7

IDENTIFICATION OF T CELL RECEPTOR αβ SEQUENCES FROM SINGLE T CELLS

Carsten Linnemann[1,4], Gavin M. Bendle[1,4], Sander Kelderman[1], Andrew D. Kaiser[1], Marja Nieuwland[2], Roelof J.C. Kluin[2], Dmitriy M. Chudakov[3], Ron M. Kerkhoven[2] and Ton N.M. Schumacher[1]

[1]Division of Immunology and [2]Central Genomics Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands.
[3]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry RAS, Moscow, Russia.
[4]These authors contributed equally

Unpublished

The composition of T cell receptor (TCR) repertoires at sites of inflammation, such as tumor lesions, has been studied with different methods that describe TCRβ chain diversity. However, a major disadvantage of these approaches is the inability to recover full TCRαβ pairs. This precludes both the further study of TCR specificities and affinities, as well as the potential clinical use of isolated TCRs by TCR gene therapy. Here we report a strategy to efficiently identify TCRαβ cDNA sequences from single T cells by next-generation sequencing. This method will allow analyses of TCR repertoires at sites of immune activity, as well as the generation of TCR libraries for potential clinical use.

| INTRODUCTION

The antigen-specificity of T cells is solely determined by the expression of the clone-specific T cell receptor (TCR) αβ heterodimer that results from the genetic re-arrangement of TCRα- and β-loci. Significant efforts have been made to describe TCR diversity at sites of infection, autoimmune disease or tumor growth, in order to better understand the role T cells play in these processes. In many studies, detection of TCR Variable (V)β genes by flow cytometry[1] and Complementarity determining region 3 (CDR3) size spectratyping[2] has been used to assess the diversity of TCR repertoires, but with relatively low resolution. The development of next-generation sequencing (NGS) has greatly increased the resolution of TCR profiling, by allowing massive parallel sequencing of TCR Vβ CDR3[3]. However, the use of V-gene specific primers in most of these studies may lead to a bias in the resulting data, in particular because of the high number of single-nucleotide polymorphisms (SNPs) in TCR V-elements[4]. While this limitation has recently been overcome by combining 5'-RACE with NGS, allowing high-resolution profiling of TCRβ repertoires[5], all these methods do not identify complete TCRαβ pairs. Such identification of TCRαβ pairs does offer two distinct advantages. First, it allows one to study the role T cells play in biological processes, e.g. at sites of infection, autoimmune disease or within tumors. Second, identified TCRαβ pairs can also be used for therapeutic interventions, such as the transfer of tumor-reactivity of T cells by TCR gene transfer[6] or the engineering of regulatory T cells[7]. Recently, a method has been reported that utilizes panels of primers against TCRα and TCRβ V-genes to obtain TCRαβ cDNA from single human T cells[8]. However, due to the substantial polymorphisms among TCR V-genes such primer sets may lead to a bias in identified TCRs[4].

Here we reported an alternative protocol for the identification of TCRαβ sequences from single T cells that is independent of TCR V-gene specific primers and allows TCRαβ sequence identification by Illumina NGS.

| RESULTS

In order to identify TCRαβ sequences directly from single T cells we modified a protocol established by Tang *et al*. for the transcriptome analysis of single mouse blastomeres by NGS[9, 10]. Most importantly, we incorporated a flow cytometry sort

to isolate single T cells and used primers specific for TCRα (TRAC) and TCRβ (TRBC) constant domains for first-strand cDNA synthesis. As reported by Tang *et al.*, Terminal deoxynucleotidyl transferase (TdT) is used to mediate a template-switch. Double-strand cDNA is subsequently amplified by PCR (**Fig. 1**).

To determine whether TCRαβ sequences could be identified with this protocol, we isolated single CD8$^+$ T cells with peptide-Major histocompatibility complex (pMHC)-multimers from both CMV pp65-specific as well as from Minor-Histocompatibility-antigen HA2-specific T cell clones. TCRαβ sequences from these T cell clones have been described previously[11]. We analyzed the cDNA obtained from sorted single CD8$^+$ T cells by real-time Taqman PCR using probes specific for the CDR3 regions of TCRα- and β-chains of these TCRs. Both TCRα and TCRβ chains were reliably detected in 73% of single cells as defined by Ct-values < 35 for both TCRα and TCRβ (166/226 of single cells analyzed; **Fig. 2a–d**), indicating that TCRαβ mRNA from single T cells was successfully reverse-transcribed to significant levels of cDNA (Average Ct-values from 166 single cells: ~22 for both TCRα and TCRβ).

Next, we tested whether the amount of cDNA obtained from single T cells was sufficient for TCR identification by Illumina NGS. To this purpose, we used cDNA isolated from five single-sorted CMV pp65-specific CD8$^+$ T cells, which had previously been analyzed by real-time PCR Taqman PCR (**Fig. 2a**). For purification of cDNA, we utilized the TCR gene capture library preparation protocols that we previously established[12]. Captured cDNA was subsequently analyzed by paired-end 250bp Illumina chemistry on a MiSeq Personal Sequencer and the resulting sequencing data was analyzed with the IMRA algorithm[13, 14] to identify TCR CDR3 sequences. Of note, identification of the

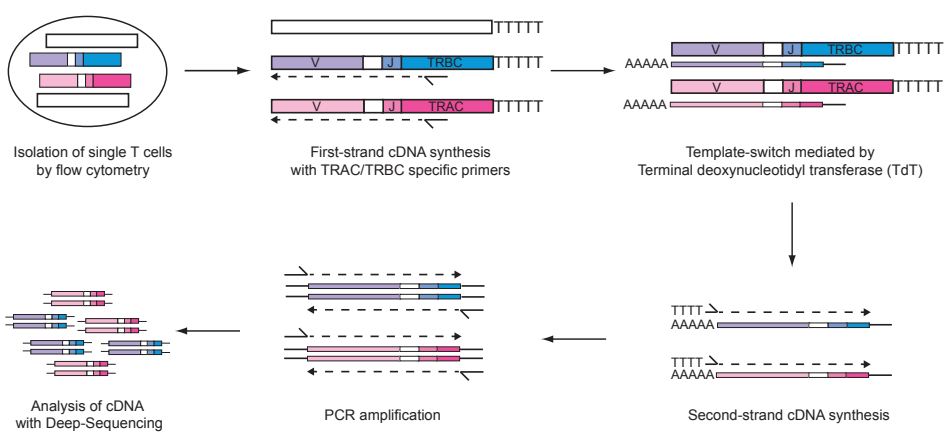

Figure 1. Schematic representation of the strategy to generate TCRαβ cDNA from single T cells. Single T cells are directly sorted into lysis buffer-containing PCR plates by flow cytometry. First-strand cDNA is reverse-transcribed from TCRαβ mRNA using TCRαβ constant domain specific primers. Free primers are degraded and the cDNA first-strand is poly-adenylated by Terminal deoxynucleotidyl Transferase allowing the subsequent synthesis of double-stranded TCRαβ cDNA. TCRαβ cDNA can subsequently be amplified by PCR and analyzed by Next-Generation-Sequencing.

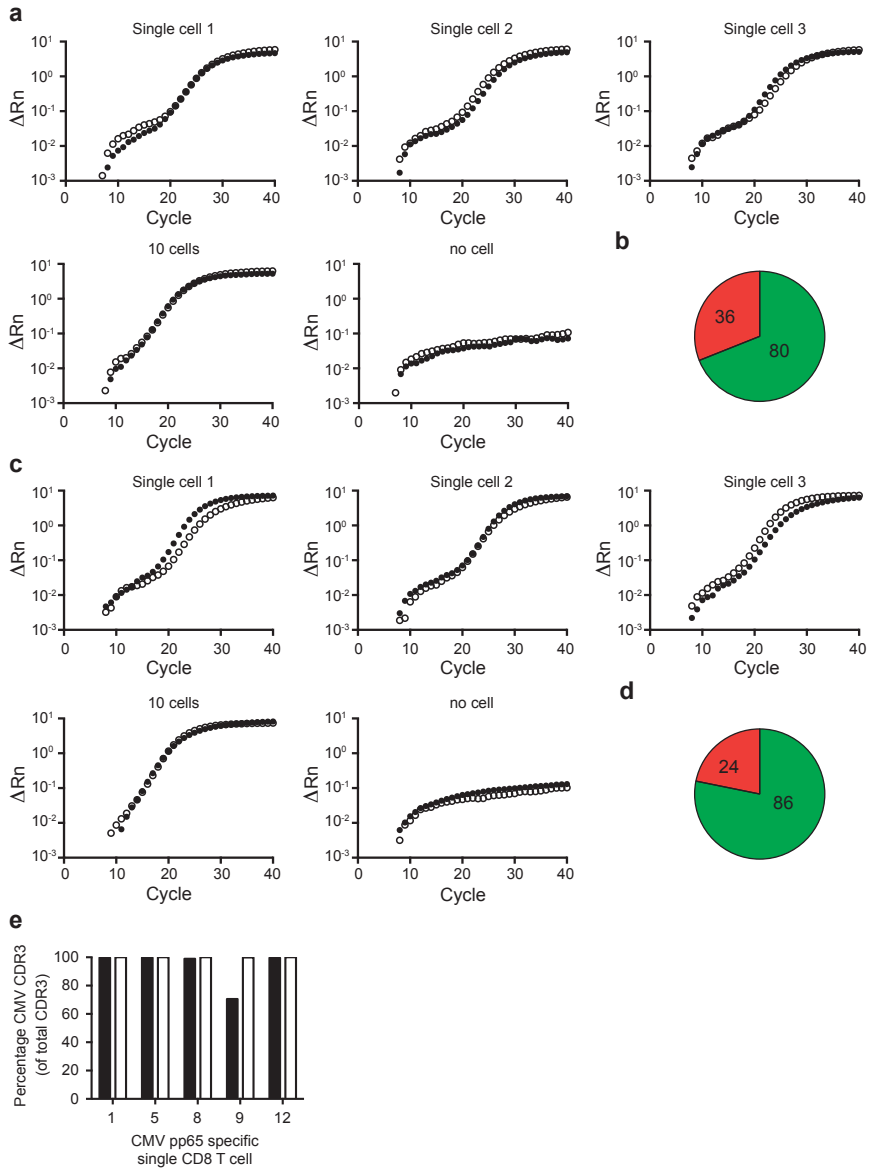

Figure 2. Identification of TCRαβ sequences in cDNA obtained from single CD8⁺ T cells. (**a**) Real-Time Taqman PCR detecting TCRα (black) and TCRβ (white) CDR3 sequences within cDNA obtained from single CMV pp65 specific CD8⁺ T cells. Representative plots out of five experiments are shown. (**b**) Overall efficiency for the detection of TCRαβ CDR3 after analysis of single CMV pp65 specific CD8⁺ T cells (n = 116). TCRαβ cDNA synthesis was considered successful if Ct-value was < 35 for both TCRα and TCRβ. (**c**) Real-Time Taqman PCR detecting TCRα CDR3 (black) and TCRβ CDR3 (white) within cDNA obtained from single HA2 specific CD8⁺ T cells. Representative plots out of three experiments are shown. (**d**) Overall efficiency for the detection of TCRαβ CDR3 after analysis of single HA2 specific CD8⁺ T cells (n = 110). TCRαβ cDNA synthesis was considered successful if Ct-value was < 35 for both TCRα and TCRβ. **(e)** Relative abundance of CMV pp65 specific TCRα (black) and TCRβ (white) CDR3s among all CDR3 sequences after Illumina Paired-end 250bp sequencing of five single, CMV pp65 specific CD8⁺ T cells.

correct TCRαβ sequences was successful for all 5 single-cells tested (**Fig. 2e**). Interestingly, while this CMV pp65-specific T cell clone contains both a functional and non-functional TCRα CDR3, a strong bias towards the in-frame recombination product was detected (>99% for 4/5 analyzed single-cells; **Fig. 2e**), potentially consistent with non-sense mediated RNA decay for the other gene product.

| DISCUSSION

In contrast to the isolation of antibody sequences from single B cells[15, 16], it has only recently become possible to efficiently isolate TCRαβ sequences from single human T cells[8]. Here, we describe an alternative protocol which 1. demonstrates the possibility to analyze single-cell derived cDNA material by NGS and 2. avoids the use of TCR V-gene specific primers. While the use of NGS, offering a higher sample throughput, may potentially also be incorporated in published protocols, the circumvention of TCR V-gene primers is unique to this protocol and highly relevant for TCR repertoire analysis. Specifically, due to the substantial degree of polymorphism within the TCR V genes[4], the use of such primers is a concern with regards to potential amplification bias.

We have recently reported an alternative strategy to identify TCR sequences that involves the capturing and sequencing of genomic DNA fragments[12]. This TCR gene capture method offers a rapid and highly reliable approach for the assessment of bulk TCR repertoires. Furthermore, as genomic DNA is used for TCR gene capture, the method is well suited to applications that involve clinical material of variable quality. In contrast, the RNA-based, single-cell TCRαβ identification protocol described here offers the possibility to perform more in-depth identification of TCRαβ pairs. As such, we believe that both technologies will be complementary.

With the development of high throughput strategies to identify TCR sequences it is now possible to interrogate TCR repertoires in human diseases. This is of obvious value to understand the formation and role of pathogen-, tumor- and autoimmune-related T cell populations. For example, the infiltration of different T cell subsets in tumors has been correlated with clinical outcome of some malignancies[17]. It is tempting to speculate that the frequency of tumor-reactive TCRs within tumors – which can be readily assessed with single-cell based methods – may at least in part explain why such a correlation has been observed for some malignancies but not others. Moreover, the isolation of TCRs from intratumoral regulatory T cells may be of interest. Specifically, while there is some evidence for recognition of tumor-antigens[18, 19], the extent of restriction of regulatory T cells to tumor-antigens has remained unclear. Isolated TCRαβ pairs from intratumoral regulatory T cells could be used to assess reactivity against autologous tumor in order to answer this question. As a final example, the method described here will likely also be of value to study lineage-relationships between T cell subsets. For instance, through analysis of the overlap of TCR repertoires between different $CD4^+$ T cell subsets within tumors combined with the subsequent analysis of the tumor-reactive potential of identified TCRs, it should be feasible to reveal kinship of

regulatory and conventional CD4 T cells for both the tumor-specific and the 'bystander' T cell population. In summary, the strategy for identification of TCRαβ pairs outlined here should enable a number of different TCR repertoires studies that will enhance our understanding of T cell activity at sites of interest.

| METHODS

Cell culture

CMV pp65 and HA2 specific T cell clones were a kind gift of Mirjam H.M. Heemskerk (LUMC, Leiden). Cells were cultured in RPMI media (Life technologies) supplemented with 10% (v/v) AB serum (Life Technologies), Penicillin/Streptomycin (Roche), GlutaMax (Life Technologies) and 50 μM 2-Mercaptoethanol (Sigma-Aldrich) and 3000 IU ml^{-1} rh-IL-2 (Novartis) for 3–7 days prior to assays.

Isolation of single $CD8^+$ T cells

Cell sorting was performed on a FACS Aria I (BD Biosciences). Live, single $CD8^+$ T-cells were sorted using PE-/APC-labeled pMHC-multimers and antibodies specific for CD8 (BD Biosciences) into 96-well PCR plates (Biorad) containing lysis buffer.

Synthesis of TCRαβ cDNA from single $CD8^+$ T cells

Synthesis of cDNA from single cells was performed as previously published[9] with the use of TCR constant domain specific primers instead of an Oligo(dT) primer for first strand synthesis.

TRAC primers (all primers read 5'->3')

TRAC1:
ATATGGATCCGGCGCGCCGTCGACG
TCTCTCAGCTGGTACACGGCAGG

TRAC2:
ATATGGATCCGGCGCGCCGTCGACTG
AGAATCAAAATCGGTGAATAGG

TRAC3:
ATATGGATCCGGCGCGCCGTCGACGA
TATACACATCAGAATCCTTACT

TRAC4:
ATATGGATCCGGCGCGCCGTCGACCT
GTTGCTCTTGAAGTCCATAGAC

TRAC5:
ATATGGATCCGGCGCGCCGTCGACAA
GGCGTTTGCACATGCAAAGT

TRBC (all primers read 5'->3') as reported previously[14]:

TRBC1:
ATATGGATCCGGCGCGCCGTCGACCA
GTATCTGGAGTCATTGA

TRBC2:
ATATGGATCCGGCGCGCCGTCGACTG
CTTCTGATGGCTCAAACAC

TRBC3:
ATATGGATCCGGCGCGCCGTCGAC
CGACCTCGGGTGGGAACA

Prior to cell sorting, lysis buffer was aliquoted into 96-well PCR plates that were used to collect sorted single $CD8^+$ T cells. Subsequent cDNA analysis was carried out as described by Tang *et al.*, using a Biorad DNA engine Thermal cycler or an Eppendorf Mastercycler Pro Thermal cycler.

Real-time Taqman PCR

Taqman probes specific for sequences of TCRαβ CDR3s of the two T cell clones were designed and obtained from Applied Biosystems. Real-time PCRs were

performed using 10 µl cDNA, according to manufacturer's guidelines on 7500 Fast Real-Time PCR System (Applied Biosystems). Ct-values were determined using 7500 Fast Real-Time PCR System software and according to standard practice considered specific if Ct-value was < 35.

Analysis of cDNA material by Illumina sequencing

The obtained cDNA from single T cells was prepared for NGS-analysis by using TruSeq DNA library preparation kit (Illumina) and protocols for TCR gene capture[12]. The resulting sequencing libraries were analysed on an Illumina MiSeq Personal Sequencer.

TCRαβ CDR3 sequences were identified using a previously reported algorithm[13, 14]. Briefly, we localized the TCR J-gene segment in each sequence read based on the identity of a short six nucleotide-motif for every TCR J-gene (http://www.imgt.org) containing the conserved Phenylalanine. TCR J-gene identity was expanded in both directions until the last matched nucleotide was encountered. At least 12 aligned nucleotides were required as a minimum for the identification of TCR J-genes. Similarly, we identified the TCR V-gene segment using the conserved Cysteine residue. TCRB D-gene segments were localized based on the identity of at least six nucleotides between TCR V- and TCR J-gene segments. The CDR3 was extracted for each read as the nucleotide sequence between the conserved TCR V-gene Cysteine and TCR J-gene Phenylalanine residues. Extracted CDR3 with identical nucleotide sequences were clustered to clonotypes.

| REFERENCES

1. Langerak,A.W. *et al.* Molecular and flow cytometric analysis of the Vbeta repertoire for clonality assessment in mature TCRalphabeta T-cell proliferations. *Blood* **98**, 165-173 (2001).
2. Gaspar,H.B. *et al.* Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* **3**, 97ra79 (2011).
3. Robins,H.S. *et al.* Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099-4107 (2009).
4. Mackelprang,R. *et al.* Sequence diversity, natural selection and linkage disequilibrium in the human T cell receptor alpha/delta locus. *Hum. Genet.* **119**, 255-266 (2006).
5. van Heijst,J.W. *et al.* Quantitative assessment of T cell repertoire recovery after hematopoietic stem cell transplantation. *Nat. Med.* **19**, 372-377 (2013).
6. Morgan,R.A. *et al.* Cancer regression in patients after transfer of genetically engineered lymphocytes. *Science* **314**, 126-129 (2006).
7. Wright,G.P. *et al.* Adoptive therapy with redirected primary regulatory T cells results in antigen-specific suppression of arthritis. *Proc. Natl. Acad. Sci. U. S. A* **106**, 19078-19083 (2009).
8. Wang,G.C., Dash,P., McCullers,J.A., Doherty,P.C., & Thomas,P.G. T cell receptor alphabeta diversity inversely correlates with pathogen-specific antibody levels in human cytomegalovirus infection. *Sci. Transl. Med.* **4**, 128ra42 (2012).
9. Tang,F. *et al.* RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* **5**, 516-535 (2010).
10. Tang,F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377-382 (2009).
11. Heemskerk,M.H. *et al.* Efficiency of T-cell receptor expression in dual-specific T cells is controlled by the intrinsic qualities of the TCR chains within the TCR-CD3 complex. *Blood* **109**, 235-243 (2007).
12. Linnemann,C. *et al.* High-throughput identification of antigen-specific TCRs by TCR gene capture. *Nat. Med.* In press (2013).
13. Bolotin,D.A. *et al.* Next generation sequencing for TCR repertoire profiling: platform-specific features and correction algorithms. *Eur. J. Immunol.* **42**, 3073-3083 (2012).

14. Mamedov,I.Z. *et al.* Quantitative tracking of T cell clones after haematopoietic stem cell transplantation. *EMBO Mol. Med.* **3**, 201-207 (2011).
15. Wrammert,J. *et al.* Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* **453**, 667-671 (2008).
16. Wardemann,H. *et al.* Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374-1377 (2003).
17. Fridman,W.H., Pages,F., Sautes-Fridman,C., & Galon,J. The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer* **12**, 298-306 (2012).
18. Wang,H.Y. *et al.* Tumor-specific human CD4+ regulatory T cells and their ligands: implications for immunotherapy. *Immunity.* **20**, 107-118 (2004).
19. Wang,H.Y., Peng,G., Guo,Z., Shevach,E.M., & Wang,R.F. Recognition of a new ARTC1 peptide ligand uniquely expressed in tumor cells by antigen-specific CD4+ regulatory T cells. *J. Immunol.* **174**, 2661-2670 (2005).