Highly reproducible TCR profiling using RNA from rhesus macaque PBMC

John Beckford*, Bryan Bell & Andrew Farmer

Takara Bio USA, Inc., San Jose, CA 95131, USA

*Corresponding Author: john_beckford@takarabio.com

Abstract

Non-human primates (NHP) such as the rhesus macaque *(Macaca mulatta)* have long been key translational models in biomedical research because of their genetic and physiological similarity to humans[1–4]. Studies using rhesus macaques have contributed significantly to our understanding of T-cell responses to vaccines, cancer, and infectious diseases. More recently, these NHP have emerged at the forefront of COVID-19 vaccine research.

Increasingly complex information can now be gleaned from immune system processes. High-throughput TCR sequencing (TCR-seq) profiles T-cell responses in exquisite detail. A comprehensive understanding of immune responses in such a closely related organism as the rhesus macaque would be a significant advance in science.

Numerous tools exist for performing TCR-seq in human samples[5–6], but equivalent tools for rhesus samples have been lacking. Because rhesus macaques often serve as surrogates in the lead-up to human studies, there is an industry need for a complete TCR-seq solution for these NHP samples.

Due to strong species homology, our SMARTer® Human TCR a/b Profiling Kit v2 (TCRv2) can generate high-quality TCR sequencing libraries using human or rhesus macaque RNA.

Methods

Library preparation from two different species

Frozen PBMC from three rhesus macaque donors were purchased from HumanCells Biosciences (Cat. # M5-100-C10M). RNA from PBMCs was extracted using NucleoSpin RNA kit (Takara Bio, Cat. # 740955.50). Human control RNA was included with the kit. Mixed *TRA/TRB* libraries were generated using the TCRv2 kit as per the user manual. Human and rhesus macaque libraries were prepared in triplicate using 10 ng or 100 ng RNA.

Library quantification and analysis

Following purification and size selection, libraries were sequenced on both the Illumina® MiSeq® (600-cycle V3 cartridge; Cat. # MS-102-3003) and Illumina NextSeq® (300-cycle v2.5 Mid Output; Cat. # 20024908).

As directed by TCRv2 protocol, sequencing data analysis was performed with the Cogent™ NGS Immune Profiler Software v1 using the human genome as the alignment reference. A parallel analysis was performed with third-party MiXCR v3 software using standard parameters and the rhesus macaque genome reference (obtained from IMGT) for read alignment.

1 TCRv2 library preparation workflow

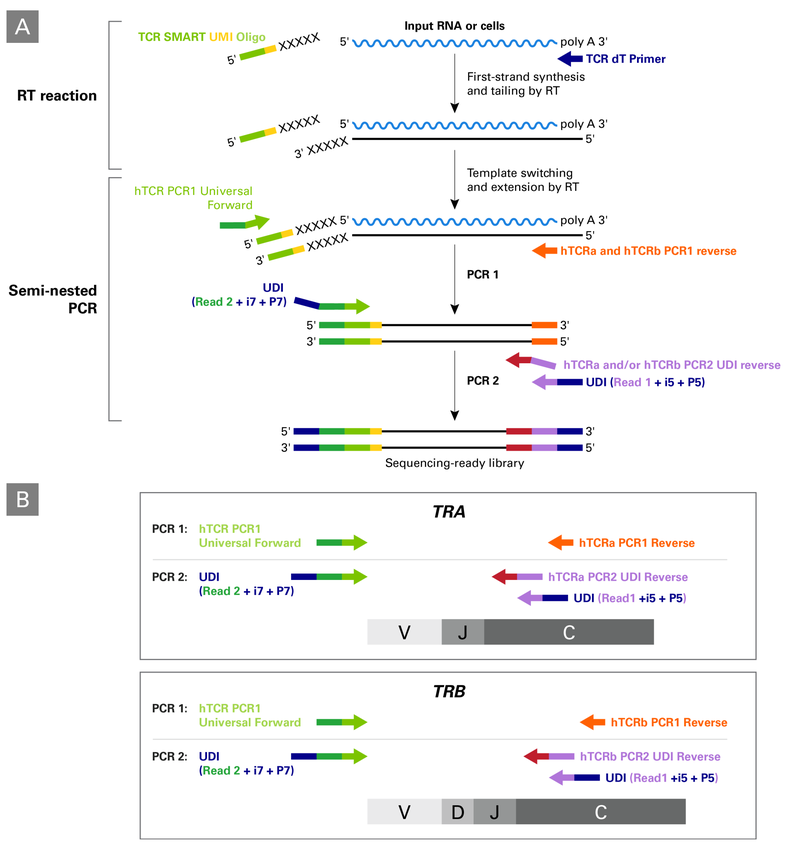

Figure 1. SMARTer Human TCR a/b Profiling Kit v2 workflow—when rhesus macaque RNA input was used, no changes were made to accommodate the change of species.

Panel A. First-strand cDNA synthesis is dT-primed and performed by the MMLV-derived SMARTScribe™ Reverse Transcriptase (RT), which adds proprietary nontemplated nucleotides (**XXXXX**) upon reaching the 5' end of each mRNA template. The TCR SMART® UMI Oligo anneals to these nontemplated nucleotides, allowing it to serve as a template for the incorporation of the SMART Sequence (**light green**) and UMI (**yellow**) at the 5' end of each cDNA molecule by the RT. This is the template-switching step. The resulting cDNA then undergoes two rounds of TCR-specific PCR amplification. In the PCR 1 step, the forward primer (hTCR PCR1 Universal Forward) anneals to the SMART Sequence (**light green**) and adds the Read 2 sequence (**dark green**). The reverse primer (hTCRa/hTCRb PCR1 Reverse, **orange**) anneals near the 3' end of the *TRA/TRB* transcript, within the constant region (see details in Panel B). The PCR 2 step uses the PCR 1 product as a template. PCR 2's forward primer has complementarity to the SMART Sequence (**light green**) and Read 2 sequence (**dark green**) and adds Unique Dual Indexes (UDIs, **dark blue**). PCR 2 uses semi-nested reverse primers. One reverse primer (hTCRa/hTCRb PCR2 UDI Reverse) anneals to an upstream portion (**red**) of the constant region relative to PCR 1's reverse primer (**orange**) and it adds the Read 1 sequence (**lavender**). The other reverse primer has complementarity to the Read 1 sequence (**lavender**) and adds UDIs (**dark blue**). Thus, a sequencing-ready library of *TRA* and *TRB* transcripts is created.

Panel B. A more detailed view shows the approximate annealing locations of all PCR 1 and PCR 2 primers within the V(D)J and constant regions of the *TRA* and *TRB* sequences. By priming from the SMART Sequence and the constant region, the PCR 1 step ensures the capture of the entire V(D)J region and specifically amplifies all *TRA/TRB* transcripts in an unbiased manner. The PCR 2 step further enriches *TRA/TRB* transcripts, adds UDIs, and decreases amplicon size, facilitating read-through on Illumina instruments. The UDIs include adapter and index sequences that are compatible with Illumina sequencing platforms and allow for multiplexing of up to 192 samples in a single flow cell lane.

Takara Bio USA, Inc.
United States/Canada: +1.800.662.2566 • Asia Pacific: +1.650.919.7300 • Europe: +33.(0)1.3904.6880 • Japan: +81.(0)77.565.6999

2 Human vs. rhesus macaque TCRv2 library size distribution comparison

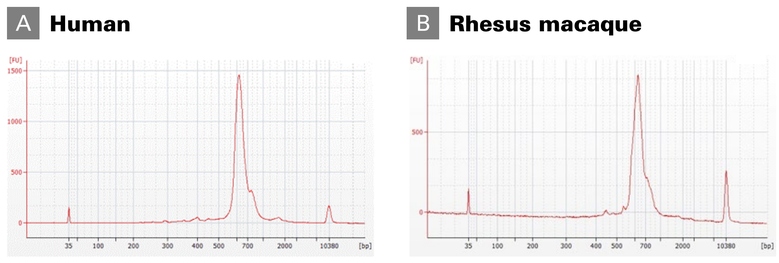

Figure 2. TCR sequencing libraries with similar size distributions are produced from human and rhesus macaque RNA inputs. *TRA/TRB* mixed libraries were prepared using 100 ng total RNA isolated from PBMC of a human (**Panel A**) or a rhesus macaque donor (**Panel B**). Completed libraries were run through the Agilent 2100 Bioanalyzer system.

TCRv2 libraries shown in Figure 2 were generated in triplicate, using high-quality RNA (RIN ≥8). Hereafter, libraries made from 10 ng RNA from rhesus macaque PBMC are referred to as **RM-10** while those made from 100 ng are **RM-100**.

3 Reproducible TCR clonotype detection in the rhesus macaque

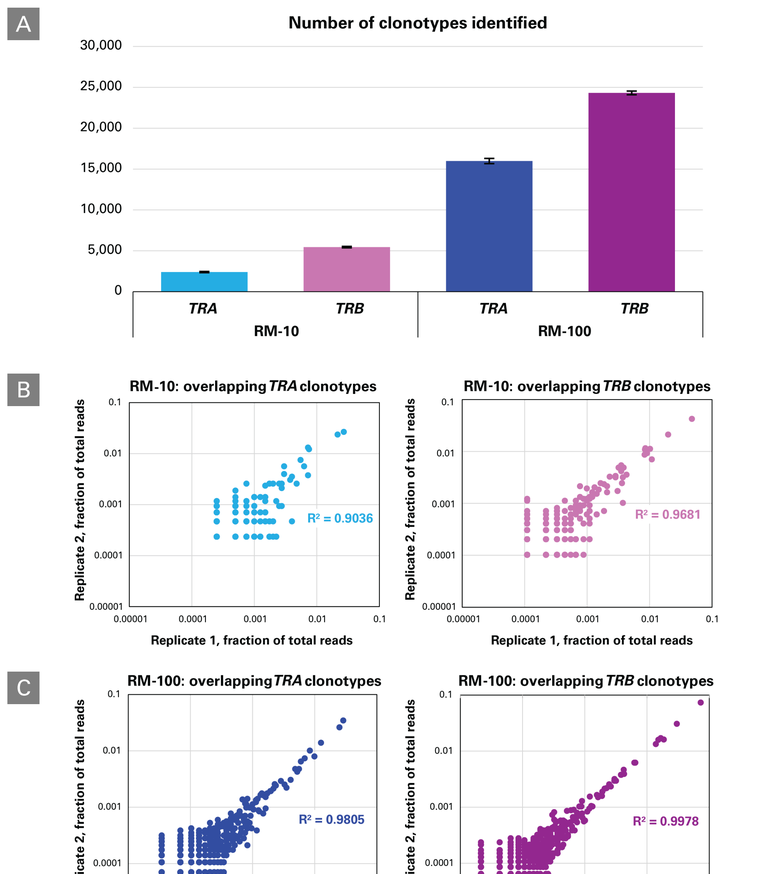

Figure 3. TCRv2 sequencing libraries generated from rhesus macaque samples show highly reproducible clonotype detection. RM-10 and RM-100 libraries were created in triplicate from PBMC of three rhesus macaque donors. Full-length *TRA* and *TRB* sequences were obtained using the Illumina MiSeq platform. Average reads/sample were 3 million for RM-10 and 1 million for RM-100. Clonotypes were identified using Cogent NGS Immune Profiler v1 software. The total number of clonotypes per input amount and gene target were identified in one representative donor (**Panel A**). Error bars represent the standard deviation among triplicates. Two replicates from a representative donor's RM-10 (**Panel B**) and RM-100 (**Panel C**) libraries were chosen for pairwise comparisons of overlapping *TRA* and *TRB* clonotypes. Scatterplots compare fraction of total reads assigned to common clonotypes between replicates.

4 Compatibility of Cogent NGS IP software with TCR-seq data from rhesus macaque samples

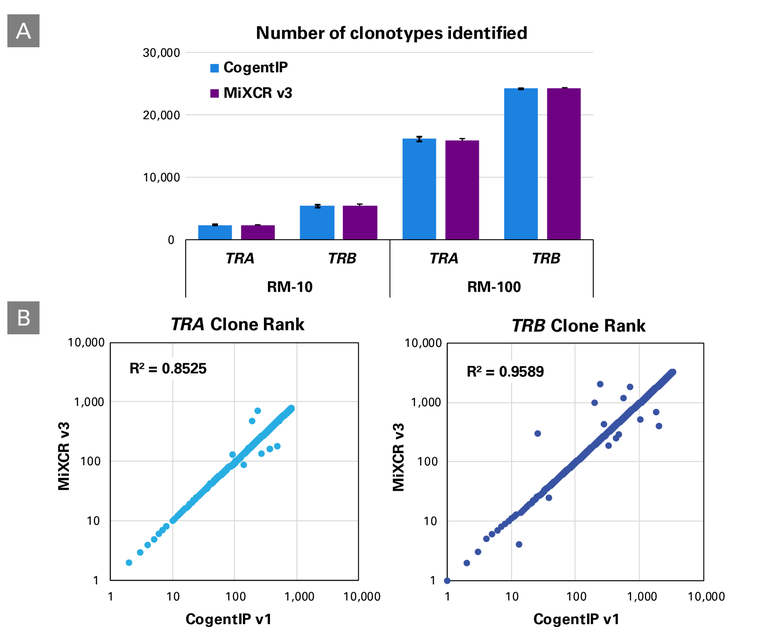

Figure 4. Cogent NGS Immune Profiler v1 and MiXCR v3 produced similar results when used to analyze TCR-seq data from rhesus macaques. RM-10 and RM-100 libraries were created in triplicate from PBMC of three rhesus macaque donors. Full-length *TRA* and *TRB* sequences were obtained using the Illumina MiSeq platform. Average reads/sample were 3 million for RM-10 and 1 million for RM-100. CogentIP and MiXCR v3 software were used in parallel to identify clonotypes from the FASTQ files generated from sequencing RM-10 and RM-100 libraries. The total number of *TRA* and *TRB* clonotypes across donors is shown in **Panel A**. Error bars represent the standard deviation among triplicate library preparations. Rankings of matched clonotypes for each TCR chain, as generated by the two platforms, were investigated. A pairwise comparison of *TRA* and *TRB* rankings within two of the three replicates from a representative donor's RM-10 libraries is shown in **Panel B**.

5 Thorough V-J pair capture in rhesus macaque

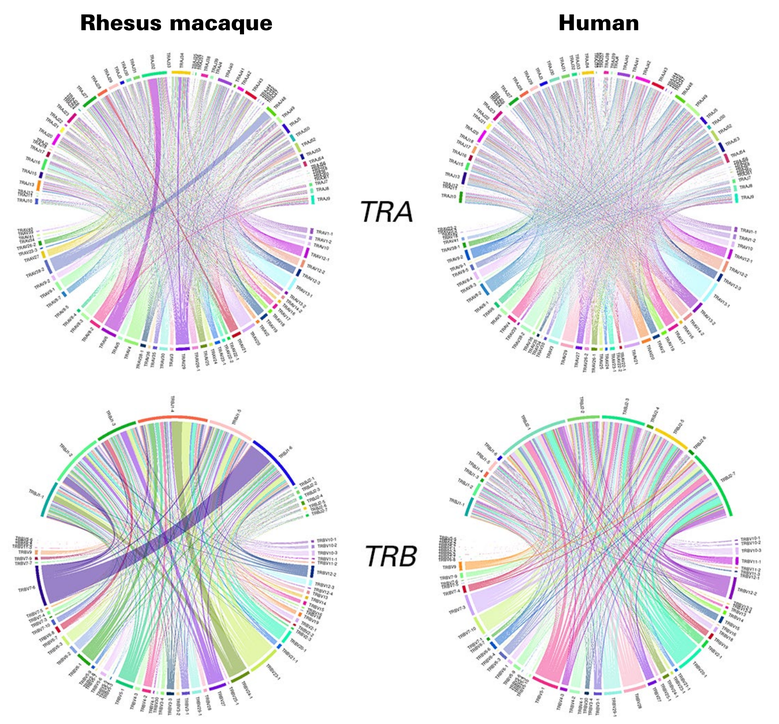

Figure 5. The full spectrum of *TRA* and *TRB* V-J pairs was captured in rhesus macaque and human samples. The RM-100 library of one rhesus macaque donor and a similar library created using 100 ng of human PBMC RNA are depicted. Each chord diagram depicts the distribution of the indicated *TRA* and *TRB* Variable-Joining (V-J) segment combinations. Each arc (on the periphery of each diagram) represents a V or J segment and is scaled lengthwise according to the relative proportion of the segment in the dataset. Each chord (connecting the arcs) represents a set of clonotypes that include the indicated V-J combination and is weighted according to the relative abundance of that combination in the dataset.

6 Flexibility in sequencing instrument choice

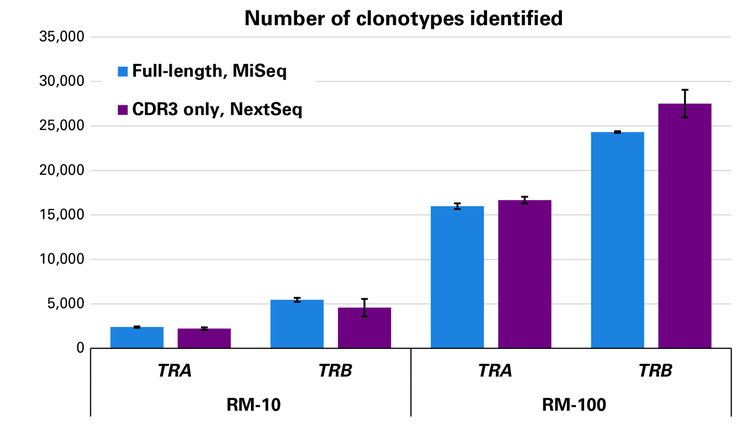

Figure 6. TCRv2 libraries from rhesus macaque samples show highly reproducible clonotype detection, whether sequencing full-length *TRA* and *TRB*, or CDR3 regions only. RM-10 and RM-100 libraries were first run on the MiSeq platform, with full-length sequencing of *TRA* and *TRB* performed at 2 x 300 bp reads. Average reads/sample were 3 million for RM-10 and 1 million for RM-100. The same libraries and genes were also sequenced using the NextSeq platform for CDR3-region sequencing at 2 x 150 bp reads. Average reads/sample were 5 million for RM-10 and 4 million for RM-100. The total number of *TRA* and *TRB* clonotypes detected for each condition in a representative donor is shown. Error bars represent the standard deviation among triplicate library preparations.

Conclusions

- The SMARTer Human TCR a/b Profiling Kit v2 (TCRv2) generates high-quality, *TRA/TRB*-mixed libraries using RNA from human or rhesus macaque. No changes to the protocol or reagents are required.
- Cogent NGS Immune Profiler v1 software (CogentIP), our bioinformatic solution for TCR data analysis, is validated for use with TCR-sequencing data from rhesus macaque samples.
- TCRv2 libraries generated from rhesus macaque are suitable for full-length *TRA/TRB* sequencing on the Illumina MiSeq platform or CDR3-only sequencing on high-throughput platforms such as Illumina NextSeq.
- Together, the SMARTer Human TCR a/b Profiling Kit v2 and Cogent NGS Immune Profiler v1 software represent an industry first: a complete TCR repertoire analysis solution for use with rhesus macaque samples.

References

1. Giudicelli, V. *et al.* IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic acids res.* **34,** Database issue D781–4 (2006).
2. Thiel, C. *et al.* Structure and diversity of the T-cell receptor alpha chain in rhesus macaque and chimpanzee. *Hum. immunol.* **43,** 85–94 (1995).
3. Levinson, G. *et al.* Sequence and diversity of rhesus monkey T-cell receptor beta chain genes. *Immunogenetics* **35,** 75–88 (1992).
4. Bolotin, D. A. *et al.* MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).
5. Zhang, Z. *et al.* A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7,** 203–14 (2000).
6. Barennes, P. *et al.* Benchmarking of T cell receptor repertoire profiling methods reveals large systematic biases. *Nat. Biotechnol.* **39,** 236–245 (2021).

Scan to download your copy of this poster and view additional posters, or visit takarabio.com/ngs-posters

800.662.2566
Visit us at takarabio.com