Journal of Visualized Experiments

T and B Cell Receptor Immune Repertoire Analysis using Next-generation Sequencing

--Manuscript Draft--

Article Type:	Methods Article - JoVE Produced Video
Manuscript Number:	JoVE61792R1
Full Title:	T and B Cell Receptor Immune Repertoire Analysis using Next-generation Sequencing
Corresponding Author:	Lael Werner, Ph.D Sheba Medical Center at Tel Hashomer Ramat Gan, Israel ISRAEL
Corresponding Author's Institution:	Sheba Medical Center at Tel Hashomer
Corresponding Author E-Mail:	werner.lael@sheba.health.gov.il
Order of Authors:	Lael Werner, Ph.D
	Chen Dor
	Naomi Salamon
	Meital Nagar
	Dror Shouval
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Ramat Gan, Israel
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please specify the section of the submitted manuscript.	Immunology and Infection
Please provide any comments to the journal here.	

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Manuscript Click here to access/download;Manuscript;61792_R1.docx

TITLE:

T and B Cell Receptor Immune Repertoire Analysis using Next-generation Sequencing

AUTHORS:

Lael Werner[1,2], Chen Dor[3,4], Naomi Salamon[1], Meital Nagar[3,4], Dror S. Shouval[1,2]

[1]Pediatric Gastroenterology Unit, Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Ramat Gan, Israel
[2]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[3]Cancer Research Center, and [4]Molecular Hematology Laboratory, Sheba Medical Center, Ramat Gan, Israel

werner.lael@sheba.health.gov.il
Chen.Gefen@sheba.health.gov.il
Naomisalamon@hotmail.com
Meital.Nagar-Marciano@sheba.health.gov.il
dror.shouval@gmail.com

Corresponding Author:

Lael Werner
wernerlael@gmail.com

Keywords:

Next-generation sequencing; TCRβ; IGH; immune repertoire; adaptive immunity; T cells; B cells.

Summary:

The current protocol describes a method for DNA isolation from blood samples and intestinal biopsies, generation of TCRβ and IGH PCR libraries for next-generation sequencing, performance of a NGS run and basic data analysis.

Abstract

Immunological memory, the hallmark of adaptive immunity, is orchestrated by T and B lymphocytes. In circulation and different organs, there are billions of unique T and B cell clones, and each one can bind a specific antigen, leading to proliferation, differentiation and/or cytokine secretion. The vast heterogeneity in T and B cells is generated by random recombination of different genetic segments. Next-generation sequencing (NGS) technologies, developed in the last decade, enable an unprecedented in-depth view of the T and B cell receptor immune repertoire. Studies in various inflammatory conditions, immunodeficiencies, infections and malignancies demonstrated marked changes in clonality, gene usage, and biophysical properties of immune repertoire, providing important insights about the role of adaptive immune responses in different disorders.

Here, we provide a detailed protocol for NGS of immune repertoire of T and B cells from blood and tissue. We present a pipeline starting from DNA isolation through library preparation, sequencing on NGS sequencer and ending with basic analyses. This method enables exploration of specific T and B cells at the nucleotide or amino-acid level, and thus can identify

dynamic changes in lymphocyte populations and diversity parameters in different diseases. This technique is slowly entering clinical practice and has the potential for identification of novel biomarkers, risk stratification and precision medicine.

Introduction

The adaptive immune system, comprised of T and B lymphocytes, utilizes immunological memory to recognize a previously encountered antigen and initiate a rapid response. Lymphocytes are generated in the bone marrow and mature in the thymus (T cells) or bone marrow (B cells). Both the T cell receptor (TCR) and B cell receptor (BCR) display unique configurations that allow recognition of specific antigens. In homeostasis, T and B cells constantly circulate and survey the trillions of different peptides presented on antigen-presenting cells. TCR or BCR ligation of a specific antigen with high affinity, together with appropriate co-stimulation, leads to cell activation, resulting in cytokine secretion, clonal expansion and generation of antibodies, in the case of B cells.

The enormous array of all the different T or B cells is collectively termed immune repertoire, enabling recognition of countless of different epitopes. In order to generate such a vast repertoire, a complex process of random assembly of different gene segments takes place, creating nearly endless combinations of receptors that can bind unique antigens[1]. This process, called V(D)J recombination, includes rearrangements of different variable (V), diversity (D) and joining (J) genes, accompanied by random deletions and insertions of nucleotides in the junctions[2].

The architecture of the adaptive immune system has interested scientists in different fields for many decades. In the past, Sanger sequencing, complementary determining region 3 (CDR3) spectratyping, and flow cytometry were used to characterize the immune repertoire, but provided low resolution. In the last decade, advances in next-generation sequencing (NGS) methods enabled in-depth insight into the characteristics and composition of an individual's TCR and BCR repertoires[3,4]. These high-throughput systems (HTS) sequence and process millions of rearranged TCR or BCR products simultaneously and permit a high-resolution analysis of specific T and B cells at the nucleotide or amino acid level. NGS provides a new strategy to study the immune repertoire in both health and disease. Studies utilizing HTS demonstrated altered TCR and BCR repertoires in autoimmune diseases[5], primary immunodeficiencies[6,7], and malignancies, such as in acute myeloid leukemia[8]. Using NGS, we and others have shown oligoclonal expansion of specific T and B cell clones, in patients with inflammatory bowel disease (IBD), including ulcerative colitis and Crohn's disease[9-14] . Overall, studies from different fields suggest changes in the repertoire have a crucial role in the pathogenesis of immune-mediated disorders.

The current protocol describes a method for isolation of DNA from intestinal biopsies and blood, generation of TCRβ and IGH PCR libraries for NGS, and performance of sequencing run. We also provide basic steps in immune repertoire data analysis. This protocol can be applied for the generation of TCRα, TCRγ, and IGL libraries as well. The method is also compatible with other organs (e.g., lymph nodes, tumors, synovial fluid, fat tissue, etc.) as long as tissue-specific digestion protocols are used.

Protocol

This study was approved by the institutional review board at Sheba Medical Center, and informed written consent was obtained from all participating subjects.

1. DNA isolation and quantification

1.1. Digestion and cell lysis of intestinal biopsies

1.1.1. Retrieve intestinal biopsies, either freshly collected or those stored at -20 °C or -80 °C. If using frozen biopsies, thaw on ice.

1.1.2. Add 600 µL of nuclei lysis solution to a sterile 1.7 mL micro-centrifuge tube, chilled on ice.

1.1.3. Place biopsy into a microcentrifuge tube with lysis solution and incubate at 65 °C for 15-30 min.

1.1.4. Add 17.5 µL of 20 mg/mL proteinase K and incubate overnight at 55 °C, with gentle shaking.

1.1.5. Allow sample to cool to room temperature for 5 min before proceeding to step 1.3.

1.2. Cell lysis of whole blood

1.2.1. Obtain 3 mL of whole blood in EDTA-, heparin-, or citrate-containing tube.

1.2.2. Gently rock tube until thoroughly mixed, and transfer blood to a sterile 15 mL centrifuge tube containing 9 mL of cell lysis solution. Invert several times during a 10 min incubation period at room temperature.

1.2.3. Centrifuge at 2,000 x g for 10 min at room temperature, and then discard as much supernatant as possible without disturbing the visible white pellet. Approximately 50–100 µL of residual liquid will remain.

1.2.4. Vortex tube vigorously for 15 s until completely resuspended. Add 3 mL of nuclei lysis solution and pipet several times. The solution should become viscous.

1.3. Protein precipitation

1.3.1. Add 200 µL of protein precipitation buffer to the tissue, or 1 mL to the blood. Vortex vigorously for 20 s and incubate on ice for 5 min. Small protein clumps may be visible after vortexing.

1.3.2. Centrifuge at 16,000 x g for 4 min. A tight pellet containing precipitated proteins should form.

1.3.3. Carefully transfer supernatant, without disturbing the pellet, to a new 1.7 mL tube.

NOTE: If pellet is not tight, consider another centrifuge at 16,000 x *g* for 4 min.

1.4. DNA precipitation and rehydration

1.4.1. Add 1 mL of 2-propanol to tube containing the supernatant, and gently mix by inverting. DNA should become visible as white floating substance.

1.4.2. Centrifuge at 16,000 x *g* for 1 min. Remove the supernatant completely using a pipette.

1.4.3. Add 1 mL of freshly prepared 70% ethanol, and invert tube several times. Centrifuge at 16,000 x *g* for 1 min at room temperature. Discard supernatant carefully.

1.4.4. Repeat step 1.4.3.

1.4.5. Place open tube inverted on clean absorbent paper for 10-15 min. Re-invert tube and confirm complete hydration. If needed, leave to air-dry for additional 10-15 min.

NOTE: It is crucial for pellet to dry completely.

1.4.6. Add 50 µL of ultra-pure water (UPW). If needed, DNA can be diluted more after quantification.

1.4.7. Incubate for 1 h on heat block at 65 °C, for DNA rehydration.

1.5. DNA quantification

NOTE: DNA was quantified using a fluorometer and designated reagents, as specified in table of materials.

1.5.1. Bring standards to room temperature and prepare the kit, including 0.5 mL designated tubes.

1.5.2. Prepare buffer and dye mix in a 15 mL tube, according to number of samples. Add 200 µL of buffer and 1 µL of dye per sample. Include 2 samples for standards. It is recommended to calculate an additional sample in order to avoid that the buffer will not be sufficient.

1.5.2.1. For the 2 standards, place 190 µL of the above-prepared mix in 0.5 mL tube. Add 10 µL of the standard.

1.5.2.2. For each sample, place 199 µL of the above-prepared mix in 0.5 mL tube. Add 1 µL of the DNA.

1.5.3. Read the samples. The result indicates DNA concentration of the sample.

1.5.4. If samples are over the limit, dilute the DNA 1:10 with UPW, and repeat read.

2. Library preparation

NOTE: The current protocol utilizes a multiplex, PCR-based assay kit compatible with NGS sequencers. The kit contains 24 different indices each targeting conserved regions within the V_β and the J_β regions. This enables a one-step PCR reaction and pooling of different samples. See table of materials for details.

2.1. Prepare DNA samples. Obtain 150 ng of DNA and add UPW for a total of 5 µL.

NOTE: It is possible to use less than 150 ng of DNA for library preparation, with minimum concentration of 10 ng/µL.

2.2. Place pipettes and tips in hood with UV light for 15 min to break down residual DNA. Meanwhile, allow different index tubes (for library preparation) and DNA polymerase to thaw on ice.

2.3. In a biological hood, add 45 µL from the different index tubes into PCR tubes. To follow, add 0.2 µL of the DNA polymerase and the 5 µL of DNA prepared in step 2.1 into PCR tubes.

2.4. Run PCR reaction using a pre-set program, according to the manufacturer's instructions.

3. Amplicon purification and quantification

3.1. Use magnetic beads for removal of excess primers, nucleotides, and enzymes.

3.1.1. Warm the beads at least 30 min to room temperature. Prepare enough fresh 80% ethanol for 400 µL per sample.

3.1.2. Add 50 µL (IGH) or 35 µL (TCRβ) of the beads to each PCR tube and mix by pipetting 10 times. Incubate 10 min at room temperature.

3.1.3. Place the mixed sample on the magnetic stand for 5 min.

3.1.4. Aspirate 95 µL (IGH) or 80 µL (TCRβ) of the clear liquid and discard. Aspirate remaining liquid using a 10 µL tip.

3.1.5. Add 200 µL of 80% ethanol to each sample. Incubate 30 s. Aspirate 195 µL of ethanol and discard. Aspirate remaining liquid using 10 µL tip.

3.1.6. Repeat step 3.1.5. Open the caps of the tubes and let air dry for 5 min.

3.1.7. Immediately after the 5 min, remove from magnetic stand and add 25 µL of elution buffer (10 mM Tris-HCl, pH 8.0). Mix by pipetting until homogeneous. Incubate at RT for 2 min.

3.1.8. Place the tubes on the magnetic stand for 5 min. Transfer 22 µL to a new PCR tube. Measure the DNA concentration (step 1.5).

3.2. Amplicon quantification

NOTE: The current protocol uses an automated electrophoresis machine for quality control of DNA concentration, size, and integrity. See **Table of Materials** for details.

3.2.1. Place reagents and screentape at room temperature for at least 30 min.

3.2.2. In a new 1.7 mL micro-centrifuge tube, make a 1 ng/µL dilution of the DNA samples quantified in step 3.1.8 with UPW for a total volume of at least 3 µL.

3.2.3. Vortex diluted DNA before use. Place 2 µL of buffer, together with 2 µL of the diluted DNA in the pre-labeled specialized PCR strips (provided in the kit) and place the caps. Place on shaker for 1 min, followed by spin down of PCR strips.

NOTE: Due to the viscosity of buffer, ensure excess buffer is removed from the tip before transfer to the sample tubes.

3.2.4. Place in the machine, the screentape and PCR strips without the caps. Open the lid of tip box inside the machine. Assign the position with appropriate labels. Start the machine.

NOTE: The output histogram should be a single peak at the desired size of the amplicons (**Figure 1A**). In samples of poor quality, additional peaks will be observed (**Figure 1B**), indicating poor bead clean-up or formation of primer dimers.

4. Next-generation sequencing

4.1. Quantification and pooling of library

4.1.1. Calculate final DNA concentration in nM, using DNA value from step 3.1.8, and the size in base pairs (bp) of the product's peak, obtained from electrophoresis machine results (see **Figure 1**).

4.1.2. For each sample, calculate DNA volume to take for a final concentration of 4 nM, in a final volume of 10-20 µL elution buffer.

4.1.3. Prepare a new dilution mix for each sample and take 2 µL from it to a new pool mix.

NOTE: For samples that have less than 4 nM, add the maximum amount of sample possible.

4.2. NGS run of library

NOTE: The current protocol uses a bench-top sequencer platform. See table of materials for details.

4.2.1. Prepare a fresh solution of 0.2 N NaOH. Add 10 µL of 0.2 N NaOH to the diluted library prepared in step 4.1.3. Add 5% PhiX to tube and vortex briefly. Spin down the tube to ensure all the solution has settled to the bottom of the tube.

4.2.2. Incubate for 5 min at room temperature to denature the double stranded DNA. Add 980 µL of pre-chilled buffer from the kit to the tube containing DNA and vortex briefly.

4.2.3. Place the diluted library on ice, and prepare the library as follows. For a mix of 17 pM, add 425 µL of DNA from step 4.2.2 and 575 µL of buffer. Invert library several times and spin down. Load 600 µL of the final library onto the designated cartridge, and place samples in the machine.

4.2.4. Start the run following the software instructions.

5. Sequencing analysis

5.1. Download sequencing metrics from sequencer, and verify that sequencing data is within the following ranges (specific for the kit used in the current protocol). Samples not meeting these criteria are subject to repeat run:
Cluster Density (Density (K/mm2)): 945K – 1800K
Percent of reads passing filter (Clusters PF (%)): >90%
Quality scores (%>=Q30): >85%
PhiX alignment (Aligned %): 1.5% - 10%
PhiX Error Rate (%): <1.0%

Representative results

Herein, we describe a method for DNA isolation from intestinal tissue and blood, preparation of libraries for NGS, and basic steps of a sequencing run for immune repertoire sequencing. The run will generate fastq files, which can be further converted to fasta files for use in the international ImMunoGeneTics (IMGT)/HighV-QUEST platform. This HTS performs and manages many analyses of tens of thousands of rearranged TCRβ and IGH sequences, at the nucleotide level[15]. IMGT/HighV-QUEST enables analysis of different TCRs and IGH repertoires in both health and disease. This can lead to identification of new "disease-specific" clones, analysis of clonal expansion and diversity parameters, delineation of differential V(D)J usage, analysis of somatic hyper-mutations, and more. The IMGT/HighV-QUEST provides CSV files, which contain specific sequences and their abundance. Using genomic DNA as a starting material yields sequence numbers that are representative of cell numbers. Thus, if original DNA quantity is equal, percentage of T cells in a given sample can also be calculated.

We include a basic analysis from representative, autologous blood and rectal samples of a patient with IL10 receptor deficiency and history of severe infantile-onset IBD, resulting from a deleterious *IL10RA* mutation. Samples were assayed for both TCRβ and IGH repertoire. The intestinal TCRβ sample yielded a total 12,450 sequences, of which 9,050 sequences were unique. The blood TCRβ sample had a total 54,880 sequences, of which 35,110 were unique. In the intestinal IGH sample, a total of 49,070 sequences were obtained, of which 23,670 were unique. In the blood IGH sample, a total of 13,710 sequences were obtained, of which

13,540 were unique. All clones can be identified by their unique sequence both at the nucleotide or amino acid level. These sequences can be compared between different patients, in search for shared clones, or in the same patient between different anatomical sites (e.g., blood vs. intestine). We present for each of the samples the 5 most frequent clones (**Table 1**).

To quantify the degree of clonal expansion different indices can be used, including Shannon's H, Gini-Simpson, entropy and clonality. As an example, Shannon's H, which takes into account the number of unique sequences (richness of the repertoire) and how evenly they are distributed was found to be decreased in the patient's intestinal IGH vs. blood (8.3 vs. 9.5), suggesting clonal expansion of B cells in the inflamed gut.

For a broad overview of the repertoire, Treemap images (www.treemap.com) were generated (**Figure 2**). Each colored square represents a different clone, and the size correlates with its frequency. These Treemap images demonstrate clonal expansion in the intestinal IGH repertoire compared with the blood. In contrast, in the TCRβ repertoire, marked clonal expansion is observed in the blood, in comparison to autologous intestine.

At the gene level, NGS provides information regarding V- D- and J- usage at the level of either gene, family, or allele, as shown in **Table 1**. Moreover, specific V(D)J combinations can be inferred from the data for TCRβ and IGH repertoires (**Figure 3A,B**), which can reveal differential gene usage patterns in various conditions. Biophysical properties of the CDR3 region such as length (**Table 1**) or hydrophobicity can also be analyzed. Importantly, CDR3 length distribution is altered in different immune-mediated disorders[10,16,17]. For example, in IL10R deficient patients, blood-derived T cells, but not B cells, have shorter CDR3 length, and differential hydrophobicity, in comparison to healthy controls[7].

Figure 1. Representative bioanalyzer data. Representative images of intestinal TCRβ samples showing optimal results (**A**) with the desired peak at 400 bp. An example of a low-quality library (**B**) with additional peaks at 179bp and 114bp (marked *). Peaks seen at ends are upper and lower markers of the electrophoresis strip.

Figure 2. Overview of T and B cell immune repertoire. Treemap diagrams of TCRβ and IGH blood and intestinal samples. Each colored square represents a different clone, and the size correlates with its frequency.

Figure 3. Representative graphs of TCRβ and IGH V-J usage. A representative intestinal sample depicting total number of sequences for each V-J combination in TCRβ (**A**) or IGH (**B**). Data is not shown for D usage.

Discussion

Changes in abundance and function of B and T lymphocytes are often encountered in different malignancies[18], chronic inflammatory disorders (e.g., ulcerative colitis and rheumatoid arthritis)[10,19], and in various immunodeficiencies[17,20]. The current method utilizes NGS to facilitate an in-depth view of TCR and BCR repertoires, enabling detection of subtle changes in T and B cell clonality, sharing of clones, V(D)J gene usage, and information on the degree of somatic hyper-mutations in the case of B cells.

The method of DNA isolation was described for intestinal biopsies and blood samples. However, with modifications of lysis and DNA extraction, the method can be applied to other tissues such as tumors, lymph nodes, synovial fluid, etc. It is important during library preparation not to cross-contaminate the different primer sets. At bead clean-up, care should be taken to leave tubes on the magnet at all times of incubation. For DNA samples of low concentrations, stocks can be concentrated at 65 °C for 10-20 minutes. Moreover, amplicons can be eluted from beads in as little as 20 µL elution buffer.

Different techniques used in the past to characterize the landscape of T and B cell composition provided only a superficial overview of the immune repertoire. One of the biggest advantages of NGS for immune repertoire analysis is the ability to identify unique clones, and consequently track them in different anatomical sites (e.g., intestine vs. blood vs. lymph node) or in different individuals. The clinical implications of such a technique are significant, and go beyond enhancing the understanding of the role of T or B cells in different immune-mediated diseases or malignancies[21]. In oncology, NGS is used to identify specific clones that can be predictive biomarkers for the patients who will most likely benefit from current immunotherapies[22-24]. In leukemia, NGS is used to confirm full eradication of cancerous cells by detection of residual leukemic clones that remain in a patient after treatment[25].

One of the main limitations of immune repertoire analysis of whole tissue or blood samples is the inability to identify repertoire changes in less frequent populations. For example, if regulatory T cells, which comprise only a few percentages of total $CD4^+$ T cells, express a unique repertoire profile, this would be missed if conducting these studies on whole blood. This problem could be addressed by conducting these studies on sorted immune populations. Alternatively, in recent years technology has developed to facilitate coupling of single cell RNA data with TCR or IgH repertoire profiles[26]. This provides another level of functionality of the T or B cells, since the transcriptional landscape of each clone can be characterized. As an example, Zemmour et al. used scRNAseq TCRseq to demonstrate that regulatory T cells from human blood, display broad heterogeneity with an activated subpopulation that is transcriptionally related to conventional T cells[27]. Similarly, in patients with hepatocellular carcinoma this method identified 11 distinct T cell subsets that infiltrate the tumor, each one with a unique transcriptional and repertoire profile[28]. Studies like these will be helpful especially when studying rare populations, and will provide important functional information of specific clones.

NGS of immune repertoire is a new research field that provides novel insights about adaptive immune function and the role of T and B cells in various diseases. In addition, it is rapidly entering the clinical world in different disciplines, by tracking disease associated clonotypes. There are currently several available kits for immune repertoire. These are based on similar methodology, regardless of whether they use RNA or DNA as the source. They use similar techniques for calibration, bias corrections, and analysis tools; and thus need to be properly chosen according to specific demands. One of the great advantages of the current kit is that it is a ready-to-use and optimized for human immune repertoire, without need for calibration. In the future we will see more studies using repertoire features as biomarkers for different disease, potentially also leading to development of targeted therapies against these clones.

We believe researchers should consider applying these methods for studying the adaptive immune landscape in different diseases.

Acknowledgments:
None.

Disclosures:
All authors have nothing to disclose

References

1 Bassing, C. H., Swat, W., Alt, F. W. The mechanism and regulation of chromosomal V(D)J recombination. *Cell.* **109 Suppl** S45-55 (2002).
2 Roth, D. B. V(D)J Recombination: Mechanism, Errors, and Fidelity. *Microbiology Spectrum.* **2** (6) (2014).
3 Heather, J. M., Ismail, M., Oakes, T., Chain, B. High-throughput sequencing of the T-cell receptor repertoire: pitfalls and opportunities. *Brief Bioinformatics.* **19** (4), 554-565 (2018).
4 Pabst, O., Hazanov, H., Mehr, R. Old questions, new tools: does next-generation sequencing hold the key to unraveling intestinal B-cell responses? *Mucosal Immunology.* **8** (1), 29-37 (2015).
5 Bashford-Rogers, R. J. M., Smith, K. G. C., Thomas, D. C. Antibody repertoire analysis in polygenic autoimmune diseases. *Immunology.* **155** (1), 3-17 (2018).
6 Lee, Y. N. et al. Characterization of T and B cell repertoire diversity in patients with RAG deficiency. *Science Immunology.* **1** (6) (2016).
7 Werner, L. et al. Alterations in T and B Cell Receptor Repertoires Patterns in Patients With IL10 Signaling Defects and History of Infantile-Onset IBD. *Frontiers Immunology.* **11** 109 (2020).
8 Zhang, J. et al. Immune receptor repertoires in pediatric and adult acute myeloid leukemia. *Genome Medicine.* **11** (1), 73 (2019).
9 Chapman, C. G. et al. Characterization of T-cell Receptor Repertoire in Inflamed Tissues of Patients with Crohn's Disease Through Deep Sequencing. *Inflammatory Bowel Diseases.* **22** (6), 1275-1285 (2016).
10 Werner, L. et al. Altered T cell receptor beta repertoire patterns in pediatric ulcerative colitis. *Clinical and Experimental Immunology.* **196** (1), 1-11 (2019).
11 Bashford-Rogers, R. J. M. et al. Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nature.* **574** (7776), 122-126 (2019).
12 Wu, J. et al. Expanded TCRbeta CDR3 clonotypes distinguish Crohn's disease and ulcerative colitis patients. *Mucosal Immunology.* **11** (5), 1487-1495 (2018).
13 Rosati, E. et al. Identification of disease-associated traits and clonotypes in the T-cell receptor repertoire of monozygotic twins affected by inflammatory bowel diseases. *Journam of Crohn's and Colitis.* 10.1093/ecco-jcc/jjz179 (2019).
14 Allez, M. et al. T cell clonal expansions in ileal Crohn's disease are associated with smoking behaviour and postoperative recurrence. *Gut.* **68** (11), 1961-1970 (2019).
15 Li, S. et al. IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. *Nature Communications.* **4** 2333 (2013).

16 H, I. J. et al. Strategies for B-cell receptor repertoire analysis in primary immunodeficiencies: from severe combined immunodeficiency to common variable immunodeficiency. *Frontiers Immunology.* **6** 157 (2015).
17 Ghraichy, M., Galson, J. D., Kelly, D. F., Truck, J. B-cell receptor repertoire sequencing in patients with primary immunodeficiency: a review. *Immunology.* **153** (2), 145-160 (2018).
18 Zhuang, Y. et al. Application of immune repertoire sequencing in cancer immunotherapy. *International Immunopharmacology.* **74** 105688 (2019).
19 Liu, X. et al. T cell receptor beta repertoires as novel diagnostic markers for systemic lupus erythematosus and rheumatoid arthritis. *Annual Rheumatic Diseases.* **78** (8), 1070-1078 (2019).
20 Wong, G. K., Heather, J. M., Barmettler, S., Cobbold, M. Immune dysregulation in immunodeficiency disorders: The role of T-cell receptor sequencing. *Journal of Autoimmunity.* **80** 1-9 (2017).
21 Delhalle, S., Bode, S. F. N., Balling, R., Ollert, M., He, F. Q. A roadmap towards personalized immunology. *NPJ System Biology and Applications.* **4** 9 (2018).
22 Laubli, H. et al. The T cell repertoire in tumors overlaps with pulmonary inflammatory lesions in patients treated with checkpoint inhibitors. *Oncoimmunology.* **7** (2), e1386362 (2018).
23 Hogan, S. A. et al. Peripheral Blood TCR Repertoire Profiling May Facilitate Patient Stratification for Immunotherapy against Melanoma. *Cancer Immunology Research.* **7** (1), 77-85 (2019).
24 Aversa, I., Malanga, D., Fiume, G., Palmieri, C. Molecular T-Cell Repertoire Analysis as Source of Prognostic and Predictive Biomarkers for Checkpoint Blockade Immunotherapy. *International Journal of Molecular Sciences.* **21** (7) (2020).
25 Hirsch, P. et al. Precision and prognostic value of clone-specific minimal residual disease in acute myeloid leukemia. *Haematologica.* **102** (7), 1227-1237 (2017).
26 De Simone, M., Rossetti, G., Pagani, M. Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges. *Frontiers Immunology.* **9** 1638 (2018).
27 Zemmour, D. et al. Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nature Immunology.* **19** (3), 291-301 (2018).
28 Zheng, C. et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell.* **169** (7), 1342-1356 e1316 (2017).

Figure 1

Click here to access/download;Figure;Figure 1.TIF

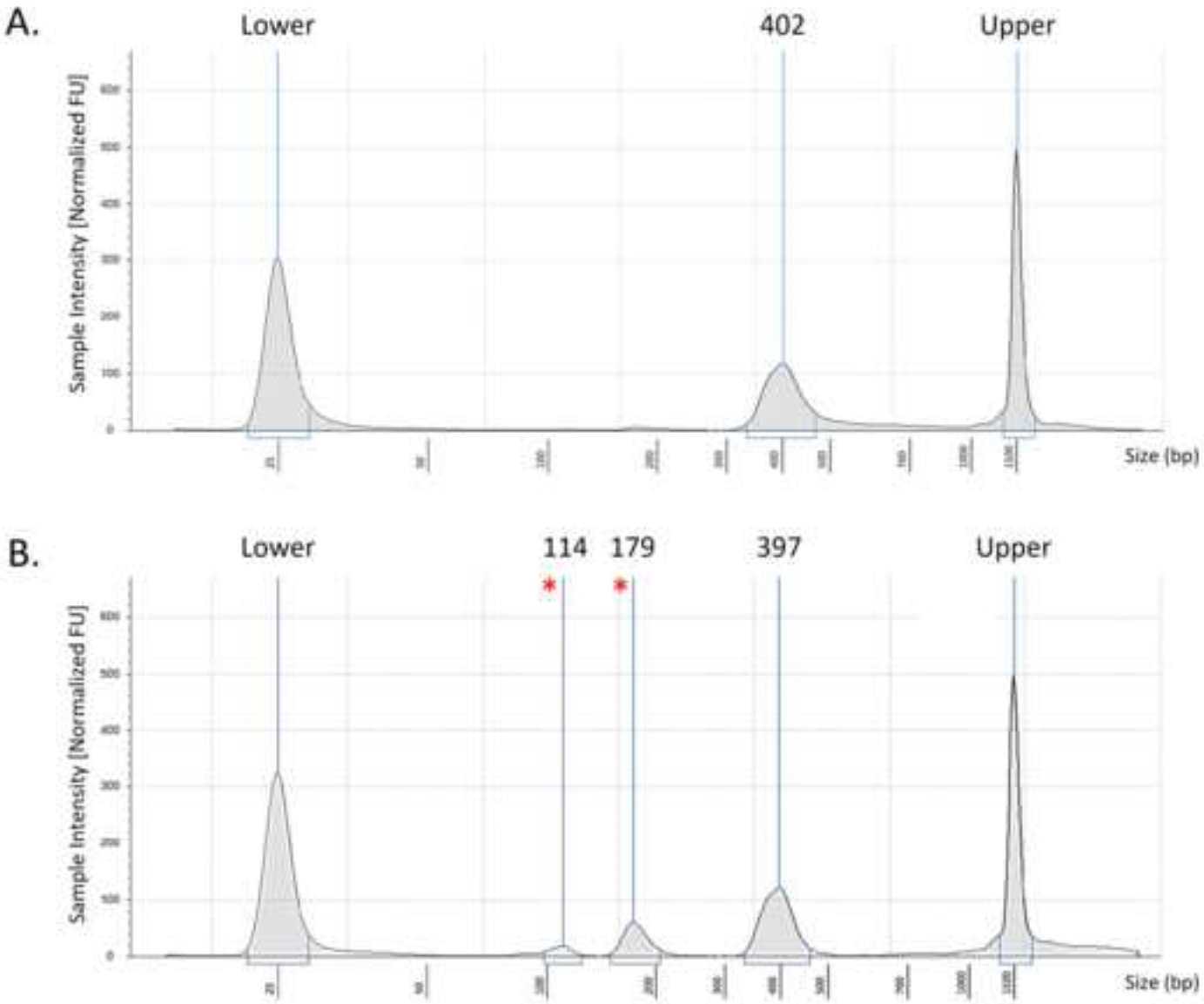
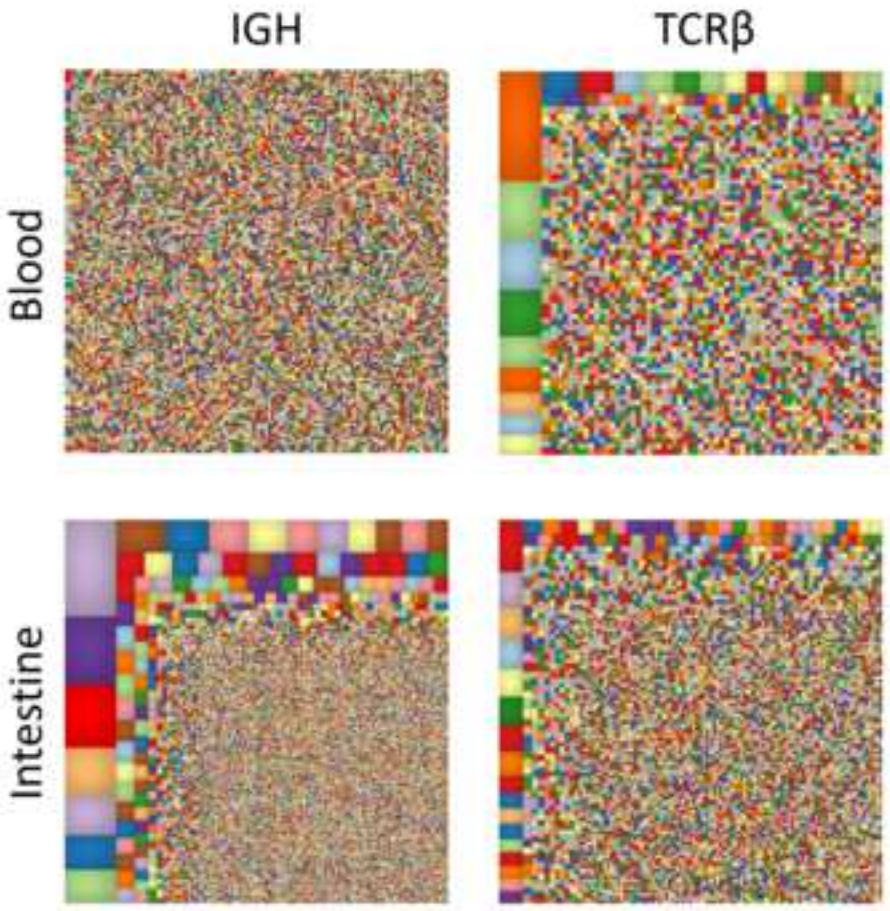


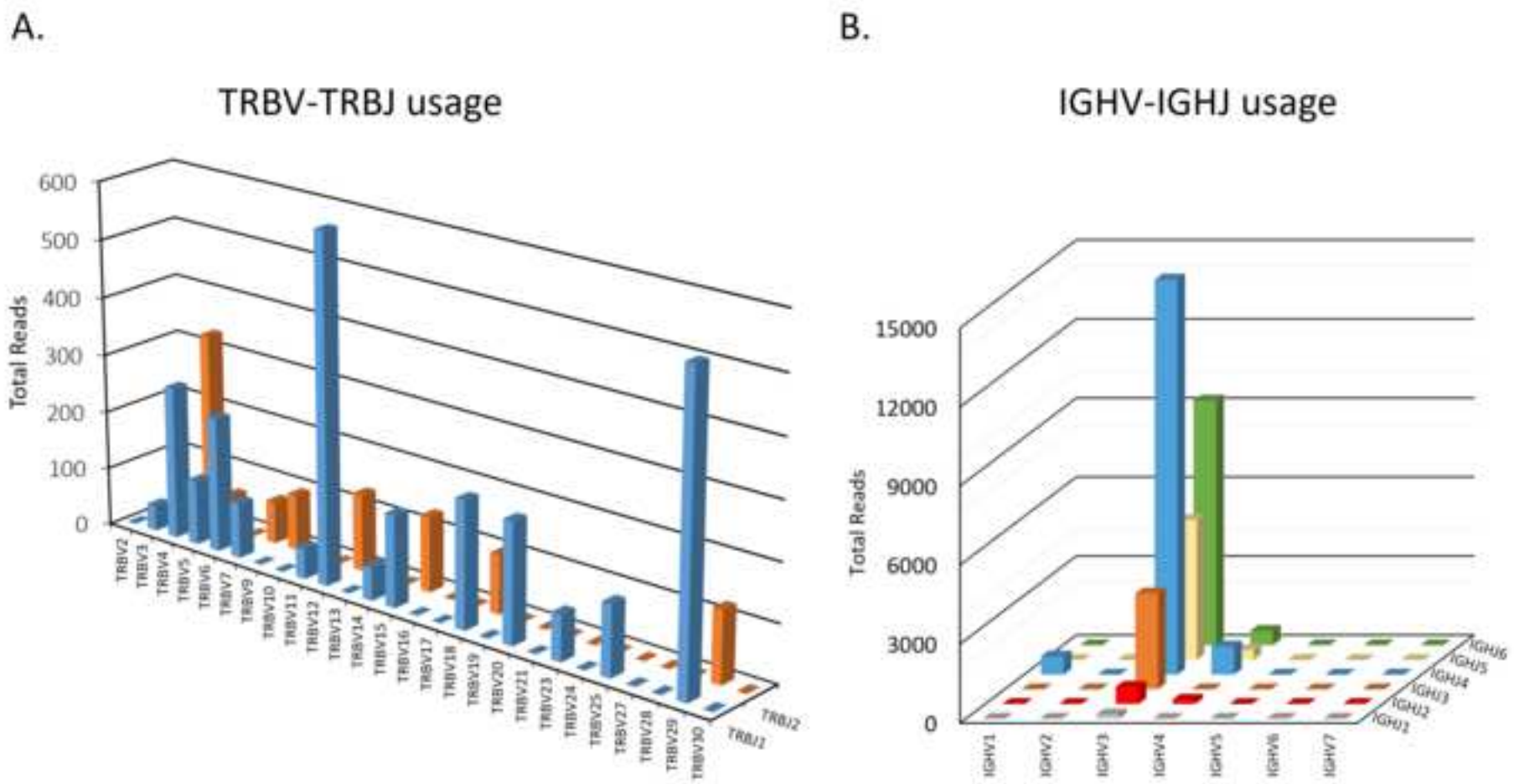
Figure 2

Click here to access/download;Figure;Figure 2.TIF



Figure

Click here to access/download;Figure;Figure 3.TIF



Click here to access/download;Table of Materials;JoVE_Materials.xlsx

Item Name	Company	Catalog Number	Comments/Description
2-propanol	Sigma	I9516-500ML	
1.7 mL micro-centrifuge tubes	Axygen	8187631104051	
15 mL centrifuge tubes	Greiner	188261	
Absolute ethanol	Merck	1.08543.0250	
Amplitaq Gold	Thermo Fisher	N8080241	
AMPure XP Beads	Beckman Coulter	A63881	
Heat block	Bioer	Not applicable	
High Sensitivity D1000 Sample Buffer	Agilent	5067-5603	For Tapestation
High Sensitivity D1000 ScreenTape	Agilent	5067-5584	For Tapestation. Tubes sold seperately
Lymphotrack Assay kit	Invivoscribe	TRB: 70-91210039 IGH: 70-92250019	Each includes 24 indexes
MiSeq Reagent Kit v2 (500 cycle)	Illumina	MS-102-2003	Includes standard flow cell type and all reagents required
MiSeq Sequencer	Illumina	SY-410-1003	
PCR strips	4titude	4ti-0792	
Proteinase K	Invitrogen	EO0491	
Qubit 4 Fluorometer	Thermo Fisher	Q33226	
Qubit dsDNA HS Assay Kit	Thermo Fisher	Q32854	Includes buffer, dye, standards, and specialized tubes
Shaker	Biosan	Not applicable	
Tapestation 2100 Bioanalyzer	Agilent	G2940CA	
ultra pure water	Bio-lab	7501	

Wizard DNA isolation kit	Promega	A1120	Includes cell lysis solution, nuclei lysis solution, and protein precipitation buffer

Rebuttal Letter
Click here to access/download;Rebuttal Letter;Response to reviewers.docx

Editorial Comments:

• Textual Overlap: Portions of the manuscript show overlap with previously published work. Please re-write steps 4.2.1, 4.2.2, 4.2.4 to avoid this overlap.

These sections were re-written and corrected accordingly.

• Protocol Language: Please ensure that all text in the protocol section is written in the imperative voice/tense as if you are telling someone how to do the technique (i.e. “Do this”, “Measure that” etc.) Any text that cannot be written in the imperative tense may be added as a “Note”, however, notes should be used sparingly and actions should be described in the imperative tense wherever possible. Some examples NOT in the imperative”

This was re-checked and verified that all text is indeed in imperative.

• Protocol Detail: Please note that your protocol will be used to generate the script for the video, and must contain everything that you would like shown in the video.
1) Please ensure that all specific details (e.g. button clicks for software actions, numerical values for settings, etc) have been added to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol.

This was done.

2) Please include an ethics statement before your numbered protocol steps indicating that the protocol follows the guidelines of your institutions human research ethics committee.

Our protocol begins with the following note: "This study was approved by the institutional review board at Sheba Medical Center, and informed written consent was obtained from all participating subjects".

• Protocol Highlight: Please double check your highlighting to ensure that it satisfies the following criteria. Please highlight ~2.5 pages or less of text (which includes headings and spaces) in yellow, to identify which steps should be visualized to tell the most cohesive story of your protocol steps.

1) The highlighting must include all relevant details that are required to perform the step. For example, if step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be included in the highlighting.
2) The highlighted steps should form a cohesive narrative, that is, there must be a logical flow from one highlighted step to the next.
3) Please highlight complete sentences (not parts of sentences). Include sub-headings and spaces when calculating the final highlighted length.
4) Notes cannot be filmed and should be excluded from highlighting.

This was done.

• Discussion: JoVE articles are focused on the methods and the protocol, thus the discussion should be similarly focused. Please ensure that the discussion covers the following in detail and in paragraph form (3-6 paragraphs): 1) modifications and troubleshooting, 2) limitations of the technique, 3) significance with respect to existing methods, 4) future applications and 5) critical steps within the protocol.

This was done.

• References:
1) Please spell out journal names.

This was done.

• Commercial Language: JoVE is unable to publish manuscripts containing commercial sounding language, including trademark or registered trademark symbols (TM/R) and the mention of company brand names before an instrument or reagent. Examples of commercial sounding language in your manuscript are Qubit, Lymhotrak Assay kit (Invivoscribe), Illumina MiSeq, Amplitaq, AMPure, Tapestation bioanalyzer (Agilent), Treemap images (www.treemap.com),
1) Please use MS Word's find function (Ctrl+F), to locate and replace all commercial sounding language in your manuscript with generic names that are not company-specific. All commercial products should be sufficiently referenced in the table of materials/reagents. You may use the generic term followed by "(see table of materials)" to draw the readers' attention to specific commercial names.

This was performed. Treemap is an online software and thus was not deleted.

• Table of Materials:
1) Please sort in alphabetical order.

This was performed.

Reviewers' comments:
Reviewer #1:
Manuscript Summary:
The Authors describe a protocol for analysis of immune repertoire. In particular the description is for T-cell receptor beta (TCR) and B-cell receptor heavy chain (IGH) but is applicable also to other TCR and BCR chains. The protocol is based on next generation sequencing (NGS), is based on RNA and applies to analysis of blood and intestinal tissue, but potentially also other tissues. Immune repertoire analyses are an expanding field in the recent years and further protocols for analysing these targets may be of interest for the scientific community.

Major Concerns:
1. I did not understand completely point 2.3 "add 45 µL from the different index tubes into PCR tubes"

are these the primers? Are they specific for IGH/TRB? And do they contain different indexes? Indexes, DNA and polymerase are added, and no buffer? No dNTPs? Are they included in the Amplitaq Gold DNA polymerase?

The Lymhotrack kit includes 24 different index tubes. We believe that dNTPs, MgCl, and buffer are all included in this master mix. This is not specified in the company's protocol, as it is propriety information. However, as we know the DNA polymerase does not include these, and the kits works, this assumption is reasonable. Please see comment #2 below for the clarification we have added to the protocol.

2. I may have missed it, but It is not clear to me weather the method is based on multiplex TCR/BCR PCR primers for the V and J region or on other approach. Please explain this further.

We appreciate this comment, as it was not clarified adequately. We have now added the following note: "The current protocol utilizes a multiplex, PCR-based assay kit compatible with NGS sequencers. The kit contains 24 different indices each targeting conserved regions within the V_β and the J_β regions. This enables a one-step PCR reaction and pooling of different samples. See table of materials for details."

3. How does this method compare with other available methods for TCR/BCR analysis? Multiplex PCR or RNA based methods? It would be good to add some comments on this and add references to previously published protocols for immune repertoire analysis. (as Mamedov et al. 2013 Preparing Unbiased T-Cell Receptor and Antibody cDNA Libraries for the Deep Next Generation Sequencing Profiling)
3.1 If the protocol is based on Multuiplex PCR, which steps has to be taken to correct for primer efficiency bias?

All current commercially available kits for immune repertoire are based on similar methodology, regardless of whether they use RNA or DNA as the source. They use similar techniques for calibration, bias corrections, and analysis tools. This has been stressed in the discussion section.

3.1 Correction for primer efficiency bias and verification of primer balance was performed by testing samples from healthy individuals and comparing their repertoire to the expected wild type repertoires (*Gra OA, et al. Analysis of T-Cell Receptor- Gene Rearrangements Using Oligonucleotide Microchip. J Mol Diagn. 2007 9(2): 249–257 and Lawnicki, C et al. "The distribution of gene segments in T-cell receptor gamma gene rearrangements demonstrates the need for multiple primer sets." The Journal of molecular diagnostics : JMD vol. 5,2 (2003): 82-7.*) However, for the user there is no need for these calibrations and "in-silico" corrections for analysis in these kits.

4. Is it possible to quantify starting T cell number with this protocol?

In the results it was written: "Using genomic DNA as a starting material yields sequence numbers that are representative of cell numbers. Thus, if original DNA quantity is equal, percentage of T cells in a given sample can also be calculated".

5. More a curiosity: Is it possible to adapt this protocol for usage with unique molecular identifiers?

This method could potentially be modified to be used with UMIs. However, the current assay design has been demonstrated in several publications to identify and characterize clonal sequences extremely well without molecular barcodes. Additionally, the use of UMIs is challenging using single-step PCR methods. Since diversity in the full >100bp sequence of gene rearrangements at TCR and IGH loci are generally diverse, the need for UMIs is significantly reduced as compared to a method targeting point mutations (SNVs).

Minor Concerns:
6.Is the method applicable also for the Illumina NovaSeq and/or NextSeq?

The current protocol is not applicable to Nextseq, as maximum length of paired-end sequencing is 150bp, while the MiSeq is 250bp. Regarding Novaseq, it is theoretically possible to work with it. However, as it is designed for 300 million reads it will be quite wasteful, even if uniting several different panels. The Lymphotrack kit is adapted to work with the Myseq. It is limited to 24 indices, and up to 30 million reads. As this is beyond the scope of the current manuscript, it was not added to the text.

7.Is nuclei lysis solution part of the Wizard DNA isolation kit?

See reply#8

8. Protein precipitation buffer is part of which kit?
All reagents for cell lysis and DNA isolation are included in the Wizard DNA isolation kit. As it is not possible to mention brand names in the main text, a comment was added to the Materials list stating: "Wizard DNA isolation kit includes cell lysis solution, nuclei lysis solution, and protein precipitation buffer".

Reviewer #2:
Manuscript Summary:
In the era of high-throughput bulk and single-cell repertoire sequencing, a paper describing a very general protocol using a commercial library preparation kit will not, most likely, attract the attention of the scientific society.

Major Concerns:
There is no reference for the nuclei lysis solution from the step 1.1.2 in the Table of Materials. Generally, referring to the specific supplier and catalog number for each reagent will make the paper more clear and easy to follow.

Please see reply#8 to reviewer #1.

For step 1.2, can authors justify the use of this approach instead of the most commonly used isolation of lymphocytes (either via PBMC isolation or direct isolation)?

The current protocol uses whole-cell lysis. As the Lymphotrack kit contains primers targeting conserved regions within TRB or IGH genes, it thus detects only TRB or IGH gene rearrangements. It is possible to apply these methods when working with PBMC, but not necessarily needed. If one wishes to study immune repertoire in specific T/B cell subsets then isolation of PBMCs is needed followed by cell sorting

Step 1.4.1, it's not clear where 2-propanol should be added.

2-propanol is added to tube containing the supernatant. This was added in step 1.4.1

Step 2, there are multiple typos in "LymphoTrack" and the catalog number is absolutely necessary here since Invivoscribe has a very broad portfolio of products for immune receptor sequencing

We have now added to the Materials Table the catalog numbers of the TRB and IGH Lymphotrack kits detailed in the manuscript. Typos were corrected.

Step 2.4. Since PCR conditions are one of the cornerstones of sequencing library preparation, it would be important to see if authors tried to compare different PCR set-ups and attempted to optimize the protocol.

One of the great advantages of the current kit is that it is a ready-to-use, optimized and calibrated for human immune repertoire. Thus, there is no need for optimization or calibration, as we clearly see in numerous studies. This important remark was added to the discussion.

Step 3.1.4. Could the authors comment on this step in the light of the fact that Beckman Coulter strongly recommends keeping some liquid on this step inside the wells to prevent beads from drying?

According to the protocol of Beckman Coulter for the AMPure Beads: "Remove the last ethanol wash as completely as possible, but add the elution buffer right away and mix the samples."

We have added to this step that elution buffer should be added **immediately** after the 5 minutes.

Step 3.1.7, which elution buffer do the authors suggest?
What concentration of PhiX control libraries did the authors use for the sequencing run?

Elution buffer used is 10 mM Tris-HCL, pH 8.0. PhiX was used at a concentration of 5%. These details were added to manuscript.

Several statements in the Discussion, for example, "It is important during library preparation…" and "At bead clean-up, care should be taken…" belong to the common knowledge and are not appropriate here.

According to the instructions for authors, the discussion should address "modifications and troubleshooting". These statements are important for proper execution of these experiments, and are the most common pitfalls encountered. Thus, we have opted to leave these comments.

Reviewer #3:
Manuscript Summary:
In this work, the authors are reporting the experimental procedure of T and B cell repertoire analysis by use of high-throughput pair-end next-generation sequencing. Although the manuscript is mostly well-written, this reviewer feels that there remain a few issues to be more clarified, particularly in the section of sequencing analysis.

Major Concerns:
1) P9L5: This "sequencing analysis" section lacks more detailed in silico methods by which the authors assign obtained sequences to specific TCR and BCR clonotypes. For example, IGHV assignment is frequently compromised by somatic hypermutation that significantly affects germline CDR1 and CDR2 sequences. Do the authors get the entire IGH sequence to get a handle with such caveats? Perhaps, the addition of a new Figure to show analytical algorithm would be helpful to the readers.

The assays were developed to minimize false negative results due to SHM and rearrangements occurring in the primer sequence by optimizing primer designs and including as many relevant primers as possible. For example, IGH FR1 targets upstream of the CDR1 and CDR2 domains that are prone to somatic hypermutation. This, combined with the proprietary bioinformatics pipeline, ensures that as many unique clones as possible are identified and reported as possible. It should be mentioned that the sequences are aligned to a stored database from IGblast. Most of the in silico workflow is proprietary but it could be summarized roughly as: Read 1 and Read 2 sequences are combined into a single long sequence based on the overlap between the two reads (For MiSeq Only). The frequency of any given sequence is measured, including any sequences containing up to 2-mismatches. Finally, the sequence is compared to a database of IGH sequences to identify the most likely V and J genes. As this is beyond the scope of this protocol, we only reference the IMGT/HighV-QUEST platform as the place for these in silico methodology and analysis.

Minor Concerns:
1) In the abstract, the estimated number of distinct T and B cell clonotypes per human body is described as trillions. However, this description requires appropriate theoretical grounds because the previous mathematical studies have reported more modest estimation (Morbach H, Clin Exp Immunol 2010; Lythe G, J Theor Biol 2016).

We thank the reviewer for this correction. By several accounts, the total number of lymphocytes in humans is estimated to be about $5x10^{11}$. We have changed the number from trillions to billions in the

abstract.

2) P3L14: Regarding the V(D) J recombination, it would be better to discriminate VJ and VDJ patterns in the context of corresponding TCR and BCR subunits.

The current protocol includes only VDJ analysis. From our experience VJ patterns are very low and minimal, and we do not include these. If needed these can be filtered.

3) P12: In the discussion section, it will be more helpful to the readers to include comments on the comparison of DNA-based and RNA-based immunosequencing concerning their merits and limitations.

Please see Reply #3. Reviewer#1

Click here to access/download;Supplemental File (Figures, Permissions, etc.);Jove NGS.Table 1.docx

		Rank	CDR3 Sequence (amino acid)	Count	% Total	V-gene	D-gene	J-gene	Length (amino acid)
TCRβ	Blood	1	CAWSGGTEAFF	1852	3.37	TRBV30	TRBD2	TRBJ1-1	33
		2	CSAIGGAYEQYF	1136	2.07	TRBV20	TRBD2	TRBJ2-7	36
		3	CASSFGPQYNQPQHF	1073	1.96	TRBV28	TRBD1	TRBJ1-5	45
		4	CASSLGGGRDYNEQFF	863	1.57	TRBV13	TRBD2	TRBJ2-1	48
		5	CASSPGGPYEQYF	721	1.31	TRBV28	TRBD2	TRBJ2-7	39
	Intestine	1	CSAFDRGNTEAFF	107	0.86	TRBV20	TRBD1	TRBJ1-1	39
		2	CATWSQGEAFF	72	0.58	TRBV15	TRBD1	TRBJ1-1	33
		3	CSAAGTSGSNEQFF	68	0.55	TRBV29	TRBD2	TRBJ2-1	42
		4	CASREEGSGETQYF	59	0.47	TRBV2	TRBD1	TRBJ2-5	42
		5	CASSIANGPYYEQYF	58	0.47	TRBV19	Unknown	TRBJ2-7	45
IGH	Blood	1	CHAIMTTMVLHYW	8	0.08	IGHV3-73	IGHD3	IGHJ4	39
		2	CASLRFLEWFSFDYW	4	0.08	IGHV3	IGHD3	IGHJ4	45
		3	CGRDFWYAFQW	4	0.06	IGHV3-66	Unknown	IGHJ4	33
		4	CAKGSGSLVADYW	4	0.08	IGHV3-23	IGHD6	IGHJ4	39
		5	CASGSGWWRYYGMDVW	3	0.03	IGHV4-34	IGHD6	IGHJ6	48
	Intestine	1	CARASYGSNSRGFYYGVDVW	1660	4.36	IGHV3	IGHD03	IGHJ6	60
		2	CEWTGSLHYDYW	1172	3.10	IGHV3-2	IGHD03	IGHJ4	36
		3	CARDSKLKWDPPRPQAATCDYW	1081	2.84	IGHV1-18	IGHD01	IGHJ4	66
		4	CARDFRGRVTRVNFYGMDVW	820	2.15	IGHV1-46	Unknown	IGHJ6	60
		5	CARIQRSYSSGWFFWYFDLW	671	1.77	IGHV2-26	IGHD06	IGHJ2	60

Table 1. Top clones of TCRβ and IGH repertoires. Frequency of the five most abundant clones in each of the blood or intestinal samples of a patients with an *IL10RA* mutation and history of severe infantile-onset inflammatory bowel disease. Additionally, V(D)J gene usage and length of the CDR3β are shown.