1 # Ligation-anchored PCR unveils immune repertoire of TCR-beta

2 # from whole blood

3

4 Fan Gao[1#] & Kai Wang[1, 2*]

5

6 [1]: Zilkha Neurogenetic Institute, University of Southern California, Los Angeles, CA 90089, USA

7 [2]: Department of Psychiatry, University of Southern California, Los Angeles, CA 90089, USA

8 [#]: Current address: The Picower Institute for Learning and Memory, MIT, Boston, MA 02139, USA

9 [*]: Correspondence should be addressed to Dr. Kai Wang, 1501 San Pablo Street, Los Angeles, CA 90089,

10 USA, kaiwang@usc.edu

11

12 Emails: Fan Gao fangao@mit.edu, Kai Wang kaiwang@usc.edu

13

14

15

1 Abstract

2 Background

3 As one of the genetic mechanisms for adaptive immunity, V(D)J recombination generates an enormous
4 repertoire of T-cell receptors (TCRs). With the development of high-throughput sequencing techniques,
5 systematic exploration of V(D)J recombination becomes possible. Multiplex PCR method has been
6 previously developed to assay immune repertoire, however the usage of primer pools has inherent bias in
7 target amplification. In our study, we developed a ligation-anchored PCR method to unbiasedly amplify
8 the repertoire.

9 Results

10 By utilizing a universal primer paired with a single primer targeting the conserved constant region, we
11 amplified TCR-beta (TRB) variable regions from total RNA extracted from blood. Next-generation
12 sequencing libraries were then prepared for Illumina HiSeq 2500 sequencer, which provided 151 bp read
13 length to cover the entire V(D)J recombination region. We evaluated this approach on blood samples
14 from patients with malignant and benign meningiomas. Mapping of sequencing data showed 64% to 91%
15 of mapped TCRV-containing reads belong to TRB subtype. An increased usage of TRBV29-1 was
16 observed in malignant meningiomas. Also distinct signatures were identified from CDR3 sequence logos,
17 with predominant subset as 42 nt for benign and 45 nt for malignant samples, respectively.

18 Conclusions

19 In summary, we report an integrative approach to monitor immune repertoire in a systematic manner.

20

21

22

23 Keywords

1   Background

2   Activation of the immune system, together with genomic alternation, including somatic hypermutation

3   and recombination, establishes innate and adaptive immunity in the immune system [1,2]. Both the

4   immunoglobulin (Ig) and the T cell receptor (TCR) loci contain many different V, D and J segments,

5   which are subject to a tightly regulated genomic rearrangement process – V(D)J recombination – during

6   early lymphoid differentiation [3-5]. For a given TCR subtype, complementarity determining region 3

7   (CDR3) is generated by the V(D)J combination at this subtype locus, with the translated protein sequence

8   forming the center of the antigen binding site [6], defining the affinity and specificity of the receptor for

9   individual peptide-MHC complexes. The CDR3 sequence of dominant clones in cancer patient may serve

10   as a signature to diagnose cancer or to classify tumor into subtypes. In addition, signatures could be

11   obtained at the time of disease diagnosis and then monitored on an ongoing basis, to assess the effects of

12   anticancer therapies or for early detection of recurrence [7,8].

13   The importance of monitoring TCR in human health and disease has been increasingly recognized.

14   Recent studies showed that TCR repertoire has been found to affect a wide range of diseases, including

15   malignancy, autoimmune disorders and infectious diseases, and, given the broad involvement of the

16   immune system in almost all of human health and disease, this reach should be expected to expand greatly

17   [9]. However, conventional methods to measure V(D)J recombination have several limitations to prevent

18   detailed characterization of immune repertoire. More recent approaches, such as multiparameter flow

19   cytometry, spectrotyping, or custom-designed real-time PCR assays, are slightly more quantitative and

20   offer higher resolution, but these methods are labor intensive and are unable to offer sequence-level

21   insights as to the exact V(D)J recombination patterns in patients. Solving this problem will enable the

22   wider application of monitoring immune repertoire in clinical settings.

23   With the development of massively parallel, single-molecule sequencing techniques, it has now become

24   feasible to assay V(D)J recombination by next-generation sequencing, as a means to exhaustively profile

25   the immune repertoire within human subjects. For example, the Roche 454 sequencers has been used to

26   measure and clinically monitor human lymphocyte clonality [10], which takes advantage of 454's ability to

27   generate longer sequencing reads that potentially covers V(D)J recombination junction points. Another

28   similar study also used Roche 454 to study human T cell subsets [11]. Note that they separated T cell

29   subpopulations and focused on TCR loci only. However, other investigators have focused on Illumina

30   Genome Analyzer or HiSeq that generates only ~50 bp to ~100 bp reads. For example, a group has

1   developed a short reads assembly strategy to first assemble 50bp sequences and then sample the CDR3

2   diversity in human T lymphocytes from peripheral blood [12,13]. The data analysis involved in such strategy

3   is much less straightforward, but the technology is more accessible and cost-effective in large scale

4   studies. Several additional TCR-Seq studies [13-16] have been extensively reviewed recently. Some of the

5   commonalities include: (1) they almost always used T-cell DNA or RNA as the first starting material; (2)

6   most of the studies use multiplex PCR reactions to enrich the V(D)J recombinations for next-generation

7   sequencing; (3) large-scale T-cell repertoire analysis has been limited to interrogation of a single TCR

8   subunit per sequencing run, although functional antigen-engaging TCRs are heterodimeric proteins

9   comprising both an α and a ß chain.

10   In the current study, we explored the feasibility of using ligation-anchored PCR reaction, rather than

11   multiplex PCR reactions, for unbiased amplification of TCR-beta repertoire. The key aspect of our

12   approach is ligating a synthesized linker oligo sequence (/5Phos/NNN…NNN/3ddC/) to the 3'end of

13   single-stranded cDNA, thus it is applicable to extracted RNA samples. As shown below, the method

14   works with frozen whole blood. Additionally, similar to our previous study examining immunoglobulin

15   repertoire [17], instead of relying on flow cytometry or magnetic beads to isolate T cells or B cell

16   populations from peripheral blood, we attempted to assay RNAs extracted from whole blood directly.

17   Finally, we evaluated sequencing data from the latest Illumina Hi-Seq2500 sequencer, which can generate

18   151bp single-end reads (~160 million reads per run), with a significantly shortened turnaround time (48

19   hours for Rapid Mode) compared to Illumina Hi-Seq2000 or Roche 454 sequencer. Despite the shorter

20   reads than 454 sequencer or Illumina MiSeq sequencer, 151bp read length is still long enough to map

21   V(D)J recombination sites, given appropriate primer design for PCR amplification.

22   To evaluate this approach, we did a pilot study on patients suffered from meningiomas, the most common

23   primary brain tumors in the United States [18,19], and compared difference of the detected TRB repertoire

24   between patients with benign (grade I) and malignant (grade III) tumors. This tumor arises from the

25   membranous layers surrounding the central nervous system (CNS), and is not subject to blood-brain

26   barrier. Several previous studies have reported the presence of both humoral [20,21] and cellular [22,23] immune

27   responses in patients with meningiomas. Indeed, it has been proposed that frequent antibody response

28   against specific antigens in even benign meningioma can serve as diagnostic targets [20]. Therefore, in

29   addition to testing the technical feasibility, our pilot study has the added value of investigating whether

30   there is increased clonality with increased severity of cancer, and whether different patients may share

31   identical or similar CDR3 sequences in clonal T cells.

1

Results

3 **Total RNA extraction and PCR amplification of TRB variable region**

4 To explore all the actively transcribed immune receptor variable regions within whole blood in an

5 unbiased manner, we developed an integrative approach to extract total RNA from whole blood, followed

6 by a PCR protocol to capture the variable region (V-region) of specific immune receptor (**Figure 1A**). In

7 this scheme, after total RNA extraction from whole blood, we generate single-stranded cDNA from

8 reverse transcription of mRNA. After ligation of a synthesized linker oligo sequence to the 3'end of

9 single-stranded cDNA, PCR reaction was carried out using a primer pair (primers P-U and P-VDJ) to

10 target the linker sequence and the consensus starting sequence of immune receptor constant region (C-

11 region), respectively. The P-VDJ primer was specifically designed such that it did not have high sequence

12 homology to any other region other than the TCR locus, and that it avoids any known SNPs. In our proof-

13 of-concept study, total RNA extracted from whole blood of a healthy donor was used to test our

14 methodology on capturing TRB variable region (**Figure 1B**). For samples that were subject to both

15 reverse transcription (either oligo-dT or random primer mix) and linker ligation, PCR reaction resulted in

16 a predominant band with fragment size around 500 bp, roughly the size expected for TRB variable region.

17 For control samples without reverse transcription, this band was not visible. Additionally, controls

18 without linker ligation did not produce this specific band in DNA gel. Thus, the protocol developed in our

19 study should be able to capture transcribed TRB variants from whole blood directly, without separation of

20 T cells first.

21

22 **TRB profiling using Illumina high-throughput sequencing**

23 To further evaluate whether the developed protocol is applicable to profile TRB repertoire of clinical

24 samples, we tested the protocol on several blood samples from patients with malignant and benign

25 meningiomas (**Supplementary Tables 1**). The fact that immune responses are present in patients with

26 meningiomas [20,21] suggests that immune receptors such as TCR may respond to this tumor type and serve

27 as a biomarker for diagnosis. We followed the same protocol to profile TRB V-region variants using

28 previously collected whole blood samples (stored at -80°C for more than one year) from two malignant

29 patients and two benign patients, respectively. Compared to controls without linker ligation, all four

1   reverse transcribed samples with ligation showed the expected predominant band (~500bp, **Figure 2A**),

2   suggesting the robustness of this approach. The amplified fragments were further used for next-generation

3   sequencing library preparation.

4   Purified PCR amplicons were barcoded with Illumina TruSeq kit for high-throughput sequencing. The

5   latest Illumina HiSeq-2500 sequencer was used for sequencing, with load of ~30% PhiX for proper

6   matrixing and phasing into the lane. Collected sequencing data was filtered, and the selected reads were

7   mapped using IgBLAST (see **Methods, Supplementary Tables 2&3**). Of note, the majority of the reads

8   with either P-U or P-VDJ end tag contained TRB fragments (64% to 91% of the mapped TCRV-

9   containing reads belong to TRB subtype). Other subtypes of TCR, such as TRA, TRD, TRG, only

10   accounted for a small portion of the TCRV-containing reads (**Figure2B**). Therefore, ligation-anchored

11   PCR approach is indeed capable of profiling TRB repertoire of clinical samples.

12

13   **V-segment usage in benign and malignant patients.**

14   We next compared the usage of different V-segments from TRB aligning reads (either end-tagged with P-

15   U or P-VDJ primers) in patients with malignant and benign meningiomas. Interesting to note, two

16   malignant samples were grouped together in both P-U and P-VDJ tagged V-segment distribution heat

17   maps, suggesting overall similarity in V-segment usage (**Figure 3A**). Although most of the V-segments

18   did not present consistent change of percentages in the total population between malignant and benign

19   samples (**Figure 3A**), one segment TRBV29-1 showed drastic increase in the malignant group (top

20   ranked gene from SAM analysis of MeV software suite, two-class unpaired, FDR<2%), revealing

21   expansion of this particular V-segment in whole blood of malignant patients. Consistent with this

22   discovery, when we further sorted out the translationally productive V(D)J reads within P-VDJ tagged

23   read pool (see Methods), we also noticed that TRBV29-1 level was elevated in malignant samples

24   (**Figure 3B**), with 5 to 7-fold increase of this population. In addition to TRBV29-1, another V-segment

25   TRBV15 also showed an increase in malignant phenotype. Therefore, the detected clonal expansion of

26   TRBV29-1 possibly reflects human immune response to malignant transformation of meningiomas.

27

Characterization of "immune signature" – CDR3

In addition to explore potential clonal expansion, we are also interested to evaluate our experimental protocol in capturing the sequences of the CDR3 regions of TCR. From high-throughput sequencing data, we deduced the unique DNA sequence logos of the FR3-CDR3 junction part (**Figure 4A**) with the last 20 nt of FR3 and the first 60 nt into CDR3 included, and evaluated if any signatures were recurrent features that can separate malignant from benign patients. Initial observation at positions 54 and 56 (**Figure 4A**) revealed that the most frequent bases were C and G in the malignant samples; however, they switched to base T in the benign samples. Positions 54 to 56 encode an amino acid that is in frame to TRB coding sequence; in malignant samples, the encoded amino acid was Gln (Q), whereas in benign group, it changed to Tyr (Y). We thus generated protein sequence logos of FR3-CDR3 junction (**Figure 4B**) for further comparison. Of note, a conserved Cys (C) at the end of FR3 region (position 6) was separated from a QXFGXGTRL motif in CDR3 by 11 and 10 amino acids in malignant and benign sample groups, respectively. In a previous study [12], two major CDR3 subsets with 42 nt and 45 nt were identified in a pooled normal human leukocyte RNA sample. In our study, the protein sequence logos of meningioma blood samples revealed predominant 42-nt subset in benign group and 45-nt subset in malignant group. It is possible that FR3-CDR3 junction presents specific sequence features of different patients with distinct profiles of immune response.

Exploration of V-J segment pairing in V(D)J recombination

To study V-J pairing in V(D)J recombination, normalized pairing frequencies between different V- and J-segments were presented in a Circos plots [24] (**Figure 5**). Contrast to only a few predominant V-J pairing events in benign samples, malignant samples showed expansion of more V-J pairing events. Furthermore, pairing of TRBV5-6 with TRBJ1-1 represented one of the major pairing events detected in all samples. Three pairing events including TRBV7-3 with TRBJ2.2, TRBV30 with TRBJ2-6, and TRBV16 with TRBJ1-3, occurred more frequently in the two malignant samples compared to benign ones, possibly representing clonal expansion of specific TCR-expressing T cells in both malignant samples. In summary, despite the use of 151bp sequencing length, our protocol enabled the delineation of the V-J segment pairing in an efficient manner.

1 Discussion and conclusions

2 In the current study, we presented an integrated approach by using single primer PCR together with next-
3 generation sequencing to interrogate immune repertoire of TCR-beta. We have demonstrated the technical
4 feasibility to use this system to infer immune repertoire, using whole blood from four meningiomas
5 samples. By aligning reads to a sequence database of germline V-genes, D-genes and J-genes, the usage
6 of different V-gene segments was quantified. Interestingly, comparison between malignant and benign
7 groups revealed an increased usage of TRBV29-1 in malignant meningiomas. Analysis of CDR3 region
8 sequence logos revealed addition of 1 amino acid between FR3-CDR3 junction and a consensus motif in
9 malignant group. However, we caution that these observations were made on a small number of samples,
10 and they may not have any biological significance. Our purpose is to use these data to demonstrate the
11 technical feasibility of "single-primer" interrogation of immune repertoire, rather than determining what
12 differs between malignant and benign tumors.

13 There are several unique aspects of our protocol, compared to previous studies. First of all, total RNA is
14 extracted directly from frozen blood samples for profiling, thus the procedure can be easily adapted for
15 clinical application. Second, by using ligation-anchored PCR for amplification, all the recombination
16 events at a particular immune gene locus can be amplified in an unbiased manner. Finally, sequencing of
17 barcoded libraries through Illumina Hi-Seq 2500 ensures fast turn-around time (less than 48 hours) and
18 good sequencing depth (~160 million reads per lane) at a relatively low cost.

19 There are several major limitations of our protocol as well. First, due to the need to add in 3'-adaptors to
20 the cDNA terminus for ligation-anchored PCR, our method relies on RNA samples, which are less readily
21 available and more vulnerable to degradation, compared to genomic DNA samples. However, in the
22 current study, we used frozen whole blood samples and still obtained satisfactory results, suggesting that
23 it is practically feasible to use this method in real-world clinical settings. Second, although Illumina Hi-
24 Seq 2500 provided longer read length (151bp) than Illumina Hi-Seq 2000 (101bp) to cover immune
25 signature region CDR3, longer read length is still needed to cover the entire variable region of immune
26 genes. Thus the latest Illumina MiSeq with 250bp chemistry should be more suitable for profiling
27 immune repertoire. Third, our sequencing data showed higher sequencing depth of malignant sample
28 libraries (010_M, 221_M) compared to benign ones (**Supplementary Table 2**). We acknowledge that
29 uneven mixing of the constructed libraries for sequencing could be the reason. To address the issue of
30 sequencing depth, we performed additional data analysis using the first 7 million sequencing reads of
31 010_M and 221_M data (**Supplementary Figure, Supplementary Table 4&5**). Of note, the results from

1   the 7-million-read data were consistent with the conclusions from the whole sequencing data, indicating

2   that the sequencing depth in our study is deep enough. Fourth, we specifically designed primers to target

3   TCR-beta isoforms. Our sequencing data contained some other TCR isoforms as well, suggesting the

4   presence of some minor cross reactions. Our future studies may design specific primers for other types of

5   isoforms as well, to investigate whether it is feasible to interrogate all major TCR isoforms in the same

6   sequencing run. Finally, the sample size in our study is small, which prevents us to draw any statistically

7   significant conclusions on clonal expansion of TRB subtypes associated with malignant transformation of

8   meningiomas; however, the results from high-throughput sequencing data analysis do support validity of

9   the developed integrated approach.

10   In conclusion, we have developed a ligation-anchored PCR approach to interrogate immune repertoire,

11   and demonstrated its technical feasibility and effectiveness in capturing the TCR-beta landscapes. Further

12   development of this technology may enable a comprehensive delineation of immune repertoire, including

13   other forms of TCRs as well as immunoglobulins.

14

Methods

15

Sample collection, RNA extraction and cDNA synthesis

16

17   Peripheral blood from patients with meningiomas from the USC Brain Tumor Bank was collected for the

18   study. For assay optimization purpose, peripheral blood from a healthy donor was also collected. The

19   Institutional Review Board (IRB) at the University of Southern California reviewed and approved the

20   study. Informed consent was obtained for all participants.

21   Total RNA was extracted from ice thawed 400μL of frozen blood (stored at -80°C) using TRIzol

22   extraction reagent (Life Technologies, Grand Island, NY). For each TRIzol extracted RNA sample, 2mL

23   of anhydrous ethanol was added before passing through an RNeasy-mini column (QIAGEN, Valencia,

24   CA) with in-column DNase I digestion performed. Purified total RNA was eluted in ddH$_2$O for further

25   cDNA synthesis using M-MuLV reverse transcriptase kit (NEB, Ipswich, MA) following the

26   recommended protocol, and stored at -20°C.

27

Linker ligation and PCR amplification of TRB variants

For each reverse transcribed product, AMPure XP magnetic beads (Beckman Coulter, Danvers, MA) was added to the mixture, followed by 80% ethanol wash to remove primers introduced in reverse transcription step. Each sample was eluted in 30μL EB buffer (QIAGEN, Valencia, CA) with 1μL RNase H (10U/μL, Epicentre, Madison, WI) and 2μL RNase A (5μg/μL, Epicentre, Madison, WI) added to digest RNA templates for 1 hour at 37C, followed by a heat-inactivation step for 5 min. at 95°C.

Ligation of the linker primer to RNase-treated cDNA sample (ssDNA) was carried out in a total volume of 60μL using T4 RNA ligation kit (Promega, Madison, WI) following vendor recommended protocol. The linker primer was synthesized (Integrated DNA Technologies, San Diego, CA) based on the sequence and modification in a previous study [25]. The mixture was incubated overnight at room temperature. For control reaction, ddH$_2$O instead of T4 RNA ligase (Promega, Madison, WI) was added.

After overnight incubation, the ligation product was incubated with AMPure XP beads for magnetic purification. Each sample was eluted in 20μL ddH$_2$O. PCR reaction was performed using HotStart PCR reagent kit (EMD Chemicals, Billerica, MA) with 2μL of ligated ssDNA as PCR template. In a reaction mixture without primer loaded, a universal primer targeting the linker (GCG GCC GCT TAT TAA CCC) was first added for 2nd strand synthesis at 95°C for 2.5 min., 70°C for 5min. and 4°C forever. Then another primer targeting the consensus sequence of TRB constant region (GAC CTC GGG TGG GAA CAC) was added for 40 cycles of PCR reaction with annealing at 65°C and extension at 70°C.

NGS library preparation and sequencing

Illumina TruSeq Sample Prep Kit (Illumina, San Diego, CA) was used for library preparation. Briefly, amplicons were purified before sequential steps of end-repair, adenylation of 3'end and adaptor ligation using the recommended protocol. Finally, the barcoded libraries were enriched with 10 cycles of PCR reactions and cleaned with AMPure XP beads. The quality of prepared libraries was assessed using Agilent High Sensitivity DNA Chip on a Bioanalyzer (Agilent Technologies, Santa Clara, CA). The bar-coded libraries were quantified with KAPA Library Quantification Kit (Kapa Biosystems, Woburn, MA) before mixing for sequencing in a single lane on the Illumina HiSeq 2500 system. The raw FASTQ data files of 151 bp single-end reads were collected for downstream analysis.

1 **Analysis of sequencing data**

2 For the raw FASTQ data collected from high-throughput sequencing, sequencing quality analysis was

3 first performed using FASTQC [26] to ensure the mean values of sequence quality (Phred Score) for each

4 base is greater than 32. Then a custom PERL script was used to select the reads with the starting

5 sequences containing either P-U (with additional 5 bp extended to the ligation site) or P-VDJ primer

6 sequence. Unlike genome or exome sequencing, the nature of sequencing data from immune gene

7 variable regions requires 1) mapping against a sequence database containing V-gene, D-gene and J-gene

8 segments; and 2) aligning the V(D)J junctions formed. For mapping of the sequenced reads, recently

9 released IgBLAST [27] was employed to identify reads that can be mapped to germline V, V-D, V-J, or V-

10 D-J segments. The default parameters were used for mapping. Due to lack of a standalone version of

11 IgBLAST for TCR gene analysis, the sequencing data (reformatted to FASTA files) were split into small

12 batches and uploaded to the IgBLAST web server for mapping. Hierarchical clustering analysis was

13 carried out using MEV, which is a component of the TM4 suite of microarray analysis tools [28]. Sequence

14 logo graphs were generated using standalone version of WebLogo (weblogo.berkeley.edu) [29]. Circos

15 maps were plotted using Circos Table Viewer web server (http://mkweb.bcgsc.ca/tableviewer/) [24]. All

16 other plots were generated using R.

17

18 **Acknowledgements**

19 We are grateful for the patients with meningiomas for donating their blood samples for research purposes.

20 We thank William J. Mack for help in obtaining whole blood samples and phenotype information.

21

22 **Conflict of Interests Statement**

23 The authors declare no conflict of interests.

24

25 **Authors' contribution**

26 FG & KW designed the study. FG performed the experiment, collected and analyzed the data. FG & KW

27 wrote the paper. All authors read and approved the manuscript.

1

References

1.   Di Noia, J.M. & Neuberger, M.S. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* **76**, 1-22 (2007).

2.   Davis, M.M. & Bjorkman, P.J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395-402 (1988).

3.   Tonegawa, S. Somatic generation of antibody diversity. *Nature* **302**, 575-81 (1983).

4.   Schatz, D.G. & Swanson, P.C. V(D)J recombination: mechanisms of initiation. *Annu Rev Genet* **45**, 167-202 (2011).

5.   Bowen, S., Wangsa, D., Ried, T., Livak, F. & Hodes, R.J. Concurrent V(D)J recombination and DNA end instability increase interchromosomal trans-rearrangements in ATM-deficient thymocytes. *Nucleic Acids Res* **41**, 4535-48 (2013).

6.   Paul, P.E. *Fundamental Immunology*, (Lippincott Williams & Wilkins, Philadelphia, PA, 2008).

7.   Cooper, Z.A. et al. BRAF inhibition is associated with increased clonality in tumor-infiltrating lymphocytes. *Oncoimmunology* **2**, e26615 (2013).

8.   Kim, J.A. et al. CDR3 spectratyping identifies clonal expansion within T-cell subpopulations that demonstrate therapeutic antitumor activity. *Surgery* **136**, 295-302 (2004).

9.   Woodsworth, D.J., Castellarin, M. & Holt, R.A. Sequence analysis of T-cell repertoires in health and disease. *Genome Med* **5**, 98 (2013).

10. Boyd, S.D. et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. *Sci Transl Med* **1**, 12ra23 (2009).

11. Wang, C. et al. High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proc Natl Acad Sci U S A* **107**, 1518-23 (2010).

12. Freeman, J.D., Warren, R.L., Webb, J.R., Nelson, B.H. & Holt, R.A. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res* **19**, 1817-24 (2009).

13. Warren, R.L. et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res* **21**, 790-7 (2011).

14. Robins, H.S. et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099-107 (2009).

15. Mariani, S. et al. Comprehensive assessment of the TCRBV repertoire in small T-cell samples by means of an improved and convenient multiplex PCR method. *Exp Hematol* **37**, 728-38 (2009).

16. Warren, R.L., Nelson, B.H. & Holt, R.A. Profiling model T-cell metagenomes with short reads. *Bioinformatics* **25**, 458-64 (2009).

17. Gao, F. et al. Characterizing immunoglobulin repertoire from whole blood by a personal genome sequencer. *PLoS ONE* **8**, e75294 (2013).

18. Pham, M.H. et al. Molecular Genetics of Meningiomas: A Systematic Review of the Current Literature and Potential Basis for Future Treatment Paradigms. *Neurosurg Focus* **30**, E7 (2011).

19. Mawrin, C. & Perry, A. Pathological classification and molecular genetics of meningiomas. *J Neurooncol* **99**, 379-91 (2010).

20. Comtesse, N. et al. Complex humoral immune response against a benign tumor: frequent antibody response against specific antigens as diagnostic targets. *Proc Natl Acad Sci U S A* **102**, 9601-6 (2005).

21. Heckel, D. et al. cDNA cloning and chromosomal mapping of a predicted coiled-coil proline-rich protein immunogenic in meningioma patients. *Hum Mol Genet* **6**, 2031-41 (1997).

1 22. Pees, H.W. & Seidel, B. Cell-mediated immune response of patients with meningiomas defined in
2   vitro by a [3H]proline microcytotoxicity test. *Clin Exp Immunol* **24**, 310-6 (1976).
3 23. Rossi, M.L., Cruz Sanchez, F., Hughes, J.T., Esiri, M.M. & Coakham, H.B. Immunocytochemical
4   study of the cellular immune response in meningiomas. *J Clin Pathol* **41**, 314-9 (1988).
5 24. Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. *Genome Res* **19**,
6   1639-45 (2009).
7 25. Troutt, A.B., McHeyzer-Williams, M.G., Pulendran, B. & Nossal, G.J. Ligation-anchored PCR: a
8   simple amplification technique with single-sided specificity. *Proc Natl Acad Sci U S A* **89**, 9823-5
9   (1992).
10 26. [FASTQC]. FASTQC.
11 27. Ye, J., Ma, N., Madden, T.L. & Ostell, J.M. IgBLAST: an immunoglobulin variable domain
12   sequence analysis tool. *Nucleic Acids Res* **41**, W34-40 (2013).
13 28. Saeed, A.I. et al. TM4 microarray software suite. *Methods Enzymol* **411**, 134-93 (2006).
14 29. Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. WebLogo: a sequence logo generator.
15   *Genome Res* **14**, 1188-90 (2004).

16

17

18

Figure legends

Figure 1. PCR method utilizing single target-specific primer to explore actively transcribed immune genes. A), The scheme of an integrative approach to explore immune repertoire from whole blood (P-U stands for a universal primer sitting on the linker, whereas P-VDJ represents a TRB-specific primer targeting the consensus sequence in the constant region right after J-segment); B), DNA gel image of PCR amplified variable region (V-region) of TRB gene from a healthy donor under different conditions.
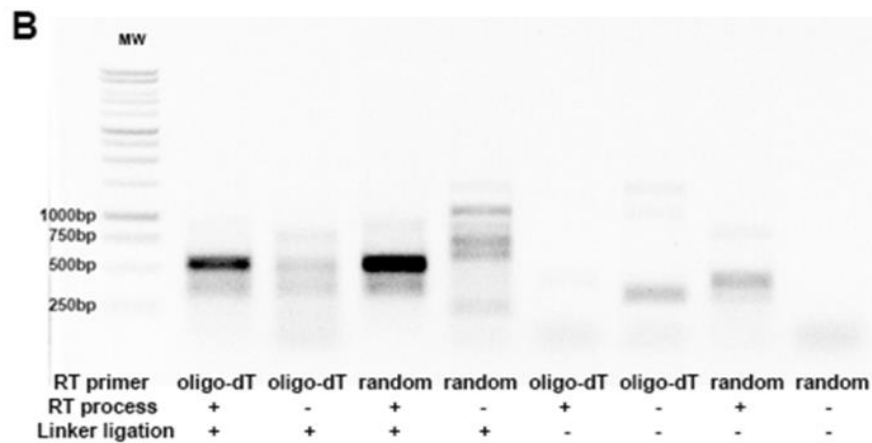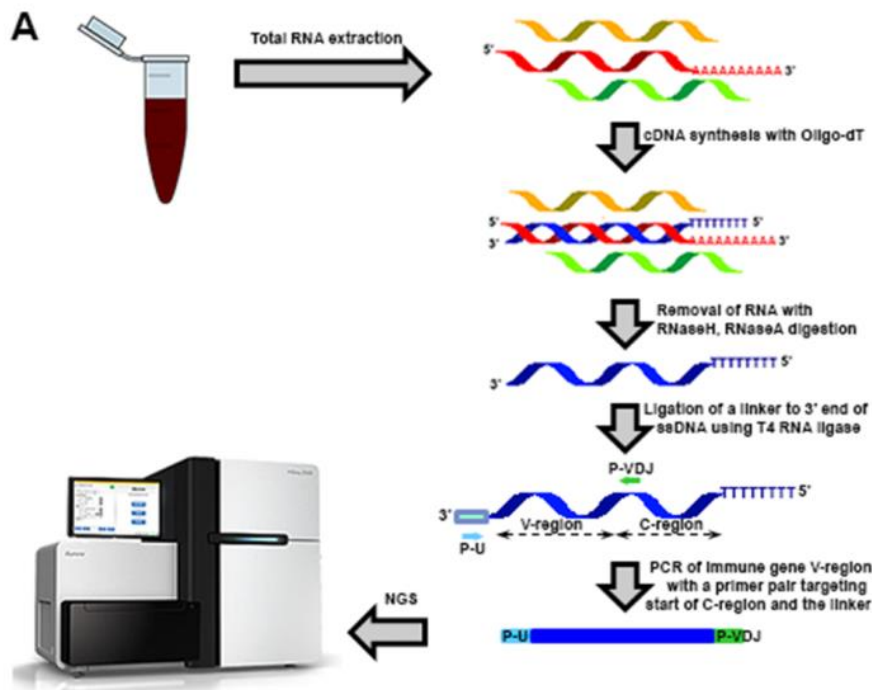
Figure 2. Specificity of the integrative approach in amplifying V-region of TRB gene from whole blood of patients with meningiomas (malignant and benign samples were labeled with M and B, respectively). A), DNA gel image of amplicons from malignant and benign patients under linker ligation or control condition; B), Subtype population of sequenced reads that are mappable to V-region of TCR in each meningioma patient (results are based on the reads with P-U or P-VDJ end tags).
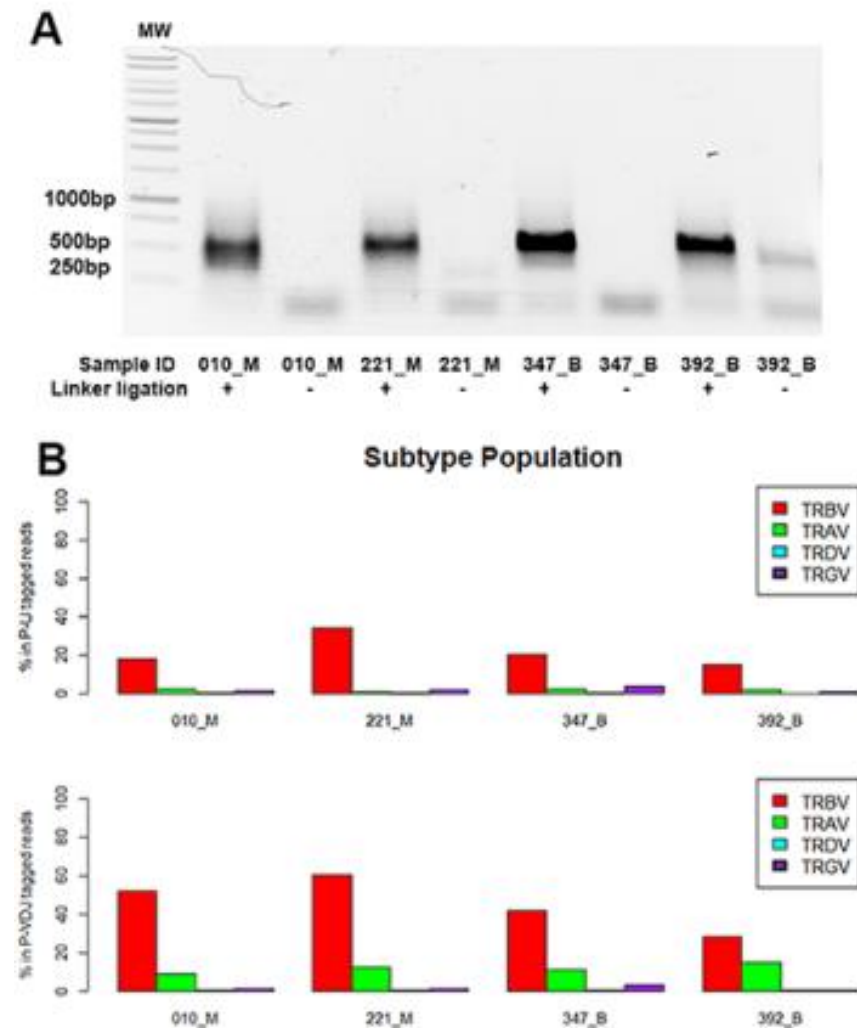
Figure 3. The usage of different V-segments in TRB gene in meningioma blood samples. A), Hierarchical clustering of V-segment usage from both P-U and P-VDJ primer-aligned reads in each patient (color scale represents % of total primer-aligned reads). B), V-segment usage from V(D)J-containing reads that are translationally productive (no stop codon). Arrows highlight expanded V-segments in malignant samples.
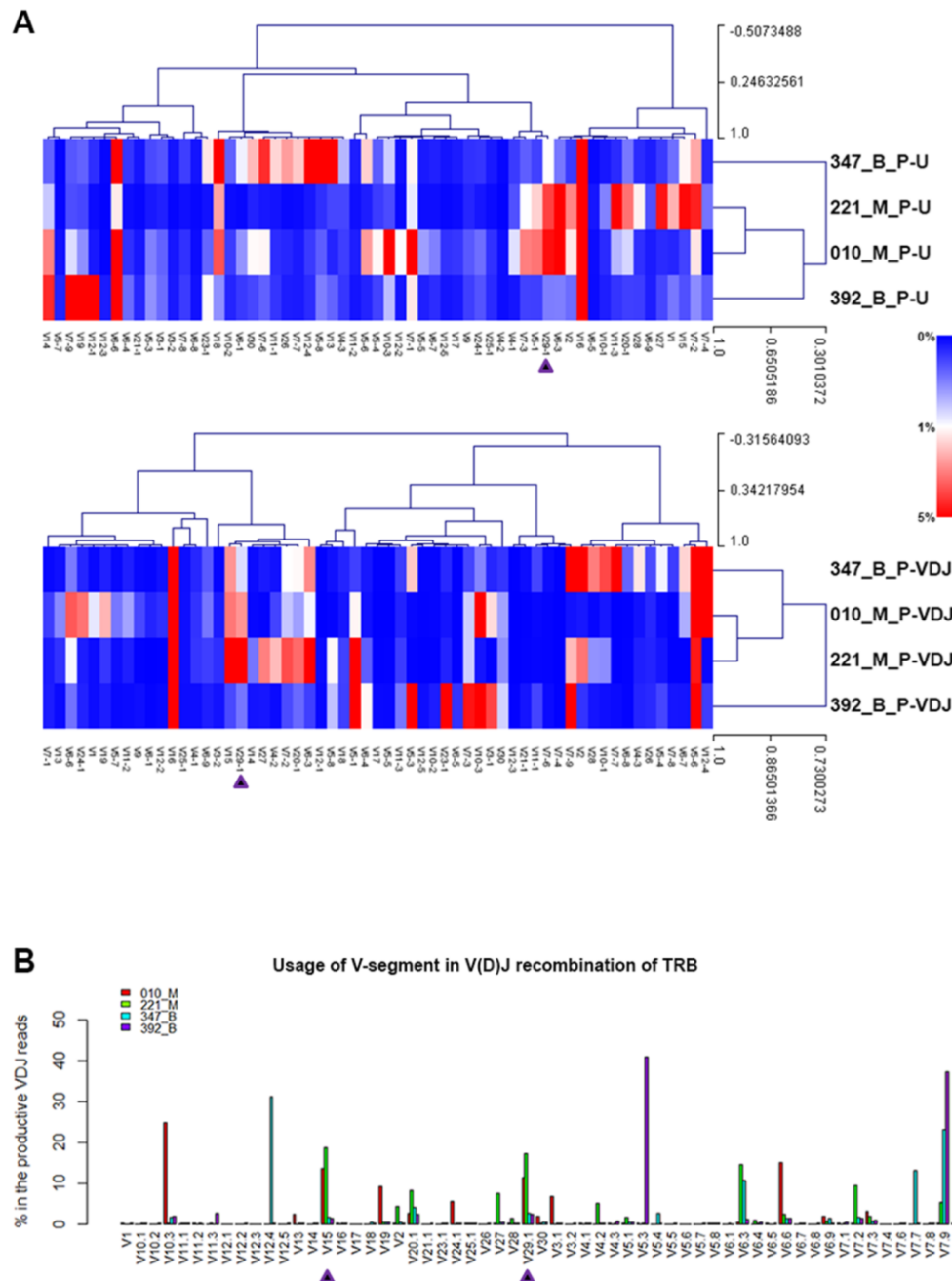
Figure 4. Sequence logos for detected FR3 and CDR3 portions of meningioma blood samples. A) Visualized in the DNA sequence logos at the last 20 nt of FR3 and the first 60 nt into CDR3. A purple-colored box highlights positions 54 to 56 in the DNA sequences. B), In the protein sequence logos, a highly conserved Cys(C) at the end of FR3 is at position 6th of the translated sequence. Purple-colored boxes enclose QXFGXGTRL motif in malignant and benign protein sequences.
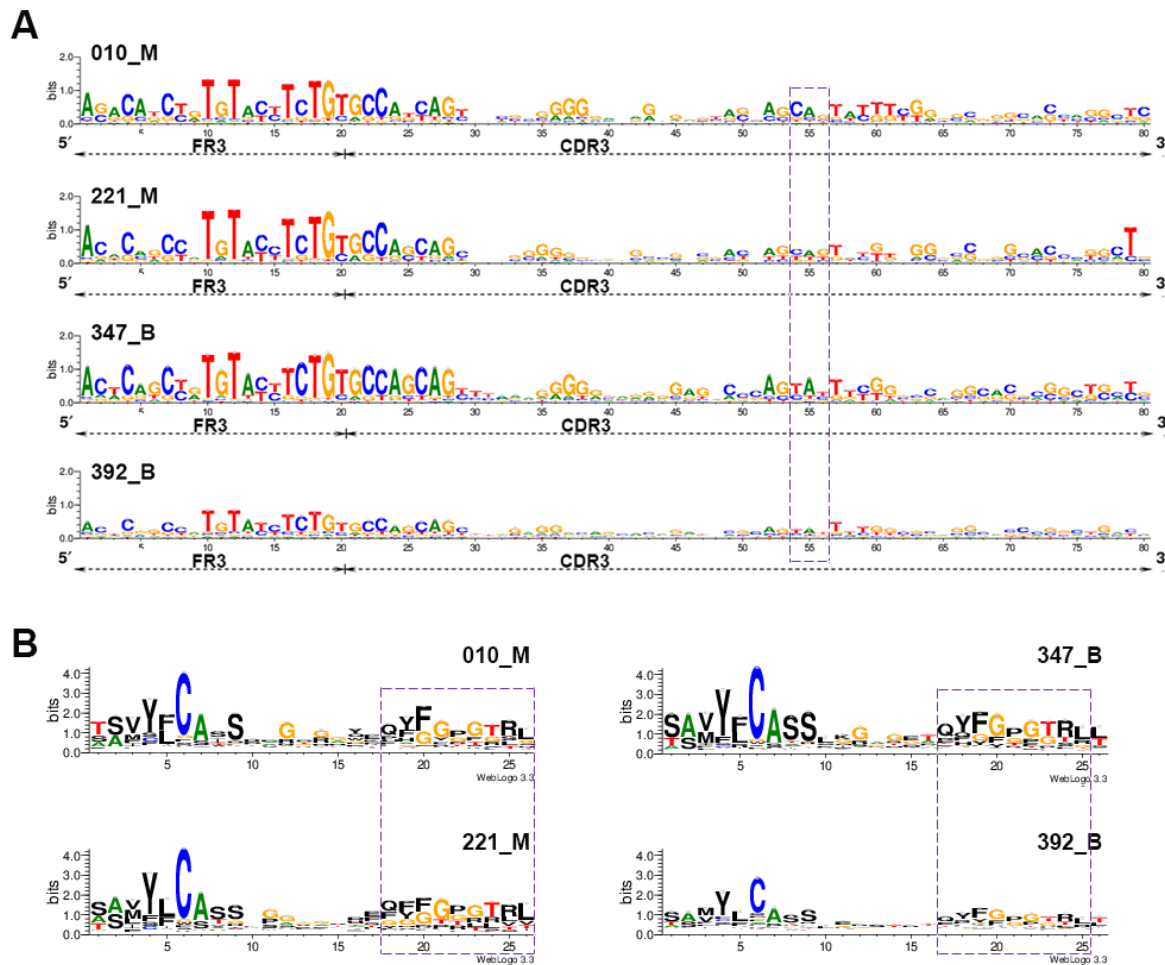
Figure 5. Circos maps illustrating pairing frequencies of V-segments and J-segments from V(D)J-containing reads (all identified segments are ranked by usage frequencies and presented in a clock-wise orientation from the top of the circle, in the order of V-segments and J-segments).