; Robins *et al.*, 2010 *Sci. Translat. Med.* 2:47ra64; Robins *et al.*, 2011 *J. Immunol. Meth.* doi:10.1016/j.jim.2011.09. 001; Sherwood *et al.* 2011 *Sci. Translat. Med.* 3:90ra61; U.S.A.N. 13/217,126, U.S.A.N. 12/794,507, WO/2010/151416, WO/2011/106738 (PCT/US2011/026373), WO2012/027503 (PCT/US2011/049012), U.S.A.N. 61/550,311, and U.S.A.N. 61/569,118, according to certain preferred embodiments the present methods involve a multiplex PCR method using a set of forward primers that specifically hybridize to the V segments and a set of reverse primers that specifically hybridize to the J segments where the multiplex PCR reaction allows amplification of all the possible VJ (and VDJ) combinations within a given population of T or B cells.

[00156] DNA or RNA can be extracted from cells in a sample, such as a sample of blood or lymph or other sample from a subject known to contain lymphoid cells, using standard methods or commercially available kits known in the art. In some embodiments, genomic DNA is used. In other embodiments, cDNA is transcribed from mRNA obtained from the cells and then used for multiplex PCR.

[00157] A multiplex PCR system can be used to amplify rearranged adaptive immune cell receptor loci from genomic DNA, preferably from a CDR3 region. In certain embodiments, the CDR3 region is amplified from a TCR α , TCR β , TCR γ or TCR δ CDR3 region or similarly from an IgH or IgL (lambda or kappa) locus. Compositions are provided that comprise a plurality of V-segment and J-segment primers that are capable of promoting amplification in a multiplex polymerase chain reaction (PCR) of substantially all productively rearranged adaptive immune receptor CDR3-encoding regions in the sample for a given class of such receptors (*e.g.*, TCR γ , TCR β , IgH, etc.), to produce a multiplicity of amplified rearranged DNA molecules from a population of T cells (for TCR) or B cells (for Ig) in the sample. Preferably and in certain embodiments, primers are designed so that each amplified rearranged DNA molecule in the multiplicity of amplified rearranged DNA molecules is less than 600 nucleotides in length, thereby excluding amplification products from non-rearranged adaptive immune receptor loci.

[00158] In the human genome, there are currently believed to be about 70 TCR V α and about 61 J α gene segments, about 52 TCR V β , about 2 D β and about 13 J β gene segments, about 9 TCR V γ and about 5 J γ gene segments, and about 46 immunoglobulin heavy chain (IGH) V_H, about 23 D_H and about 6 J_H gene segments. Accordingly, where genomic sequences for these loci are known such that specific molecular probes for each of them can

be readily produced, it is believed according to non-limiting theory that the present compositions and methods relate to substantially all (*e.g.*, greater than 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%) of these known and readily detectable adaptive immune receptor V-, D- and J-region encoding gene segments.

[00159] The TCR and Ig genes can generate millions of distinct proteins via somatic mutation. Because of this diversity-generating mechanism, the hypervariable complementarity determining regions (CDRs) of these genes can encode sequences that can interact with millions of ligands, and these regions are linked to a constant region that can transmit a signal to the cell indicating binding of the protein's cognate ligand. The adaptive immune system employs several strategies to generate a repertoire of T- and B-cell antigen receptors with sufficient diversity to recognize the universe of potential pathogens. In $\alpha\beta$ and $\gamma\delta$ T cells, which primarily recognize peptide antigens presented by MHC molecules, most of this receptor diversity is contained within the third complementarity-determining region (CDR3) of the T cell receptor (TCR) α and β chains (or γ and δ chains).

[00160] The assay technology uses two pools of primers to provide for a highly multiplexed PCR reaction. The first, "forward" pool (*e.g.*, by way of illustration and not limitation, V-segment oligonucleotide primers described herein can in certain preferred embodiments be used as "forward" primers when J-segment oligonucleotide primers are used as "reverse" primers according to commonly used PCR terminology, but the skilled person will appreciate that in certain other embodiments J-segment primers can be regarded as "forward" primers when used with V-segment "reverse" primers) includes an oligonucleotide primer that is specific to (*e.g.*, having a nucleotide sequence complementary to a unique sequence region of) each V-region encoding segment ("V segment) in the respective TCR or Ig gene locus. In certain embodiments, primers targeting a highly conserved region are used, to simultaneously capture many V segments, thereby reducing the number of primers required in the multiplex PCR. Similarly, in certain embodiments, the "reverse" pool primers anneal to a conserved sequence in the joining ("J") segment.

[00161] Each primer can be designed so that a respective amplified DNA segment is obtained that includes a sequence portion of sufficient length to identify each J segment unambiguously based on sequence differences amongst known J-region encoding gene segments in the human genome database, and also to include a sequence portion to which a J-segment-specific primer can anneal for resequencing. This design of V- and J-segment-specific primers enables direct observation of a large fraction of the somatic rearrangements present in the adaptive immune receptor gene repertoire within an individual. This feature in

turn enables rapid comparison of the TCR and/or Ig repertoires (i) in individuals having a particular disease, disorder, condition or other indication of interest (*e.g.*, cancer, an autoimmune disease, an inflammatory disorder or other condition) with (ii) the TCR and/or Ig repertoires of control subjects who are free of such diseases, disorders conditions or indications.

[00162] In some embodiments, as used herein, the term “gene” refers to the segment of DNA involved in producing a polypeptide chain such as all or a portion of a TCR or Ig polypeptide (*e.g.*, a CDR3-containing polypeptide); it includes regions preceding and following the coding region “leader and trailer” as well as intervening sequences (introns) between individual coding segments (exons), and can also include regulatory elements (*e.g.*, promoters, enhancers, repressor binding sites and the like), and can also include recombination signal sequences (RSSs) as described herein.

[00163] The nucleic acids of the present embodiments, also referred to herein as polynucleotides, and including oligonucleotides, can be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA can be double-stranded or single-stranded, and if single stranded can be the coding strand or non-coding (anti-sense) strand. A coding sequence which encodes a TCR or an immunoglobulin or a region thereof (*e.g.*, a V region, a D segment, a J region, a C region, *etc.*) for use according to the present embodiments can be identical to the coding sequence known in the art for any given TCR or immunoglobulin gene regions or polypeptide domains (*e.g.*, V-region domains, CDR3 domains, *etc.*), or can be a different coding sequence, which, as a result of the redundancy or degeneracy of the genetic code, encodes the same TCR or immunoglobulin region or polypeptide.

[00164] In one embodiment, the present disclosure provides a plurality of V segment primers and a plurality of J segment primers, wherein the plurality of V segment primers and the plurality of J segment primers amplify substantially all combinations of the V and J segments of a rearranged immune receptor locus. In some embodiments, the method provides amplification of substantially all of the rearranged AIR sequences in a lymphoid cell, and capable of quantifying the diversity of the TCR or IG repertoire of at least 10^6 , 10^5 , 10^4 , or 10^3 unique rearranged AIR sequences in a sample. “Substantially all combinations” refers to at least 95%, 96%, 97%, 98%, 99% or more of all the combinations of the V and J segments of a rearranged immune receptor locus. In certain embodiments, the plurality of V segment primers and the plurality of J segment primers amplify all of the combinations of the V and J segments of a rearranged immune receptor locus.

[00165] In general, a multiplex PCR system can use at least 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25, and in certain embodiments, at least 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, or 39, and in other embodiments 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 65, 70, 75, 80, 85, or more forward primers, in which each forward primer specifically hybridizes to or is complementary to a sequence corresponding to one or more V region segments. The multiplex PCR system also uses at least 3, 4, 5, 6, or 7, and in certain embodiments, 8, 9, 10, 11, 12 or 13 reverse primers, in which each reverse primer specifically hybridizes to or is complementary to a sequence corresponding to one or more J region segments. Various combinations of V and J segment primers can be used to amplify the full diversity of TCR and IG sequences in a repertoire. For details on the multiplex PCR system, including primer oligonucleotide sequences for amplifying TCR and IG sequences, see, *e.g.*, Robins *et al.*, 2009 *Blood* 114, 4099; Robins *et al.*, 2010 *Sci. Translat. Med.* 2:47ra64; Robins *et al.*, 2011 *J. Immunol. Meth.* doi:10.1016/j.jim.2011.09.001; Sherwood *et al.* 2011 *Sci. Translat. Med.* 3:90ra61; U.S.A.N. 13/217,126, U.S.A.N. 12/794,507, WO/2010/151416, WO/2011/106738 (PCT/US2011/026373), WO2012/027503 (PCT/US2011/049012), U.S.A.N. 61/550,311, and U.S.A.N. 61/569,118, which are each incorporated by reference in its entirety.

[00166] Oligonucleotides or polynucleotides that are capable of specifically hybridizing or annealing to a target nucleic acid sequence by nucleotide base complementarity can do so under moderate to high stringency conditions. For purposes of illustration, suitable moderate to high stringency conditions for specific PCR amplification of a target nucleic acid sequence would be between 25 and 80 PCR cycles, with each cycle consisting of a denaturation step (*e.g.*, about 10-30 seconds (s) at greater than about 95°C), an annealing step (*e.g.*, about 10-30s at about 60-68°C), and an extension step (*e.g.*, about 10-60s at about 60-72°C), optionally according to certain embodiments with the annealing and extension steps being combined to provide a two-step PCR. As would be recognized by the skilled person, other PCR reagents can be added or changed in the PCR reaction to increase specificity of primer annealing and amplification, such as altering the magnesium concentration, optionally adding DMSO, and/or the use of blocked primers, modified nucleotides, peptide-nucleic acids, and the like.

[00167] In certain embodiments, nucleic acid hybridization techniques can be used to assess hybridization specificity of the primers described herein. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide as provided herein with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM

EDTA (pH 8.0); hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, *e.g.*, to 60°C-65°C or 65°C-70°C.

[00168] In certain embodiments, the primers are designed not to cross an intron/exon boundary. The forward primers in certain embodiments anneal to the V segments in a region of relatively strong sequence conservation between V segments so as to maximize the conservation of sequence among these primers. Accordingly, this minimizes the potential for differential annealing properties of each primer, and so that the amplified region between V and J primers contains sufficient TCR or Ig V sequence information to identify the specific V gene segment used. In one embodiment, the J segment primers hybridize with a conserved element of the J segment, and have similar annealing strength. In one particular embodiment, the J segment primers anneal to the same conserved framework region motif.

[00169] Oligonucleotides (*e.g.*, primers) can be prepared by any suitable method, including direct chemical synthesis by a method such as the phosphotriester method of Narang *et al.*, 1979, *Meth. Enzymol.* 68:90-99; the phosphodiester method of Brown *et al.*, 1979, *Meth. Enzymol.* 68:109-151; the diethylphosphoramidite method of Beaucage *et al.*, 1981, *Tetrahedron Lett.* 22:1859-1862; and the solid support method of U.S. Pat. No. 4,458,066, each incorporated herein by reference. A review of synthesis methods of conjugates of oligonucleotides and modified nucleotides is provided in Goodchild, 1990, *Bioconjugate Chemistry* 1(3): 165-187, incorporated herein by reference.

[00170] A primer is preferably a single-stranded DNA. The appropriate length of a primer depends on the intended use of the primer but typically ranges from 6 to 50 nucleotides, or in certain embodiments, from 15-35 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template nucleic acid, but must be sufficiently complementary to hybridize with the template. The design of suitable primers for the amplification of a given target sequence is well known in the art and described in the literature cited herein.

[00171] As described herein, primers can incorporate additional features which allow for the detection or immobilization of the primer but do not alter the basic property of the primer,

that of acting as a point of initiation of DNA synthesis. For example, primers can contain an additional nucleic acid sequence at the 5' end which does not hybridize to the target nucleic acid, but which facilitates cloning, detection, or sequencing of the amplified product. The region of the primer which is sufficiently complementary to the template to hybridize is referred to herein as the hybridizing region.

[00172] As used herein, a primer is “specific” for a target sequence if, when used in an amplification reaction under sufficiently stringent conditions, the primer hybridizes primarily to the target nucleic acid. Typically, a primer is specific for a target sequence if the primer-target duplex stability is greater than the stability of a duplex formed between the primer and any other sequence found in the sample. One of skill in the art will recognize that various factors, such as salt conditions as well as base composition of the primer and the location of the mismatches, will affect the specificity of the primer, and that routine experimental confirmation of the primer specificity will be needed in many cases. Hybridization conditions can be chosen under which the primer can form stable duplexes only with a target sequence. Thus, the use of target-specific primers under suitably stringent amplification conditions enables the selective amplification of those target sequences which contain the target primer binding sites.

[00173] In particular embodiments, primers for use in the methods described herein comprise or consist of a nucleic acid of at least about 15 nucleotides long that has the same sequence as, or is substantially complementary to, a contiguous nucleic acid sequence of the target V or J segment. Longer primers, *e.g.*, those of about 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, 45, or 50 nucleotides long that have the same sequence as, or sequence complementary to, a contiguous sequence of the target V or J segment, will also be of use in certain embodiments. Various mismatches (1, 2, 3, or more) to the target sequence can be contemplated in the primers, while preserving complementarity to the target V or J segment. All intermediate lengths of the aforementioned primers are contemplated for use herein. As would be recognized by the skilled person, the primers can have additional sequence added (*e.g.*, nucleotides that cannot be the same as or complementary to the target V or J segment), such as restriction enzyme recognition sites, adaptor sequences for sequencing, bar code sequences, and the like (see *e.g.*, primer sequences provided herein and in the sequence listing). Therefore, the length of the primers can be longer, such as 55, 56, 57, 58, 59, 60, 65, 70, 75, or 80 nucleotides in length or more, depending on the specific use or need. For example, in one embodiment, the forward and

reverse primers are both modified at the 5' end with the universal forward primer sequence compatible with a DNA sequencing nucleic acid sequence.

[00174] Also contemplated for use in certain embodiments are adaptive immune receptor V-segment or J-segment oligonucleotide primer variants that can share a high degree of sequence identity to the oligonucleotide primers. Thus, in these and related embodiments, adaptive immune receptor V-segment or J-segment oligonucleotide primer variants can have substantial identity to the adaptive immune receptor V-segment or J-segment oligonucleotide primer sequences disclosed herein, for example, such oligonucleotide primer variants can comprise at least 70% sequence identity, preferably at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or higher sequence identity compared to a reference polynucleotide sequence such as the oligonucleotide primer sequences disclosed herein, using the methods described herein (*e.g.*, BLAST analysis using standard parameters). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding ability of an oligonucleotide primer variant to anneal to an adaptive immune receptor segment-encoding polynucleotide by taking into account codon degeneracy, reading frame positioning and the like. Typically, oligonucleotide primer variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the annealing ability of the variant oligonucleotide is not substantially diminished relative to that of an adaptive immune receptor V-segment or J-segment oligonucleotide primer sequence that is specifically set forth herein. As also noted elsewhere herein, in preferred embodiments adaptive immune receptor V-segment and J-segment oligonucleotide primers are designed to be capable of amplifying a rearranged TCR or IGH sequence that includes the coding region for CDR3.

[00175] According to certain embodiments contemplated herein, the primers for use in the multiplex PCR methods of the present disclosure can be functionally blocked to prevent non-specific priming of non-T or B cell sequences. For example, the primers can be blocked with chemical modifications as described in U.S. patent application publication US2010/0167353. According to certain herein disclosed embodiments, the use of such blocked primers in the present multiplex PCR reactions involves primers that can have an inactive configuration wherein DNA replication (*i.e.*, primer extension) is blocked, and an activated configuration wherein DNA replication proceeds. The inactive configuration of the primer is present when the primer is either single-stranded, or when the primer is specifically hybridized to the target DNA sequence of interest but primer extension remains blocked by a chemical moiety that is linked at or near to the 3' end of the primer.

[00176] The activated configuration of the primer is present when the primer is hybridized to the target nucleic acid sequence of interest and is subsequently acted upon by RNase H or another cleaving agent to remove the 3' blocking group, thereby allowing an enzyme (*e.g.*, a DNA polymerase) to catalyze primer extension in an amplification reaction. Without wishing to be bound by theory, it is believed that the kinetics of the hybridization of such primers are akin to a second order reaction, and are therefore a function of the T cell or B cell gene sequence concentration in the mixture. Blocked primers minimize non-specific reactions by requiring hybridization to the target followed by cleavage before primer extension can proceed. If a primer hybridizes incorrectly to a sequence that is related to the desired target sequence but which differs by having one or more non-complementary nucleotides that result in base-pairing mismatches, cleavage of the primer is inhibited, especially when there is a mismatch that lies at or near the cleavage site. This strategy to improve the fidelity of amplification reduces the frequency of false priming at such locations, and thereby increases the specificity of the reaction. As would be recognized by the skilled person, reaction conditions, particularly the concentration of RNase H and the time allowed for hybridization and extension in each cycle, can be optimized to maximize the difference in cleavage efficiencies between highly efficient cleavage of the primer when it is correctly hybridized to its true target sequence, and poor cleavage of the primer when there is a mismatch between the primer and the template sequence to which it can be incompletely annealed.

[00177] As described in US2010/0167353, which is incorporated by reference in its entirety, a number of blocking groups are known in the art that can be placed at or near the 3' end of the oligonucleotide (*e.g.*, a primer) to prevent extension. A primer or other oligonucleotide can be modified at the 3'-terminal nucleotide to prevent or inhibit initiation of DNA synthesis by, for example, the addition of a 3' deoxyribonucleotide residue (*e.g.*, cordycepin), a 2',3'-dideoxyribonucleotide residue, non-nucleotide linkages or alkane-diol modifications (U.S. Pat. No. 5,554,516). Alkane diol modifications which can be used to inhibit or block primer extension have also been described by Wilk *et al.*, (1990 *Nucleic Acids Res.* 18 (8):2065), and by Arnold *et al.* (U.S. Pat. No. 6,031,091). Additional examples of suitable blocking groups include 3' hydroxyl substitutions (*e.g.*, 3'-phosphate, 3'-triphosphate or 3'-phosphate diesters with alcohols such as 3-hydroxypropyl), 2', 3'-cyclic phosphate, 2' hydroxyl substitutions of a terminal RNA base (*e.g.*, phosphate or sterically bulky groups such as triisopropyl silyl (TIPS) or tert-butyl dimethyl silyl (TBDMS)). 2'-alkyl silyl groups such as TIPS and TBDMS substituted at the 3'-end of an oligonucleotide are described by Laikhter *et al.*, U.S.

patent application Ser. No. 11/686,894, which is incorporated herein by reference. Bulky substituents can also be incorporated on the base of the 3'-terminal residue of the oligonucleotide to block primer extension.

[00178] In certain embodiments, the oligonucleotide can comprise a cleavage domain that is located upstream (*e.g.*, 5' to) of the blocking group used to inhibit primer extension. As examples, the cleavage domain can be an RNase H cleavage domain, or the cleavage domain can be an RNase H2 cleavage domain comprising a single RNA residue, or the oligonucleotide can comprise replacement of the RNA base with one or more alternative nucleosides. Additional illustrative cleavage domains are described in US 2010/0167353, which is incorporated by reference in its entirety.

[00179] Thus, a multiplex PCR system can use 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, 45, 50, 55, 60, 65, 70, 75, 80, 85, or more forward primers, wherein each forward primer is complementary to a single functional TCR or Ig V segment or a small family of functional TCR or Ig V segments, *e.g.*, a TCR V β segment, or (see *e.g.*, the TCR primers as set forth in the Sequence Listing), and, for example, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more reverse primers, each specific to a TCR or Ig J segment, such as TCR J β segment (see *e.g.*, Sequence Listing). In another embodiment, a multiplex PCR reaction can use four forward primers each specific to one or more functional TCR γ V segment and four reverse primers each specific for one or more TCR γ J segments. In another embodiment, a multiplex PCR reaction can use 84 forward primers each specific to one or more functional V segments and six reverse primers each specific for one or more J segments. Accordingly, various combinations of V and J primers can be used in a multiplex PCR reaction.

[00180] In some embodiments, the V and J segment primers are used to produce a plurality of amplicons from the multiplex PCR reaction. In certain embodiments, the amplicons range in size from 10, 20, 30, 40, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500 to 1600 nucleotides in length. In preferred embodiments, the amplicons have a size between 50-600 nucleotides in length.

[00181] According to non-limiting theory, these embodiments exploit current understanding in the art (also described above) that once an adaptive immune cell (*e.g.*, a T or B lymphocyte) has rearranged its adaptive immune receptor-encoding (*e.g.*, TCR or Ig) genes, its progeny cells possess the same adaptive immune receptor-encoding gene rearrangement, thus giving rise to a clonal population that can be uniquely identified by the presence therein of rearranged (*e.g.*, CDR3-encoding) V- and J-gene segments that can be amplified by a

specific pairwise combination of V- and J-specific oligonucleotide primers as herein disclosed.

[00182] The practice of certain embodiments of the present invention will employ, unless indicated specifically to the contrary, conventional methods in microbiology, molecular biology, biochemistry, molecular genetics, cell biology, virology and immunology techniques that are within the skill of the art, and reference to several of which is made below for the purpose of illustration. Such techniques are explained fully in the literature. *See, e.g.,* Sambrook, *et al., Molecular Cloning: A Laboratory Manual* (3rd Edition, 2001); Sambrook, *et al., Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); Maniatis *et al., Molecular Cloning: A Laboratory Manual* (1982); Ausubel *et al., Current Protocols in Molecular Biology* (John Wiley and Sons, updated July 2008); *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, Greene Pub. Associates and Wiley-Interscience; Glover, *DNA Cloning: A Practical Approach*, vol. I & II (IRL Press, Oxford Univ. Press USA, 1985); *Current Protocols in Immunology* (Edited by: John E. Coligan, Ada M. Kruisbeek, David H. Margulies, Ethan M. Shevach, Warren Strober 2001 John Wiley & Sons, NY, NY); *Real-Time PCR: Current Technology and Applications*, Edited by Julie Logan, Kirstin Edwards and Nick Saunders, 2009, Caister Academic Press, Norfolk, UK; Anand, *Techniques for the Analysis of Complex Genomes*, (Academic Press, New York, 1992); Guthrie and Fink, *Guide to Yeast Genetics and Molecular Biology* (Academic Press, New York, 1991); *Oligonucleotide Synthesis* (N. Gait, Ed., 1984); *Nucleic Acid Hybridization* (B. Hames & S. Higgins, Eds., 1985); *Transcription and Translation* (B. Hames & S. Higgins, Eds., 1984); *Animal Cell Culture* (R. Freshney, Ed., 1986); Perbal, *A Practical Guide to Molecular Cloning* (1984); *Next-Generation Genome Sequencing* (Janitz, 2008 Wiley-VCH); *PCR Protocols (Methods in Molecular Biology)* (Park, Ed., 3rd Edition, 2010 Humana Press); *Immobilized Cells And Enzymes* (IRL Press, 1986); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); Harlow and Lane, *Antibodies*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1998); *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and CC Blackwell, eds., 1986); Riott, *Essential Immunology*, 6th Edition, (Blackwell Scientific Publications, Oxford, 1988); *Embryonic Stem Cells: Methods and Protocols* (Methods in Molecular Biology) (Kurstad Turksen, Ed., 2002); *Embryonic Stem Cell Protocols: Volume I: Isolation and*

Characterization (Methods in Molecular Biology) (Kurstad Turksen, Ed., 2006); *Embryonic Stem Cell Protocols: Volume II: Differentiation Models* (Methods in Molecular Biology) (Kurstad Turksen, Ed., 2006); *Human Embryonic Stem Cell Protocols* (Methods in Molecular Biology) (Kurstad Turksen Ed., 2006); *Mesenchymal Stem Cells: Methods and Protocols* (Methods in Molecular Biology) (Darwin J. Prockop, Donald G. Phinney, and Bruce A. Bunnell Eds., 2008); *Hematopoietic Stem Cell Protocols* (Methods in Molecular Medicine) (Christopher A. Klug, and Craig T. Jordan Eds., 2001); *Hematopoietic Stem Cell Protocols* (Methods in Molecular Biology) (Kevin D. Bunting Ed., 2008) *Neural Stem Cells: Methods and Protocols* (Methods in Molecular Biology) (Leslie P. Weiner Ed., 2008).

VIII. Computer System

[00183] FIG. 12 is a high-level block diagram illustrating an example of a computer 1200 for use in analyzing molecular analytes, in accordance with one embodiment. Illustrated are at least one processor 1202 coupled to a chipset 1204. The chipset 1204 includes a memory controller hub 1220 and an input/output (I/O) controller hub 1222. A memory 1206 and a graphics adapter 1212 are coupled to the memory controller hub 1220, and a display device 1218 is coupled to the graphics adapter 1212. A storage device 1208, keyboard 1210, pointing device 1214, and network adapter 1216 are coupled to the I/O controller hub 122. Other embodiments of the computer 1200 have different architectures. For example, the memory 1206 is directly coupled to the processor 1202 in some embodiments.

[00184] The storage device 1208 is a non-transitory computer-readable storage medium such as a hard drive, compact disk read-only memory (CD-ROM), DVD, or a solid-state memory device. The memory 1206 holds instructions and data used by the processor 1202. The pointing device 1214 is used in combination with the keyboard 1210 to input data into the computer system 1200. The graphics adapter 1212 displays images and other information on the display device 1218. In some embodiments, the display device 1218 includes a touch screen capability for receiving user input and selections. The network adapter 1216 couples the computer system 1200 to the network. Some embodiments of the computer 1020 have different and/or other components than those shown in Figure 12. For example, the server can be formed of multiple blade servers and lack a display device, keyboard, and other components.

[00185] The computer 1200 is adapted to execute computer program modules for providing functionality described herein. As used herein, the term “module” refers to computer program instructions and other logic used to provide the specified functionality. Thus, a module can be implemented in hardware, firmware, and/or software. In one embodiment,

program modules formed of executable computer program instructions are stored on the storage device 1208, loaded into the memory 1206, and executed by the processor 1202.

[00186] In some embodiments, the computer 1200 is designed to execute a machine learning algorithm for predicting an immune response of a test subject. The system 1200 enables software to carry out actions for a computer-implemented method for determining an immunological status of a test subject.

[00187] In some embodiments, the computer-implemented method includes steps for storing data for a control subject obtained from a plurality of samples at various timepoints, said data comprising for each sample, nucleic acid sequence information for a plurality of unique rearranged nucleic acid sequences in said sample, an AIR sequence diversity score for said sample, a frequency of occurrence of each unique rearranged nucleic acid sequence in said sample, and a determined immunological status for said subject.

[00188] The computer-implemented method includes steps for determining rules by a processor for assessing an immunological status of a test subject based on said data of said control subject; inputting data for a test subject for a plurality of samples obtained at various timepoints before and after immunotherapy, said data comprising for each sample, nucleic acid sequence information for a plurality of unique rearranged nucleic acid sequences in said sample, an AIR sequence diversity score for said sample, and a frequency of occurrence of each unique rearranged nucleic acid sequence in said sample; and receiving a determination of an immunological status of said test subject.

[00189] The computer-implemented method comprises determining a predicted response to immunotherapy of said test subject. In one embodiment, the data for said control subject comprises nucleic acid sequence information obtained from said control subject at a timepoint prior to immunotherapy treatment. In another embodiment, the data for said control subject comprises nucleic acid sequence information obtained from said control subject at a timepoint after immunotherapy treatment.

EXAMPLES

EXAMPLE 1: T CELL RECEPTOR DIVERSITY AFTER UMBILICAL CORD BLOOD TRANSPLANT AS PREDICTOR OF MORTALITY FROM INFECTION

[00190] This example describes a clinical study in which 34 patients with high risk hematological malignancies were myeloablated and then transplanted with double umbilical cord blood (CB) units. Blood samples were collected at 0, 28, 56, 100, 180, and 360 days post transplant. At each time point, Immunoseq™ high-throughput T cell receptor (TCR)

sequencing assay (Adaptive Biotechnologies Corp., Seattle, WA) was applied to all samples. The Immunoseq™ data were used to assay the adaptive immune system at unprecedented depth, so that T cell clonal expansion and contraction of hundreds of thousands of T cell clones were tracked over time and TCR repertoire diversity was directly measured. Using the ability to track clones, the adaptive immune system reconstitution was shown to oscillate wildly with an almost entirely new repertoire appearing at least monthly after CB transplant. The largest clones from the prior blood draw dropped to below detectable levels within weeks, contrasting with the control data where the top clones in healthy patients were not only all observed at the following time point, but remained the highest frequency clones. Also described herein is a demonstration that diversity of the T cell repertoire is a measure of immunocompetence, as a clinical application of high-throughput sequencing. Of the 34 patients, six patients died between Day 100 and Day 360 of infectious cause. At both Day 56 and Day 100, the diversity of the T cell repertoire of each of these six patients was far lower than the average of the remaining patients (P-value = 0.015). By two months after transplant, TCR diversity accurately predicted risk of death due to infection.

[00191] Patients undergoing hematopoietic cell transplantation (HCT) are at increased risk of early post-transplant morbidity and mortality from infectious complications secondary to the prolonged period of pancytopenia and immune dysregulation that results from the conditioning regimen. Cord blood transplant (CBT) recipients appear to be at even greater risk of early transplant related mortality; in fact a recent study demonstrated that non-relapse mortality (NRM) is highest in double CBT (dCBT) recipients when compared to matched and mismatched unrelated donor recipients. Multiple studies have also demonstrated that immune reconstitution following CBT is significantly delayed relative to conventional donor stem cell sources (PB and BM), further contributing to the increased risk of life-threatening infectious complications in this patient population, and consequently CBT recipients have a higher incidence of opportunistic infections (OI) in the first year after transplant [6-8]. Further complicating matters is a dearth of assays that can adequately measure reconstitution of the adaptive immune system, thereby making it difficult to directly address the role of delayed immunity on CBT outcomes in the setting of many other contributing variables (*e.g.*, age, HLA match, intensity of the regimen, concurrent immunosuppressive therapy for prevention and treatment of graft-versus-host disease (GVHD)). This Example demonstrates the ability to more accurately measure functional immune reconstitution in patients undergoing HCT and thus determine the consequent risk of mortality from infectious complications, which will positively impact direct medical decision-making aimed at

reducing this risk, especially in the setting of immunosuppressive therapy for the prevention and treatment of GVHD.

[00192] In the blood of a healthy adult, an individual T cell primarily expresses one of millions of different TCRs, and a clone is the set of T cells expressing the same TCR^{1,2}. Diversity of the TCR repertoire is known to be necessary for adequate protection against foreign pathogens. This is evident in humans with primary or acquired immunodeficiency diseases (*e.g.*, SCIDS, CVID, and HIV), in aging, and following hematopoietic cell transplantation where loss of TCR diversity has been implicated in the increase in morbidity and mortality from infection that is observed in these patients.

[00193] Due to the large number of different T cell clones in the human body, estimates of diversity of the TCR repertoire have generally been extrapolated only indirectly, and with low sensitivity. As such, conventional estimates of diversity are insufficient for clinical decision making or in assessment of the health of the cellular adaptive immune system, for instance, as a guide to patient risk for infectious complications related to a level of immunocompetence, or for determining the level of immunocompetence *per se*. In order to address these shortcomings of prior approaches, herein is described application of a high-throughput method to sequence millions of TCRs from a single sample, allowing direct determination of the diversity of a T cell repertoire for use in directly measuring the health of the cellular adaptive immune system. This information provides a guide to clinical decision making in the setting of acquired or congenital immunosuppression.

[00194] As described herein, T cell repertoire diversity provides a direct measurement of immune reconstitution after myeloablative CBT. TCR diversity was measured in CBT recipients at time points 28, 56, and 100 days post transplant, and shown to be predictive of non-relapse mortality (NRM). Additionally, to better understand the dynamics of immune reconstitution, quantitative TCR sequence diversity and distribution data were obtained at each time point and used to track the expansion and contraction of hundreds of thousands of T cell clones simultaneously.

[00195] METHODS

[00196] Study Design. Patients undergoing a myeloablative single or double CBT were eligible for this retrospective analysis (of data collected prospectively). All patients provided signed consent to participate in the study which was approved by study's Institutional Review Board.

[00197] Patients, Treatment Regimens and Post-Transplant Supportive Care. Patients with hematologic malignancy, aged ≤ 45 years old, received a myeloablative CBT if they lacked a

suitably HLA-matched related or unrelated donor. The patients' underlying disease was categorized as standard or high-risk based upon previously described criteria [17]. Patients received a single or double CB graft as determined by institutional priority criteria. All CB units were HLA-typed at the intermediate resolution level for HLA-A and HLA-B and allele-level (high resolution) for HLA-DRB1, and all CB units were required to be matched to the recipient at ≥ 4 of the 6 HLA loci. Patients without pre-transplant blood samples stored for TCR analysis or who died before day 28 were excluded.

[00198] Myeloablative conditioning consisted of either cyclophosphamide (Cy) (total 120mg/kg), hyperfractionated total body irradiation (TBI) over 4 days (total of 13.2 Gy), and fludarabine (Flu) (total 75 mg/m²), or Treosulfan (Treo) (total 42 gm/m²), Flu (total 150mg/m²), and a single fraction of 2 Gy TBI. All patients received GVHD immunoprophylaxis with cyclosporine-A (CSA) and mycophenolate mofetil (MMF) beginning on day -3. All patients received standard prophylactic antimicrobial and antifungal agents during follow-up [18] and remained at the institution for a minimum of 100 days post-transplant. Patients were seen at least once per week for clinical assessment and follow-up. Acute GVHD was graded using standard criteria based on stages of organ involvement and categorized as acute GVHD grades 0–IV [16].

[00199] ImmunoSeq™ Assay and Evaluation of Immune Reconstitution Post Transplant. Peripheral blood was collected on days 28, 56, 80-100, 180, and one and two years post transplant, as well as pre-transplant, for retrospective analysis of immune recovery utilizing the ImmunoSeq™ assay for high-throughput sequencing of TCR β (TCRB) from genomic DNA extracted from peripheral blood mononuclear cells (PBMCs). Sequencing was performed of the CDR3 region of TCR β genes from approximately 250,000 PBMCs from each time point in surviving patients enrolled in the trial. The TCR β CDR3 region was defined according to the IMGT collaboration³, beginning with the second conserved cysteine encoded by the 3' portion of the V β gene segment and ending with the conserved phenylalanine encoded by the 5' portion of the J β gene segment. The number of nucleotides between these codons determined the length and therefore the frame of the CDR3 region. TCR β CDR3 regions were amplified and sequenced using previously described protocols (Robins *et al.*, 2009 *Blood* 114, 4099; Robins *et al.*, 2010 *Sci. Translat. Med.* 2:47ra64; Robins *et al.*, 2011 *J. Immunol. Meth.* doi:10.1016/j.jim.2011.09.001; Sherwood *et al.* 2011 *Sci. Translat. Med.* 3:90ra61; U.S.A.N. 13/217,126, U.S.A.N. 12/794,507, WO/2010/151416, WO/2011/106738 (PCT/US2011/026373), WO2012/027503 (PCT/US2011/049012). Briefly, a multiplexed PCR method was employed to amplify all possible rearranged genomic

TCR β sequences using 52 forward primers, each specific to a TCR V β segment, and 13 reverse primers, each specific to a TCR J β segment. Sequence reads of length 60 bp were obtained using the Illumina HiSeq™ System (Illumina, Inc., San Diego, CA). Raw HiSeq™ sequence data were preprocessed to remove errors in the primary sequence of each read, and to compress the data. A nearest neighbor algorithm was used to collapse the data into unique sequences by merging closely related sequences, to remove both PCR and sequencing errors.

[00200] Statistical Considerations. Due to sample size limitations (34 patients, including six NRM after day 56), the standard Cox proportional hazards model would not be informative. Given the sample size, the study was powered to robustly test the hypothesis that in blood samples from patients undergoing a myeloablative single or double CBT, a lower diversity of the patient's T cell repertoire at Day 56 (or Day 100) was predictive of death from NRM. Because the distribution of TCR repertoire size among patients was not well-characterized, a one-tailed nonparametric Mann-Whitney U test was used to determine whether patients who eventually suffered NRM had lower values of TCR repertoire diversity. All possible confounding factors could not be assessed in a multivariate model, but each possible confounding factor was analyzed separately. A two-tailed Fisher's exact test was used for binary data, and a two-tailed Mann-Whitney U test was used for continuous data.

[00201] RESULTS

[00202] Study cohort. Patients (N= 34) and donor characteristics are outlined in Table 2. Patients were transplanted for treatment of high risk hematologic malignancies, primarily acute lymphoblastic or myeloblastic leukemias (N= 26), with 12 patients in CR1 and 14 in CR2 or greater. Table 2 summarizes the patient demographics, diagnoses, and clinical course. The treatment regimen for all patients was similar for the first 56 days post transplant. Of the 34 patients, 31 had GVHD, with six grade III-IV; accordingly all patients were prophylactically treated with corticosteroids.

Table 2: Cohort Characteristics of 34 patients included in the analysis

Patient Age, Median Years (range)	27 (1 – 58)
Follow-up, Median days (range)	369 (34 – 1657)
Median Recipient Weight in Kg (range)	75.3 (22.5 – 114.4)
Disease, No (%)	
AML	16 (47)
ALL	10 (29)

MDS/MPD	6 (18)
Other	2 (6)
Minimal Residual Disease (MRD)	17 (50)
Transplant Type, No (%)	
FLU/CY/TBI 1320 cGy †	24 (71)
TREO/FLU/TBI 200 cGy ††	10 (29)
GVHD Prophylaxis, No (%)	
Cyclosporine/Mycophenolate	34 (100)
Recipient CMV Serostatus, No (%)	
Positive	22 (64)
Negative	12 (36)
HLA disparity, No (%) §	
4/6	20 (60)
5/6	11 (31)
6/6	3 (9)
Number of CB units infused per patient	
1	2 (6)
2	32 (94)
Acute GVHD	
Grade II or less	25 (73)
Grade III-IV	6 (18)

[00203] In Table 2, ALL = acute lymphoblastic leukemia; AML = acute myeloid leukemia; MDS/MPD = myelodysplastic/myeloproliferative diseases; FLU = fludarabine; CY = cytoxan; TBI = total body irradiation; TREO = treosulfan; GVHD = graft-versus-host-disease; CMV = cytomegalovirus; CB= Cord blood; CR= Complete remission; MRD= Minimal residual disease. † = Fludarabine 75 mg/m², Cytoxan 120 mg/kg, TBI 1320 cGy. †† = Treosulfan 42 gm/m², Fludarabine 150 mg/m², TBI 200 cGy. § = HLA matching reflects the lowest HLA-match of the 2 units.

[00204] Sample collection. All patients included in this study had blood drawn pre- and post-transplant on days + 28, 56, 80-100, 180, and at one year, for those surviving. Aliquots of over one million PBMCs from each blood draw were sent to Adaptive Biotechnologies (Seattle, WA) for deep sequencing of the TCR DNA using the ImmunoSeq™ assay. A fixed

amount of DNA was input into the Immunoseq™ assay to allow for direct comparison between samples. In addition, the percent and absolute numbers of T, B and NK cells, and immunoglobulin levels were assessed for each sample.

[00205] Dynamics of immune reconstitution of the cellular adaptive immune system. Given the enormous number of potential rearrangements for TCR genes, it was implied that each clone had a virtually unique TCR sequence. These sequences were molecular identifiers for each clone, so that clonal expansion and contraction could be tracked over time, in addition to determining the properties of the clonal distribution at each fixed timepoint.

[00206] Overlap with pre-treatment repertoire. Using the TCR sequence as a molecular tag, hundreds of thousands of T cell clones were tracked across time. All of the 34 patients underwent a myeloablative procedure prior to transplant, so it was assumed that a very limited subset of the pre-treatment repertoire would persist after myeloablation. Surprisingly, the proportion of clones remaining after myeloablative treatment varied widely; in three of the 34 patients, 10% or more of the TCR repertoire at day 180 was derived from the pre-treatment (host) immune system.

[00207] In most patients, however, few or no pre-existing clones remained after transplant. Figure 1 shows the proportion of TCR repertoire carried over after transplant across patients and time points. The fraction of pre-transplant clones persisting post-transplant was stable over time, as evidenced by a strong correlation between the proportion of pre-transplant clones observed at each time point post-transplant (*e.g.*, $r = 0.9$ between the proportion of holdovers at day 28 and day 100), and for most patients was roughly 0.001, indicating that on average about one T-cell per thousand was part of a clone observed pre-transplant. Analysis of clones using the Immunoseq™ high throughput TCR sequencing assay yielded different results than microchimerism as measured by previous methods. The process of TCR rearrangement made it unlikely that substantial clonal identity at the DNA level could be observed by chance in the absence of surviving host T cells. Therefore, these results suggested that the Immunoseq™ high quality TCR sequencing method detected microchimerism in some patients who were previously thought not to harbor any host T cells.

[00208] Tracking clones post transplant. In order to assess the stability of the reconstituting adaptive immune system over time, the persistence of TCR clones found at early time-points was investigated in later samples. Using only patients with samples collected and sequenced at 28, 56, 100, 180 and 365 days post-transplant, the top 10 TCR clones were determined by frequency in each patient at the 28, 56, 100 and 180 day time-points and the sequences of each of these clones was classified as either persistent or transient. A top-ten TCR clone that

was observed (at any frequency) at a later time-point was considered persistent, and clones that were never again observed in samples from the same patient were considered transient.

[00209] Figure 2 shows the median number of transient TCR clones in the top 10, at each time-point post-transplant. At 28 and 56 days post-transplant, dynamic and highly unstable TCR repertoires were observed in which many TCR clones that were present at high frequency in an early sample were never again observed subsequently. Starting at 100 days post-transplant, this pattern began to subside and patients' TCR repertoires became more stable. To confirm that this pattern was highly unusual, PBMC samples were sequenced from four healthy control subjects over the same length of time. The median number of transient TCR clones in the top 10 was 0 for these healthy controls at each time-point, confirming the assumption that the high prevalence of transient TCR clones following transplant was indicative of an unusually unstable TCR repertoire. It has been shown previously that a high rate of apoptosis prevailed in T cells following stem cell transplant.⁴ Here, it was demonstrated that this apoptotic process was not random with respect to the T cells constituting a clone, but occurred as rapid expansion and contraction of entire clones.

[00210] Changes in T cell clonal diversity post transplant. The distribution of T cell clones was used to estimate the lower bound on the diversity in the full blood using an unseen species analysis (Robins *et al.*, 2009 *Blood* 114, 4099). The diversity estimate was computed for each time point.

[00211] Figure 3 presents a summary of immune reconstitution as measured by TCR repertoire during the first year post-transplant in this cohort. The geometric mean of the diversity metric is shown at each time point and is illustrative of the general course of reconstitution. Myeloablative conditioning regimens resulted in a large drop in TCR diversity from pre-transplant values. Diversity decreased from pre-transplant values to day 28, which was close to the mean time to engraftment for most patients (engraftment time ranged from 7 to 45 days with a mean of 24). In this cohort of patients, TCR diversity reached its lowest value at 56 days post-transplant before beginning a slow recovery to a substantial increase in TCR repertoire diversity by one year post-transplant. However, despite this recovery, patient TCR repertoires still had a much lower diversity than healthy repertoires by the end of the one-year study.

[00212] Some of the observed changes in TCR diversity can be explained by variations in absolute T cell counts. In the extreme, if a patient had very few T cells, the TCR diversity of such a patient was limited. At day 56, several of the patients had very low CD3 counts and, therefore, low diversity. However, the correlation between diversity and absolute CD3

counts was weak for the remaining cohort ($r = 0.05$, Figure 5). For example, there were patients in this cohort who had higher absolute CD3 counts but little diversity, secondary to highly oligoclonal TCR repertoire (a small number of highly expanded clones).

[00213] TCR Diversity as a predictor of mortality from infection. Of the 34 patients in the present study, 15 died in the first year post transplant. The sole cause of death in 6 of these patients was relapsed disease, with death primarily from infectious causes in the remaining nine patients. Infectious complications post transplant were not unexpected or uncommon. However, prior to the present disclosure, there was no concrete measurement having strong predictive value to assess which patients were at increased risk of dying from life threatening infections post transplant. Such predictive capability can change the medical management of patients post stem cell transplant. Therefore, the present direct measurement of TCR diversity was analyzed for its potential as an informative predictor of the ability of the adaptive immune system to fight infection. The diversity of the T cell repertoire at early time points post transplant (Day 56 and Day 100) was indeed a strong predictor of mortality from non-relapse causes. The measurement of T cell diversity pre-transplant trended toward predictive value but did not reach statistical significance in this cohort. Day 28 data also lacked significant predictive value, presumably due to the fact that TCR repertoire diversity at this time point was likely dominated by T cells from the graft that were differentiated prior to infusion and thus did not reflect true immune reconstitution. However, patients who eventually died of non-relapse causes beyond day 100 did display significantly reduced TCR diversity as early as 56 days post-transplant (*See* Figure 4). Data from those patients who died of NRM (non relapse mortality) consistently yielded a lower estimate of TCR diversity at each time point measured, but this relationship only reached statistical significance (calculated using a one-tailed Mann-Whitney U test, $p = 0.0153$) at day 56, when patients' immune reconstitution had commenced in earnest following engraftment. A significant difference in the TCR diversity of patients who would go on to die from NRM persisted at 100 days post-transplant ($p = 0.0153$). It is worth noting that five of the six patients who were alive at day 56 but eventually died from infectious causes survived through day 180, indicating that clinical identification of high-risk status between 28 and 56 days post transplant would allow ample time for prophylactic clinical intervention.

[00214] Other factors as predictors of mortality from infection. Post transplant immune recovery is influenced by many factors, most significantly the use of immunosuppressive therapy (IST) for the prevention and treatment of GVHD, in particular the use of steroids. In order to determine the utility of direct measurement of TCR diversity in predicting risk of

infection, the roles of possible confounding factors were evaluated, focused on differences in treatment with IST and total absolute CD3 counts. There were 27 patients who developed GVHD at a median of 23 days post transplant. These patients were initially treated with prednisone.

[00215] Another major confounding factor in the correlation of TCR diversity measurements with clinical outcome was the recovery of total CD3+ cells. A correlation of TCR diversity with clinical outcome would be of limited utility if it simply reiterated a result obtained by counting total T cells in peripheral blood. The kinetics of T cell recovery as measured by the absolute CD3+ cells/ μ l were measured in the peripheral blood at the same time as the measurement of TCR diversity. Figure 5 shows that absolute CD3+ count was very weakly correlated with estimated TCR repertoire diversity, indicating that these two metrics were essentially independent. This correlation was driven mostly by samples near the lower limit of absolute CD3+ cell counts, since substantially depressed total T-cell counts would necessarily result in determination of a TCR repertoire having low diversity. However, at higher absolute CD3+ counts, it remained possible that an oligoclonal T cell population would nonetheless have had very low TCR diversity and so would have failed to provide protection against opportunistic pathogens.

[00216] Table 3 summarizes the results obtained when analyzing CD3+ counts alongside the TCR diversity metric, using data from day 56 and 100 post-transplant. In Table 3, each sample was assigned to a high (at or above median) or low (below median) group for both metrics, and the number of eventual deaths from NRM falling into each category were compared (out of the six that survived until day 56). In addition, Table 3 presents the results of a one-tailed p-value for a Mann-Whitney U test of the hypothesis that patients dying from NRM tended to have lower values of absolute CD3+ counts or TCR diversity. The TCR diversity metric was a predictor of clinical outcome as early as day 56 post transplant while absolute CD3+ counts were uninformative at that time. Furthermore, diversity was more significantly correlated to clinical outcome than CD3+ counts at 100 days post-transplant. The repertoires of the six patients who died from relapse alone were also analyzed. Neither absolute CD3 count nor TCR repertoire diversity predicted death from relapse in this cohort.

Table 3. Comparison of CD3+ (total T-cell) counts and estimated TCR repertoire size at days 56 and 100 post-transplant.

Non-relapse mortalities, CD3+ counts v. TCR repertoire size

CD3+ cells/cm³	High	Low	p-value
56 days post-transplant	3/15	3/14	0.139
100 days post-transplant	1/15	5/14	0.021
Estimated TCR repertoire size	High	Low	p-value
56 days post-transplant	1/15	5/14	0.015
100 days post-transplant	1/15	5/14	0.015

[00217] Numbers for each metric represent the number of non-relapse mortalities (n=6) that fall into the high (at or above median) or low (below median) categories. At right is shown the p-value of a one-tailed Mann-Whitney U test. Total T-cell counts were not predictive of clinical outcome at 56 days post-transplant, and were a weaker predictor of clinical outcome than estimated TCR repertoire size at 100 days post-transplant.

[00218] In addition to GVHD treatment and total CD3+ counts, the correlation of TCR diversity measurement with clinical outcome could also have been driven by any number of other variables. Table 3 presents a comparison of the 6 patients who lived to day 56 but eventually died of non-relapse causes versus the 25 other patients who were alive at day 56. Several metrics were evaluated between these groups in the same fashion as the TCR diversity metric (Fisher's exact test was used for categorical data, and a Mann-Whitney U test for continuous data). None reached statistical significance in this cohort, with many appearing completely uncorrelated with NRM. The 6 NRM patients were somewhat older than the others ($p = 0.051$), which can indicate a weak correlation to NRM with or without TCR diversity acting as an intermediary. These results indicated that in this cohort, in which TCR repertoire diversity was a statistically significant predictor of non-relapse mortality, non-relapse mortality could not be easily predicted by any of the other variables we examined.

[00219] This study used a direct TCR sequence-based measure of immune reconstitution that correlated with adverse clinical outcomes, particularly the increased risk of infectious complications in patients undergoing myeloablative cord blood transplantation. Recipients of CBT were at increased risk of delayed hematopoietic and immune recovery, and improvement in overall survival for these patients was dependent on strategies that can enhance the kinetics of neutrophil and immune system recovery. Direct measures of hematopoietic recovery are simple and well established by obtaining complete blood counts. However, a direct measure of immune system recovery, especially with respect to T cell function as opposed to T cell numbers, has been lacking prior to the present disclosure. Thus,

there are standards of care regarding medications for the prevention and treatment of GVHD. However, not all patients are at equal risk of developing complications post transplant and not all patients will respond equally to medical interventions. Novel measurements of immune recovery, as described herein using the ImmunoSeq™ assay, permit tailoring the medical management of individual patients not only with respect to anti-microbial prophylaxis, but also with respect to managing IST if patients are identified as being at higher risk of NRM from infections. A direct measure of this risk would dramatically change medical management.

[00220] Here, a significant correlation was demonstrated between measurement of immune reconstitution using high-throughput T cell receptor sequencing and non-relapse mortalities in a cohort of 34 patients. Consistent with this result is the understanding that delayed immune reconstitution, as measured by low diversity of TCR rearrangements in circulating T cells, puts patients at high risk of complications from infectious disease. These results cannot be trivially replicated by CD3+ cell counts or by any of several other variables that were measured in the small cohort. The present results were obtained by analyzing the entire T cell compartment, while it is known that the kinetics and clinical implications of immune reconstitution differ between T cell subsets (*e.g.*, CD4+ v. CD8+ T cells).⁵ Hence, investigation of the reconstitution of T cell subsets using HTS can yield additional insights.

[00221] The present results demonstrated that high TCR diversity was associated with better outcomes in blood samples taken from patients undergoing myeloablative cord blood transplantation. As such, TCR diversity is a highly useful measure with which to stratify patients soon after transplant based on the risk of future infectious complications. Thus, if a patient has not met a threshold level of TCR diversity by two to three months post-transplant, that patient can benefit by coming off IST more rapidly as tolerated, and/or can be treated more aggressively with anti-microbial prophylaxis, and/or can be kept under observation for a longer period until immune reconstitution has reached adequate levels. Given that decreased TCR diversity preceded non-relapse mortality by several months in the present study, this measure should allow ample time for such prophylactic measures. In addition to identifying high-risk patients, however, the presently described robust measurement of immune reconstitution can also help determine when patients have achieved sufficient immune reconstitution to discontinue prophylactic treatment, rather than administering a regimen of the same duration to each patient.

**EXAMPLE 2: T CELL RECEPTOR
REPertoire DISTRIBUTION AS PREDICTOR OF IMMUNOTHERAPY
RESPONDERS**

[00222] T cell receptor diversity and distribution were determined as described above in blood and solid tumor samples, obtained prior to and after initiation of immunotherapy, from cancer patients who were candidates to receive either a CTLA-4 inhibitor or a PD-1 inhibitor. The efficacy of each immunotherapy agent was independently assessed by standard oncology clinical criteria (categorizing subjects as responders or non-responders) and the relative ability of each patient's adaptive immune system to respond beneficially to the immunotherapy was shown to be predicted by a modified entropy calculation of the distribution of the TCR repertoire prior to immunotherapy.

[00223] Before the initiation of immunotherapy (anti-CTLA-4 mAb), responders exhibited relatively higher TCR sequence diversity in lymphocytes present in blood and tumor samples, and higher TCR sequence distribution entropy, observed as a flatter TCR distribution profile, relative to non-responders.

[00224] As shown in Figure 6, the results of quantitatively sequencing TCR encoding DNA from blood samples show that low TCR repertoire clonality, indicative of higher TCR sequence diversity and higher TCR sequence distribution entropy, was a predictor of immunotherapy (anti-CTLA-4 mAb) responder status. Figure 6 also shows that high TCR repertoire clonality, indicative of lower TCR sequence diversity and lower TCR sequence distribution entropy, was a predictor of immunotherapy non-responder status.

[00225] Figure 7 shows the results from quantitative sequencing of TCR encoding DNA from tumor tissue samples. Figure 7 illustrates an increase in TCR clonality in lymphocytes present in solid tumor tissue samples obtained after administration of immunotherapy (anti-CTLA-4 mAb) (DT) relative to the level of TCR clonality detected in tumor samples obtained prior to immunotherapy (AT). The sample was obtained from a responder subject. Figure 7 demonstrates that the immunotherapy treatment had a noticeable impact on the subject's T cell repertoire.

[00226] In Figure 8, the dynamics of individual TCR clonal representations were investigated as a function of time in blood samples and in tumor samples obtained prior to immunotherapy (treatment with ipilimumab (an anti-CTLA-4 mAb)) and post immunotherapy. Timepoints A, B, and C are timepoints taken from blood samples. Timepoint A is before immunotherapy, and timepoints B and C are two timepoints after starting the immunotherapy regimen. Timepoints AT, BT, CT are paired tumor samples (AT is before immunotherapy, and timepoints BT and CT are two timepoints after starting the immunotherapy regimen). Each line of datapoints follows a single clone. The arrow at timepoint CT indicates a single clone that was unremarkable in its relative abundance prior to

therapy, but that increased in relative frequency in tumor samples post-therapy, to account for 10% of the repertoire at timepoint CT. Two clones (X1 and X2) that each accounted for 7-9% of TCR sequences in blood at timepoints A and B subsequently declined significantly in relative abundance, while several T cell receptor sequences that initially had very low frequencies in blood increased significantly by timepoint C. The three most numerous clones in tumor samples at timepoint A decreased significantly in their subsequent relative representation, as determined at later timepoints.

[00227] In a separate study with the same immunotherapeutic agent used for Figures 6-8, TCR sequence diversity and distribution entropies were determined in blood samples obtained prior to (timepoint A) and after (timepoints B and C) initiation of immunotherapy. The results are summarized in Figure 9, which shows dynamics of individual TCR clonal representations over time. The arrow indicates a single clone that was not highly represented prior to immunotherapy but that increased in preponderance post-therapy to account for greater than 10% of the repertoire at timepoint C.

[00228] As shown in Figures 8 and 9, a side effect of an immunotherapy treatment can be a proliferation of a single or few clones in the blood of a subject, such that the frequency of occurrence of the single or few clones is statistically significantly greater than the frequencies of occurrence of the other clones in the repertoire. In some embodiments, the frequency of occurrence of a single clone is determined to be greater than a predetermined threshold, such as greater than the top quartile of frequencies of occurrence of the clones in the repertoire. In Figures 8 and 9, the single clone that accounts for greater than 10% of the repertoire after immunotherapy treatment is statistically significantly different and is an indicator of poor response by the subject. For example, a clone frequency that is less than 1 % in frequency of occurrence before immunotherapy and spikes in frequency to greater than 1% of frequency of occurrence in the repertoire is an indicator of poor outcome in the subject. This presents a situation of a expansion of a single clone (high clonality) and low diversity of the repertoire, leading to poor response and outcome. In situations where a single or few clones exhibit a statistically significantly greater frequency of occurrence compared to the repertoire, adjustments can be made to the treatment of the subject, including use of corticosteroids or immunosuppressants to decrease immune response (e.g., inflammation, etc.).

EXAMPLE 3: T CELL RECEPTOR CLONALITY AS A PREDICTOR OF IMMUNOTHERAPY RESPONSE

[00229] In another example, a study was performed using high-throughput sequencing of the TCRB gene locus to characterize the repertoire of tumor-infiltrating lymphocytes (TILs)

in late-stage metastatic melanoma patients undergoing immunotherapy (treatment with an anti-PD-1 antibody). The goal of the study was to determine whether characterization of the intratumoral T cell repertoire by high-throughput sequencing is sufficient to predict clinical outcome (i.e., drug response) using immunological profiling (by TCRB sequencing) of a pre-treatment tumor biopsy.

[00230] T cell receptor diversity and distribution were determined as described above in solid tumor samples, obtained prior to initiation of immunotherapy, from metastatic melanoma patients who were candidates to receive a PD-1 inhibitor (Lambrolizumab). PD-1 (Programmed cell death protein 1) is a type 1 membrane protein, a member of the immunoglobulin superfamily, and thought to play a role in B cell differentiation.

[00231] The efficacy of the immunotherapy treatment was independently assessed by standard oncology clinical criteria. Subjects were characterized as follows: responders (separated into “partial response” indicating a reduction in patient tumor burden and “stable disease” indicating lack of progression without decreased tumor burden) or non-responders (continued disease progression). The relative ability of each patient’s adaptive immune system to respond beneficially to the immunotherapy was shown to be predicted by a modified entropy calculation of the distribution of the TCR repertoire prior to immunotherapy.

[00232] A modified entropy calculation (“clonality”) was used in which each tumor sample’s TCR sequence distribution entropy was normalized to the range (0 – 1) by accounting for the number of unique TCR rearrangements observed in that tumor sample and inverted so that a high normalized entropy becomes a low clonality and vice versa. Before the initiation of immunotherapy, patients who would respond to the immunotherapy exhibited relatively higher TCR sequence distribution clonality in lymphocytes present in tumor biopsy samples, relative to non-responders.

[00233] Briefly, frozen tissue samples from malignant lesion biopsies were prepared for 12 patients before administration of an anti-PD-1 antibody (administered as an immunotherapeutic agent for patients with metastatic melanoma). Tissue samples (biopsies from melanoma lesions) were also collected during and after anti-PD-1 immunotherapy. Subsequently, genomic DNA was extracted from these tissue samples. Using the ImmunoSEQ platform for high-throughput TCRB sequencing, the repertoire of TILs in each sample was characterized to determine (1) the extent of intratumoral lymphocyte infiltration, and (2) the clonal structure of the intratumoral lymphocyte repertoire.

[00234] Of the 12 patients studied, 8 responded to the treatment (stable disease or partial response), while 4 did not respond (disease progression). The results of quantitatively sequencing TCR encoding DNA from frozen tumor tissue biopsies are shown in Figure 10A.

[00235] Figure 10A shows that high intratumoral TCR repertoire clonality, indicative of a TCR repertoire characterized by a small number of highly-expanded T cell clones (low AIR sequence diversity), was a statistically-significant predictor of immunotherapy responder status in this retrospective study of a 12-patient cohort. The mean and standard deviation of TCR repertoire clonality (a modified metric based on TCR sequence distribution entropy normalized to the range (0 – 1) by accounting for the number of unique TCR sequences present in each sample) are presented according to response to immunotherapy. TCR sequence distribution clonality was significantly higher in the 8 patients who responded compared to the 4 patients who did not respond ($p = 0.015$ by two-tailed unpaired t-test).

[00236] In Figure 10B, T cell infiltration was assessed in the cohort of 12 patients. The mean and standard deviation of T cell infiltration (measured as T cell receptor rearrangements per diploid genome) are presented according to response to immunotherapy. Levels of T lymphocyte infiltration were higher in the 8 patients who responded compared to the 4 patients who did not respond ($p = 0.056$ by two-tailed unpaired t-test).

[00237] Data from the first cohort (12 patients with pre- and post-treatment tumor biopsies) suggested that two factors from pre-treatment tumor biopsies (low levels of T cell infiltration and a highly diverse infiltrating T cell repertoire) were associated with failure to respond to treatment (i.e., progressed as opposed to stable or improving disease state; Figs. 10A, 10B).

[00238] In a second study, a confirmatory cohort of 13 additional patients was assessed, in which only pre-treatment tumor biopsies were sequenced. In Figure 11A, TCR repertoire clonality and response to immunotherapy were assessed for cohorts 1, 2 and the combined cohort. The mean and standard deviation of TCR repertoire clonality (a modified metric based on TCR sequence distribution entropy normalized to the range (0 – 1) by accounting for the number of unique TCR sequences present in each sample) is shown according to response to immunotherapy in cohort 1 (plain square), cohort 2 (slant striped square) and in the combined data (straight striped square). TCR sequence distribution clonality was higher in the patients who responded compared to the patients who did not respond ($p = 0.00065$ in the combined data by a two-tailed unpaired t-test).

[00239] In Figure 11B, T cell infiltration was assessed in the patient groups. The mean and standard deviation of T cell infiltration (measured as T cell receptor rearrangements per diploid genome) is shown according to response to immunotherapy in cohort 1 (plain square),

cohort 2 (slanted stripe square) and in the combined data (straight stripe square). Levels of T lymphocyte infiltration are higher in the patients who responded compared to the patients who did not respond ($p = 0.0015$ in the combined data by a two-tailed unpaired t-test).

[00240] Figure 11C shows a comparison of each patient's rank (out of 25 patients total) for level of T cell infiltration and TCR repertoire clonality. Compared to responders (diamond), non-responders (circle) simultaneously tend toward low TCR repertoire clonality and low levels of infiltrating T lymphocytes.

[00241] The data from the second cohort of 13 patients were in near-perfect agreement with the initial data from the cohort of 12 patients (Figures 11A-C). However, these data together demonstrate that the results of sequencing T cell infiltrates in melanoma tumor biopsies before immunotherapy are highly correlated with patient response to immunotherapy, providing a potential biomarker to predict the effect of immunotherapy on a patient-by-patient basis.

[00242] The various embodiments described above can be combined to provide further embodiments. All of the U.S. patents, U.S. patent application publications, U.S. patent applications, foreign patents, foreign patent applications and non-patent publications referred to in this specification and/or listed in the Application Data Sheet are incorporated herein by reference, in their entirety. Aspects of the embodiments can be modified, if necessary to employ concepts of the various patents, applications and publications to provide yet further embodiments. These and other changes can be made to the embodiments in light of the above-detailed description.

[00243] In general, in the following claims, the terms used should not be construed to limit the claims to the specific embodiments disclosed in the specification and the claims, but should be construed to include all possible embodiments along with the full scope of equivalents to which such claims are entitled. Accordingly, the claims are not limited by the disclosure.

List of References

1. Laughlin MJ, Eapen M, Rubinstein P. Outcomes after transplantation of cord blood or bone marrow from unrelated donors in adults with leukemia. *N Engl J Med.* 2004;351:2265-2275.
2. Wagner JE, Barker JN, Defor TE, *et al.* Transplantation of unrelated donor umbilical cord blood in 102 patients with malignant and nonmalignant diseases: influence of CD34 cell dose and HLA disparity on treatment related-mortality and survival. *Blood* 2001;100:1611-1618.
3. Rocha V, Cornish J, Sievers EL, *et al.* Comparison of outcomes of unrelated bone marrow and umbilical cord blood transplants in children with acute leukemia. *Blood* 2001;97:2962-2971.
4. Grewal SS, Barker JN, Wagner JE, *et al.* Unrelated donor hematopoietic cell transplantation: marrow or umbilical cord blood? *Blood* 2003;101:4233-4244.
5. Rocha V, Labopin M, Sanz G, *et al.* Transplants of umbilical-cord blood or bone marrow from unrelated donors in adults with acute leukemia. *N Engl J Med.* 2004;351:2276-2285.
6. Thomson BG, Robertson KA, Gowan D, *et al.* Analysis of engraftment, graft-versus-host-disease, and immune recovery following unrelated donor cord blood transplantation. *Blood* 2000;96:2703-2711.
7. Barker JN, Hough RE, van Burik JA, *et al.* Serious infections after unrelated donor transplantation in 136 children: impact of stem cell source. *Biol Blood Marrow Transplant.* 2005;11:362-37.
8. Merindol N, Charrier E, Duval M, Soudeyins H. Complementary and contrasting roles of NK cells and T cells in paediatric umbilical cord blood transplantation. *J Leukocyte Biol.* Prepublished on March 2nd, 2011.
9. Brunstein CG, Barker JN, Weisdorf DJ, *et al.* Umbilical cord blood transplantation after nonmyeloablative conditioning: impact on transplantation outcomes in 110 adults with hematologic disease. *Blood.* 2007;110(8):3064-3070.
10. Cornetta K, Laughlin M, Carter S, *et al.* Umbilical cord blood transplantation in adults: results of the prospective Cord Blood Transplantation (COBLT). *Biology of Blood and Marrow Transplantation.* 2005;11:149-160.
11. Long GD, Laughlin M, Madan B, *et al.* Unrelated umbilical cord blood transplantation in adult patients. *Biology of Blood and Marrow Transplantation.* 2003;9:772-780.

12. Brown JA, Boussiotis VA. Umbilical cord blood transplantation: basic biology and clinical challenges to immune reconstitution. *Clinical immunology*. 2008;127:286-297.
13. Sauter C, Abboud M, Barker JN, et al. Serious infection risk and immune recovery after double-unit cord blood transplantation. *Biology of Blood and Marrow Transplantation*. 2011;17:1460-71.
14. Guérin-El Khourouj V, Dalle JH, Pédrón B, et al. Quantitative and Qualitative CD4 T Cell Immune Responses Related to Adenovirus DNAemia in Hematopoietic Stem Cell Transplantation. *Biol Blood Marrow Transplant*. 2011;17:476-85.
15. Tormo N, Solano C, Benet I, et al. Reconstitution of CMV pp65 and IE-1-specific IFN- γ CD8(+) and CD4(+) T-cell responses affording protection from CMV DNAemia following allogeneic hematopoietic SCT. *Bone Marrow Transplant*. Prepublished on January 1, 2011.
16. Przepiorka D, Smith TL, Folloder J, et al. Risk factors for acute graft-versus-host disease after allogeneic blood stem cell transplantation. *Blood* 1999;94:1465-1470.
17. Mielcarek M, Storer BE, Boeckh M, et al. Initial therapy of acute graft-versus-host disease with low-dose prednisone does not compromise patient outcomes. *Blood* 2009;113:2888-2894.
18. Nakamae H, Kirby KA, Sandmaier BM, et al. Effect of conditioning regimen intensity on CMV infection in allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2009;15:694-703.

The claims defining the invention are as follows:

1. A method for predicting an immunological response of a test subject to an immunotherapy, comprising:
 - 5 amplifying nucleic acid sequences from one or more solid tumor samples comprising lymphoid cells, wherein the samples are obtained from the test subject at one or more time points prior to and after treatment with said immunotherapy, and wherein the nucleic acid sequences are amplified in a multiplex PCR assay using a plurality of adaptive immune receptor (AIR) V-segment oligonucleotide primers and either a plurality of AIR J-segment oligonucleotide primers or one or more AIR C-segment oligonucleotide primers;
 - 10 obtaining nucleic acid sequence information from the amplified nucleic acid sequences, wherein said nucleic acid sequence information comprises sequences for a plurality of unique rearranged nucleic acid sequences, each encoding an adaptive immune receptor (AIR) polypeptide;
 - 15 quantifying an AIR sequence diversity score for said one or more solid tumor samples based on the total number of unique rearranged DNA sequences in said one or more solid tumor samples, wherein said AIR sequence diversity score is quantified using said nucleic acid sequence information;
 - 20 quantifying an AIR sequence distribution score for said one or more solid tumor samples by calculating a frequency of occurrence of a plurality of unique rearranged DNA sequences in said one or more solid tumor samples, wherein said AIR sequence distribution score is quantified using said nucleic acid sequence information;
 - 25 determining a test subject rating score for said one or more solid tumor samples based on said AIR sequence diversity score and said AIR sequence distribution score;
 - 30 comparing said test subject rating score with a control subject rating score obtained from a sample from a control subject with a known response to said immunotherapy; and
 - predicting an immunological response for said test subject to said immunotherapy based on said comparison.
2. The method of claim 1, wherein said nucleic acids comprise genomic DNA, cDNA or
35 mRNA.

3. The method of claim 1 or claim 2, wherein determining a test subject rating score comprises quantifying an AIR sequence distribution score for said test subject by: determining a number of unique rearranged AIR sequences that account for combined frequency of occurrence of up to 50%, 40%, 30%, 20% or 10% of the total number of observed rearranged sequences in said sample; and characterizing said AIR sequence distribution score as a low score if the number of unique rearranged AIR sequences that have said combined frequency of occurrence of the total number of observed rearranged sequences in said sample is less than or equal to a predetermined threshold.
4. The method of claim 1, wherein said test subject is (i) predicted to have the same outcome of response to immunotherapy as compared to the control subject, or (ii) predicted to have a different outcome of response to immunotherapy as compared to said control subject.
5. The method of claim 1, wherein said immunotherapy comprises:
- (i) an immunotherapeutic antibody, a cytokine, an immunosuppressive agent, or a vaccine;
 - (ii) an inhibitor of a negative regulator of the immune system, optionally wherein either (a) said negative regulator is selected from a group consisting of CTLA-4, LAG3 and PD-1, or (b) said inhibitor is an anti-PD-1 antibody or an anti-CTLA-4 antibody selected from the group consisting of ipilimumab and tremelimumab; or
 - (iii) an agent that targets a potentiator of an immune response, optionally wherein said potentiator of an immune response is selected from 41BB/CD137, OX40/CD134 and CD40.
6. The method of claim 1, further comprising determining a presence of a clonal expansion of at least one clone in a sample obtained from said test subject after said immunotherapy treatment, wherein said at least one clone has a frequency of occurrence that is statistically significantly different from a mean frequency of occurrence of a set of remaining clones in said sample and previously did not have a frequency of occurrence that was statistically significantly different from a mean frequency of occurrence of a set of remaining clones in a sample obtained from said test subject prior to said immunotherapy.

7. The method of claim 6, wherein said set of remaining clones comprise clones each having a frequency of occurrence that is in the top 50%, 40%, 30%, 20% or 10% of the total clones in said sample.
- 5 8. The method of claim 6, wherein said clonal expansion of said at least one clone is indicative of a poor response of said test subject to said immunotherapy treatment.
9. The method of claim 1, wherein determining a test subject rating score comprises determining an AIR sequence distribution score of said one or more samples by
10 selecting a set of unique rearranged sequences in one of said samples having a frequency of occurrence that is statistically significantly higher compared with a threshold frequency of occurrence set by other unique rearranged sequences in said one or more samples.
- 15 10. The method of claim 1, further comprising:
categorizing said test subject based on said test subject rating score in said one or more samples as a non-responder or responder to immunotherapy in comparison to the control subject; and
stratifying a patient population of test subjects according to relative likelihood of
20 responding to immunotherapy.
11. The method of claim 1, wherein: (i) determining said test subject rating score comprises extrapolating based on a mathematical model a total AIR repertoire diversity of said test subject by sequencing said nucleic acid sequences from one of said samples and
25 determining a test subject rating score from said total AIR repertoire diversity, optimally wherein said mathematical model is an unseen species model; or (ii) determining said test subject rating score comprises quantifying an AIR sequence distribution score, wherein said quantifying said AIR distribution score comprises calculating a Shannon entropy score using said nucleic acid sequence information from said one or more
30 samples.
12. The method of claim 1, wherein said adaptive immune receptor (AIR) polypeptide is a mammalian AIR polypeptide and is selected from a T cell receptor-gamma (TCRG) polypeptide, a T cell receptor-beta (TCRB) polypeptide, a T cell receptor-alpha (TCRA)
35 polypeptide, a T cell receptor-delta (TCRD) polypeptide, an immunoglobulin heavy-chain (IGH) polypeptide, and an immunoglobulin light-chain (IGL) polypeptide,

5 optionally wherein: (i) said IGH polypeptide is selected from an IgM, an IgA polypeptide, an IgG polypeptide, an IgD polypeptide and an IgE polypeptide; (ii) said IGL polypeptide is selected from an IGL-lambda polypeptide and an IGL-kappa polypeptide; or (iii) said mammalian AIR polypeptide is selected from a human AIR polypeptide, a non-human primate AIR polypeptide, a rodent AIR polypeptide, a canine AIR polypeptide, a feline AIR polypeptide and an ungulate AIR polypeptide.

10 13. The method of claim 1, wherein said test subject is a subject having or suspected of having a malignant condition, optionally wherein: (i) said malignant condition is selected from a hematologic malignancy, a melanoma, a sarcoma and a carcinoma; or (ii) said malignant condition is selected from malignant melanoma, small cell lung cancer, non-small cell lung cancer, renal cell carcinoma, pancreatic cancer, breast cancer, ovarian cancer and prostate cancer.

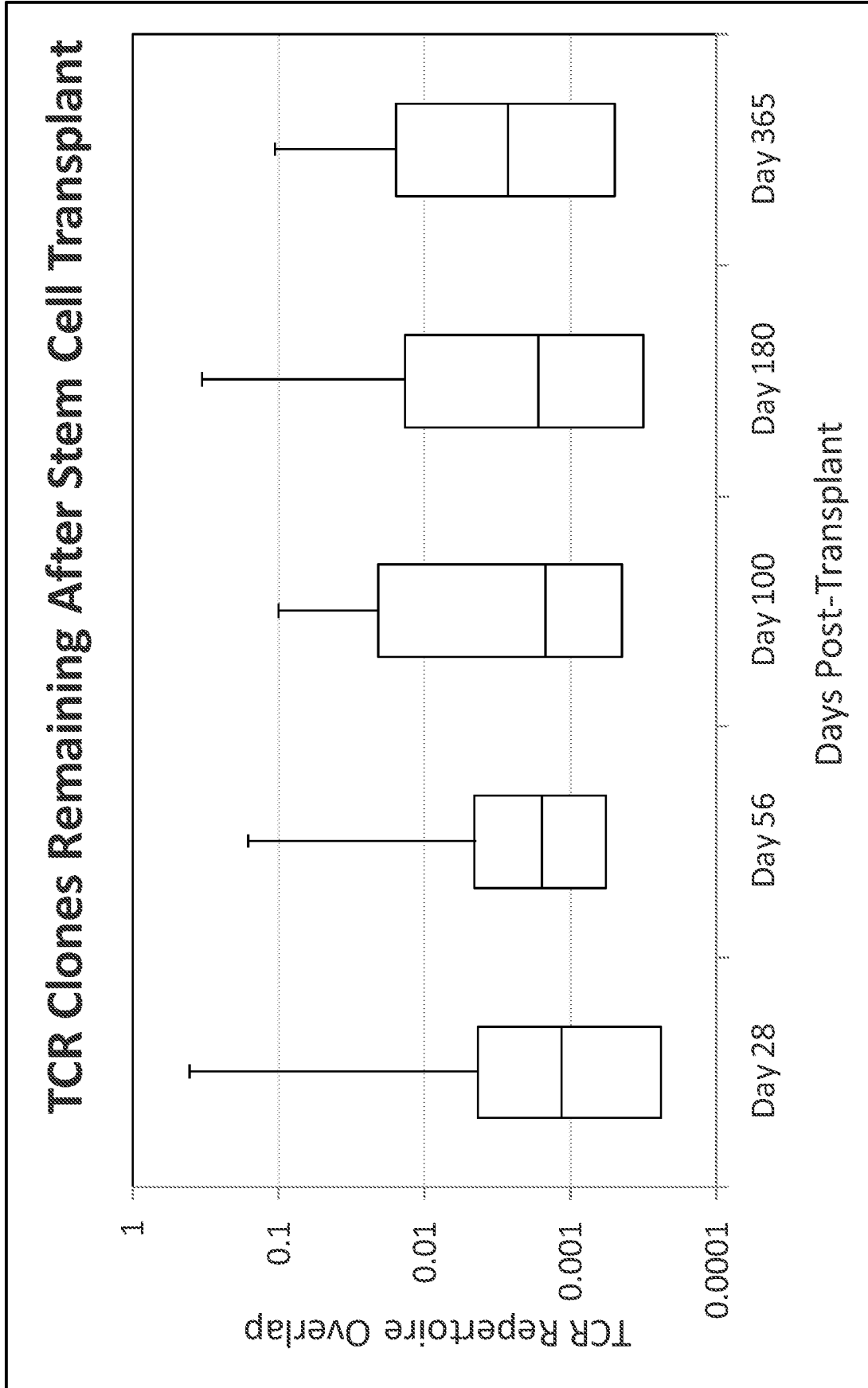


FIGURE 1

Transience of Top TCR Clones Post-Transplant

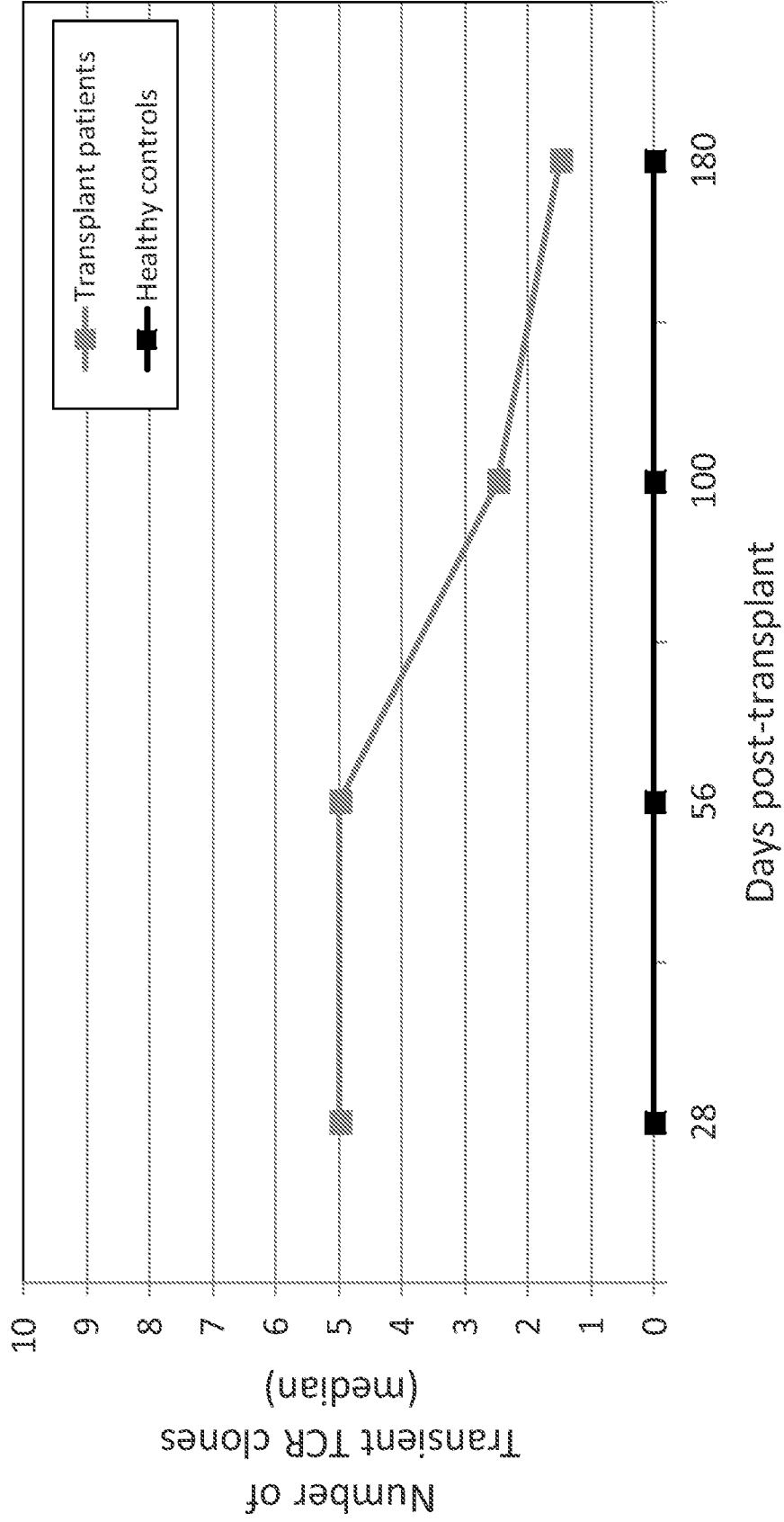


FIGURE 2

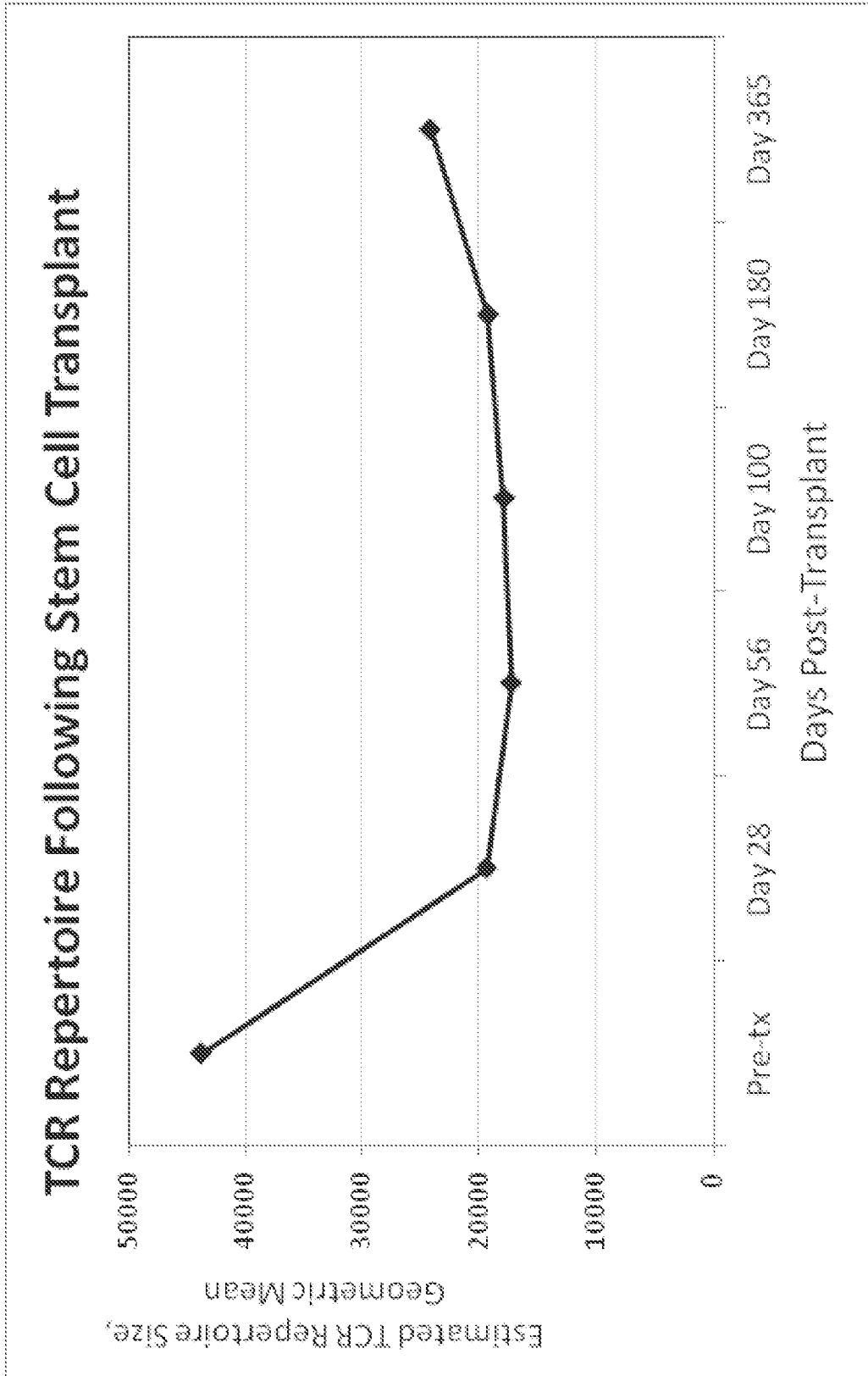


FIGURE 3

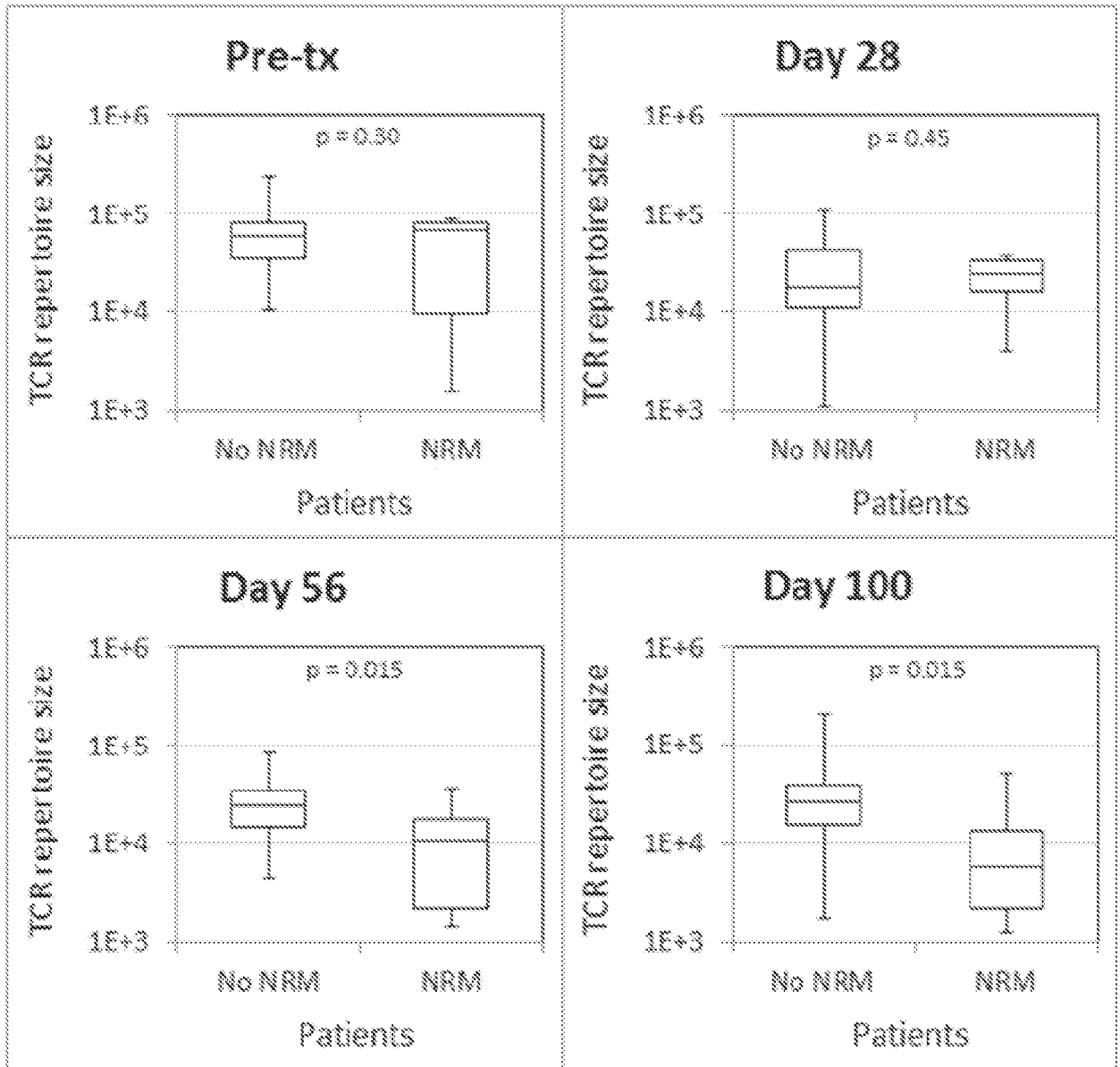


FIGURE 4

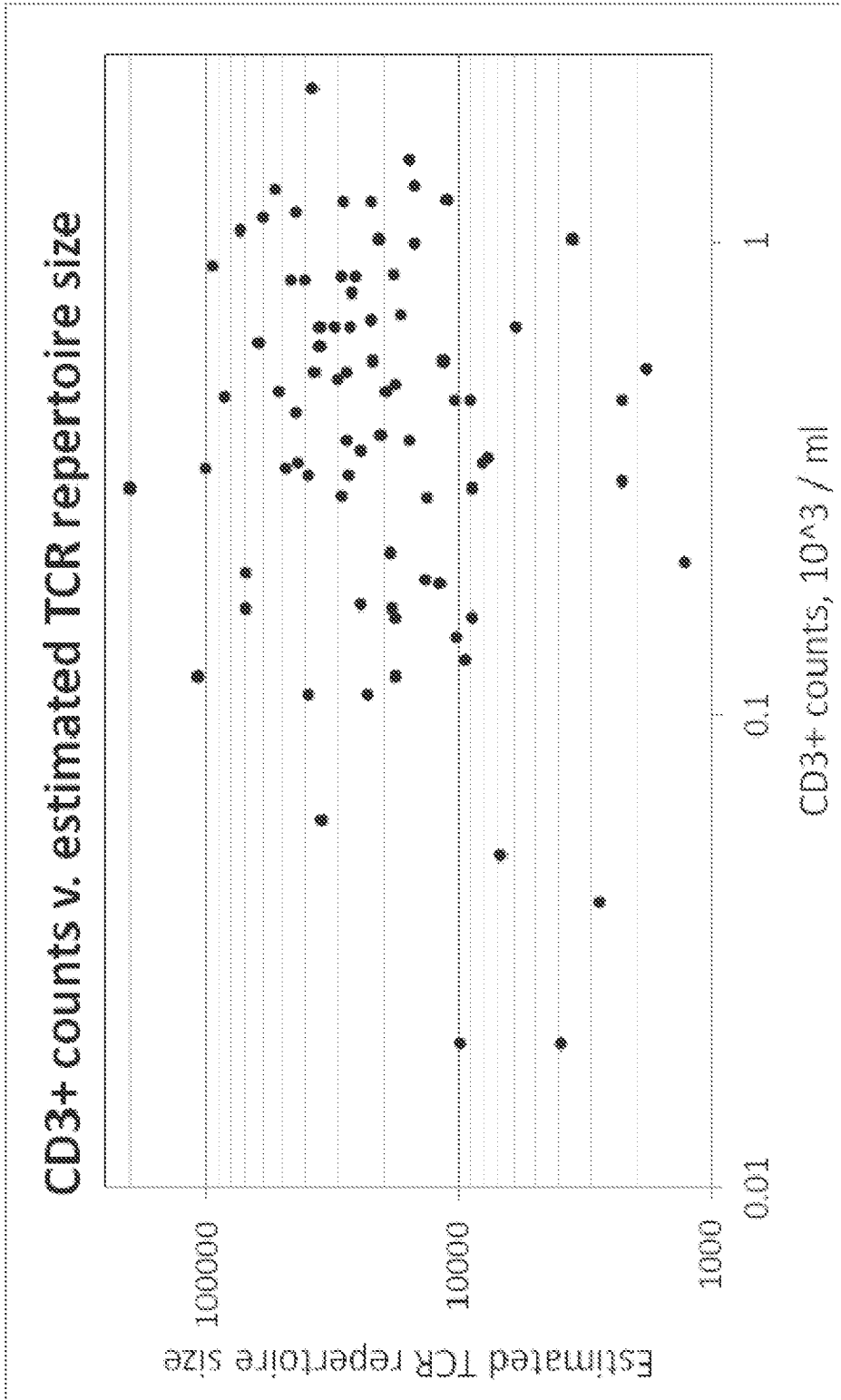
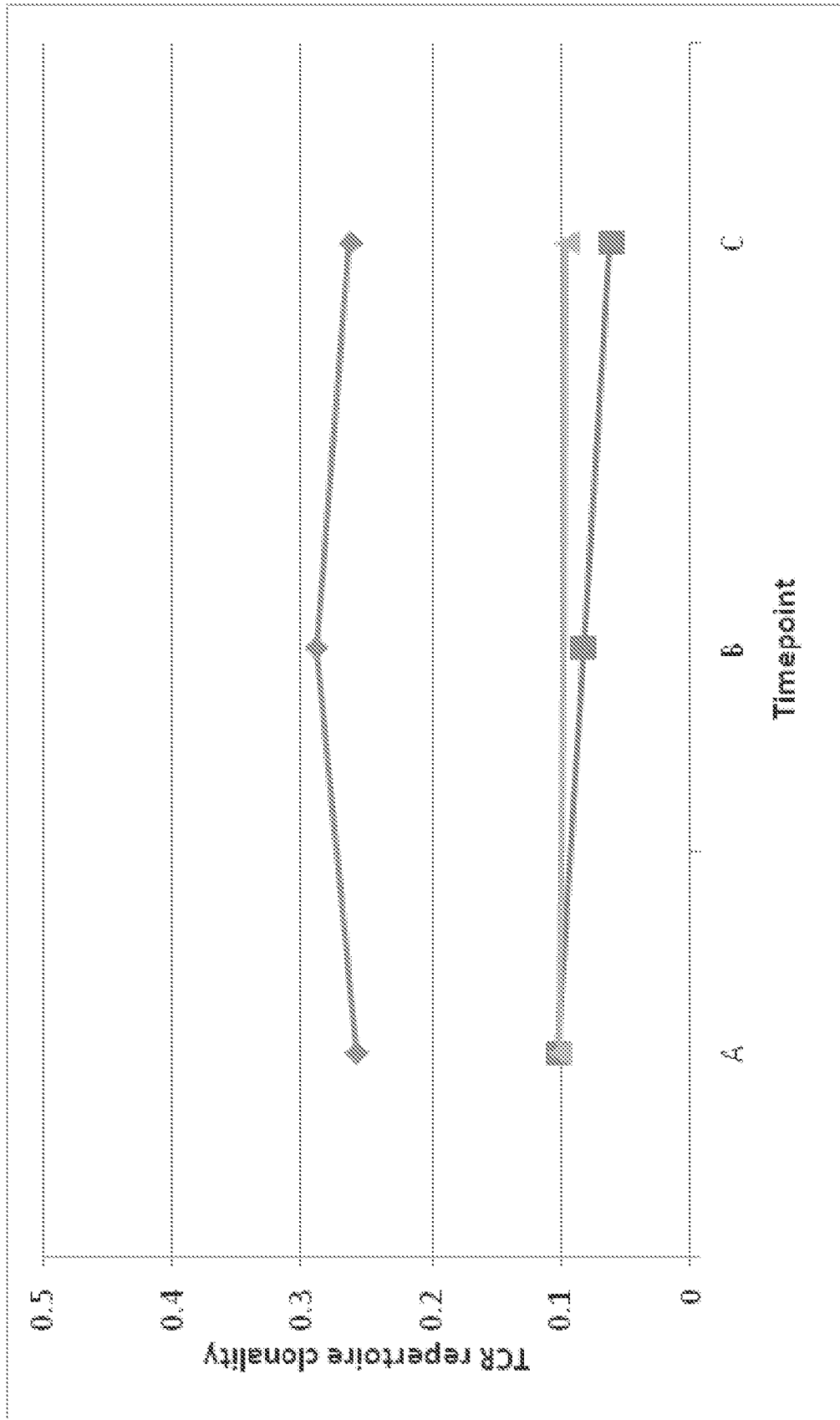


FIGURE 5



◆ = non-responder
■, ▲ = responders

FIGURE 6

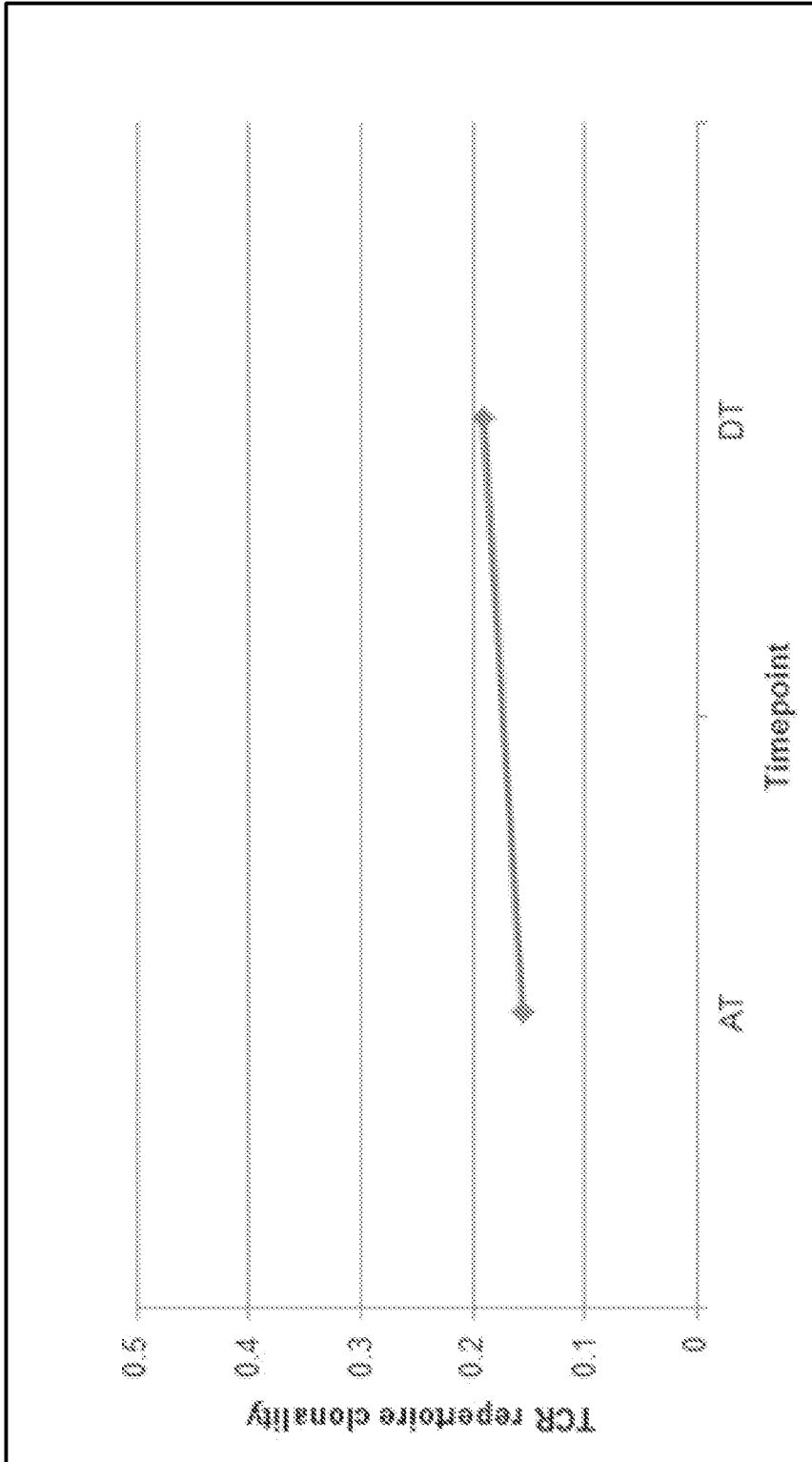
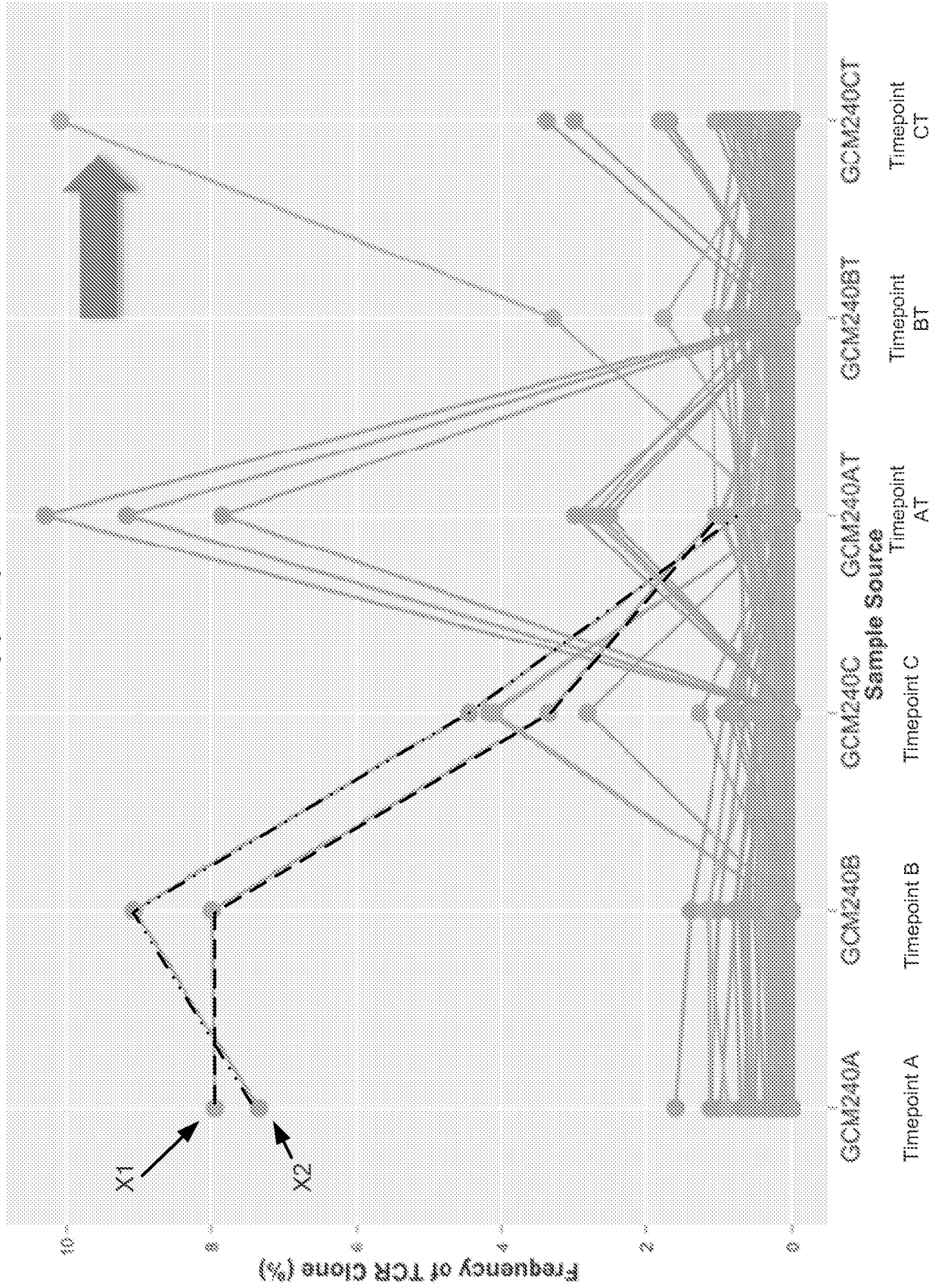


FIGURE 7

FIGURE 8



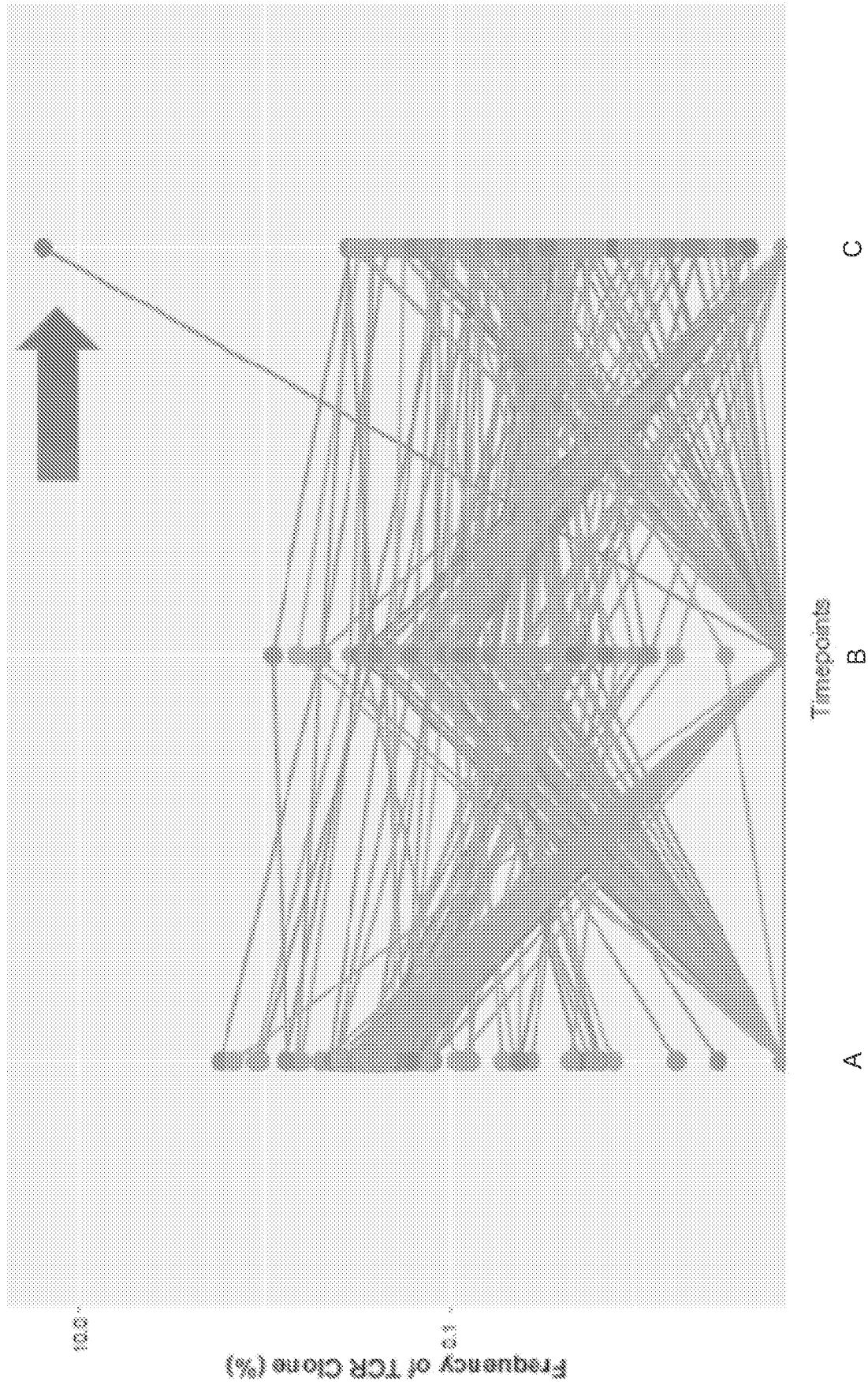


FIGURE 9

FIGURE 10A

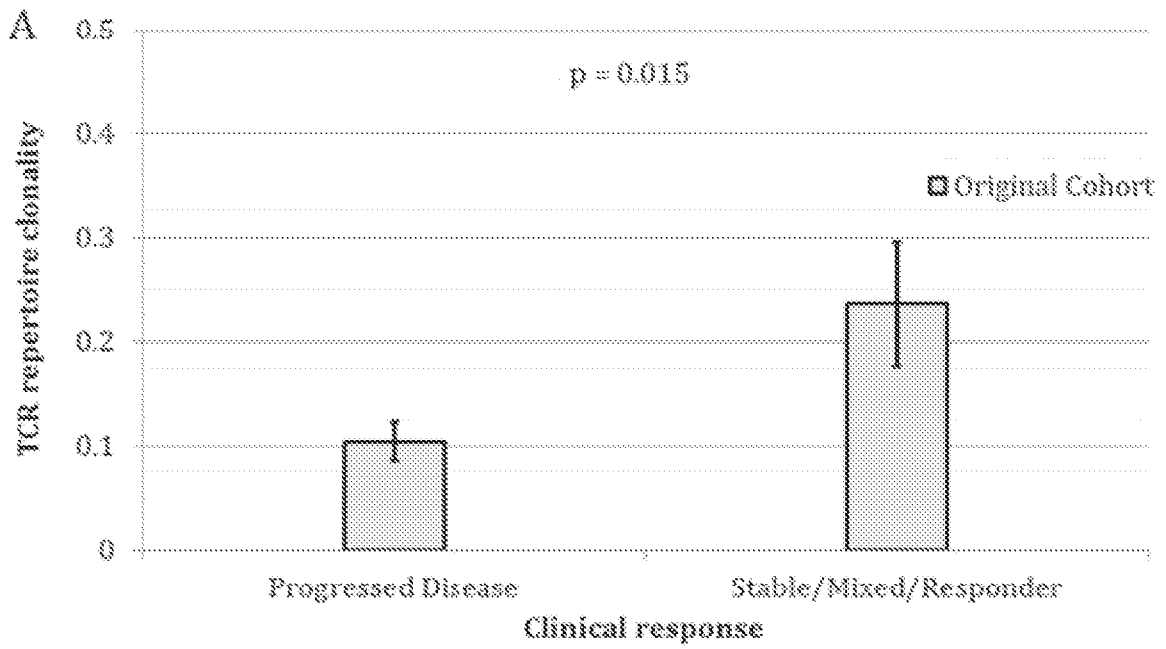
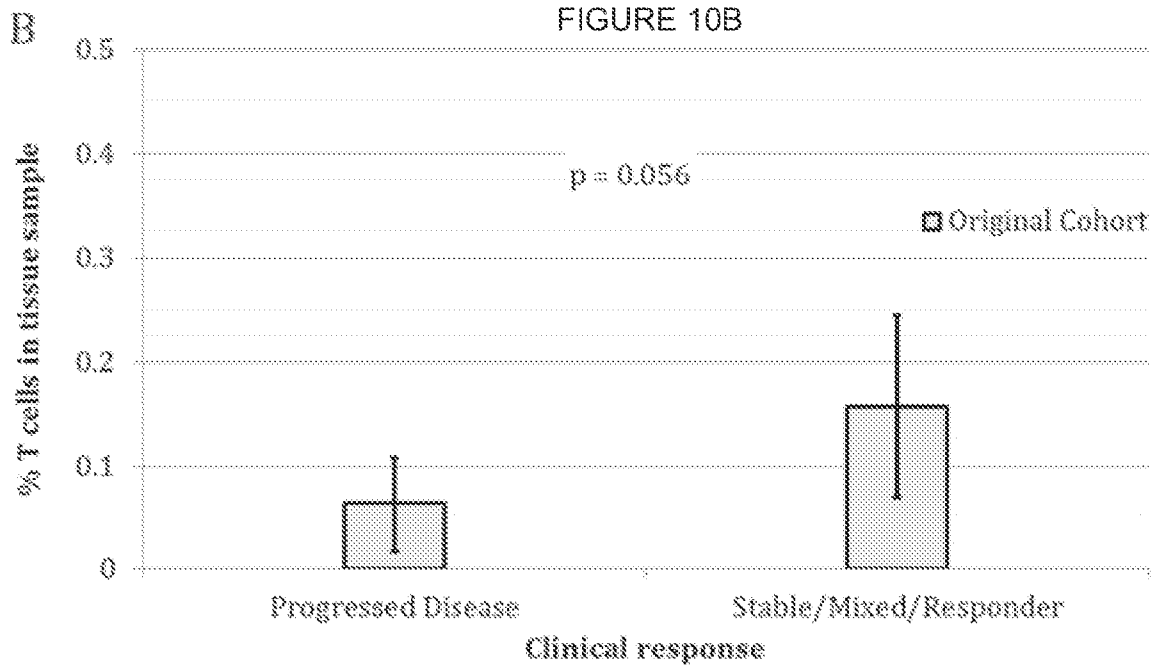


FIGURE 10B



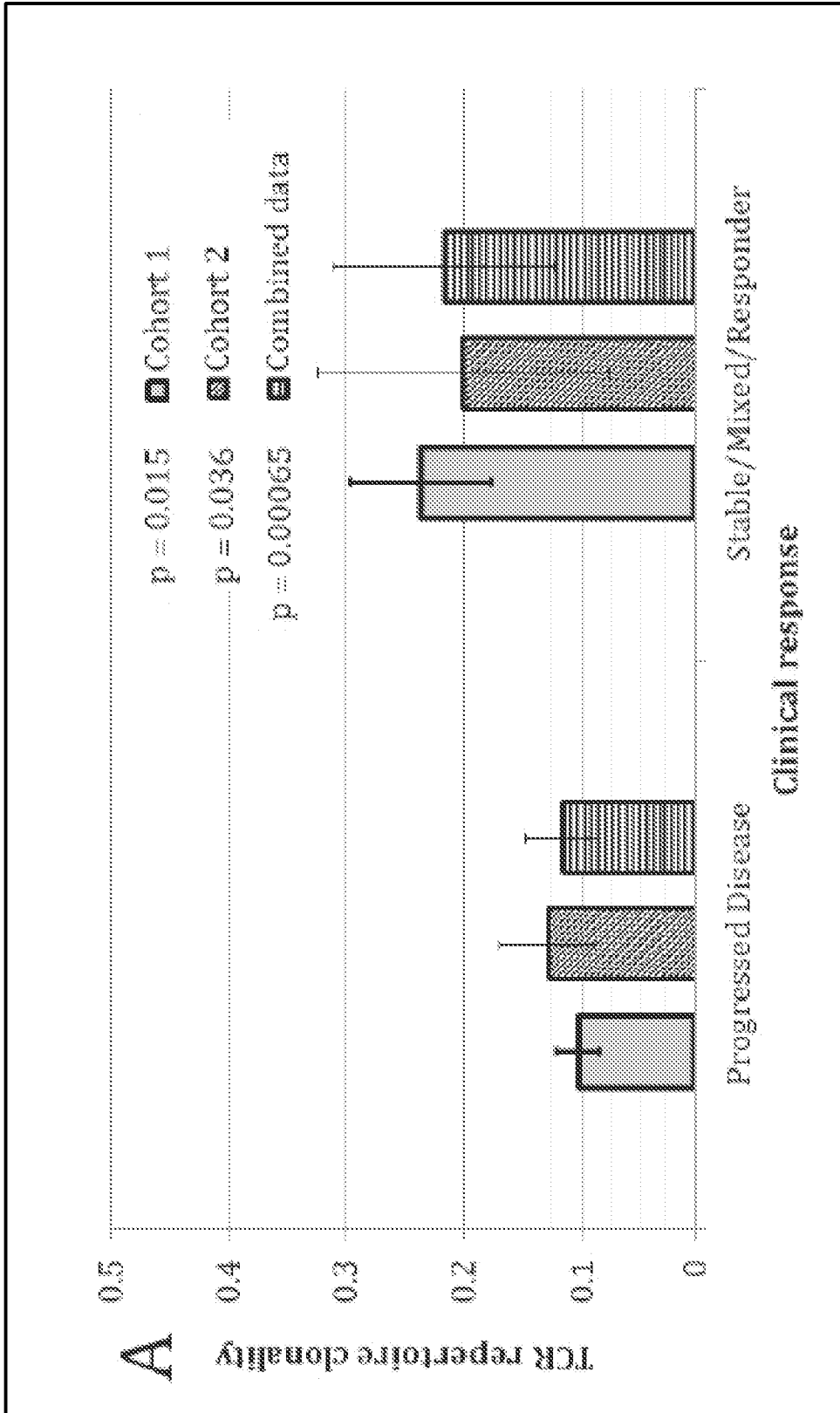


FIGURE 11A

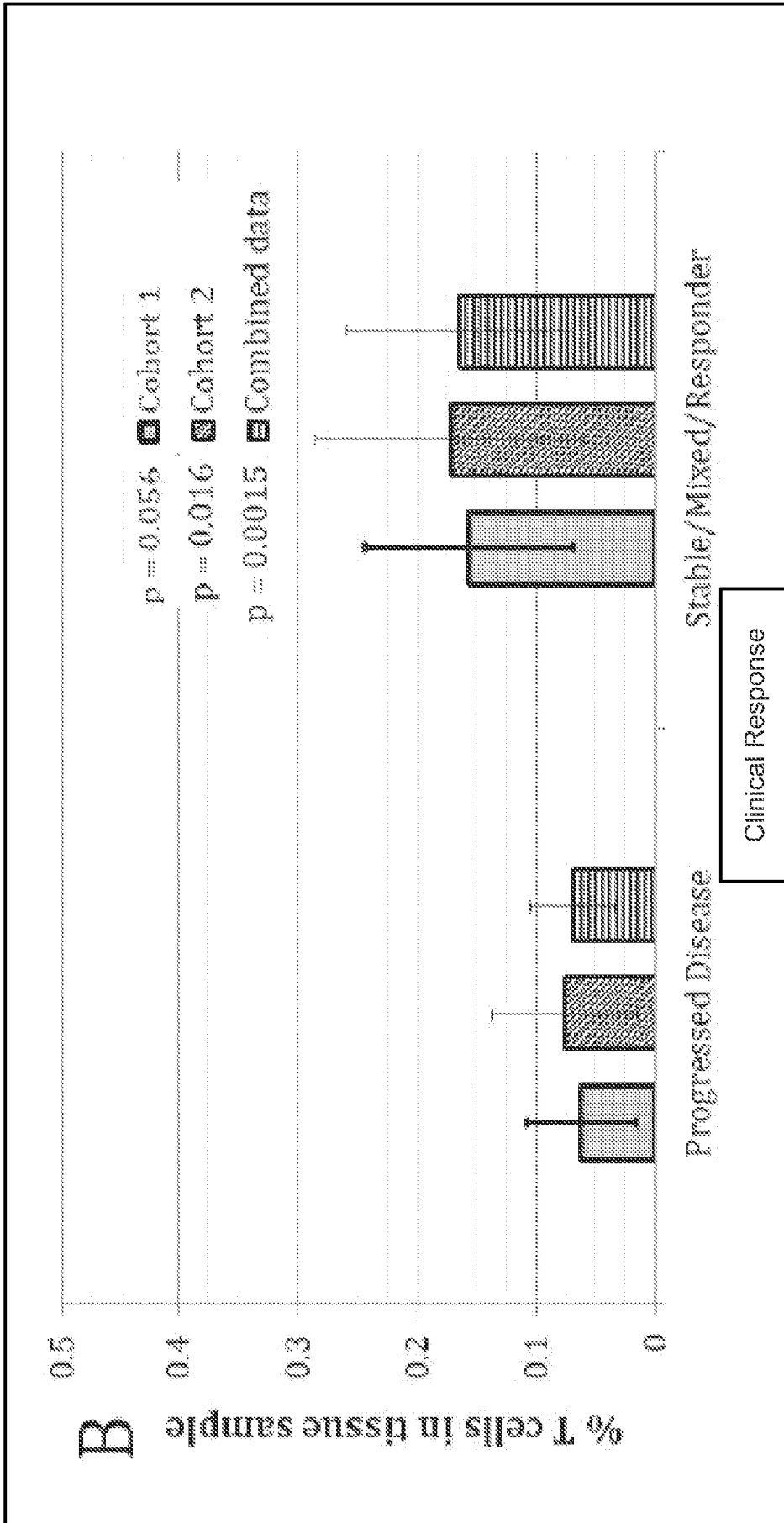


FIGURE 11B

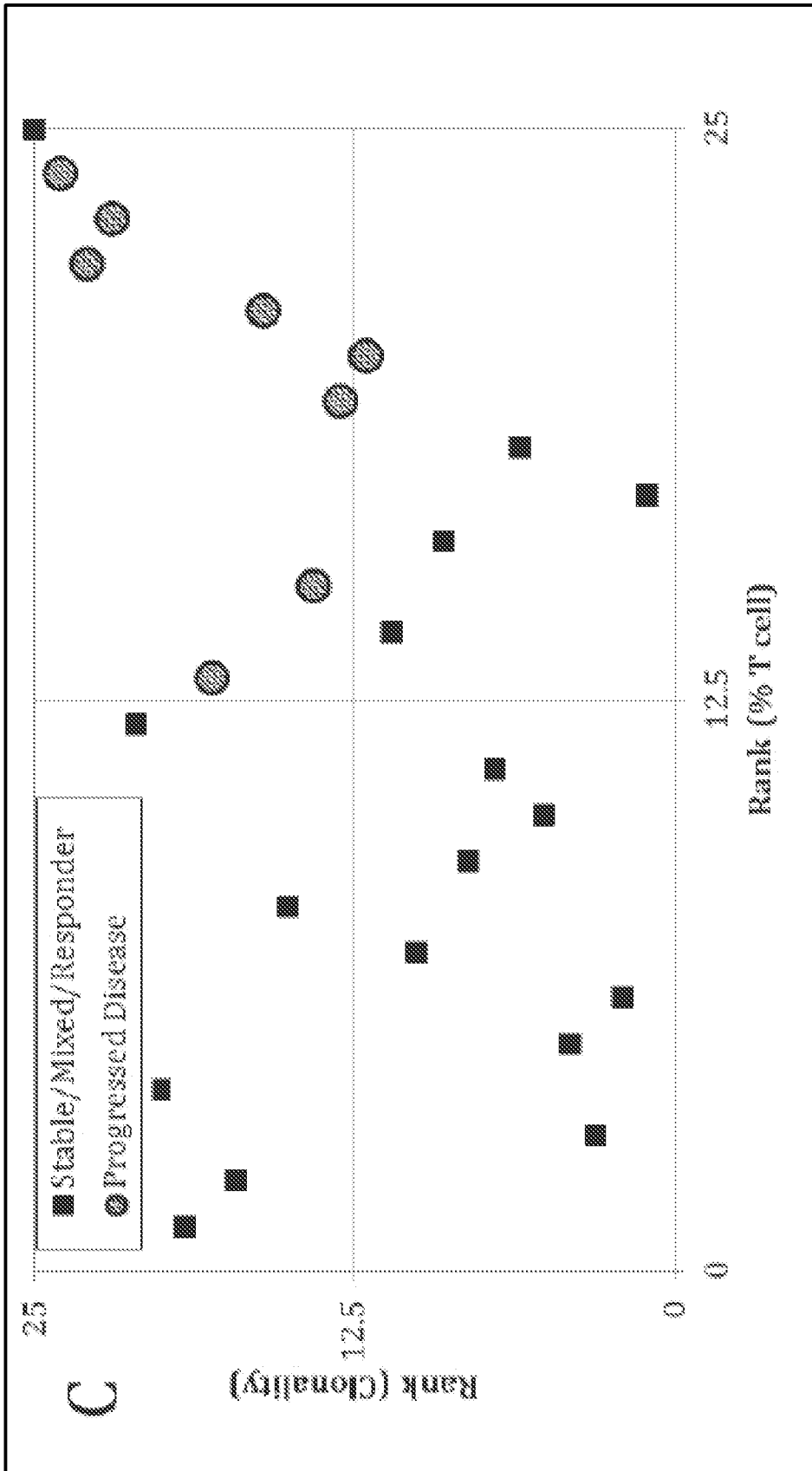


FIGURE 11C

