



(43) International Publication Date
17 September 2015 (17.09.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/138746 A1

(51) International Patent Classification:
C40B 30/04 (2006.01)

(21) International Application Number:
PCT/US2015/020215

(22) International Filing Date:
12 March 2015 (12.03.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/951,908 12 March 2014 (12.03.2014) US

(71) Applicant: **THE BOARD OF TRUSTEES OF THE LE-
LAND STANFORD JUNIOR UNIVERSITY** [US/US];
3000 El Camino Real, Building 5, Suite 300, Palo Alto,
California 94306-2100 (US).

(72) Inventors; and

(71) Applicants : **QUAKE, Stephen R.** [US/US]; 318 Campus
Drive Clark Center E350Q, Stanford, California 94305
(US). **DE VLAMINCK, Iwijn** [—/US]; 318 Campus
Drive, Stanford, California 94305 (US). **VOLLMERS,
Christopher S.** [DE/US]; 318 Campus Drive, Stanford,
California 94305 (US).

(74) Agent: **SHERWOOD, Pamela J.**; Bozicevic, Field &
Francis LLP, 1900 University Avenue, Suite 200, East
Palo Alto, California 94303 (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, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, 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, ST, 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).

Published:

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

(54) Title: MONITORING IMMUNOCOMPETANCE USING IMMUNE REPERTOIRE SEQUENCING

(57) Abstract: Methods, machine readable media and kits are provided for analyzing an antibody repertoire of a subject to determine the immunocompetence of the subject. The method may include analyzing the subject's antibody repertoire. Analyzing may include identifying the composition of abundant related IgH sequences of the subject's antibody repertoire. The method may further include providing an assessment of the immunocompetence of the subject, based on the analysis.



WO 2015/138746 A1

MONITORING IMMUNOCOMPETANCE USING IMMUNE REPERTOIRE SEQUENCING

GOVERNMENT RIGHTS

[0001] This invention was made with Government support under contract AI092673 awarded by the National Institutes of Health. The Government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0002] Predictions of the state or competence of the immune system are important in the elderly, newborns, vaccine recipients, organ acceptors, patients suffering from autoimmune disease, cancer patients, and any such patients receiving immunosuppressive therapy. For example, immune suppression may be used to treat a patient with an autoimmune disease, or to reduce risk of organ rejection in an organ transplant recipient. However, a balance must be considered, as over immune suppression can make the subject vulnerable to infection.

[0003] Immunosuppression is frequently achieved by administration of calcineurin inhibitors, which directly inhibit T-cell activation, resulting in a modification of the composition of T- and B-cell populations. T-cells are required for B-cell maturation through the process of affinity maturation and class switching. While naïve B cells express germline sequence IgM antibodies at low levels, after T cell activated mature B lineage cells, such as plasmablasts, express affinity selected antibodies with non-germline sequences, usually of the IgG and IgA isotypes. Through B cell activation and maturation, the expressed antibody heavy chain changes fundamentally not only in abundance but also in sequence. Certain immunosuppressive drugs, e.g. Cyclosporin and Tacrolimus, lower immune competence by inhibiting T cells activation. This can result in an alteration of the predominant antibody isotypes present in the blood.

[0004] Organ transplants are expensive and donor organs are rare. Maintaining the health of a donor organ for the longest possible duration is of therefore of the highest priority following the transplantation. A finely balanced level of immunosuppression protects both organ and organ recipient following organ transplants, such as heart transplants. While drug levels can be tightly monitored and adjusted, drugs have varied effects on individuals and thus equal levels of an immunosuppressive drug, such as Cyclosporin or Tacrolimus, do not equate equal levels of immunosuppression.

[0005] Fine-tuning of the state of the immune system may require accurate and predictive measurements of the immunosuppressive state of the subject. A single assay that accurately determines the net state of immunosuppression in a given transplant recipient is highly desirable, as it would allow titration of one or more immunosuppressive drugs to achieve a favorable balance between the competing risks of rejection and infection.

Publications

- [0006] Vollmers C, Sit RV, Weinstein JA, Dekker CL, & Quake SR (2013) Genetic measurement of memory B-cell recall using antibody repertoire sequencing. Proc Natl Acad Sci U S A 110(33):13463-13468.

SUMMARY OF THE INVENTION

- [0007] Methods, machine readable media and kits are provided for analyzing an antibody repertoire of a subject to determine the immunocompetence of the subject, e.g. where the subject is treated with an immunosuppressive regimen. The status of immunocompetence thus determined may be provided to the individual or to a health care giver. The methods of the invention also include providing an immunosuppressive agent in accordance with the immunocompetence status, where, for example, an individual showing high levels of immunosuppression may be administered a low dose of an immunosuppressive agent, and an individual showing low levels of immunosuppression may be administered a high dose of an immunosuppressive agent.
- [0008] Methods of the invention include, without limitation, analysis of an antibody repertoire in detail, e.g. including determination of variable region sequences for a plurality of B cells present in a biological sample. Analysis is also performed to identify the isotype composition of antibodies expressed by a plurality of B cells in a biological sample obtained from the subject. In some embodiments, the analysis is performed on mRNA transcripts from a plurality of B cells present in the biological sample. Analyzing may include identifying the composition of abundant related IgH sequences in a biological sample obtained from the subject. The method may further include providing an assessment of the immunocompetence of the subject, based on the analysis. In some embodiments, the immunosuppressive regimen of the subject is modified in accordance with the assessment of immunocompetence.
- [0009] In certain embodiments, the method includes providing an assessment of the immunocompetence of the subject. For example, providing the assessment may include providing an assessment that the subject is not adequately immunosuppressed when the relative amount of IgG, IgA, IgD, IGE, mutated IgM or a combination thereof (e.g. all non-naïve IgM) is above a pre-determined threshold relative to the total number of abundant similar IgH sequences; or when the relative amount of naïve-IgM is below a predetermined threshold. Alternatively or in addition, the step of providing the assessment may include providing an assessment that the subject is overly immunosuppressed when the relative amount of IgG, IgA, IgD, IGE, mutated IgM or a combination thereof (e.g. all non-naïve IgM) is below a predetermined threshold or when the relative amount of naïve-IgM is above a predetermined threshold. In some aspects, the assessment includes a prediction of the likelihood of allograft rejection. In certain aspects, the assessment may further include an assessment of the level of

donor-derived cfDNA in the subject's blood, a biopsy of a transplanted organ, or a combination thereof.

[0010] The subject can be immunosuppressed through any suitable immunosuppressive regimen. The immunosuppressive regimen may involve administering one or more immunosuppressive drugs, such as a T-helper cell suppressant. In certain embodiments, the immunosuppressive regimen involves administering a calcineurin inhibitor, e.g. cyclosporine, tacrolimus, sirolimus, etc.; myriocin; mycophenolate; or a combination thereof as known in the art. Alternatively or in addition, the immunosuppressive regimen may comprise administering one or more steroids, thymocyte depleting antibodies, or other agents suitable to an immunosuppressive regimen.

[0011] In certain embodiments, the subject is the recipient of an organ transplant. The organ transplant may be a heart transplant. In other embodiments, the organ transplant may be an eye, pancreas, liver, kidney, heart, brain, bowel, lung or skin transplant. In certain embodiments, the subject is a mammal, such as a human.

[0012] In certain embodiments, the method includes obtaining a sample comprising B cells from the subject. For example, the sample may be any of a blood sample, synovial fluid, CSF, lymph node biopsy, spleen tissue, bone marrow tissues, extracted peripheral blood mononucleated cells (PBMCs), extracted B-cells, extracted plasma cells, extracted plasmablasts, and so forth.

[0013] The method may include preparing a library of IgH encoding molecules from the sample, prior to the step of analyzing the sequence representation. Preparing the library may include isolating RNA from cells in the sample, and may further comprise amplifying isolated RNA transcripts by cDNA synthesis followed by PCR to obtain IgH encoding molecules. The amplification product will generally encompass a region of the IgH sequence sufficient to identify the isotype. The amplification product optionally includes a region sufficient to identify the sequence of the variable region in addition to the isotype. Amplification may be performed through any suitable technique, including without limitation Rapid Amplification of cDNA Ends (RACE). One or more primers use for cDNA synthesis may each include a unique identifier nucleotide sequence (UID). Transcripts originating from the same RNA transcript may share a unique identifier nucleotide sequence (UID).

[0014] The method may include sequencing the IgH encoding molecules in the library to obtain IgH molecule reads. The IgH molecule reads may be obtained by aligning one or more shared variable (V), diversity (D) and joining (J) segments of shorter sequence reads. In certain embodiments, the sequencing may be paired-end sequencing.

[0015] The step of analyzing may involve classifying IgH encoding molecule reads sharing the same UID as belonging to the same IgH sequence. The step of analyzing may further involve identifying IgH sequences having more than a predetermined threshold of IgH encoding

molecules as “abundant related IgH sequences”. Alternatively, the step of analyzing may involve classifying similar IgH sequences as belonging to the same IgH lineage, and optionally further, classifying IgH lineages having more than a predetermined threshold of IgH sequences (or IgH encoding molecules) as abundant related IgH sequences.

[0016] The step of identifying the composition of abundant related IgH sequences may involve identifying the relative amount of one or more isotypes, such as the relative amount of IgA, relative amount of IgG or subclasses thereof, relative amount of IgM (e.g., naïve, non-naïve, or total IgM), or a combination thereof. For example, the step of identifying the composition of abundant related sequences may involve identifying the relative amount of abundant ‘non naïve-IgM’ sequences” (ABS).

[0017] The method may include analyzing an additional characteristic of the antibody repertoire. The additional characteristic may be isotype distribution, number of abundant related IgH sequences shared between time points, number of abundant related IgH sequences not present in an earlier time point, changes in the abundance of one or more isotypes, an abundance of mutations in related sequences of one or more isotypes, mutation rate, or a combination thereof.

[0018] The step of providing an assessment of the immunocompetence of the subject may include providing an assessment that the subject is not adequately immunosuppressed when the relative amount of IgG is above a certain threshold, when the relative amount of IgA is above a certain threshold, when the relative amount of non-naïve IgM or total IgM is above a certain threshold, when the relative amount of naïve-IgM is below a certain threshold, when the relative amount of ABS is above a certain threshold, or any combination thereof. Alternatively or in addition, the step of providing the assessment may include providing an assessment that the subject is overly immunosuppressed when the relative amount of IgG is below a certain threshold, when the relative amount of IgA is below a certain threshold, when the relative amount of non-naïve IgM or total IgM is below a certain threshold, when the relative amount of naïve-IgM is above a certain threshold, when the relative amount of ABS is below a certain threshold, or any combination thereof. In some aspects, the assessment may be a prediction of the likelihood of allograft rejection. In certain aspects, the assessment may further include an assessment of the level of donor-derived cfDNA in the subject’s blood, a biopsy of a transplanted organ, or a combination thereof.

[0019] Any of the above methods may be performed on multiple samples obtained at distinct time points. The assessment may be based on changes in the composition of abundant related IgH sequences of the subject’s antibody repertoire, any of the above mentioned additional characteristics, or a combination thereof, over multiple time points.

[0020] Methods are also provided for determining or modifying an immunosuppressive regimen of the subject based on the assessment of the immunocompetence of the subject. Methods for

modifying an immunosuppressive regimen may include any of the above described steps and embodiments. In addition, a method for modifying an immunosuppressive regimen may include adjusting the dosage of an immunosuppressive drug. The dosage may be the amount of the immunosuppressive drug administered at one time, the frequency of administration, the level of the drug to maintain in the subject's blood, and so forth. In one example, the dosage may be increased based on an assessment that the subject is not adequately immunosuppressed. In another example, the dosage may be decreased based on an assessment that the subject is overly immunosuppressed. The method may further include discontinuing treatment with an immunosuppressive drug based on the assessment. Alternatively or in addition, the method may further include treating the subject with an immunosuppressive drug not previously part of the immunosuppressive regimen, based on the assessment.

[0021] Also included are machine readable media including software configured to perform any of the above mentioned steps of sequencing, analyzing, and/or providing an assessment. In addition, kits are provided for performing any of the above methods. The kit may include primers for performing cDNA synthesis or PCR amplification on IgH sequences (e.g., as described in Example 1). One or more primers may each comprise a unique identifier nucleotide sequence (UID). The kit may additionally comprise a machine readable media that includes the above mentioned software.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022] **Figs. 1A-I** provides a consensus read approach validation. **(A)** Experimental setup of Experiment 1. Blood is drawn from an individual. Two separate PBMC aliquots are prepared, and RNA is extracted. Several sequencing libraries are prepared and sequenced. **(B)** Venn diagrams illustrating the IgH sequence overlap between L1 D7 and Lib. Rep. for all IgH sequences or only abundant IgH sequences represented by more than 5 IgH molecules. **(C)** The abundances of IgH sequences shared between L1 D7 and Lib. Rep. are shown as a scatterplot. **(D-I)** IgH sequences in L1 D7 are ordered by abundance. **(D)** Abundance (IgH molecules/IgH sequence) of sequence. In bins of 300 sequences from left to right, percent of shared sequences between L1 D7 and indicated samples [**(E)** Seq. Rep., **(F)** Lib. Rep., and **(G)** Bio. Rep.] are shown. **(H)** Isotype distribution of sequences. **(I)** Mutation rate (%) of L1 D7 sequences separated by isotypes (IgM, IgG, and IgA).

[0023] **Figs. 2A-B** provides a repertoire size estimation. **(A)** Visualization of lineages shared between replicate sample pairs (R1 and R2) from individuals (L1, T6, T7, and T8). Data were subsampled to 50,000 IgH molecules to normalize for variability in sampling depth. Lineages of each time point are plotted on the circumference of the circle, with the gray area representing abundance of the respective lineages (logarithmic). Lineages present in two time points are

connected with lines colored according to their isotype. **(B)** Capture–recapture estimation of different IgH lineage populations. Average ($\pm$ SD) of four individuals is shown.

[0024] **Figs. 3A-B** characterizes influenza vaccination response. **(A)** Change in the number of abundant lineages (containing at least 5 IgH molecules) in response to vaccination by LAIV or TIV is shown for 14 individuals each. Data were subsampled to 50,000 molecules to normalize for variability in sampling depth. Data are shown for each isotype separately. P value is determined by a Mann–Whitney u test. **(B)** IgH repertoire at days 0, 7, and 28 for individual T4. Each sample was subsampled to 50,000 IgH molecules to normalize for variability in sampling depth. Sequences were ordered by abundance, and each sequence is shown as a vertical bar [the height depicts its abundance (logarithmic)]. The color of each bar indicates the isotype of the sequence.

[0025] **Figs. 4A-B** characterizes TIV recall response. **(A)** Number of lineages shared between indicated time points normalized to the overall amount of lineages. Average of five individuals is shown as heat maps for IgG, IgA, and IgM. **(B)** Median abundance of IgG, IgA, and IgM lineages shared between the indicated time points.

[0026] **Figs. 5A-C.** characterizes lineage recall across different time points. **(A and B)** Visualization of lineages shared between time points. Data were subsampled to 50,000 IgH molecules for each time point to normalize for variability in sampling depth. Lineages of each time point are plotted on circumference of the circle, with the gray area representing the abundance of the respective lineages (logarithmic). Lineages present in two time points are connected with lines colored according to their isotype. All shared lineages are shown in A, and shared IgG, IgA, and IgM lineages are shown separately in B. **(C)** Shared IgG lineages between Y1 D7 and Y2 D7 time points of three individuals are shown as in A.

[0027] **Figs. 6A-B** is a visualization of the consensus read library generation and sequencing approach. **(A)** The human immunoglobulin heavy chain locus undergoes somatic recombination to form its variable region. After transcription, the isotype-determining constant region is spliced to the variable region to form IgH RNA molecules. IgH RNA molecules are reverse transcribed using primers specific for the 5' end of the constant region using pool isotype-specific primers. Second-strand synthesis uses a primer pool specific to the framing region 3 (FR3) of IgH V segments. Both second-strand and reverse transcription primers contain 8 random nt and partial sequencing adapter sequences. A PCR is used to amplify the IgH sequences and complete the sequencing adapters. After sequencing, 16 random nt are used to build unique identifier (UID) groups of raw sequencing reads that all originated from the same original IgH RNA molecule. High-quality consensus reads are generated for each IgH molecule, and identical IgH molecules are grouped into IgH sequences. **(B)** Schematic of how IgH molecules were grouped into IgH sequences, which were then clustered into IgH lineages. Identical IgH molecules are grouped into IgH sequences. Clustering of IgH sequences requires 90%

similarity in the junctional region but allows for differences in the other parts of the IgH sequences, like mutations (red stars) or different isotypes.

[0028] **Fig. 7A-C** provides a consensus read validation. **(A)** Histogram showing the number of consensus groups containing the indicated number of reads. **(B)** Estimated error rates for consensus reads and raw Illumina reads are shown. **(C)** The number of unique IgH sequences among ~500,000 consensus reads, ~500,000 raw reads, or ~2,500,000 raw reads is shown.

[0029] **Fig. 8** shows primer sequences used in Experiment 1. Partial paired end (PE) 1 adapters are in red, and partial PE2 adapters are in blue. Random nucleotides are in pink, and Illumina barcodes are in turquoise. Gene-specific nucleotides are in black, and complete PE1 and PE2 are in green.

[0030] **Fig. 9** provides a visualization of the antibody sequences in Peripheral Blood Mononuclear Cells (PBMC), as well as sorted naive B cells and plasmablasts, as determined by Consensus Read immune repertoire sequencing (CRIS). Sequences are plotted on the circumferences of the circle with the colored area representing their expression level. Sequences shared between two populations are connected. Naive B cells lowly express IgM sequences while plasmablasts highly express IgA and IgG sequences. The PBMC fraction contains all B cells in the blood and therefore presents a mix of both naive B cells and plasmablasts. Sequences shared by the PBMC and the plasmablast samples can be identified by their high expression level. Therefore, the contribution of naive B cell and plasmablast populations to the overall B cell pool can be determined by assaying PBMC samples.

[0031] **Figs. 10A-B** provides scatter plots characterizing immunosuppressive state and rejection events in heart transplant subjects. **(A)** shows the ratio of highly expressed IgA sequences and **(B)** Tacrolimus level after transplantation. Every individual is plotted with a different color. In **(A)** rejection events are plotted as solid circles. In **(B)** the average Tacrolimus levels at every time point are shown as solid black circles.

[0032] **Fig. 11** provides a schematic of the experimental design for measuring transplant organ health and immunosuppression in Example 3. Organ transplant recipients undergo immunosuppressive therapy. Tacrolimus is used to inhibit T cell activation. This leads to less active cytotoxic T cells (T_C) damaging the donor organ but also less active helper T cell (T_H) activating naive B cells (B_N) to differentiate into plasmablast (B_{PB}). We measured donor organ health in by both endomyocardial Biopsy as well as donor DNA levels in the blood (cfDNA-Seq). Immunosuppression was measured using Immune Repertoire Sequencing (RNA IgH-Seq).

[0033] **Figs. 12A-C** characterizes the use of Immune Repertoire Sequencing to measure Immunosuppression. **(A)** IgH Sequence abundance (Molecules/Sequence) is shown as histogram. Bar colors indicate contribution of different isotypes. Insert shows rescaled histogram. **(B)** Abundance of non-singleton, non-naive IgM sequences (ABS) as a function of

Tacrolimus levels (Average + SEM). **(C)** Median ABS levels across all individual in the study (n=12) shown for the first year post transplant. Colors indicate isotype contribution.

[0034] **Figs. 13A-D** shows lowered Immunosuppression is indicative of organ rejection. ABS levels are shown for 2 individuals with **(C, D)** and 2 individuals without **(A, B)** moderate to severe rejection events (top). Colors indicate isotype contribution.

[0035] cfDNA levels are shown for the same individuals (bottom). Biopsy scores are shown as vertical blue lines (bottom). Thin lines indicate mild (1R), thick lines indicate moderate to severe (>2R) rejection events.

[0036] **Figs. 14A-C** shows ABS level correlate well with cfDNA level and Biopsy scores. **(A)** ABS levels are plotted for all participants in the study. Rejection events as diagnosed by Biopsy or cfDNA levels are shown in color. **(B)** Samples are binned into Non-, Mild-, Moderate/Severe-Rejecting, and elevated cfDNA. As Biopsy scores are independent of cfDNA measurements, samples in the elevated cfDNA and Mild or Moderate/Severe Biopsy bins overlapped. **(C)** ROC curves for ABS levels and cfDNA (red) or Biopsy scores (blue). Samples with negative Biopsy scores and non-elevated cfDNA levels were treated as true negatives.

DETAILED DESCRIPTION

[0037] Methods, machine readable media and kits are provided for analyzing an antibody repertoire of a subject to determine immunocompetence. The method includes analyzing the subject's antibody repertoire, which may include identifying the composition of abundant related IgH sequences of the subject's antibody repertoire. The method may further include providing an assessment of the immunocompetence of the subject, based on the analysis.

[0038] Before the subject invention is described further, it is to be understood that the invention is not limited to the particular embodiments of the invention described below, as variations of the particular embodiments may be made and still fall within the scope of the appended claims. It is also to be understood that the terminology employed is for the purpose of describing particular embodiments, and is not intended to be limiting. In this specification and the appended claims, the singular forms "a," "an" and "the" include plural reference unless the context clearly dictates otherwise.

[0039] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range, and any other stated or intervening value in that stated range, is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated

range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

[0040] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs. Although any methods, devices and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, illustrative methods, devices and materials are now described.

[0041] All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the subject components of the invention that are described in the publications, which components might be used in connection with the presently described invention.

[0042] "Subject" or "patient" shall mean any animal, such as a human, non-human primate, mammal, e.g. mouse, rat, guinea pig, rabbit, dog, horse, cow, cat, etc.

[0043] "Sample", as used herein, refers to a sample from an individual or a pool of individuals, that contains cells expressing antibodies of interest. Such cells comprise, without limitation, lymphocytes expressing immunoglobulin, e.g. pre-B cells, B-cells, e.g. memory B cells, and plasma cells. Samples of interest may be complex, e.g. a blood sample, a lymph node sample, synovial fluid sample, CSF fluid, biopsy sample, spleen biopsy, sample from joints, lymphatic vessels, Peyer's patch, lymph fluid, purified peripheral blood mononucleated cells (PBMCs), purified B-cells, purified plasma cells, purified plasmablasts, etc., comprising more than 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} or more different antibody sequences, where each cell typically expresses a distinct sequence.

[0044] For isolation of cells from tissue, an appropriate solution can be used for dispersion or suspension. Such solution will generally be a balanced salt solution, e.g. normal saline, PBS, Hank's balanced salt solution, etc., conveniently supplemented with fetal calf serum or other naturally occurring factors, in conjunction with an acceptable buffer at low concentration, generally from 5-25 mM. Convenient buffers include HEPES, phosphate buffers, lactate buffers, etc. The separated cells can be collected in any appropriate medium that maintains the viability of the cells, usually having a cushion of serum at the bottom of the collection tube. Various media are commercially available and may be used according to the nature of the cells, including dMEM, HBSS, dPBS, RPMI, Iscove's medium, etc., frequently supplemented with fetal calf serum.

[0045] The term "inflammatory" response is the development of a humoral (antibody mediated) and/or a cellular (mediated by antigen-specific T cells or their secretion products) response to an immunogen. An "immunogen" is capable of inducing an immunological response against itself on administration to a mammal or due to autoimmune disease.

[0046] The term “autoimmune disease” refers to conditions in which there is an undesirable immune response directed at endogenous molecules. Autoimmune diseases may be primarily T cell mediated, antibody mediated, or a combination of both. The following listing of specific conditions is intended to be exemplary, not comprehensive. Autoimmune diseases include rheumatoid arthritis, a chronic autoimmune inflammatory synovitis affecting 0.8% of the world population.

[0047] The response of the host immune system to a graft, or of a graft towards the host (GVHD) is indicated by alterations in the immune repertoire. Grafts include the transplantation of cells, tissues and organs, such as the transfusion of blood or blood components, the grafting of bone, skin, bone marrow, etc., and the transplantation of tissues of the eye, pancreas, liver, kidney, heart, brain, bowel, lung, etc. Of interest are transplantation of solid organs, such as a heart transplant, kidney transplant, liver transplant, pancreas transplant, lung transplant, intestine transplant and skin transplant. As used herein, a graft recipient is an individual to whom tissue or cells from another individual (donor), commonly of the same species, has been transferred, particularly where one or more of the Class I MHC antigens are different in the donor as compared to the recipient. However, in some instances xenogeneic, e.g. pig, baboon, etc., tissue, cells or organs will be involved. Tissue rejection resulting from immune response of the host immune system to a graft may range from mild to severe, e.g., as diagnosed by the presence of cell-free donor DNA (cfDNA) in the blood and the results of a biopsy of the donated tissue.

[0048] A subject’s “immunosuppressive state” or “immunocompetence” as used herein refers to the ability of the subject’s immune system to mount an immune response to a pathogen or tissue (e.g., such as a transplanted organ).

[0049] An “immunosuppressive drug”, “immunosuppressant” and the like refer to any drug that reduces the activity, proliferation and/or survival of one or more immune cell types. Such cell types include any T or B lymphocyte populations. A “T-helper cell suppressant” refers to any immunosuppressant that acts on T-helper cells. Examples of T-helper cell suppressants include but are not limited to cyclosporine, tacrolimus, sirolimus, myriocin, mycophenolate, and so forth.

[0050] An “immunosuppressive regimen” involves the administration or prescription of one or more immunosuppressive drugs to a subject. Adjustments to a drug regimen may include adjusting the dose, frequency of administration, level of a drug in the subject’s blood, and/or which drugs are used in the regimen. The immunosuppressive regimen may include steroids and/or thymocyte depleting antibodies in addition to immunosuppressive drugs.

[0051] It will be appreciated that as used herein, the term “immunoglobulin” or “antibody” includes, but is not limited to, any antigen-binding protein product of a vertebrate, e.g. mammalian, immunoglobulin gene complex, including human immunoglobulin isotypes IgA,

IgD, IgM, IgG and IgE. For example, a map of the human immunoglobulin VH locus in the telomeric region of chromosome 14q is described by Cook et al. (1994) *Nature Genetics* 7:162-168; for which the sequence is publicly available at Genbank, for example see Matsuda et al. (1998) *J Exp Med.* 188(11): 2151–2162 and including the D segment locus as disclosed Corbett et al. (1997) *JMB* 270:587-597; for a sequence analysis of the human light chain, see, for example, Kawasaki et al. (1997) *Genome Research* 7:250-261; each of which is herein specifically incorporated by reference for the sequence information provided and referenced in the publication.

[0052] In general, as used herein, an “antibody” (or immunoglobulin) is a protein that includes two molecules, each molecule having two different polypeptides, the shorter of which functions as the light chains of the antibody and the longer of which polypeptides function as the heavy chains of the antibody. Normally, as used herein, an antibody will include at least one variable region from a heavy or light chain. Additionally, the antibody may comprise combinations of variable regions. Through processes of genetic recombination, somatic hypermutation, and junctional changes a very large repertoire of different sequences can be generated encoding the variable regions of these proteins. In addition, isotype switching (also referred to as class switching and class switch recombination (CSR)), occurs after activation of the B-cell and results in a change in the sequence encoding the constant region of the antibody.

[0053] Antibody repertoire analysis may be performed on the total number of individual specificities present in a sample, or may be performed on a selected subset. Repertoire analysis may be performed on a region of an antibody sequence (e.g., heavy chain, light chain, variable region (V), diverse region (D), joining region (J), V/J, VDJ, constant region (C), VJC, VDJC, etc.), or on the entire sequence of the antibody sequence.

[0054] “Primers”, as used herein, are short oligonucleotides that specifically hybridize to a sequence of interest, and which initiate polynucleotide synthesis, for example reverse transcriptase synthesis, synthesis during amplification, synthesis for sequencing purposes, etc. The requirements for primers may vary according to the particular synthetic reaction that is being primed. Primers may be universal, e.g. short highly degenerate primers; oligo-dT primers for reverse transcriptase, etc.; or selective, e.g. binding only to a single V region segment, or a set of V region segments.

[0055] The term “Rapid Amplification of cDNA Ends” (or “RACE”) as used herein refers to the PCR amplification of a cDNA strand from a known sequence to either the 3’ or 5’ end of the cDNA strand.

[0056] To “Specifically hybridize” to a nucleic acid means, with respect to a first nucleic acid, that the first nucleic acid hybridizes to a second nucleic acid with greater affinity than to any other nucleic acid.

[0057] The term “cfDNA” refers to cell-free donor DNA, such as may be found in the blood of a subject with GVHD. A “cfDNA library” refers to a collection of cfDNA sequences.

[0058] The terms “antibody repertoire library” and “immune repertoire library” as used herein refer to a collection of antibody sequences (e.g., RNA or DNA sequences) of fragments thereof. An antibody repertoire library may have 102, 103, 104, 105, 106, 107, 108, 109, 1010, 1011, 1012 or more distinct sequences.

[0059] The term “unique identifier” (or “UID”) as used herein refers to a unique nucleotide sequence on primers involved in reverse transcription, PCR, and/or products thereof. In the subject methods, each UID corresponds to DNA sequences derived from the same IgH RNA molecules. The UID may be any number of nucleotides long sufficient to distinguish the UID from other UIDs. For example, a UID may be anywhere from 4 to 30 nucleotides long, such as 8 to 24, 12 to 20 and so forth.

[0060] The terms “IgH encoding molecule” and “IgH molecule read” refer to the sequence of an immunoglobulin heavy chain (IgH) of an antibody in a subject’s antibody repertoire. Each consensus read represents one IgH encoding molecule, and may comprise multiple IgH molecule reads of the same IgH encoding molecule. IgH molecules of the same “sequence” refer to different consensus reads (reads having different UIDs) with otherwise identical sequences. IgH molecules of the same “lineage” refer to different consensus reads that have similar sequences, as further characterized in the Detailed Description.

[0061] A “consensus sequence” or “IgH sequence” is the sequence of an original IgH RNA molecule as determined by matching IgH molecule reads that share the same unique identifier and have identical or near-identical sequences. The consensus sequence reduces error in the high throughput screens discussed herein.

[0062] An “IgH lineage” is the collection of IgH encoding molecules belonging to IgH sequences that share high identity outside of the UID nucleotide sequence. For example, IgH sequences in an IgH lineage may share a percent identity of more than 99%, 98%, 95%, 90%, 85%, 80%, 75%, 70% and so forth with one or more (or all) of the other IgH sequences of that lineage. In certain embodiments, the percent identity may be shared between V, D, J, or C regions, or a combination thereof. In certain aspects, the IgH sequences in an IgH lineage may comprise one or more different constant regions (may be of one or more different isotypes), or may all have the same constant region and may only differ in mutations of the V, D, and/or J regions. A visual representation of IgH sequences and lineages may be found in Fig. 6B.

[0063] A “Consensus Read Immune repertoire Sequencing” (or “CRIS”) assay is a high-throughput screen that utilizes the UID defined above to identify consensus sequences. The CRIS assay is further characterized in Experiment 1.

[0064] “Abundant related IgH sequences” and “highly expressed sequences” refers to IgH sequences or lineages (depending on the context) having more than a threshold number of IgH

encoding molecules. The threshold may be a specific number (e.g., 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, and so forth) or a relative amount of the total number of IgH molecules (e.g., 0.00001%, 0.0002%, 0.0005%, 0.001%, 0.002%, 0.005%, 0.01%, 0.02%, 0.05%, 0.1%, 0.2%, 0.5%, 1%, and so forth).

[0065] The term “abundant ‘non naïve-IgM’ sequences” (or “ABS”) refers to abundant related IgH sequences that do not encode not naïve-IgM. The term “Abundant Isotype” refers to abundant related IgH sequences encoding an antibody of a particular isotype.

[0066] The terms “relative abundance”, “abundance”, “relative amount” and “ratio”, in the context of amounts of an abundant isotype or abundant ABS, typically refers to the ratio of the isotype or ABS to the total number of abundant related IgH sequences.

[0067] Unless otherwise apparent from the context, all elements, steps or features of the invention can be used in any combination with other elements, steps or features.

[0068] General methods in molecular and cellular biochemistry can be found in such standard textbooks as Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook et al., Harbor Laboratory Press 2001); Short Protocols in Molecular Biology, 4th Ed. (Ausubel et al. eds., John Wiley & Sons 1999); and Cell and Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998).

[0069] It will be appreciated by those of skill in the art that, in light of the present disclosure, numerous modifications and changes can be made in the particular embodiments exemplified without departing from the intended scope of the invention. The subject methods are used for diagnostic and prognostic purposes, and find use in guiding prophylactic or therapeutic methods. As used herein, the term “treating” is used to refer to both prevention of conditions, prevention of relapses, and treatment of pre-existing conditions. For example, the prevention of autoimmune disease may be accomplished by administration of the agent prior to development of a relapse.

Methods

Producing Antibody Repertoire Library and Sequencing

[0070] The methods of the invention provide a means for analyzing an antibody repertoire of a subject to determine the immunocompetence of the subject. The method includes analyzing the subject’s antibody repertoire, which may include identifying the composition of abundant related IgH sequences of the subject’s antibody repertoire as well as the isotype abundance. The method may further comprise providing an assessment of the immunocompetence of the subject, based on the analysis.

[0071] The cells in a sample for analysis may have been separated or enriched prior to analysis, or a sample, e.g. a clinical sample, may be analyzed in the absence of any

enrichment. To obtain the sequence information, the cells present in the sample are lysed and nucleic acids of interest (e.g., genomic DNA, mRNA, cDNA, cRNA, etc.) are collected. Where mRNA is being analyzed, it will generally be converted to cDNA by reverse transcription. Primers for cDNA synthesis, as described above, may be selective for the specific IgH regions of interest. The immune receptor sequences are then amplified with a set of primers selective for specific IgH regions.

[0072] To this end, blood samples are drawn from patients. Either RNA is isolated from cells extracted from the blood or directly from the blood. The antibody RNA transcripts expressed by B cells are amplified using PCR and a molecular biology protocol that allows for the bias and error correction in the downstream analysis. The amplified DNA is sequenced using a high throughput sequencer. This enables the analysis of the amounts of antibody RNA of different isotypes in the sample. The antibody transcript composition further contains features like clonality and mutation level that can provide additional information.

[0073] The amplified pool (or, in some cases, a pool that has not been amplified) of nucleic acids may then be subjected to high throughput sequencing (e.g., massively-parallel sequencing). Sequencing platforms that can be used in the present disclosure include but are not limited to: pyrosequencing, sequencing-by-synthesis, single-molecule sequencing, second-generation sequencing, nanopore sequencing, sequencing by ligation, or sequencing by hybridization. Preferred sequencing platforms are those commercially available from Illumina (RNA-Seq) and Helicos (Digital Gene Expression or "DGE"). "Next generation" sequencing methods include, but are not limited to those commercialized by: 1) 454/Roche Lifesciences including but not limited to the methods and apparatus described in Margulies et al., *Nature* (2005) 437:376-380 (2005); and US Patent Nos. 7,244,559; 7,335,762; 7,211,390; 7,244,567; 7,264,929; 7,323,305; 2) Helicos BioSciences Corporation (Cambridge, MA) as described in U.S. application Ser. No. 11/167046, and US Patent Nos. 7,501,245; 7,491,498; 7,276,720; and in U.S. Patent Application Publication Nos. US20090061439; US20080087826; US20060286566; US20060024711; US20060024678; US20080213770; and US20080103058; 3) Applied Biosystems (e.g. SOLiD sequencing); 4) Dover Systems (e.g., Polonator G.007 sequencing); 5) Illumina as described US Patent Nos. 5,750,341; 6,306,597; and 5,969,119; and 6) Pacific Biosciences as described in US Patent Nos. 7,462,452; 7,476,504; 7,405,281; 7,170,050; 7,462,468; 7,476,503; 7,315,019; 7,302,146; 7,313,308; and US Application Publication Nos. US20090029385; US20090068655; US20090024331; and US20080206764. All references are herein incorporated by reference. Such methods and apparatuses are provided here by way of example and are not intended to be limiting.

[0074] Sequence reads may be combined into full-length IgH sequences (IgH molecules) by aligning sequences based on identity of the V, D, and/or J regions (e.g., as seen in Example 1).

Alignment may be pairwise alignment (e.g. a quality-score-weighted Smith-Waterman algorithm), or any suitable algorithm.

[0075] Bias and error in amplification and/or sequencing may be reduced by identification of consensus sequences (i.e., IgH sequence). In certain aspects, IgH molecules sharing a unique identification nucleotide sequence (UID) may be identified (e.g. classified) as belonging to the same consensus sequence. Consensus sequences may be used to average out error from the amplification and/or sequencing steps (e.g. as validated in Fig. 7B-C).

Characterization of Abundant related IgH sequences

[0076] IgH molecules may be classified as belonging to the same consensus sequence (IgH sequence) or IgH lineage as described above and, for example, as seen in Fig. 6B. Abundant related IgH sequences may be classified as IgH sequences or IgH lineages having more than a threshold amount of IgH encoding molecules, as discussed above. In certain embodiments, the number of reads may be downsampled, e.g., to decrease computing time.

[0077] The composition of the abundant related IgH sequences may be the relative amounts of one or more antibody isotypes of the total abundant related IgH sequences.

Diagnostics and prognostics

[0078] The invention finds use in the prevention, treatment, detection, diagnosis, prognosis, or research into any condition or symptom of any condition relating to immune state, including cancer, inflammatory diseases, autoimmune diseases, allergies and infections of an organism. The organism may be a human subject or any non-human mammals. Examples of non-human mammals include, but are not limited to, non-human primates (e.g., apes, monkeys, gorillas), rodents (e.g., mice, rats), cows, pigs, sheep, horses, dogs, cats, or rabbits.

[0079] In some aspects, the sample may be taken before or after an event, such as an organ transplant, vaccination, or beginning of an immunosuppression regimen. Multiple samples may be taken at different time points.

[0080] In certain embodiments, the method includes providing an assessment of the immunocompetence of the subject. For example, providing the assessment may include providing an assessment that the subject is not adequately immunosuppressed when the relative amount of IgG, IgA, IgD, IGE, mutated IgM or a combination thereof (e.g. all non-naïve IgM) is above a certain threshold (e.g., the threshold may be 0.1%, 0.2%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10% or more of the total number of abundant similar IgH sequences; or the threshold may be two, three, four, five or more times the relative amount measured at a previous time point; etc...), or when the relative amount of naïve-IgM is below a certain threshold (e.g., the threshold may be less than: 99.9%, 99.8%, 99.5%, 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90% of the total number of abundant similar IgH sequences).

Alternatively or in addition, the step of providing the assessment may include providing an assessment that the subject is overly immunosuppressed when the relative amount of IgG, IgA, IgD, IGE, mutated IgM or a combination thereof (e.g. all non-naïve IgM) is below a certain threshold (e.g., the threshold may be 0.1%, 0.2%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10% or more of the total number of abundant similar IgH sequences; or the threshold two, three, four, five or more times the relative amount measured at a previous time point; etc...) or when the relative amount of naïve-IgM is above a certain threshold (e.g., the threshold may be more than: 99.9%, 99.8%, 99.5%, 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90% or more of the total number of abundant similar IgH sequences). In some aspects, the assessment may be a prediction of the likelihood of allograft rejection. Fig. 10, 13 and 14 provide data on how isotype composition of the abundant related IgH sequences correlates with increased risk of allograft rejection. In certain aspects, the assessment may further include an assessment of the level of donor-derived cfDNA in the subject's blood, a biopsy of a transplanted organ, or a combination thereof.

[0081] Methods are also provided for modifying an immunosuppression regimen, by analyzing the immune repertoire in a sample, and based on that information, selecting the appropriate therapy, dose, treatment modality, etc. that is optimal for stimulating or suppressing a targeted immune response, while avoiding unnecessary toxicity or overly compromising the immune system. The treatment is optimized by selection for a treatment that minimizes undesirable toxicity, while providing for effective activity.

Computer Based Embodiments

[0082] As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the information of the present invention. The minimum hardware of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention. The data storage means may comprise any manufacture comprising a recording of the present information as described above, or a memory access means that can access such a manufacture.

[0083] Software products (or components) may be tangibly embodied in a machine-readable medium, and comprise instructions operable to cause one or more data processing apparatus to perform any of the sequencing, analysis, and assessment steps of the above methods.

[0084] In some examples, a software product (or component) includes instructions for performing any of the above method steps, in particular the steps of sequencing, analyzing, and providing an assessment.

Reagents and Kits

[0085] Also provided are reagents and kits thereof for practicing one or more of the above-described methods. The subject reagents and kits thereof may vary greatly. Reagents of interest include reagents specifically designed for use in production of the above described immune repertoire analysis. For example, reagents can include primer sets for cDNA synthesis, for PCR amplification and/or for high throughput sequencing of sequences encoding IgH or segments thereof. Examples of primers can be found in Fig. 8.

[0086] The kits can further include a software package for statistical analysis, and may include a reference database for calculating the probability of a match between two repertoires. The kit may include reagents employed in the various methods, such as primers for generating target nucleic acids, dNTPs and/or rNTPs, which may be either premixed or separate, one or more uniquely labeled dNTPs and/or rNTPs, such as biotinylated or Cy3 or Cy5 tagged dNTPs, gold or silver particles with different scattering spectra, or other post synthesis labeling reagent, such as chemically active derivatives of fluorescent dyes, enzymes, such as reverse transcriptases, DNA polymerases, RNA polymerases, and the like, various buffer mediums, e.g. hybridization and washing buffers, prefabricated probe arrays, labeled probe purification reagents and components, like spin columns, etc., signal generation and detection reagents, e.g. streptavidin-alkaline phosphatase conjugate, chemifluorescent or chemiluminescent substrate, and the like.

[0087] In addition to the above components, the subject kits may further include instructions for practicing the subject methods. These instructions may be present in the subject kits in a variety of forms, one or more of which may be present in the kit. One form in which these instructions may be present is as printed information on a suitable medium or substrate, e.g., a piece or pieces of paper on which the information is printed, in the packaging of the kit, in a package insert, etc. Yet another means would be a computer readable medium, e.g., diskette, CD, etc., on which the information has been recorded. Yet another means that may be present is a website address which may be used via the internet to access the information at a removed, site. Any convenient means may be present in the kits.

Utility

[0088] Here, we present methods, kits and computer related embodiments for use in comprehensively monitoring the status of the recipient immune system. The methods herein allow for the investigation of the IgH antibody repertoire in a highly accurate, quantitative, and reproducible manner. A consensus read method allows us to determine the abundance, sequence, and isotype of IgH mRNA. Using this approach, we have shown that we can identify abundant sequences in a blood sample. These abundant sequences are enriched for IgG, IgA, and mutated (non-naïve) IgM, and are shown herein to correspond to the activated B-cell

compartment. The identification of B-cell compartments (naive vs. memory/plasmablasts) and isotype distributions of IgH sequences directly from PBMCs obviates the distortion and additional work associated with cell sorting procedures, and the need for more invasive procedures to establish the immunocompetence of a subject.

[0089] This assay can be applied to monitor the immune system's activity following solid organ transplant of any transplant type and following bone marrow transplants. The methods are also applicable in a wider range of applications where predictions of the state or competence of the immune system are important, e.g. in the elderly, in newborns, in patients suffering from autoimmune diseases treated with drugs with immune suppressing effects, in cancer patients.

[0090] The goal of post-transplantation therapy is to establish a delicate balance between the prevention of graft rejection via immune suppression and the protection of the transplant patient against complications related to over-immunosuppression, in particular infections and malignancies. Given that the available window in transplantation therapy is narrow, and given that significant inter-individual variability exists in the sensitivity to immunosuppressants, there is a strong need for informative measurements of the state of the immune system. While immunosuppressant drugs like Tacrolimus primarily target T cells, we assumed the B cells would be targeted indirectly as well. Here, we have implemented a barcode-based immune-repertoire sequencing method to assay the temporal response of the B cell immune repertoire to immunosuppression. We found that immune repertoire sequencing enables the measurement of the extent of immunosuppression as well as the diagnosis of organ rejection events manifesting in increased immune activity.

[0091] A current gold standard in post-transplant care is the grading of biopsies by pathologist. This assay has high inter-operator variability and can only diagnose acute reaction events. The results of our assay are independent of the operator, are able to predict rejection events months in advance and have high positive predictive value. The methods herein can be combined with assays such as measuring donor DNA in the blood to diagnose risk of organ rejection. However, in comparison to assaying donor DNA in the blood, which only enters the blood when the organ is already being rejected, the methods disclosed herein are able to predict rejection and allow pharmacological intervention before the organ is significantly damaged.

[0092] Through measurement of the amount and properties of antibody transcript sequences in blood samples, we aim to determine the composition of B-cell populations in heart transplant recipients and to therefore quantify their net state of immunosuppression.

[0093] By quantifying antibody sequences expressed by activated B cells (ABS) we leverage a cumulative readout conveniently provided by the immune system. The activation of B cells is a process taking several days and requiring activated T cells, which are the targets of the most commonly used immunosuppressive drugs used for maintenance after organ transplantation,

Cyclosporin and Tacrolimus. As such, the ABS levels we quantify provide an easy to interpret measure on the activity of the immune system over the last few days. This enabled us to measure the level of immunosuppression and we provide proof of concept for detecting organ rejection events. We believe that future, larger studies will show that the combination of immune repertoire and cfDNA measurements will make it possible to replace organ biopsies as the gold standard for post-operative organ health monitoring.

[0094] Going forward, immune repertoire sequencing might be used to also improve the management of graft versus host disease after bone marrow transplant or the management of auto-immune diseases.

[0095] The immune competence assay can easily be modified in the several ways to adapt the subject methods to a number of applications and constraints. First, the preparation of the sample can be modified. In the examples provided herein, reverse transcription and PCR is used to amplify the antibody RNA transcripts. This could easily be modified by incorporating methods such as "Rapid Amplification of cDNA Ends" (or "RACE"). Primers for reverse transcription and PCR can be located on different parts of the transcripts. The high throughput sequencing platform used in the examples provided herein are Illumina platforms, but other sequencers like the Ion Torrent platforms may be used. Also, digital or quantitative qPCR might also be used to analyze isotype distribution in a sample. Second, the analysis of the data can be modified. The most promising measure of immune competence is the distribution of isotypes (IgM, IgD, IgG, IgA, and IgE). Other measurements might include analysis of the structure of the repertoire, e.g. distribution of clone sizes and mutation levels.

EXAMPLES

The following examples are offered by way of illustration and not by way of limitation.

Example 1

[0096] Here, we present a consensus read method (i.e., Consensus Read Immune repertoire Sequencing, or CRIS assay) which allows us to determine the abundance, sequence, and isotype of IgH mRNA. Using this approach, we shown that we can identify all abundant sequences in a blood sample. These abundant sequences are enriched for IgG, IgA, and mutated IgM, and they likely correspond to the activated B-cell compartment.

[0097] While the methodology and results of this example relate directly to vaccine recipients, one of skill in the art could easily apply the methods in this example to assess the immune state of other subjects, such as allograft recipients.

Methods

[0098] *Study Volunteers, PBMC Isolation, and RNA Extraction.* All study protocols were approved by the Institutional Review Boards at Stanford University. Informed consent was obtained from participants. Blood was taken before (day 0), 1 wk after (day 7), and 1 mo after (day 28) vaccination in the first year (2011) and after vaccination (day 7) in the second year (2012). Volunteers were 28 young adults (18–30 y) in generally good health and vaccinated with one dose of either seasonal LAIV (FluMist intranasal vaccine; MedImmune) or TIV (Fluzone; Sanofi-Pasteur); 60 mL peripheral blood were taken from individuals and heparinized. PBMCs were extracted from 10 mL blood using a Ficoll-Gradient and frozen in 10%(vol/vol) DMSO/40%(vol/vol) FBS according to the Protocols of the Stanford Human Immune Monitoring Center (HIMC). After thawing, total RNA was extracted from 5 million PBMCs using the Qiagen AllPrep Kit.

[0099] *Library Preparation.* Five hundred nanograms total RNA were used as input for library preparation. Reverse transcription was performed according to the manufacturer's instructions using SuperScript III Enzyme (Life) and primers for all five isotypes containing 8 random nt and partial Illumina adapters containing Illumina barcodes (Fig. 8). Second-strand synthesis was done using Phusion Polymerase (NEB) and primers containing 8 random nt and partial Illumina adapter sequences covering all V segments with a maximum of one mismatch (Fig. 8) (98 °C for 2 min, 52 °C for 2 min, and 72 °C for 10 min). Double-stranded cDNA was purified two times using Ampure XP beads at a ratio of 1:1. Double-stranded cDNA was amplified with Platinum Hifi Polymerase (Life) with two primers completing Illumina adapter sequences (95 °C for 2 min, 27 cycles of 95°C for 30 s, 65 °C for 30s, and 68 °C for 2 min, and 68 °C for 7 min). Final sequencing libraries were generated by purifying the PCR product using Ampure XP beads at a ratio of 1:1.

[00100] *Sequencing and Data Analysis.* Libraries were sequenced on the Illumina HiSeq2000 using a custom 100 × 120-bp protocol. The sequencing runs yielded 3 and 15 million raw reads per sample. Raw reads were split into UID groups with unique 16-bp identifiers, and a consensus read was built for each UID group with at least two raw reads.

[00101] For the generation of consensus reads, raw bases with an Illumina Quality Score under 20 were discarded. Each base in the consensus read was determined by a majority call of raw bases, which was weighted by quality scores. UID groups containing any majority base making up <66% of all bases at that position were discarded. This requirement avoided draws and discarded UID groups with only two raw reads if they were not completely identical. Final quality filtering discarded low-quality consensus reads. Alignments to variable (V), diversity (D), and joining (J) segments as well as mutation analysis were then performed by the pipeline described in ref. 23 using IMGT references for V, D, J, and C segments. Furthermore, bases

templated by V-segment primers as well as most of the constant region were truncated. This process generated ~50,000–500,000 high-quality consensus reads per sample. Each consensus read represents one initial IgH molecule. Identical IgH molecules were grouped into IgH sequences, which were then clustered into IgH lineages as described below (Fig. 6B). Abundance was then defined as the number of IgH molecules that an IgH sequence or IgH lineage contained. IgH sequences or lineages were defined as abundant if they contained at least 5 IgH molecules.

[00102] *Lineage Clustering.* IgH sequences were clustered into IgH lineages according to similarity in their junctional region. Lineages were created according to the following steps.

[00103] A lineage is formed and populated with one IgH sequence (seed). Then, all IgH sequences in the lineages (initially only the seed) are compared with all other IgH sequences of the same length using the same V and J segments. If their junctional regions (untemplated nucleotides and D segments) are at least 90% identical, the IgH sequence is added to the lineage. This process is repeated until the lineage does not grow.

[00104] *Subsampling.* For some figures, data were subsampled randomly to 50,000 IgH molecules/sample. These 50,000 IgH molecules were then grouped into IgH sequences and clustered into IgH lineages as for the nonsubsampled data.

Results

[00105] *Labeling of RNA Molecules with Random Nucleotide Unique Identifiers.* The sequencing approach that we used relied on labeling each RNA molecule during cDNA synthesis and preserving this nucleotide label throughout PCR amplification. Using these labels, we could identify group reads originating from the same RNA molecule. Therefore, both isotype- and V segment-specific primers were designed to include a stretch of 8 random nt followed by a partial paired-end adapter sequence on their 5' end (Fig. 6). We used total RNA extracted from the B cell containing peripheral blood mononuclear cells (PBMCs) as input; reverse transcription and subsequent primer extension resulted in a pool of double-stranded cDNA, in which initial IgH RNA molecules were labeled with a 16-nt unique identifier (UID). This pool was then amplified in a PCR where the primers completed the sequencer adapter sequences (Fig. 6). The resulting libraries were multiplexed and sequenced on the Illumina HiSeq2000 Sequencer. We used a paired-end sequencing protocol, sequencing 100 bp on the first read and 120 bp on the second read to cover the whole amplicon. The sequencing reads included the 16-nt UID, making it possible to retrieve the unique labels of the initial IgH RNA molecules. The high raw read number enabled us to sequence uniquely labeled IgH RNA molecules multiple times and build a highly accurate consensus before downstream analysis (Fig. 7A).

[00106] IgH molecules with the same nucleotide sequence were grouped into IgH sequences. The abundance of an IgH sequence is defined by the number of unique UIDs that share the

same consensus sequence. Biologically, the abundance of an IgH sequence is determined by the number B cells that express this sequence convolved with the distribution of expression levels across all these cells and represents a proxy for overall antibody abundance.

[00107] *Characterization of the Consensus Read Approach.* To validate the sequencing protocol, a blood sample was collected from a young adult individual (L1) taken 7 d after vaccination with the LAIV vaccine. A sequencing library (L1 D7) (Fig. 1A) was prepared and then sequenced. We determined the reproducibility of this IgH sequence abundance information by preparing a second sequencing library (Lib. Rep.) (Fig. 1A) from a separate aliquot of the same RNA extraction and sequenced it on an independent run. We leveraged the exceptional accuracy of our approach to find IgH sequences shared between the sequenced libraries (L1 D7 and Lib. Rep.) and compare their abundances (Fig. 1 B and C); 36,308 of 66,307 IgH sequences were shared between L1 D7 and Lib. Rep. and showed very strong similarity in their abundance ($R^2 = 0.96$) (Fig. 1C), illustrating the robustness of the approach. The overlap between the libraries was almost complete when we focused on abundant sequences, defined to be those sequences with more than 5 IgH molecules per IgH sequence (abundance ≥ 5); 5,351 of 5,637 abundant sequences in L1 D7 were shared by Lib. Rep. (Fig. 1B).

[00108] The lower rate of sharing among sequences that were not abundant was caused by either subsampling from the total RNA pool during library construction or insufficient raw read coverage during the sequencing reaction and data analysis quality filtering. To identify the main cause, we resequenced the initial sequencing library (Seq. Rep.). We then visualized the rate at which sequences were shared between these samples as a function of their abundances (Fig. 1 D–H). The rates at which sequences were shared between L1 day 7 and either Lib. Rep. or Seq. Rep. were very similar and clearly dependent on sequence abundance, ranging from ~50% for sequences with abundance of one to almost 100% for sequences with an abundance more than five (Fig. 1 E and F). The fact that the construction of a separate library (Lib. Seq.) had only a small effect on the rate of shared sequences implicated raw read coverage and quality filtering, but not library construction, as the main cause for the reduced sharing rates among low-abundance sequences. IgH sequences represented by fewer IgH molecules have a higher chance that none of their molecules meet the raw read coverage or quality thresholds and are consequently dropped completely.

[00109] The resequencing of the initial library (Seq. Rep.) also allowed us to measure the sequencing accuracy of our approach. The unique labeling of molecules enabled us to use discrepancies between IgH molecules with barcodes that appeared in both L1 D7 and Seq. Rep. to determine the sequencing error of consensus reads to be ~1/50,000 errors per 1 bp or Q47 (Fig. 8B). We used discrepancies between raw reads in UID groups in L1 D7 to determine the raw Illumina sequencing error to be ~1/500 errors per 1 bp or Q27. This 100-fold higher raw

error rate would, in the absence of barcodes, dramatically inflate the apparent number of unique IgH sequences (Fig. 8C), thereby illustrating the problem of determining the complexity of the antibody repertoire using a large number of error-prone raw sequencing reads.

[00110] These experiments showed that the consensus read approach that we developed is highly accurate and quantitative and detects practically all abundant sequences in a PBMC sample.

[00111] *Identifying the Activated B-Cell Compartment.* To determine how representative a PBMC sample is of the complete peripheral blood and whether the abundance of an IgH sequence might help us identify sequences expressed by memory B cells and plasmablasts, we generated an additional sequencing library from a PBMC sample purified from blood taken at the same visit as L1 D7 (Bio. Rep.) (Fig. 1A).

[00112] Recapturing an IgH sequence in two separate PBMC aliquots requires at least two cells of the same B-cell clone expressing this IgH sequence to exist and partition into the separate aliquots. Because activated B cells have undergone clonal expansion, we expected the sequences shared between L1 D7 and Bio. Rep. to be enriched for sequences expressed by memory B cells and plasmablasts, which are known to make up ~30% and 2% of peripheral B cells, respectively (15).

[00113] The Bio. Rep. sample shared 1,401 of 5,637 (25%) of its abundant sequences (>5 IgH molecules/IgH sequence) but only 974 of 66,307 (1.5%) low-abundance sequences (<5 IgH molecules/IgH sequence) with L1 D7.

[00114] We again visualized the rate at which sequences were shared between L1 D7 and Bio. Rep. as a function of their abundances. This visualization showed that the sequences shared between different PBMC samples were heavily biased to more abundant sequences (Fig. 1G). This bias to more abundant sequences was much more pronounced than in the Lib. Rep. sample, which was constructed from the same PBMC RNA as L1 D7.

[00115] Abundant sequences also showed a distinct isotype distribution. Although low-abundance sequences were mostly IgM, abundant sequences were mostly IgG and IgA (Fig. 1H). Furthermore, although abundance did not correlate with the mutation rate of class-switched IgG and IgA sequences, the mutation rate of IgM increased in abundant sequences to levels only slightly lower than IgG and IgA mutation rates (Fig. 1I).

[00116] The enrichment for IgG and IgA as well as mutated IgM in abundant sequences together with the heavily biased recapture by the Bio. Rep. sample indicated strongly that the abundant IgH sequences are heavily enriched for sequences expressed by activated B cells. Together, this data shows that we could use isotype and abundance information to identify activated B-cell sequences.

[00117] *Identifying Vaccine-Induced Lineages.* Next, we wanted to investigate whether the analysis of the abundant lineages (at least 5 IgH molecules/IgH lineage) representing the

activated B-cell compartment might allow us to identify lineages activated by seasonal influenza vaccines. We analyzed the IgH repertoire of 14 LAIV and 14 TIV recipients at three time points. Blood was collected at day 0 (right before vaccination) and days 7 and 28 after vaccination. Day 7 is known to show the most pronounced vaccine response; 84 resulting samples were prepared and analyzed as described above. To directly compare vaccine effects on the number of abundant lineages between samples, we limited the analysis to 50,000 randomly selected IgH molecules.

[00118] TIV is known to trigger a strong hemagglutinin inhibition (HAI) antibody response, IgG plasmablast release, and isotype class switching, whereas the effects of LAIV are less pronounced. We showed that TIV vaccination indeed strongly increases numbers of abundant IgG lineages (Fig. 3A). LAIV vaccination caused a much weaker response and on average, did not increase the numbers abundant lineages of IgG, IgA, and IgM isotypes (Fig. 3A).

[00119] To visualize these effects in more detail, we ordered the IgH lineages of one TIV recipient (T4) by rank and then plotted abundance (height) and isotype (color). Distinct patterns of vaccine response become clear in the day 7 sample, where several highly abundant IgG lineages and hundreds of moderately abundant IgG lineages dominate the TIV response (Fig. 3B). A closer look at the most abundant lineages reveals that they can consist of thousands of IgH molecules grouped into hundreds IgH sequences often from more than one isotype. Overall, focusing on abundant lineages revealed isotype distributions distinct to TIV, with the highly abundant IgG lineages observed at day 7 likely corresponding to memory-derived IgG plasmablasts known to be released into the periphery at that time.

[00120] *Recall Response to TIV Vaccination.* We went on to analyze the recall response to repeated TIV vaccination. Because of its more pronounced effect on individuals (Fig. 3A), we chose to investigate the recall response to TIV. The individuals T1–T5 were vaccinated again with TIV in the year after their initial vaccination. The formulation of TIV was changed for the influenza A (H3N2) and influenza B strains, with only the A(H1N1)pdm09 strain remaining unchanged. Blood samples were again collected at days 0, 7, and 28 (Y2 D0, Y2 D7, and Y2 D28) relative to the second year vaccination.

[00121] In all individuals, any pair of samples shared hundreds of lineages. Overall, the highest number of lineages was shared between samples taken in the same year, probably because of lineages that persist in the periphery during the duration of the blood draws. Fewer lineages were shared between samples taken in different years (Fig. 4A). When normalized to the overall number of lineages, the Y1 D7/Y2 D7 pair of samples showed the highest rate of shared lineages of IgG, IgA, and IgM isotypes (Fig. 4A). Additionally, compared with Y1 D0/Y2 D0 and Y1 D28/Y2 D28 pairs, the shared IgG lineages but not the shared IgA and IgM lineages were more abundant in all individuals (Fig. 4B). This finding highlighted two interesting findings. First, in the Y1 D0/Y2 D0 prevaccine sample, pairs were collected 12mo apart but still shared IgH

lineages, suggesting that the B-cell clones expressing these lineages either are reactivated by recurrent infection unrelated to influenza or persisted unchanged in the periphery for up to 12 mo in the absence of antigen, which would be consistent with the recent finding that germinal center reactions can persist for over 8 mo in mice. Second, highly abundant IgG lineages are preferably shared in the Y1 D7/Y2 D7 post-vaccine sample pairs, likely representing vaccine-induced recalled lineages.

[00122] We visualized the shared lineages between time points in the individual with the strongest putative recall response (Fig. 5 A and B). These connection plots clearly show a surplus in shared highly abundant IgG lineages in the Y1 D7/Y2 D7 pair. This trend is also visible in the other individuals, albeit to a lesser extent. If these shared IgG lineages had arisen independently in the 2 y sampled, one would expect IgG lineages to also be shared between individuals, assuming that different individuals have the same chance of generating the same antibody. The fact that IgG lineages show practically no overlap between individuals (Fig. 5C) suggests that our data provide direct genetic evidence for memory B-cell recall.

Example 2

[00123] Methods for library preparation, sequencing, and data analysis were similar to those described for Example 1 above. Using the Consensus Read Immune repertoire Sequencing (CRIS) assay described in Example 1, that accurately determines the sequences, mutational load, isotypes and expression levels of antibody mRNA expressed by hundreds of thousands of B-cells in a single blood sample, we tested the assumption that immunosuppression reduces the number of highly expressed and class-switched antibody sequences.

[00124] Acute rejection of the transplanted organ can result from insufficient immunosuppression. In preliminary studies, two acute rejection events were detected via endomyocardial biopsy. Concurrently-drawn blood samples from these patients demonstrated elevated levels of highly expressed IgA sequences (**Fig. 9A**). In one case, IgA levels were elevated 6 months prior to the rejection event, suggesting that it may be possible to predict impending rejection and therefore modify immunosuppression in order to abort the event and prevent graft damage. Preliminary data generated from blood samples of heart transplant recipients shows that the level of highly abundant IgA sequences drops dramatically after the introduction of immunosuppressive drugs (**Fig. 9A and 9B**).

Example 3

[00125] Immunosuppression in adult transplant patients is induced using a combination of steroids, thymocyte depleting antibodies and T cell inhibitors like Tacrolimus. Tacrolimus is then used further to maintain the immune suppression. While the Tacrolimus levels in the blood of

organ recipients are tightly controlled, measuring the level of immunosuppression directly would be strongly beneficial as even at same levels, as drug effectiveness differs between individuals.

[00126] We aimed to quantify the level of immunosuppression by leveraging the unique features of the antibody heavy chain gene expressed by B cells by using immune repertoire sequencing. We assumed that B cell activation would be hampered by post-transplant immunosuppression as Tacrolimus inhibits T cell activation. Inhibiting T helper and T killer cells minimizes organ rejection but also reduces T helper cell dependent B cell activation. Therefore Tacrolimus indirectly affects the composition of the B cell compartment by inhibiting B cell activation. Depending on the level of immunosuppression, the B cell compartment in the peripheral blood would therefore be composed of different amounts of naïve B cells, memory B cells, and plasmablast.

[00127] By using the CRIS assay (similar as described above), we can determine the composition the B cell compartment by measuring the sequence and abundance of antibody heavy chain transcripts in a blood sample with unprecedented accuracy and throughput. The antibody heavy chain gene is unique as its expression changes fundamentally not only in abundance but also in sequence when a B-cell is activated.

[00128] B cells can undergo hypermutation and class-switch recombination when they are activated. Activated B cells like plasmablasts express high levels of mutated antibodies of IgG and IgA isotypes while naïve B cells express non mutated IgM antibodies at low levels (Fig. 11). A lower number of abundant class-switched sequences would indicate a lower number of activated B cells and consequently a higher level of immunosuppression.

[00129] We tested this approach on a cohort of 13 patients that had received heart transplants and found that by using immune repertoire sequencing, we were able to measure immune suppression and detect organ rejection events as determined by cfDNA levels or Biopsies.

Methods

[00130] *Details of the Post-transplant Therapeutic Protocol* The 13 recipients of donor hearts transplants received post-transplant immunosuppression. Post-transplant immunosuppression consisted of methylprednisolone 500 mg administered immediately postoperatively followed by 125 mg every 8 hr for three doses. Antithymocyte globulin (rATG) 1 mg/kg was administered on postoperative days 1, 2, and 3.

[00131] Maintenance immunosuppression consisted of prednisone 20 mg twice daily starting on postoperative day 1 and tapered to <0.1 mg/kg/day by the 6th postoperative month and tapered further if endomyocardial biopsies showed no evidence of cellular rejection. Tacrolimus was started on postoperative day 1 and dosing was further adjusted to maintain a level of 10–15 ng/ml during months 0–6, 7–10 ng/ml during months 6–12, and 5–10 ng/ml thereafter.

Mycophenolatemofetil was started at 1,000 mg twice daily on postoperative day 1 and dose adjustments were made, if required, in response to leukopenia.

[00132] One of the 13 recipients was excluded from analysis because the recipient suffered several persistent infections and underwent a modification of the standard immunosuppression protocol as a result.

[00133] All heart transplant recipients were monitored for acute cellular rejection by surveillance endomyocardial biopsies performed at scheduled intervals after transplant: weekly during the first month, biweekly until the 3rd month, monthly until the 6th month, and then at months 9, 12, 16, 20, and 24. Biopsies were graded according to the ISHLT 2004 revised grading scale (0, 1R, 2R, 3R). If blood sampling and endomyocardial biopsies were performed on the same day, care was taken to ensure that blood was collected prior to the biopsy procedure.

[00134] *Whole Blood Processing and Nucleic Acid Extraction.* Plasma and buffy coats were extracted from whole-blood samples by sequential spins within three hours of sample collection and stored at -80 C. When required for analysis, plasma samples were thawed and circulating DNA was immediately extracted from 0.5–1 ml plasma using the QIAamp Circulating Nucleic Acid Kit (Qiagen).

RNA was extracted from Buffy coat samples using TRizol reagent (Life) followed by RNeasy column (Qiagen) extraction.

[00135] *cfDNA Library Preparation.* Sequencing libraries were prepared from the purified patient plasma DNA using the NEBNext DNA Library Prep Master Mix Set for Illumina with standard Illumina indexed adapters, or using a microfluidics-based automated library preparation platform (Mondrian SP Ovation SP Ultralow library system). Libraries were characterized using the Agilent 2100 Bioanalyzer (high-sensitivity DNA kit) and quantified by qPCR.

[00136] *Immune Repertoire Library. Preparation* 0.5-1 micrograms total RNA were used as input for library preparation. Reverse transcription was using SuperScript III Enzyme (Life) and primers for all five isotypes containing 8 random nt and partial Illumina adapters containing Illumina barcodes. Second-strand synthesis was done using Phusion Polymerase (NEB) . Double-stranded cDNA was purified two times using Ampure XP beads at a ratio of 1:1. Double-stranded cDNA was amplified with Platinum Hifi Polymerase (Life). Final sequencing libraries were generated by purifying the PCR product using Ampure XP beads at a ratio of 1:1.

[00137] *Sequencing and Data Analysis.* Immune repertoire libraries were sequenced on the Illumina HiSeq2000 using a custom 2x100bp protocol and custom dual barcode strategy to pool ~50 libraries per lane. The sequencing runs yielded between 1 and 10 million raw reads per sample. Raw reads were split into UID groups with unique 16-bp identifiers, and a consensus read was built for each UID group with at least two raw reads. Alignments to variable (V),

joining (J) segments as well as constant regions were done using BLAST. Quantification was done using custom scripts available upon request.

Results

[00138] We collected 156 blood samples longitudinally from 13 heart transplant recipients for this proof-of-concept study. The recipients were a subsample from a previous study and were selected to contain both non-rejecting and rejecting individuals. Starting on the day after transplantation, samples were collected following organ transplants at regular time points with decreasing frequency as the study progressed. Concurrently with the immune repertoire measurement, transplant organ health was assessed using cfDNA levels and endomyocardial Biopsies (Fig. 11). We used DNA extracted from the plasma fraction of whole blood samples for the cfDNA measurement and RNA extracted from the buffy coat for the immune repertoire sequencing protocol. The extracted RNA was barcoded and amplified as described previously and sequenced on the HiSeq2000 or MiSeq (2x100) and abundance, mutation level and isotype were determined for each antibody transcript.

[00139] *Immunosuppression causes shift in isotype expression.* First we investigated whether we could detect immunosuppression using immune repertoire sequencing of the antibody heavy chain transcript. We separated antibody transcripts by abundance where sequences represented by only one molecule (singletons) are likely expressed by naïve or inactive B cells and abundant sequences represented by more molecules likely stem from activated B cells like plasmablasts. We compared data pooled from all study participants at the beginning of (day 1) and six months into (day 180) immunosuppressive therapy. While 16% of sequences were abundant at day 1, 12% of sequences were abundant at day 180. A much larger change was observable in the composition of abundant sequences. 37.5% of abundant sequences were naïve IgM sequences at day 1, increasing to 89.2% at day 180. Naive IgM sequences therefore replaced all other isotypes (including mutated IgM and IgD) after the induction of immunosuppression (Fig. 12A). These abundant 'non naive-IgM' sequences (ABS) absent at day 180 were likely expressed by T cell activated B cells, which number was reduced by immunosuppression.

[00140] Consequently, we utilized the relative level of ABS as a measure of immunosuppression. Plotting the relative abundance of ABS as a function of time after organ transplant (Fig. 12C) we found that the decrease in ABS occurred rapidly in the first month followed by a slow decline over the following months and reaching their lowest levels of about 15% of the original level after 10 month. Further, plotting of the ABS levels as a function of Tacrolimus levels we showed that ABS levels are negatively correlated to the levels of Tacrolimus. This demonstrates that immune repertoire sequencing can be used to gauge the level of immunosuppression.

[00141] *Elevated ABS levels indicate rejection events.* Next, we wanted to evaluate the use of ABS levels to diagnose organ rejection events which occur when the level of immunosuppression is insufficient. We plotted ABS levels in 4 individuals, 2 of which underwent organ rejection events as diagnosed by either cfDNA levels or Biopsy scores. The ABS levels of all 4 individuals declined during the first months of treatment. While the ABS levels of individuals without rejection remained at baseline throughout the duration of the study, ABS levels in individuals undergoing rejection increased dramatically up to 8-fold at or before the rejection event. Interestingly, in one example both cfDNA and ABS levels were elevated, yet Biopsy scores did not indicate a rejection (Fig. 14A). This discrepancy suggested that having an additional independent measure of rejection might help confirm cfDNA measurements when they disagree with Biopsy scores.

[00142] *Assay Characteristics.* Next we wanted to investigate the performance of the immune repertoire sequencing assay when compared to cfDNA and Biopsy measurements. We plotted the ABS levels of all individuals in the study, highlighting rejection events as diagnosed by cfDNA and Biopsy scores. The induction of immune suppression reduced ABS levels to a base line after about 1 month. Rejection events appeared to be elevated above this base line. Separating samples into non-rejecting (negative cfDNA and Biopsy Score), mild rejecting (1R Biopsy score) , moderate/severe rejecting (>2R Biopsy score), and elevated cfDNA levels, we showed that all rejecting subgroups were significantly different from the non-rejecting subgroups. The elevated cfDNA group showed the smallest p-value at a 3-fold median increase followed by the moderate to severe rejecting samples at a 5-fold median increase. We determined the Area under the Curve for the ROC curves for ABS level cutoffs which showed that ABS level showed a higher AUC with cfDNA (0.83) than Biopsy scores (0.78).

WHAT IS CLAIMED IS:

1. A method of determining the immunocompetence of a mammalian subject, the method comprising;
 - (i) analyzing the subject's antibody repertoire, wherein analyzing comprises identifying the composition of abundant related IgH sequences in a biological sample obtained from the subject; and
 - (ii) providing an assessment of the immunocompetence of the subject, based on the analysis.
2. The method of Claim 1, wherein the mammalian subject is being treated with an immunosuppressive regimen.
3. The method of Claim 2, wherein the immunosuppressive regimen comprises administering a T-helper cell suppressant.
4. The method of Claim 3, wherein the T-helper cell suppressant is cyclosporine.
5. The method of Claim 3, wherein the T-helper cell suppressant is tacrolimus.
6. The method of any of Claims 1 to 5, wherein the subject is the recipient of an organ transplant.
7. The method of Claim 6, wherein the organ transplant is a heart transplant.
8. The method of any of Claims 1 to 7, wherein the sample is a blood sample.
9. The method of any of Claims 1 to 7, wherein the sample comprises purified PBMCs.
10. The method of any of Claims 1 to 9, further comprising preparing a library of IgH molecules from the sample, prior to the step of analyzing.
11. The method of Claim 10, wherein the step of preparing the library comprises isolating RNA from cells in the sample.

12. The method of Claim 11, wherein the step of preparing the library further comprises amplifying isolated RNA transcripts by cDNA synthesis followed by PCR.
13. The method of Claim 12, wherein primers used for cDNA synthesis comprise unique identifier nucleotide sequences (UIDs).
14. The method of Claim 12, wherein amplified transcripts originating from the same RNA transcript share a UID.
15. The method of any of Claims 1 to 14, further comprising sequencing IgH encoding molecules in the library to obtain IgH molecule reads.
16. The method of Claim 15, wherein the IgH molecule reads are obtained by aligning one or more V, D, and J segments of shorter sequence reads.
17. The method of any of Claims 1 to 16, wherein the step of analyzing comprises classifying IgH molecule reads sharing the same unique identifier nucleotide sequence (UID) as belonging to the same IgH sequence.
18. The method of Claim 17, wherein the step of analyzing further comprises identifying IgH sequences having more than a specific threshold of IgH molecules as abundant related IgH sequences.
19. The method of Claim 17, wherein the step of analyzing further comprises identifying IgH sequences belonging to the same IgH lineage.
20. The method of Claim 19, wherein the step of analyzing further comprises identifying IgH lineages having more than a specific threshold of IgH molecules as abundant related IgH sequences.
21. The method of any of Claims 1 to 20, wherein the step of identifying the composition of abundant related IgH sequences comprises identifying the relative amount of one or more isotypes.

22. The method of Claim 21, wherein the one or more isotypes are IgA, IgG, IgM, naïve IgM, non-naïve IgM, or a combination thereof.

23. The method of Claim 21, wherein the one or more isotypes are non naïve-IgM sequences (ABS).

24. The method of any of Claims 1 to 23, wherein the step of analyzing includes analyzing an additional characteristic of the antibody repertoire.

25. The method of Claim 24, wherein the additional characteristic is isotype distribution, number of abundant related IgH sequences shared between time points, number of abundant related IgH sequences not present in an earlier time point, changes in the abundance of a one or more isotypes, an abundance of mutations in related IgH sequences of a one or more isotypes, mutation rate, or a combination thereof.

26. The method of any of Claims 1 to 25, wherein the step of providing the assessment comprises providing an assessment that the subject is not adequately immunosuppressed when the relative amount of IgG is above a certain threshold, when the relative amount of IgA is above a certain threshold, when the relative amount of non-naïve IgM or total IgM is above a certain threshold, when the relative amount of naïve-IgM is below a certain threshold, when the relative amount of ABS is above a certain threshold, or a combination thereof.

27. The method of any of Claims 1 to 26, wherein the step of providing the assessment comprises providing an assessment that the subject is overly immunosuppressed when the relative amount of IgG is below a certain threshold, when the relative amount of IgA is below a certain threshold, when the relative amount of non-naïve IgM or total IgM is below a certain threshold, when the relative amount of naïve-IgM is above a certain threshold, when the relative amount of ABS is below a certain threshold, or any combination thereof.

28. The method of any of Claims 1 to 27, wherein the assessment is a prediction of the likelihood of allograft rejection.

29. The method of any of Claims 1 to 28, wherein the assessment is further based on the level of cfDNA in the subject's blood.

30. The method of any of Claims 1 to 29, further comprising modifying an immunosuppressive regimen of the subject base on the assessment.
31. The method of Claim 30, wherein modifying the immunosuppressive regimen comprises adjusting the dosage of an immunosuppressive drug.
32. The method of Claim 31, wherein the dosage is increased based on an assessment that the subject is not adequately immunosuppressed.
33. The method of Claim 31, wherein the dosage is decreased based on an assessment that the subject is overly immunosuppressed.
34. The method of any of Claims 31 to 33, further comprising discontinuing treatment with an immunosuppressive drug based on the assessment.
35. The method of any of Claims 31 to 33, further comprising treating the subject with an immunosuppressive drug not previously part of the immunosuppressive regimen, based on the assessment.
36. The method of any of Claims 1 to 35, wherein the subject is human.
37. A machine readable medium comprising software configured to perform the method of Claim 1 or any of Claims 15 to 29.
38. A kit comprising reagents for performing the method of any of Claims 1 to 36.
39. The kit of Claim 38, comprising primers that specifically hybridize to the 5' end of the one or more IgH RNA molecule constant regions and comprising a first adapter sequence.
40. The kit of Claim 38 or 39, comprising a second set of primers that specifically hybridize to one or more V segments and comprising a second adapter sequence.

41. The kit of Claim 40, wherein the first and second set of primers comprise unique identifier nucleotide sequences.
42. The kit of Claim 41, further comprising a third set of primers for PCR, wherein one of the primers in the third set comprises the first adapter sequence and another primer in the third set comprises the second adapter sequence.
43. The kit of any of Claims 38 to 42, further comprising the machine readable medium of Claim 37.
44. The kit of any of Claims 38 to 43, further comprising a reverse transcriptase.
45. The kit of any of Claims 38 to 44, further comprising a DNA polymerase.

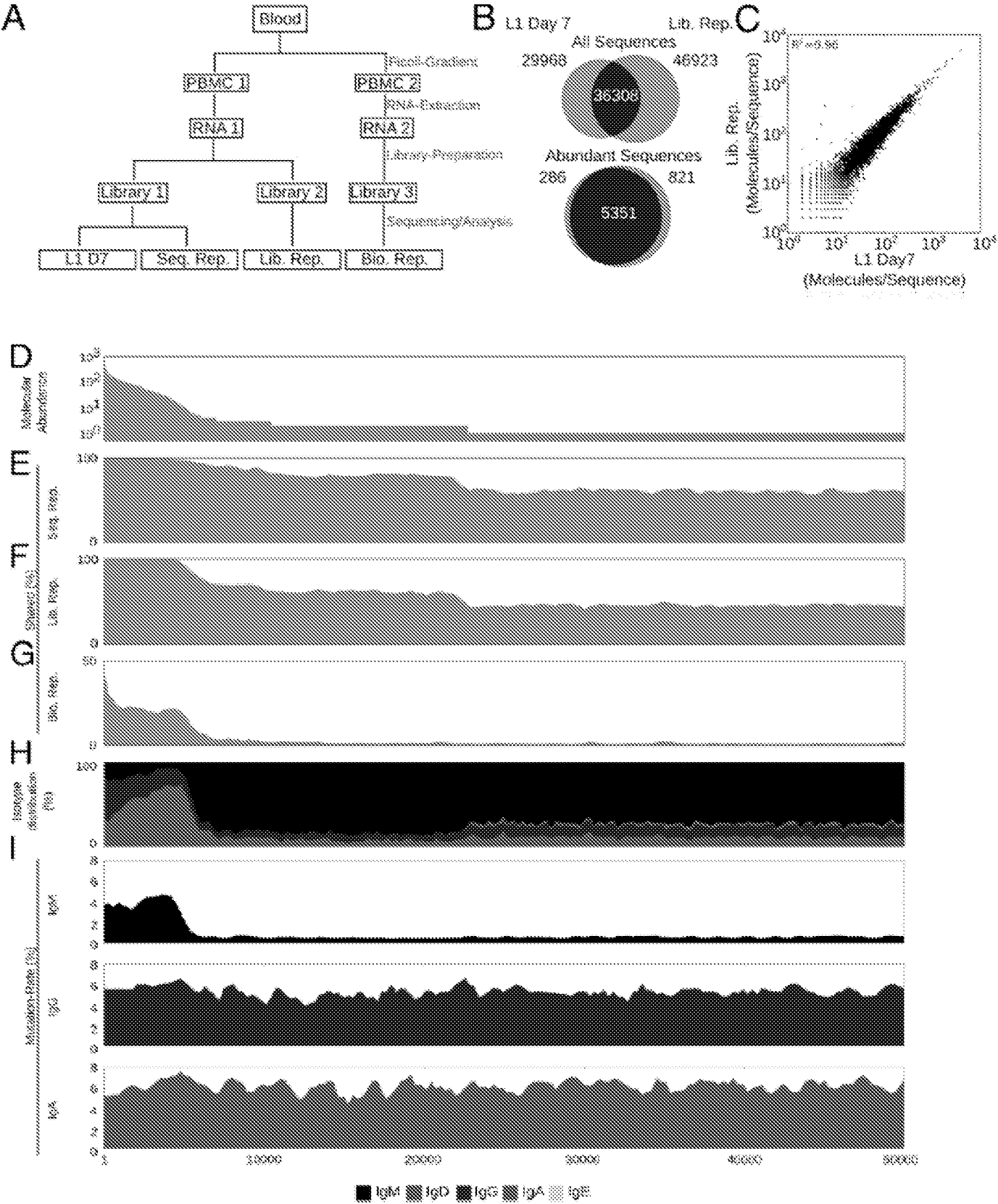


Figure 1

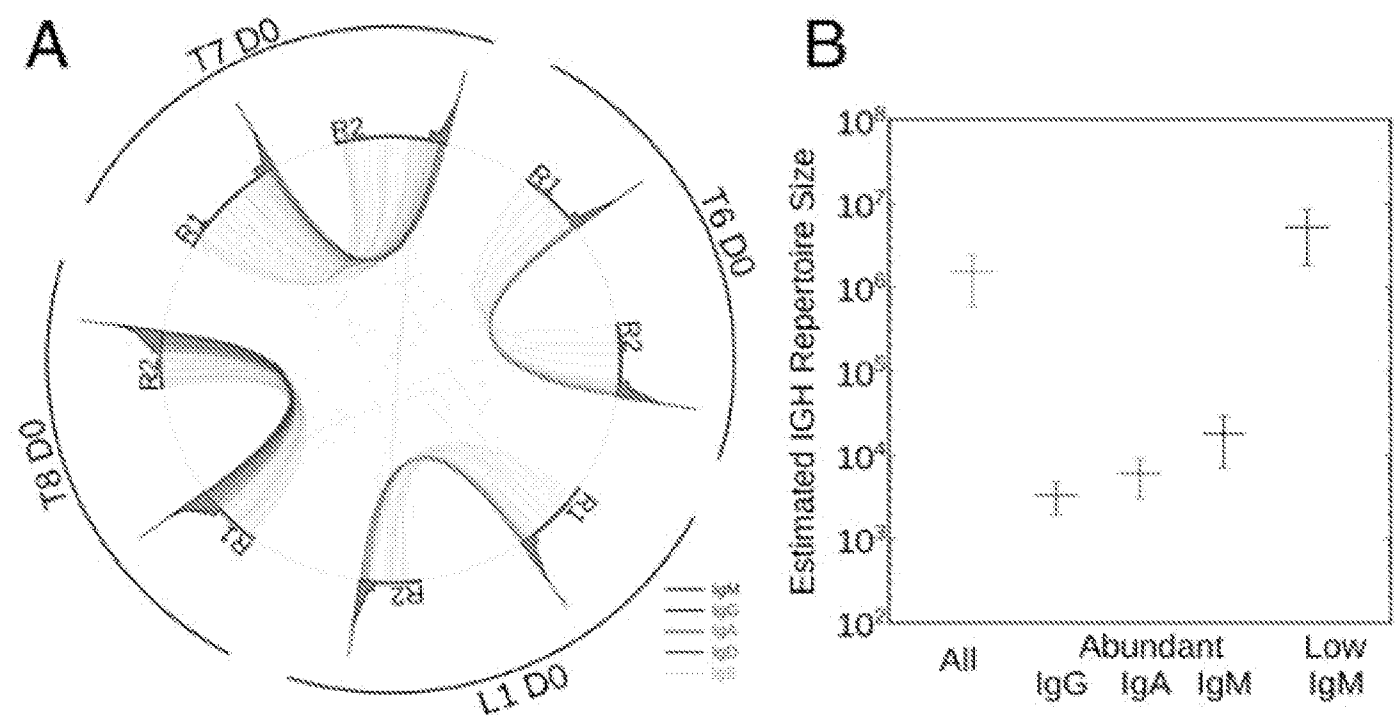


Figure 2

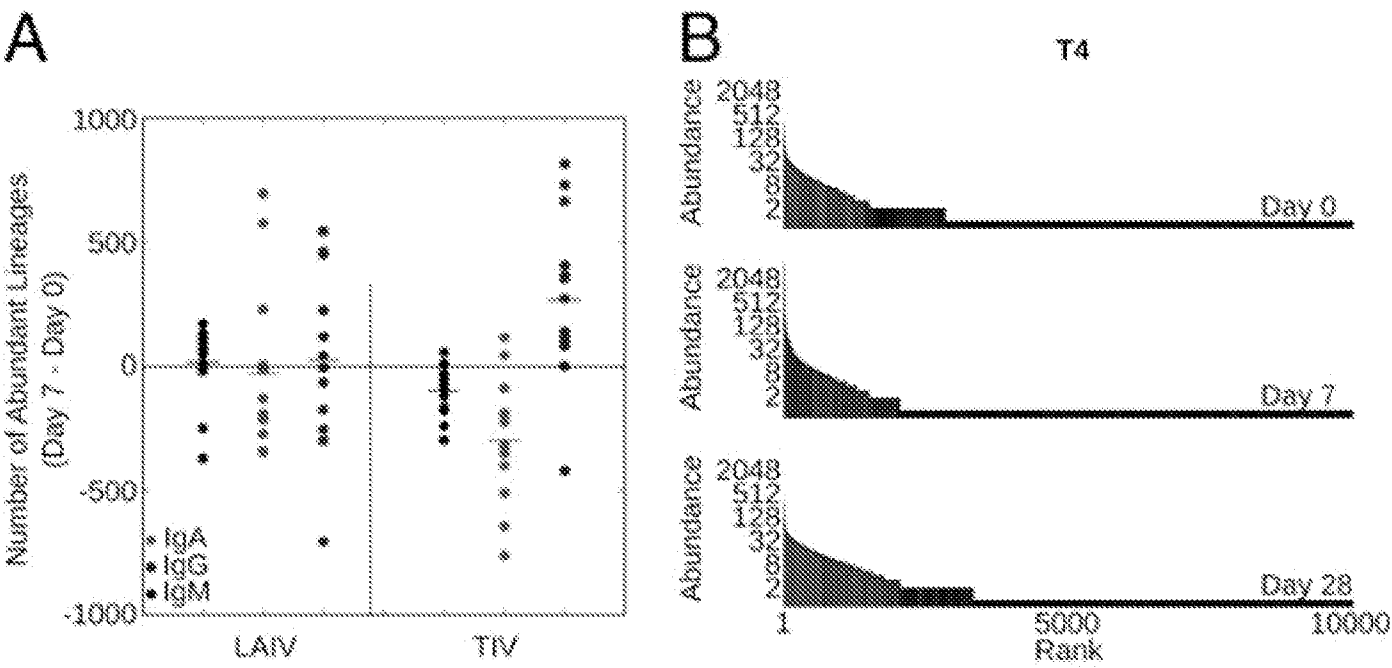


Figure 3
3/15

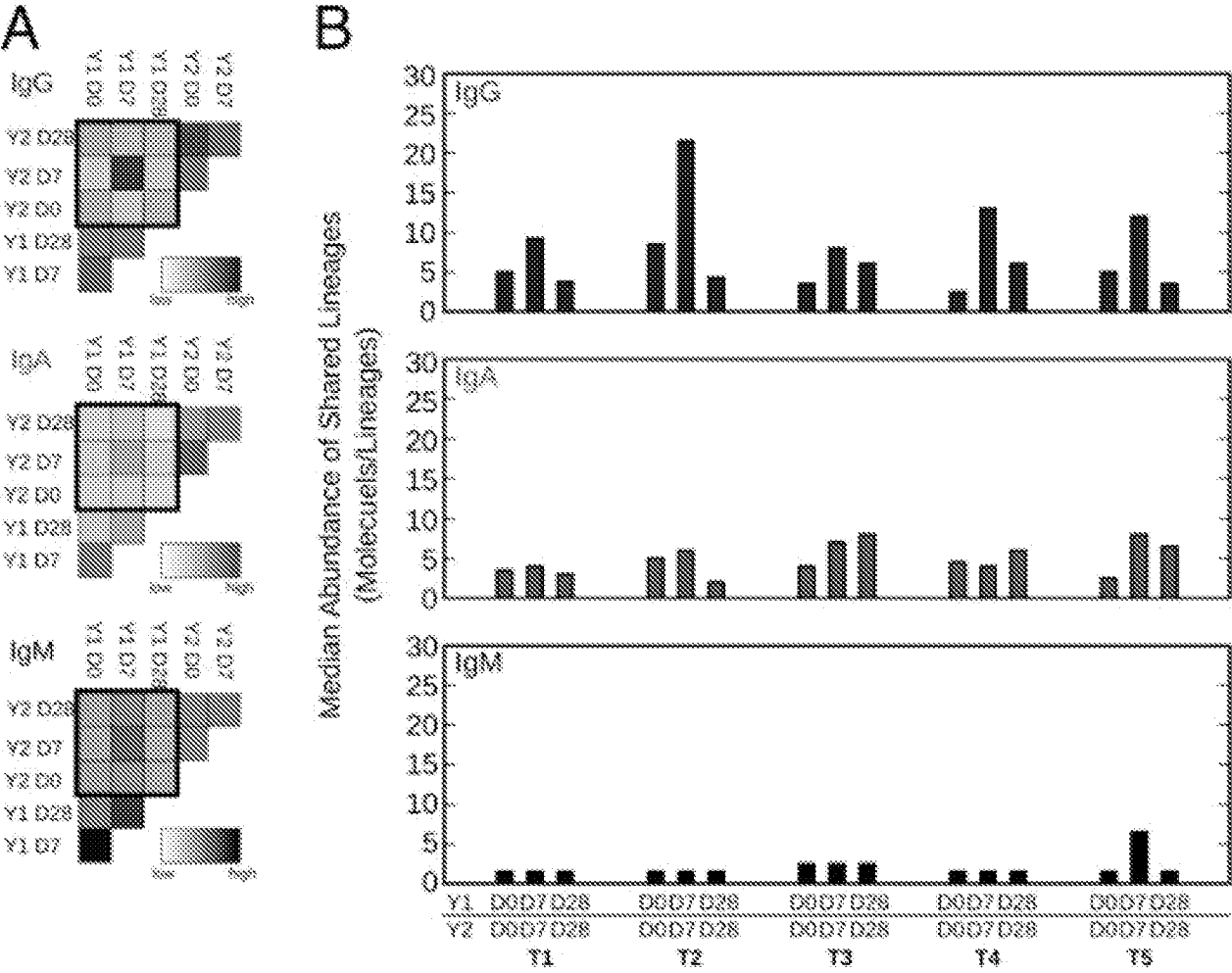


Figure 4

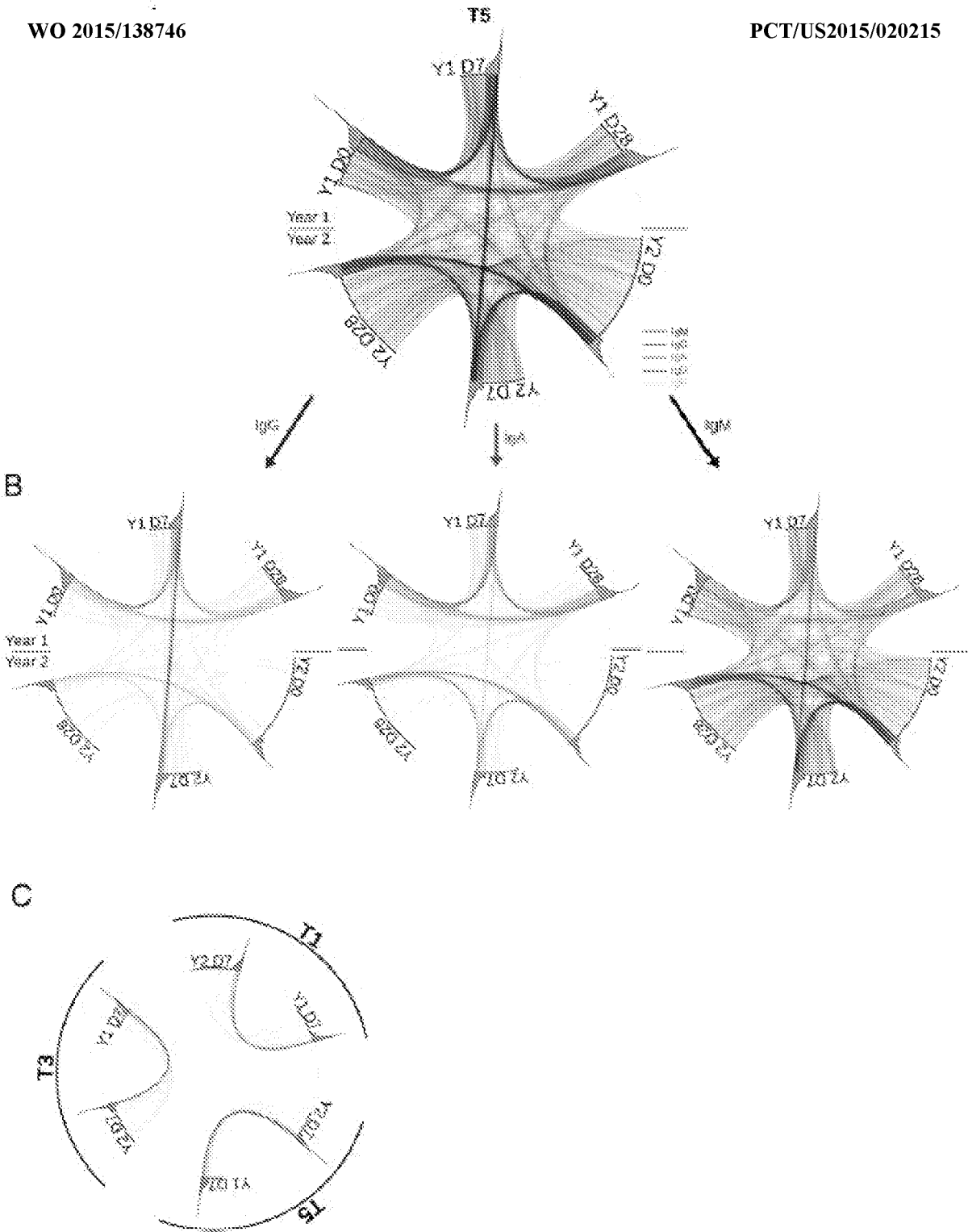
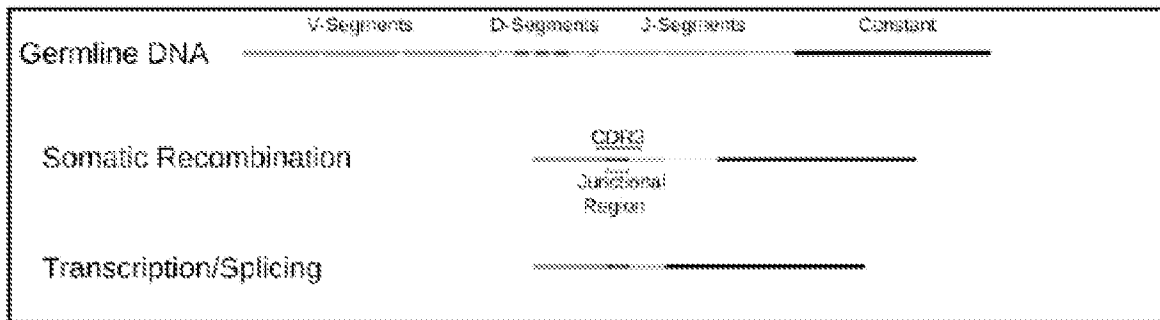
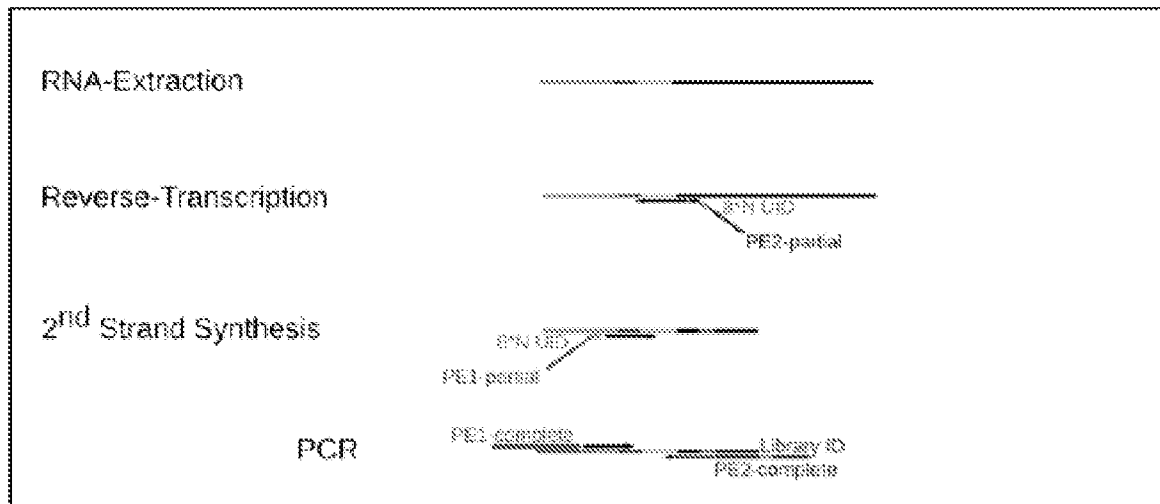


Figure 5

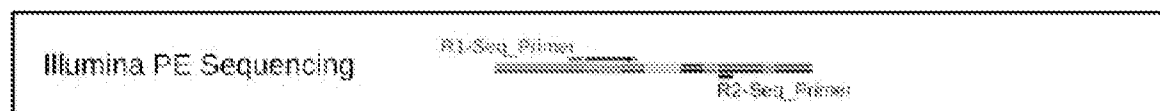
IGH-Locus



Amplification



Sequencing



Analysis

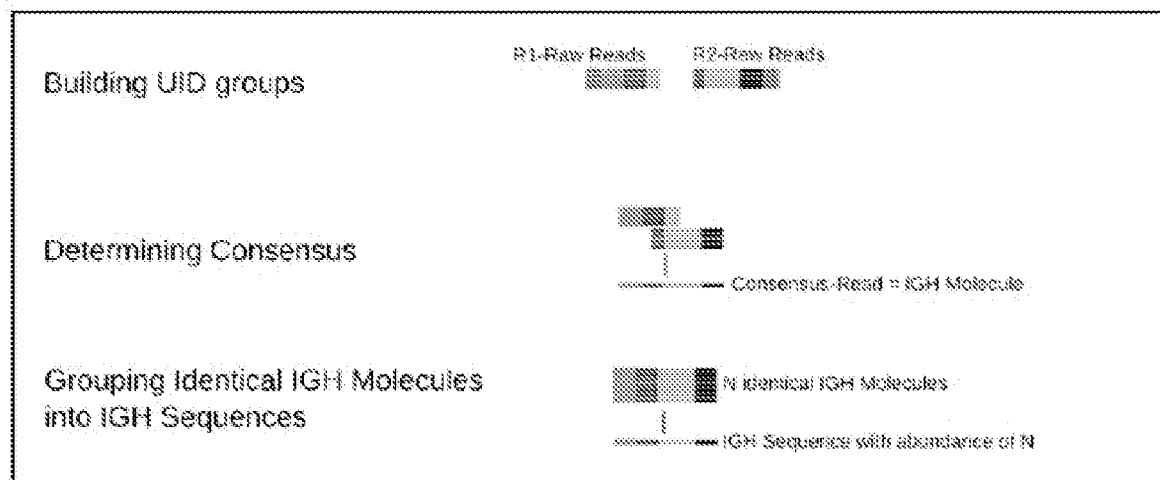
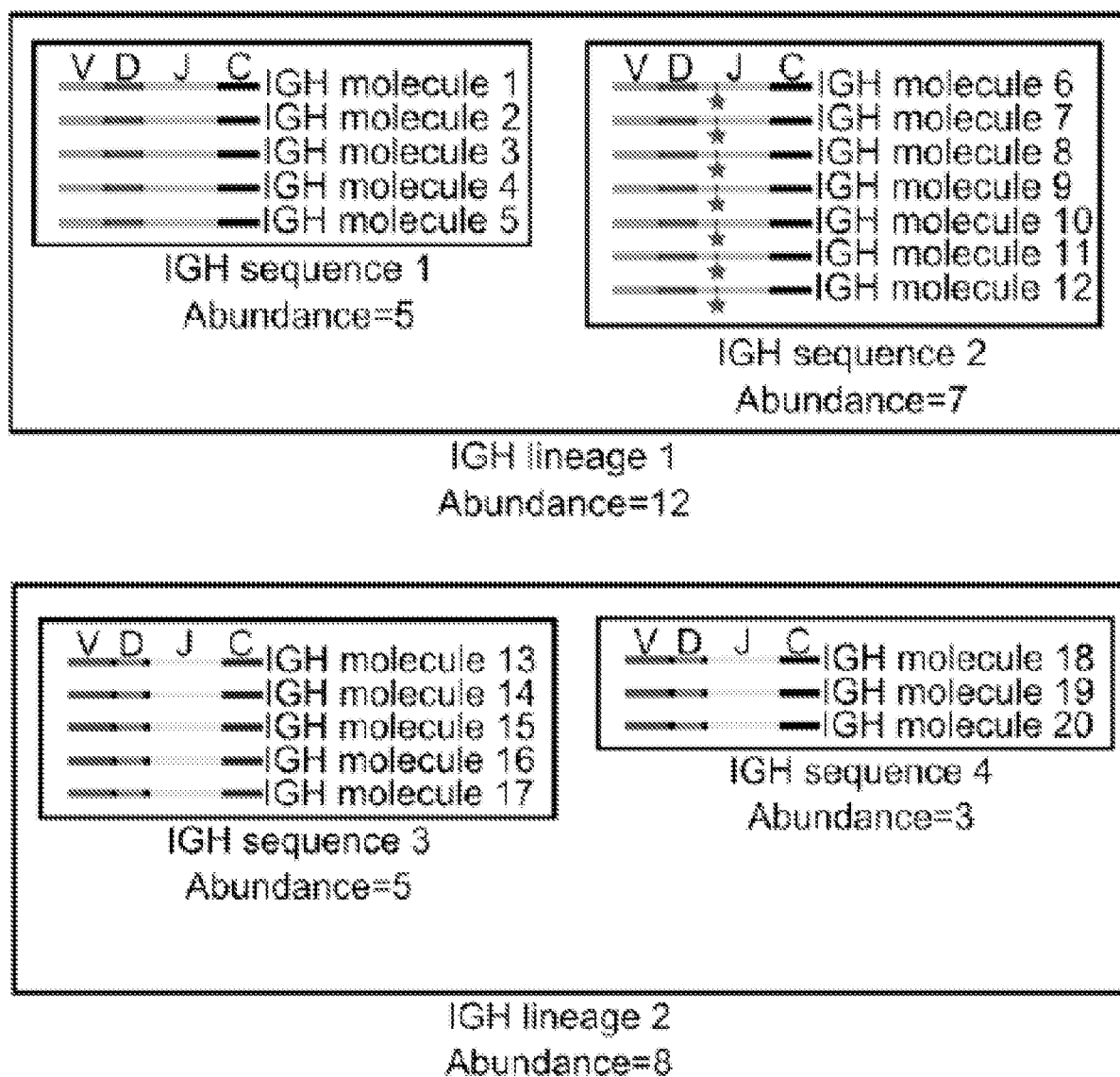
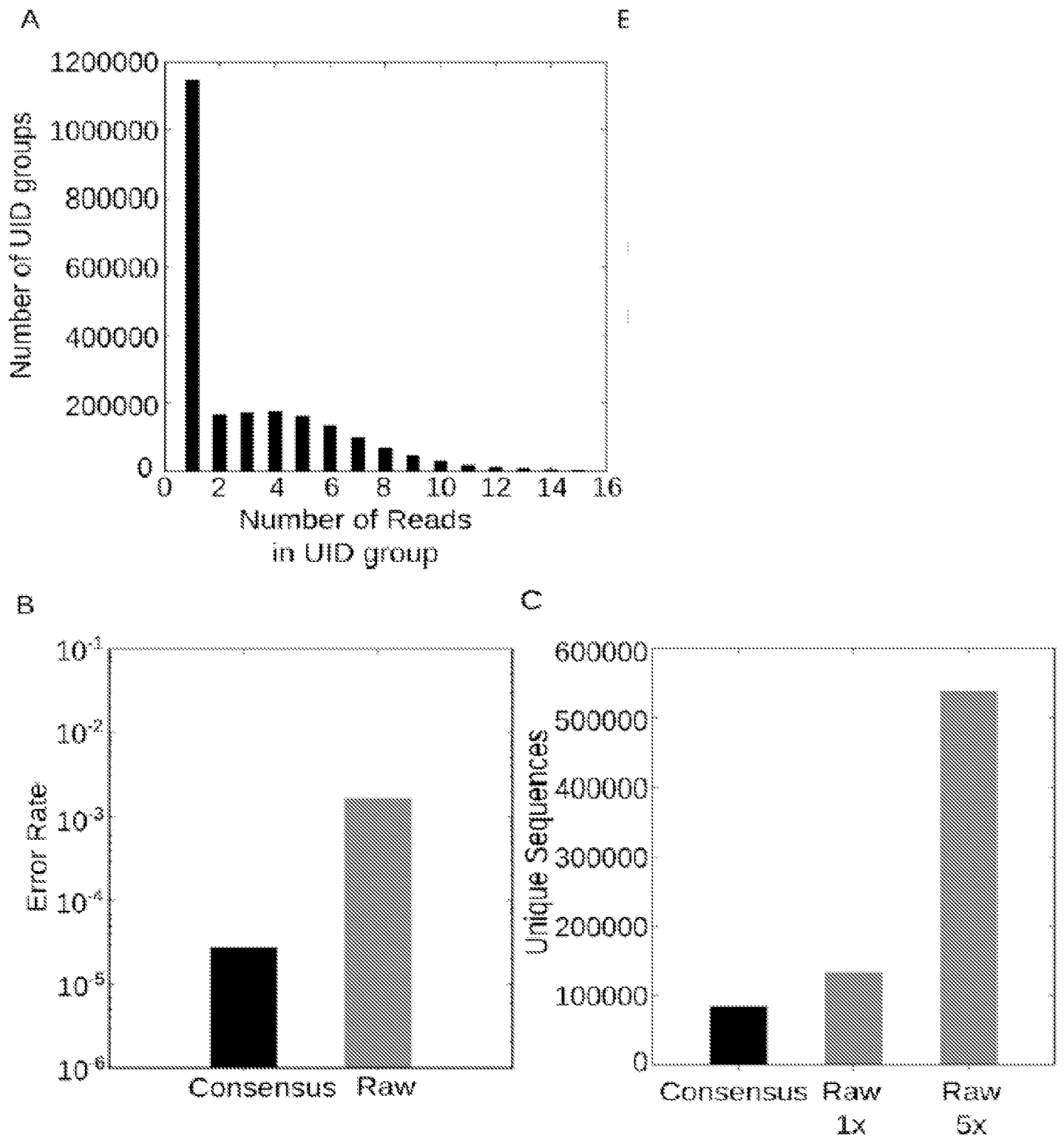


Figure 6A

**Figure 6B**

**Figure 7**

Step/relative amount in pool	Name	Sequence (5'→3')
Reverse transcription		
1	G	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNAAGACCGATGGGCCCTTG
1	A	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNGAAGACCTTGGGGCTGGT
1	M	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNGGGAATTCTCAGGAGACG
1	D	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNGGGTGTCTGCACCCTGATA
1	E_1	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNGAAGACGGATGGGCTCTGT
1	E_2	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNTTGCAGCAGCGGCTCAAGGG
Second-strand synthesis		
1	V1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNAGCCTACATGGAGCTGAGC
1	V2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNAGGTGGTCCCTACAATGACCAAC
1	V3_1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNTCTGCAAATGAACAGCCTGA
0.2	V3_2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNTGTTCAAATGAGCAGTCTGAGAG
0.2	V3_3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNTCTGCAAATGGGCAGCCTGA
1	V4/6	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNTTCTCCCTGAAGCTGAACTCTG
1	V5	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNGCGTACCTGCAGTGGAGCAG
1	V6	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNTTCTCCCTGCAGCTGAACTCTG
1	V7_1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNGCATATCTGCAGATCAGCAGC
1	V7_2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNCAGATCAGCAGCCTAAAGGC
PCR		
1	PE1	AATGATACGGCGACCAACGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
1	PE2	CAAGCAGAAGACGGCATACGAGATAACGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATT

Figure 8

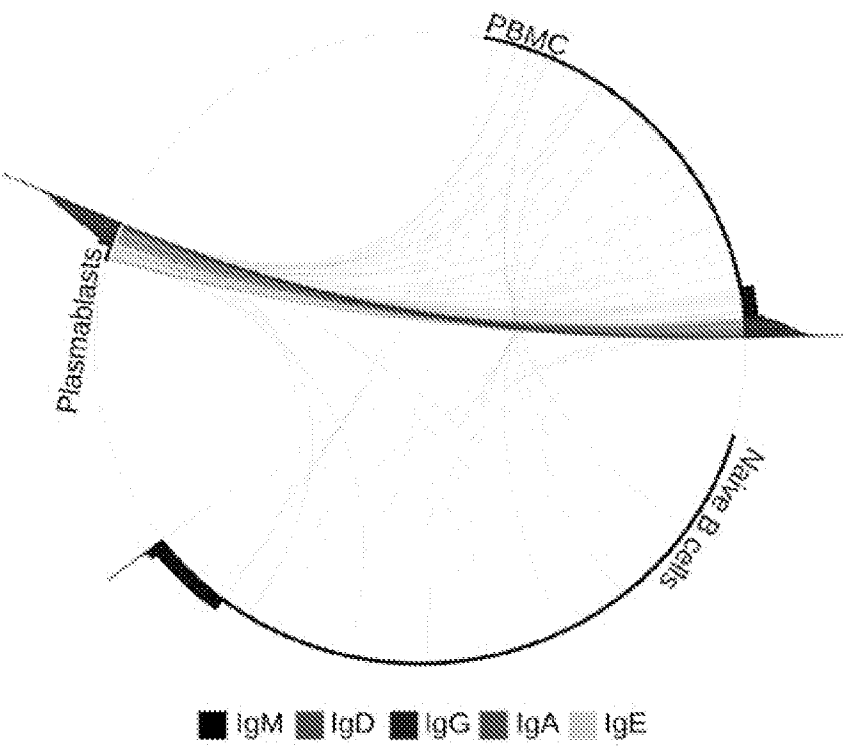
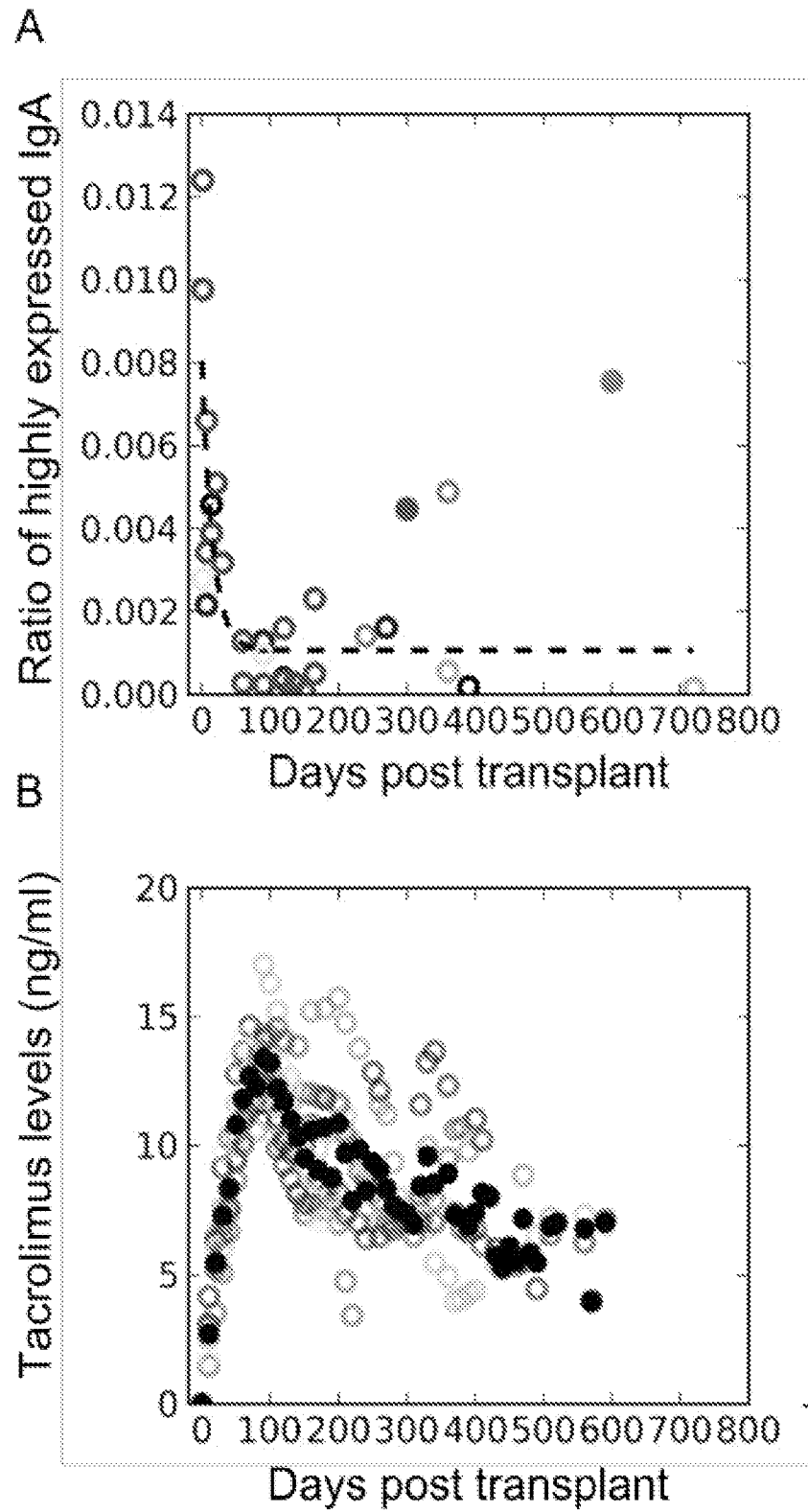
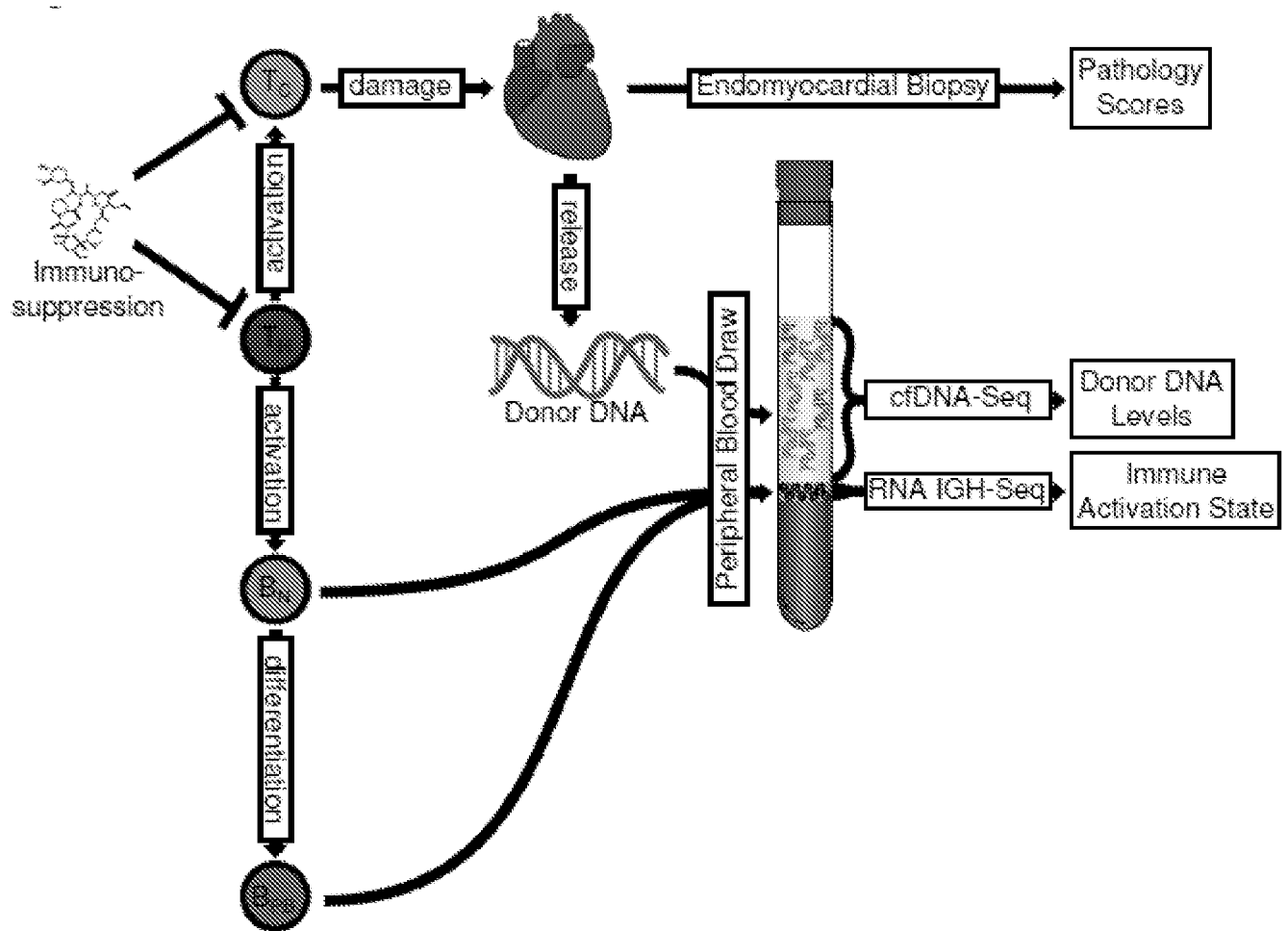
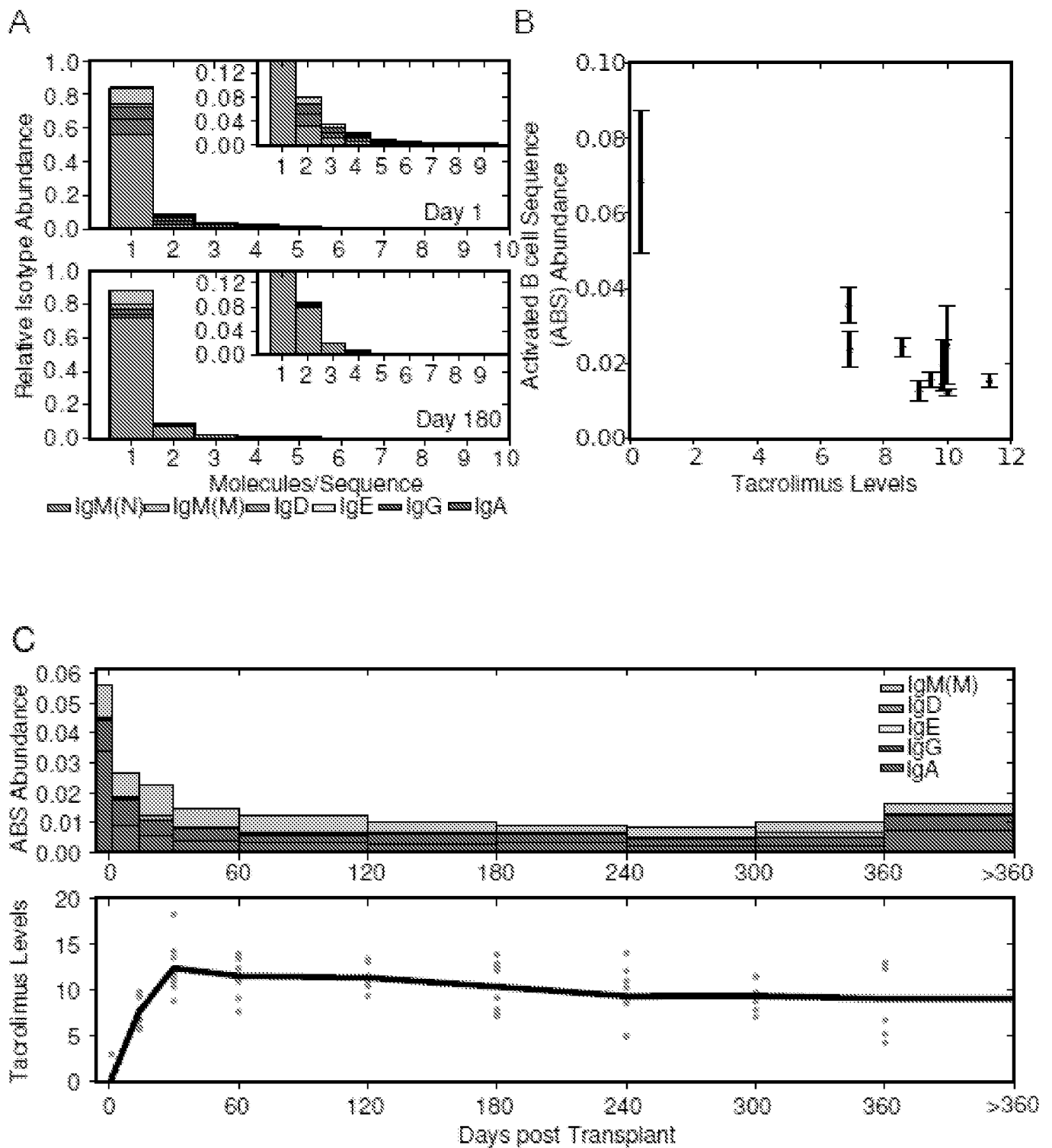


Figure 9
10/15

**Figure 10**

**Figure 11**

**Figure 12**

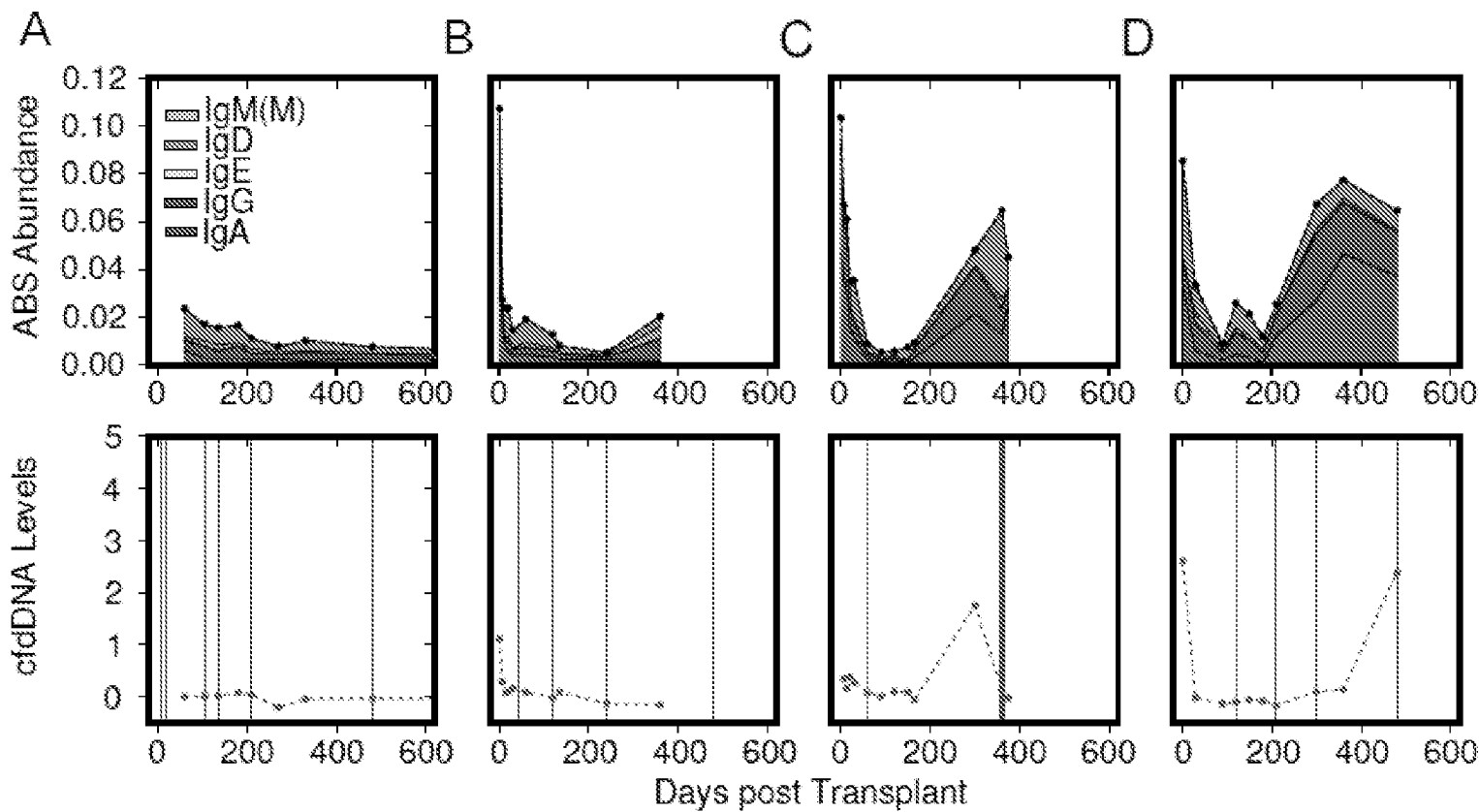
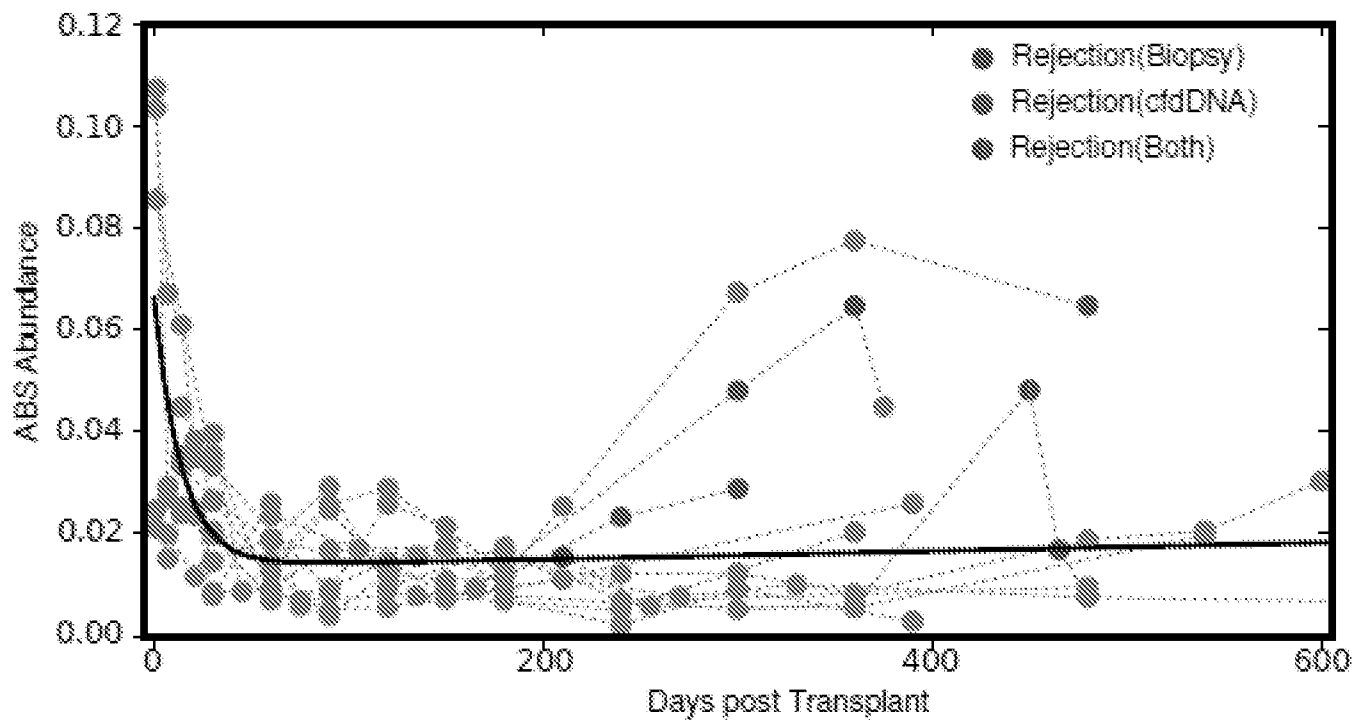
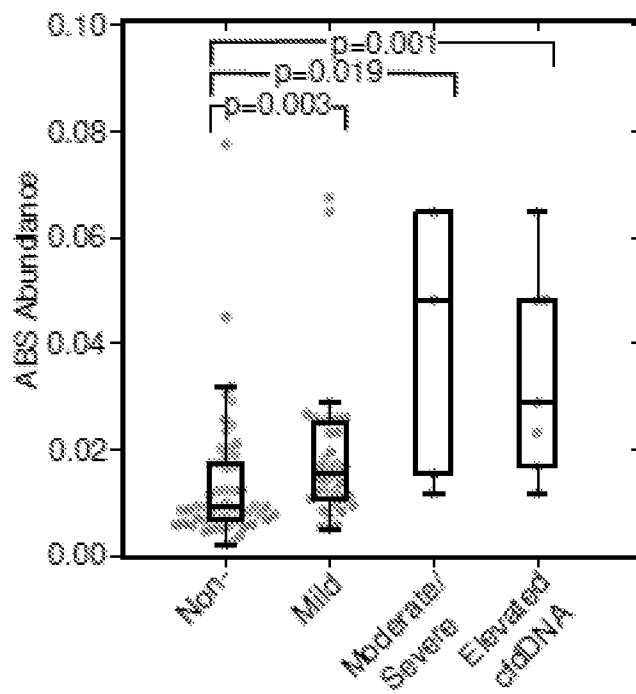


Figure 13
14/15

A



B



C

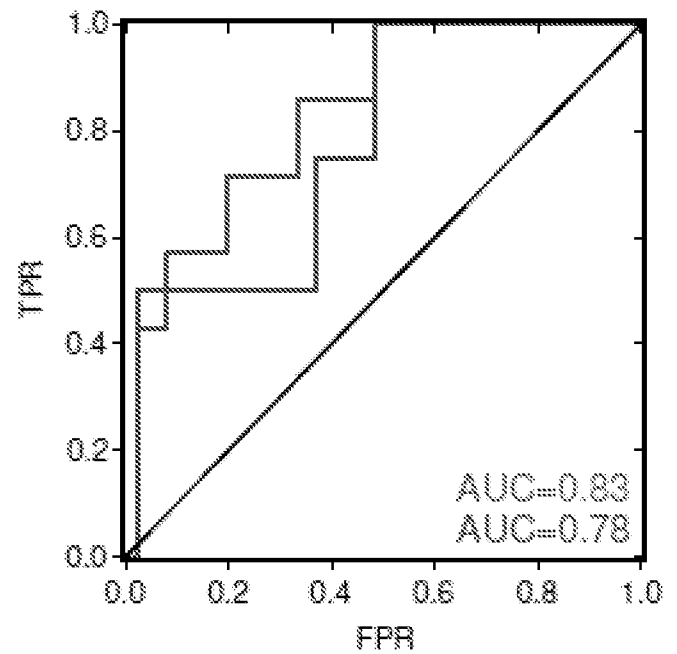


Figure 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 15/20215

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C40B 30/04 (2015.01)

CPC - B01J 2219/00585

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC (8): C40B 30/04 (2015.01)

CPC: B01J 2219/00585

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

IPC (8): C12Q 1/68 (2015.01)

CPC: C12Q 1/6869

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Google patents, Google scholar, Google web, PatBase, Proquest, Dialog

Determine/identify/analyze/assess/monitor; immunocompetence; antibody/immunoglobulin; repertoire/profile; IgH/Immunoglobulin heavy chain; immunosuppressive; T-helper cell suppressant/cyclosporine/tacrolimus; organ/heart/issue; transplant/graft

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	VOLLMERS et al. Genetic measurement of memory B-cell recall using antibody repertoire sequencing. Proc Natl Acad Sci U S A. 13 August 2013. Vol. 110. No. 33. pp 13463-13468, especially, Abstract; page 13468, Col. 1, para 1; page 13463, Col. 2, para 2; page 13465, Col. 1, para 3-9; page 13465, Col. 2, para 4; Fig. 2; page 13464, Col. 1, para 2; Fig. 1; page 13466, Col. 1, para 3; page 13466, Col. 2, para 1; page 13466, Col. 2, para 5, pg 13467, col 1, para 2	1 2-7
Y	US 2010/0151471 A1 (FAHAM et al.) 17 June 2010 (17.06.2010) para [0065-0066]; claim 20; para [0327]; para [0435]; para [0257]; para [0422-0424])	2-7

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

22 May 2015 (22.05.2015)

Date of mailing of the international search report

24 JUN 2015

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 15/20215

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 8-45
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.