


# SCIENTIFIC REPORTS



OPEN

## *Mycobacterium tuberculosis* peptide E7/HLA-DRB1 tetramers with different HLA-DR alleles bound CD4<sup>+</sup> T cells might share identical CDR3 region

Yichuan Gan<sup>1,2,3</sup>, Cong Wang<sup>1,2,3</sup>, Yimin Fang<sup>4</sup>, Yanan Yao<sup>1,2,3</sup>, Xiaoxin Tu<sup>1,2,3</sup>, Jiao Wang<sup>1,2,3</sup>, Xi Huang<sup>2,5</sup>, Yaoju Tan<sup>4</sup>, Tao Chen<sup>6</sup>, Kouxing Zhang<sup>7</sup>, Yanming Shen<sup>4</sup>, Lin Zhou<sup>6</sup>, Jianxiang Liu<sup>4</sup> & Xiaomin Lai<sup>1,2,3</sup> 

Human CD4<sup>+</sup> T cells play an important role in the immune response to *Mycobacterium tuberculosis* (MTB). However, little is known about the spectratyping characteristics of the CD4<sup>+</sup> T-cell receptor (TCR)  $\alpha$ - and  $\beta$ -chains CDR3 region in tuberculosis (TB) patients. We sorted MTB peptide E7-bound CD4<sup>+</sup> T cells by using E7/HLA-DR tetramers constructed with different HLA-DRB1 alleles and extracted the CDR3 amino-acid sequences of TCR  $\alpha$ - and  $\beta$ -chains. The results showed that the CDR3 sequences of E7-bound CD4<sup>+</sup> T cells were completely or partially identical in a single patient. The sequences of MTB peptide C5-bound CD4<sup>+</sup> T cells shared another, and non-peptide bound CD4<sup>+</sup> T cells, as well as unbound CD4<sup>+</sup> T cells with tetramers were different from each other. Specifically, diverse CDR3 sequences of E7-bound CD4<sup>+</sup> T cells displayed similar protein tertiary structure in one TB patient. In summary, the TCR  $\alpha$ - and  $\beta$ -chains of CDR3 lineage of CD4<sup>+</sup> T cells in TB patients apparently drifted, and the predominant CDR3 sequences of TCR  $\alpha$ - and  $\beta$ -chains that recognized the MTB antigen exhibited peptide specificity, and certain HLA-DR restriction was also established. This study elucidates the possible causes and mechanisms of peptide-specific CD4<sup>+</sup> T-cell-related presentation against MTB.

Tuberculosis (TB) is one of the most widespread chronic infectious diseases mainly presented as a respiratory infection caused by *Mycobacterium tuberculosis* (MTB). While the mechanisms leading to loss of immune defense and disease reactivation are yet unknown, it is well established that CD4<sup>+</sup> T cells are critical in controlling TB infection<sup>1–5</sup>. T cells recognize antigens by their T-cell receptors (TCRs), a process limited by the histocompatibility complex (MHC)<sup>6</sup>. CD4<sup>+</sup> T cells are activated by recognizing the antigen peptide/MHC class II complex and induce a series of immunological reactions<sup>7</sup>. The human leukocyte antigen (HLA) gene is a set of complicated complex components and structures, the most important feature of which is its high polymorphism. The HLA-II class loci are in the HLA-D region, which includes HLA-DQ, DP, and DR sub-regions. The complex of HLA-DR

<sup>1</sup>Department of Microbiology, Zhongshan School of Medicine, Sun Yat-sen University, 74 Zhongshan Road II, Guangzhou, 510080, China. <sup>2</sup>China Ministry of Education Key Laboratory of Tropical Diseases Control, Tuberculosis Research Institute, Zhongshan School of Medicine, Sun Yat-sen University, 74 Zhongshan Road II, Guangzhou, 510080, China. <sup>3</sup>Gangdong Provincial Department of Education Key Laboratory of Functional Molecules from Marine Microorganisms, Gangdong Provincial Research Center for Severe Infectious Disease Prevention and Control Technology, Zhongshan School of Medicine, Sun Yat-sen University, 74 Zhongshan Road II, Guangzhou, 510080, China. <sup>4</sup>State Key Laboratory of Respiratory Disease of China, Guangzhou Chest Hospital, 62 Hengzhigang Road, Guangzhou, 510095, China. <sup>5</sup>Department of Immunology, Zhongshan School of Medicine, Sun Yat-sen University, 74 Zhongshan Road II, Guangzhou, 510080, China. <sup>6</sup>Tuberculosis Control Center of Guangdong Province, 485 West Huangpu Avenue, Guangzhou, 510630, China. <sup>7</sup>Third Affiliated Hospital of Sun Yat-sen University, 600 Tianhe Road, Guangzhou, 510630, China. Yichuan Gan, Cong Wang and Yimin Fang contributed equally to this work. Correspondence and requests for materials should be addressed to L.Z. (email: [13332864810@vip.163.com](mailto:13332864810@vip.163.com)) or J.L. (email: [Ljxr64@qq.com](mailto:Ljxr64@qq.com)) or X.L. (email: [laixm@mail.sysu.edu.cn](mailto:laixm@mail.sysu.edu.cn))

Tetramer No.	Peptide-HLA-DRA <sup>A</sup>	HLA-DRB1 alleles <sup>B</sup>
1	E7	DRB1*08032
2	E7	DRB1*0818
3	E7	DRB1*0404
4	E7	DEB1*160201
9	E7	DRB1*090102
10	E7	DRB1*130201
11	E7	DRB1*150101
12	E7	DRB1*1503
6	C5	DRB1*0818
15	C5	DRB1*150101
16	C5	DRB1*1503
AK2	No peptide	DRB1*0818
AK11	No peptide	DRB1*150101
AK12	No peptide	DRB1*1503

**Table 1.** Components of MTB peptide/non-peptide HLA-DR tetramers constructed with different HLA-DRB1 alleles. <sup>A</sup>Peptide fused with HLA-DRA. <sup>B</sup>HLA-DRB1 alleles of tetramers.

(human leukocyte antigen-antigen D-related) and its ligand, a peptide composed of nine or more amino-acids, constitutes a ligand for the TCR. T cells expressing  $\alpha\beta$  receptors play an important role in immunization to MTB. High expansions of  $\alpha\beta$  T cells within the TCR repertoire also have been shown to occur in various diseases, such as cancers and immunological disorders, as well as in inflammatory and infectious diseases<sup>8–10</sup>. Furthermore, each V $\alpha$ - (variable region of  $\alpha$ -chain) or V $\beta$ -chain constitutes three loops called the complementarity determining regions (CDRs) 1, 2, and 3, which interact with the peptide/MHC molecule. CDR1 and CDR2 regions recognize and bind to the side walls of an antigen in antigen-binding-groove of MHC molecules, whereas the CDR3 region is directly combined with antigenic peptides. Thus, the TCR specificity is mainly determined by its CDR3 region. In other words, analyzing and comparing the length and sequence of the CDR3 region can be judged as an indicator of T-cell clones<sup>11</sup>.

In 1996, Altman established a peptide/MHC tetramer assay using the principle of biotin-avidin cascade amplification, and greatly increased the intermolecular affinities and stability of MHC/peptide-TCR binding<sup>12</sup>. In our previous work, we showed the MTB peptide E7 is an ideal CD4<sup>+</sup> T cell-response antigen which was detected with IFN $\gamma$ -ELISPOT test and was authorized for China patent CN 101446585 A<sup>13</sup>. E7 comes from the early secretory antigenic target 6 (ESAT6), a type of secreted protein purified and separated from the early culture medium of MTB. Peptide C5 is obtained from the 10-kDa culture filtrate protein (CFP-10) of MTB which is known as ESAT-6-like protein encoded by the *esxB* gene. We extracted an MTB peptide/HLA-DR monomer from the stable cell lines established that were already expressing soluble biotinylated MTB peptide/HLA-DR monomer. E7, C5, or non-peptide tetramers constructed with different HLA-DRB1 alleles (Table 1) were used to analyze the peptide-bound CD4<sup>+</sup> T cells in pleural fluid (PLF) from TB patients by MACS.

In addition, we observed a phenomenon of drift binding between HLA-DR alleles and TCR in our previous examinations. We used E7/HLA-DRB1\*08032 and E7/HLA-DRB1\*0818 tetramers to combine and detect the tetramer-bound CD4<sup>+</sup> T cells in TB patients. Our findings revealed a positive rate that was not as high as expected in TB patients with a HLA-DRB1\*08 allele, while it had a clearly available detection in TB patients with non-HLA-DRB1\*08 allele<sup>14</sup>. Besides, there have been many reports of the binding drift phenomenon between HLA-DR allele and TCR<sup>15,16</sup>. For instance, D'Orsogna discovered the use of HLA-B\*0801/EB virus tetramer to sort CD8<sup>+</sup> T cells specifically and those cells could combine with artificial APC with HLA-B\*4402 and HLA-B\*5501 allele<sup>17</sup>. Therefore, in the present study, we studied the spectratyping characteristics of the MTB peptide tetramers-bound CD4<sup>+</sup> TCR $\alpha$ - and  $\beta$ -chains CDR3 region in TB patients and intended to explore the possible causes and mechanisms of drift binding and antigen presentation between HLA-DRB1 alleles and CD4<sup>+</sup> TCR against MTB.

## Results

**CDR3 amino-acid sequences of E7-bound CD4<sup>+</sup> T cells were completely or partially identical in one of the patients.** We obtained E7-bound CD4<sup>+</sup> TCR $\alpha$  CDR3 amino-acid sequences from four TB patients (PLFs 1, 2, 5, and 12) and TCR $\beta$  CDR3 sequences from nine TB patients (PLFs 2, 3, 4, 5, 6, 7, 8, 11, and 12). The sequences contained two termini with conserved sequences at the N and C termini, i.e., YLC and FG, respectively, and the inner region exhibited diverse sequences. The comparison of nucleotide and amino-acid sequences showed that E7-bound CD4<sup>+</sup> T cells with the same function shared the identical TCR  $\alpha$ - or  $\beta$ -chain CDR3 sequence in one TB patient (Tables 2 and 3). In different patients, it was diverse or just shared a similar amino-acid motif. Irrespective of the patient examined, the CDR3 amino-acid sequences of both E7 and non-peptide tetramers unbound CD4<sup>+</sup> T cells were quite different (Table S1). These results suggested that E7-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles might display clonal expansion in one single individual.

PLF No.	Tetramer <sup>A</sup>	Clones	V $\alpha$ <sup>B</sup>	CDR3 Region <sup>C</sup>	J $\alpha$ <sup>D</sup>	AAN <sup>E</sup>
1	1/2/4/9/10	20	TACATCTGT	GCTGTGAGAGATGAAAACATATCAGTTAATCTGGGGC	GCTGGG	12
			YIC	AVRDQNYQLIWG	AG	
2	2/3/4/9/10/11/12	28	TACCTCTGT	GCAAGCAGTGGTATCAACGCAGAGTACTGGATATTGCACCATTCC	TTGGGG	15
			YLC	ASSGINAEYSILHHS	LG	
5	1/2/3/4/9/11/12	28	TACCTCTGC	GCAATGAGCGCGGGGAACCGAGGAGGCAAGCTTATC	TTTGGA	12
			YLC	AMSAGNQGGKLI	FG	
12	2	5	TACCTCTGT	GCCAGTGCAGACAGCTTTTCTACGAGCAGTAC	TTGGGG	11
			YLC	ASADSFYSYEQY	FG	

**Table 2.** Nucleotide and amino-acid sequences of TCR $\alpha$  CDR3 of E7/HLA-DR-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles. <sup>A</sup>Tetramers 1/2/3/4/9/10/11/12 were E7/HLA-DR tetramers with different HLA-DR alleles. <sup>B</sup>Nucleotide and amino-acid sequences of V region terminal. <sup>C</sup>Nucleotide and amino-acid sequences of CDR3 region. <sup>D</sup>Joining nucleotide and amino-acid sequences. <sup>E</sup>CDR3 amino-acid sequences' length.

PLF No.	Tetramer <sup>A</sup>	Clones	V $\beta$ <sup>B</sup>	CDR3 Region <sup>C</sup>	J $\beta$ <sup>D</sup>	AAN <sup>E</sup>
2	1/2/3/4/9/10/11/12	40	TACTTCTGT	GCCAGCAGTGGGGACAGGAACTTTACGAGCAGTAC	TTGGGG	12
			YFC	ASSGDRNSYEQY	FG	
3	1/2/3/4/9/10/11/12	40	TACCTCTGT	GCCAGCAGGGGGGGGGCCAGAAGAGGACCAGTAC	TTGGGG	12
			YLC	ASRGGGPEEDQY	FG	
4	1/2/3/4/9/10/11/12	40	TACCTATGT	GCTGCCAGTTCACCTGCTGCACTCCAAG	TTCCAC	10
			YLC	AAQFLLHHSK	FH	
5	1/2/3/10	20	TATCTCTGC	AGCGTTGAAGGCCAGAGGAGGTTCCCTCTTTC	TTTGGA	11
			YLC	SVEGQRRVPSF	FG	
6	2/3/11/12	20	TATTTCTGT	GCCAGCAGCTCGTCCAGCTTGAACACTGAAGCTTTC	TTTGGA	12
			YFC	ASSSSLNTEAF	FG	
7	1/9	10	TATCTCTGC	AGCGTTGAAGGGCAGGGGGCTAGTGGCTACACC	TTGGGT	11
			YLC	SVEGQGASGYT	FG	
8	10/12	10	TATCTCTGC	AGGTTGAAGATGGGACAGGGGGCCAGAGTGGCTACACC	TTGGGT	13
			YLC	SVEDTGGQSGYT	FG	
11	2/11/12	15	TACTTCTGT	GCCAGCAGTGGGGACAGGAACTCTACGAGCAGTAC	TTGGGG	12
			YFC	ASSGDRNSYEQY	FG	
12	2	6	TACCTCTGT	GCTGTGAATGAGGACACGGGCAGGAGAGCACTTACT	TTTGGA	12
			YLC	AVNEDTGRRALT	FG	

**Table 3.** Nucleotide and amino-acid sequences of TCR $\beta$  CDR3 of E7/HLA-DR-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles. <sup>A</sup>Tetramers 1/2/3/4/9/10/11/12 were E7/HLA-DR tetramers with different HLA-DR alleles. <sup>B</sup>Nucleotide and amino-acid sequences of V region terminal. <sup>C</sup>Nucleotide and amino-acid sequences of CDR3 region. <sup>D</sup>Joining nucleotide and amino-acid sequences. <sup>E</sup>CDR3 amino-acid sequences' length.

### The clonal expansion of E7-bound CD4<sup>+</sup> T cells was due to peptide specificity rather than HLA-DR specificity.

We continually studied the reason for the phenomenon that E7-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles exhibited clonal expansion in one patient. In PLF11 (Table 4), E7-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles shared an identical CDR3 amino-acid sequence (ASSGDRNSYEQY), while C5-bound CD4<sup>+</sup> T cells with different HLA-DR alleles shared another one (SGLAGGAYEQY). The terminal sequences "AS" and "SA" also included in conserved sequences which translated by different primer pairs designed before. In contrast, CDR3 sequences of non-peptide bound CD4<sup>+</sup> T cells were different from each other. In addition, the sequences of E7 or non-peptide/HLA-DR tetramers unbound CD4<sup>+</sup> T cells were not the same each other at all (Table S1). Thus, the clonal expansion of peptide-bound CD4<sup>+</sup> T cells was mostly due to peptide specificity rather than HLA-DR specificity.

### TCR V $\alpha$ and TCR V $\beta$ of E7-bound CD4<sup>+</sup> T cells in TB patients perform dominant distribution.

We chose three TB patients (PLFs 5, 6, and 11) whose E7-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles were clonal expansion and amplified their full-length nucleotide sequences of TCR  $\alpha$ - and  $\beta$ -chains. BLAST software was used to analyze their V, J, and D family. The dominant TCR V $\alpha$  and TCR V $\beta$  of E7-bound CD4<sup>+</sup> T cells in these patients were diverse from each other (Table 5). PLF5 was mostly on TRAV12-3 and TRBV12-3. In contrast, PLF6 was mainly on TRBV14, and PLF11 was entirely on TRBV10-1. Besides, TCR J and D family in these patients were more consistent. To sum up, different TB patients exhibited various TCR V $\alpha$  and TCR V $\beta$

PLF No.	Tetramer <sup>A</sup>	Peptide <sup>B</sup>	HLA-DRB1 allele <sup>C</sup>	Clones	Vβ <sup>D</sup>	CDR3 Region <sup>E</sup>	Jβ <sup>F</sup>	AAN <sup>G</sup>
11	2	E7	DRB1*0818	2	TACTTCTGT	GCCAGCAGTGGGGACAGGAACTCCTACGAGCAGTAC	TTCGGG	12
					YFC	ASSGDRNSYEYQ	FG	
11	11	E7	DRB1*150101	4	TACTTCTGT	GCCAGCAGTGGGGACAGGAACTCCTACGAGCAGTAC	TTCGGG	12
					YFC	ASSGDRNSYEYQ	FG	
11	12	E7	DRB1*1503	4	TACTTCTGT	GCCAGCAGTGGGGACAGGAACTCCTACGAGCAGTAC	TTCGGG	12
					YFC	ASSGDRNSYEYQ	FG	
11	6	C5	DRB1*0818	5	TACTTCTGT	GCCAGCAGCGGACTAGCGGGAGGTGCCTACGAGCAGTAC	TTCGGG	13
					YFC	ASSGLAGGAYEQY	FG	
11	15	C5	DRB1*150101	2	TACATCTGC	AGTGTAGTGGACTAGCGGGAGGTGCCTACGAGCAGTAC	TTCGGG	13
					YIC	SASGLAGGAYEQY	FG	
11	16	C5	DRB1*1503	4	TACCTCTGT	AGTGTAGTGGACTAGCGGGAGGTGCCTACGAGCAGTAC	TTCGGG	13
					YLC	SASGLAGGAYEQY	FG	
11	16	C5	DRB1*1503	1	TATTTCTGT	GCCAGCAGTGGGGACAGGAACTCCTACGAGCAGTAC	TTCGGG	12
					YFC	ASSGDRNSYEYQ	FG	
11	AK2		DRB1*0818	2	TATTTCTGT	GCCAGCAGTGAGGGGGGGGGGACAATGAGCAGTTC	TTCGGG	12
					YFC	ASSEGGDNEQF	FG	
11	AK2		DRB1*0818	1	TATCTCTGC	AGCGTTAGGTCCGGGACGGGACAAAATGAGCAGTAC	TTCGGG	12
					YLC	AVRSGTGQNEQY	FG	
11	AK11		DRB1*150101	2	TATTTCTGT	GCCAGCAGTGGGGACAGGAACTCCTACGAGCAGTAC	TTCGGG	12
					YFC	ASSGDRNSYEYQ	FG	
11	AK11		DRB1*150101	2	TATCTCTGC	AGCGTTAGGTCCGGGACGGGACAAAATGAGCAGTAC	TTCGGG	12
					YLC	AVRSGTGQNEQY	FG	
11	AK12		DRB1*1503	3	TATCTCTGC	GCCAGCAGCTTACTGAGCTCCACCGGGAGCTGTTT	TTTGA	12
					YLC	ASSLLSSTGELF	FG	
11	AK12		DRB1*1503	1	TATTTCTGT	GCCAGCAGTGGGGACAGGAACTCCTACGAGCAGTAC	TTCGGG	12
					YFC	ASSGDRNSYEYQ	FG	
11	AK12		DRB1*1503	1	TATCTCTGC	AGCGTTAGGTCCGGGACGGGACAAAATGAGCAGTAC	TTCGGG	12
					YLC	AVRSGTGQNEQY	FG	

**Table 4.** Nucleotide and amino-acid sequences of TCRβ CDR3 of E7/C5/non-peptide bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles in PLF11. <sup>A</sup>Tetramers 2/11/12 were E7/HLA-DR tetramers. Tetramers 6/15/16 were C5/HLA-DR tetramers. Tetramers AK2/AK11/AK12 were non-peptide HLA-DR tetramers. <sup>B</sup>Peptide fused with HLA-DRA. <sup>C</sup>HLA-DRB1 allele of tetramers. <sup>D</sup>Nucleotide and amino-acid sequences of V region terminal. <sup>E</sup>Nucleotide and amino-acid sequences of CDR3 region. <sup>F</sup>Joining nucleotide and amino-acid sequences. <sup>G</sup>CDR3 amino-acid sequences' length.

profiles. Even within the same patient, the profile of each E7-bound CD4<sup>+</sup> T cell with different HLA-DRB1 alleles was significantly different. Some TCR Vα and TCR Vβ gene families from TB patient showed oligoclonal.

**The structures of TCRβ CDR3 region of E7-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles were similar.** Since we had also observed completely different CDR3 sequences of E7-bound CD4<sup>+</sup> T cells in some single one TB patient, so considered to analyze their protein tertiary structures. Taken PLF13 as a typical example, input raw amino-acid sequences containing CDR3 region into Swiss-Model Workspace (<https://www.swissmodel.expasy.org/interactive>) to predict protein tertiary structure, area between the arrows in the upper left was CDR3 region protein tertiary structure (Fig. 1a–c). Framework sequences were conserved, N-terminal (REFV or REFI) and C-terminal (PGTRLTVTEDLKNVF) amino-acid residues of CDR3 loop were composed of almost the same amino-acid residues. On the premise of E7 combining, in spite of the different amino-acid sequences, the CDR3 protein tertiary structure showed similar each other in PLF13 (Fig. 1a). Furthermore, we also observed C5-bound CD4<sup>+</sup> TCR CDR3 protein tertiary structures showed another consistent shape but was distinct from those of E7-bound CD4<sup>+</sup> TCR CDR3 region (Fig. 1b). In contrast, non-peptide tetramers bound CD4<sup>+</sup> TCR CDR3 tertiary structures were entirely different each other (Fig. 1c). These results suggested that in some TB patients E7/HLA-DR tetramers with different HLA-DRB1 alleles were capable of recognizing and binding CDR3 fragments with different sequences but a similar structure and function.

**The most common length of TCRβ CDR3 of E7-bound CD4<sup>+</sup> T cells is 12 amino-acid sequences.** There were a few CDR3 sequences of TCR α-chain in CD4<sup>+</sup> T cells so that it was not enough to statistic. Therefore, we counted the length of TCRβ CDR3 amino-acid sequences of E7/HLA-DR-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles in eleven TB patients (clones, n = 252). The counting result showed that the maximal length of CDR3 sequence consisted of 12 amino-acids, accounting for 40.873% (103 of 252 clones) (Fig. 1d). E7-bound

PLF No.	Tetramer <sup>A</sup>	Clones	V $\alpha$ /V $\beta$ <sup>B</sup>	CDR3 Region <sup>C</sup>	J $\alpha$ /J $\beta$ <sup>D</sup>	V Family <sup>E</sup>	J Family <sup>F</sup>	D Family <sup>G</sup>
5	1/2/3/4/9/11/12	28	YLC	GCAATGAGCGCGGGGAACCAGGGAGGAAAGCTTATC	FG	TRAV12-3	TRAJ23	
				AMSAGNQGGKLI				
5	10	4	YLC	GCAATGAGCGCGGGGAACCAGGGAGGAAAGCTTATC	FG	TRAV12-1	TRAJ20	
				AMSAGNQGGKLI				
5	1/2/10	15	YLC	AGCGTTGAAGGCCAGAGGAGGGTCCCCTCTTTC	FG	TRBV12-3	TRBJ2	TRBD1
				SVEGQRRVPSF				
5	3	5	YLC	AGCGTTGAAGGCCAGAGGAGGGTCCCCTCTTTC	FG	TRBV18-1	TRBJ2	TRBD1
				SVEGQRRVPSF				
6	2/11	10	YFC	GCCAGCAGCTCGTCCAGCTTGAACACTGAAGCTTTC	FG	TRBV14	TRBJ2	TRBD2
				ASSSSSLNTEAF				
6	12	5	YFC	GCCAGCAGCTCGTCCAGCTTGAACACTGAAGCTTTC	FG	TRBV14	TRBJ2	TRBD1
				ASSSSSLNTEAF				
6	3	5	YFC	GCCAGCAGCTCGTCCAGCTTGAACACTGAAGCTTTC	FG	TRBV29-1	TRBJ2	TRBD2
				ASSSSSLNTEAF				
11	2/11/12	15	YFC	GCCAGCAGTGGGGACAGGAACCTCTACGAGCAGTAC	FG	TRBV10-1	TRBJ2	TRBD1
				ASSGDRNSYEYQ				

**Table 5.** The TCR V $\alpha$  and V $\beta$  of E7-bound CD4<sup>+</sup> T cells in TB patients performed dominant distribution. <sup>A</sup>Tetramers 1/2/3/4/9/10/11/12 were E7/HLA-DR tetramers with different HLA-DR alleles. <sup>B</sup>Amino-acid sequences of V region terminal. <sup>C</sup>Nucleotide and amino-acid sequences of CDR3 region. <sup>D</sup>Joining nucleotide and amino-acid sequences. <sup>E</sup>V $\alpha$  gene family or V $\beta$  gene family. <sup>F</sup>J $\alpha$  gene family or J $\beta$  gene family. <sup>G</sup>D $\beta$  gene family.

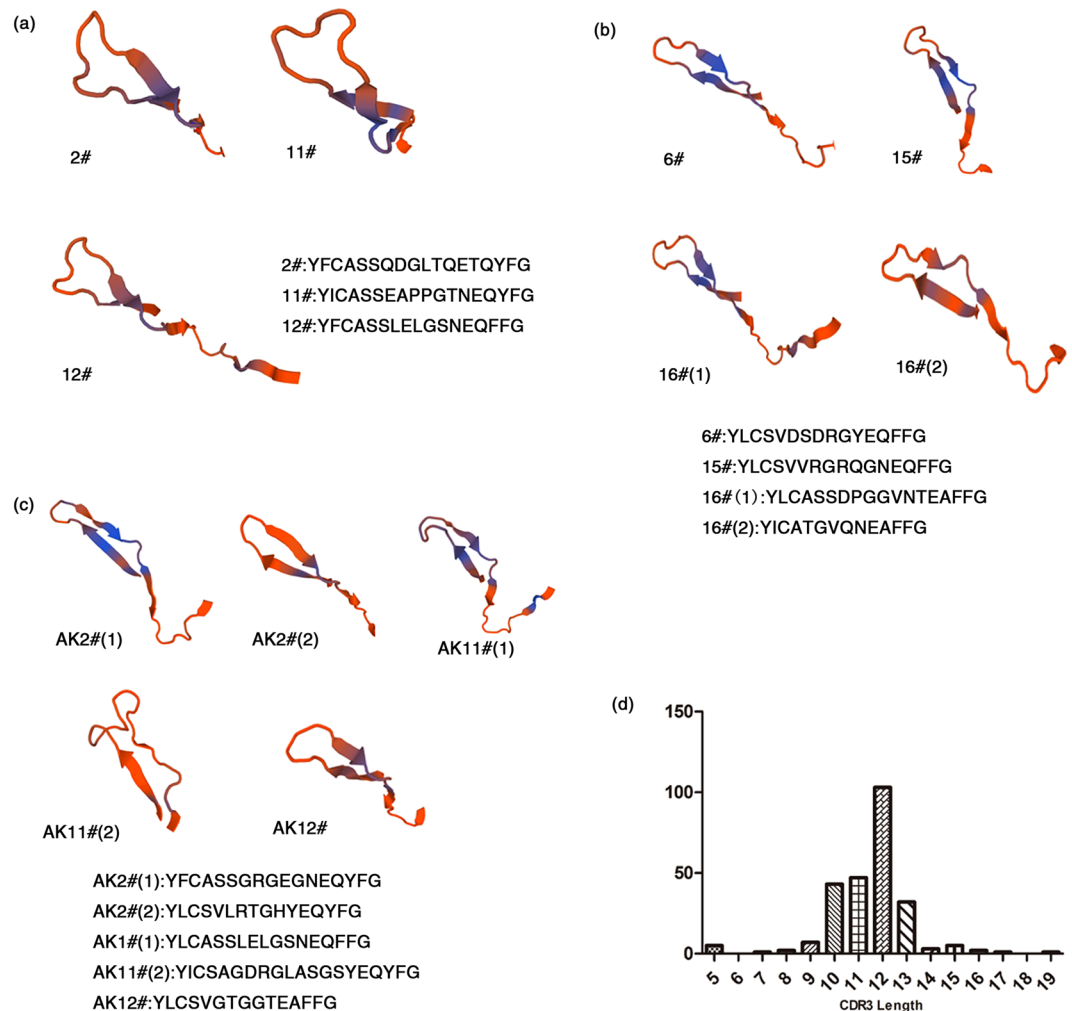
CD4<sup>+</sup> T cells displayed TCR $\beta$  CDR3 amino-acid sequences' length dominant distribution in immune reaction against specific MTB peptide E7.

**Molecular chaperones CD3 and CD4 play a supporting role in the binding force between E7/HLA-DR tetramers and CD4<sup>+</sup> T cells.** CD3 and CD4 are responsible for conducting the TCR activation signal in the process of T-cell recognition of antigens. We used E7/HLA-DR tetramers to sort CD4<sup>+</sup> T cells from two TB patients (PLFs 9 and 10), followed by treatment with CD3 and CD4 goat polyclonal antibodies to block them before sorting. The results showed that E7-bound CD4<sup>+</sup> T cells blocked with CD3 and CD4 polyclonal antibodies exhibited disorganized CDR3 amino-acid sequences (Table 6), also different with unbound CD4<sup>+</sup> T cells (Table S2). It suggested that the suppression of function of CD3 and CD4 apparently affected the process of T cell recognizing antigen.

## Discussion

The CDR3 region of TCR varies among different T-cell populations and individuals. The rearrangement of the TCR CDR3 is random in healthy individuals without antigen stimulation, but specific T cells can be clonally activated and expanded by a foreign specific antigen, such as the MTB protective 16-kDa antigen and the MTB purified protein derivative (PPD) RT23<sup>18</sup>. Specific recognition of a peptide on a certain T cell can induce the clonal expansion of T cells which express a particular TCR V gene product. Analysis of the size and distribution pattern of CDR3 can be used to define the degree of clonality of T cells and has been widely exploited in the research of many diseases, such as myelodysplastic syndrome, cancer, parasitosis, and autoimmune diseases<sup>19–23</sup>. In our previous studies, we prepared peptide E7, C5, or non-peptide tetramers constructed with different HLA-DRB1 alleles. The non-peptide DRB1 tetramers in fact contain no peptides. In the case of HLA-DR typing of the defined was detected not match with the genes of tetramers, there was a certain different level of detection<sup>14</sup>, which revealed that the phenomenon of binding drift between HLA-DR and CD4<sup>+</sup> TCR could lead to elucidating the possible causes and mechanisms of CD4<sup>+</sup> T cells related presentation against MTB.

The CDR3 regions of TCR  $\alpha$ - and  $\beta$ -chains are responsible for recognition of the antigenic peptides presented to the CD4<sup>+</sup> T cell by HLA class II. As antigen-MHC complex, MTB peptide/HLA-DRB1 tetramers bind to CDR3 region of CD4<sup>+</sup> TCR. MTB peptide directly combined with the middle region of CDR3, while the conserved flanking regions bind to the side walls of antigen-binding-groove of MHC molecules. Consequently, there needs to be a high level of diversity in CDR3 region, which is generated by varying both the sequences and the number of amino-acids. In this study, we found lengths of E7-bound CD4<sup>+</sup> TCR $\beta$  CDR3 amino-acid sequences mainly on 12. Some publications also reported the characteristics of the CDR3 lengths' distribution of T cells or B cells in diseases, such as cancer, myelodysplastic syndrome, infectious, and autoimmune disorders<sup>24–27</sup>. Additionally, several reports indicated that the CDR3 sequence of antigen-specific T cells recognizing the same peptide shared identical CDR3 amino-acid motifs<sup>28–30</sup>, but no evidence has been reported that the CDR3 sequences of the same MTB peptide-bound CD4<sup>+</sup> T cells sorted with the peptide tetramers with different HLA-DRB1 alleles are the same. In this study, we found that E7-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles shared one specific CDR3 sequence of TCR  $\alpha$ - or  $\beta$ -chains in a single individual. We speculated that the E7/HLA-DR tetramers with different HLA-DR alleles could be identified by the CD4<sup>+</sup> T cells that had the same function and an identical



**Figure 1.** The structures of TCR $\beta$  CDR3 region of E7, C5 or non-peptide/HLA-DR tetramers-bound CD4<sup>+</sup> T cells in PLF13 and comparison of amino-acid sequences' length of TCR $\beta$  CDR3 of E7/HLA-DR-bound CD4<sup>+</sup> T cells. **(a)** The protein tertiary structures of E7/HLA-DR tetramers 2#, 11#, and 12# bound CD4<sup>+</sup> TCR $\beta$  CDR3 region, area between the arrows. Although amino-acid sequences are different, the CDR3 protein tertiary structure showed similar each other in PLF13 as long as the peptide E7 binds; **(b)** The protein tertiary structures of C5/HLA-DR tetramers 6#, 15#, and 16# bound CD4<sup>+</sup> TCR $\beta$  CDR3 region. The structure of C5-bound CD4<sup>+</sup> TCR $\beta$  CDR3 region showed another shape of structure, which was distinct from those of E7-bound CD4<sup>+</sup> T cells; **(c)** The protein tertiary structures of E7/HLA-DR tetramers AK2#, AK11#, and AK12# bound CD4<sup>+</sup> TCR $\beta$  CDR3 region. Non-peptide tetramers bound CD4<sup>+</sup> T cells TCR $\beta$  CDR3 tertiary structures were entirely different each other. In a, b, and c, all structures were oriented with CDR3 region in the top left corner. According to the number of tetramers, structures, from left to right, corresponded to their respective amino-acid sequences; **(d)** Eleven TB patients were included in this study (clones = 252). Frequency histograms of TCR $\beta$  CDR3 amino-acid sequences were shown. For patients, the count of TCR $\beta$  CDR3 amino-acid sequences' length of E7-bound CD4<sup>+</sup> T cells in total sequence results was calculated (Y axis), which focused on 5 to 19 amino-acid residues (X axis), mainly distributed in 12 amino-acid sequences.

sequence of the CDR3 region in one TB patient. Nevertheless, that the TCR CDR3 amino-acid sequence was diverse in different patients suggested that E7/HLA-DR tetramers with different HLA-DR alleles might be identified by those CD4<sup>+</sup> T cells with the same function but the different CDR3 sequences in different patients. Overall, the same MTB peptide bound CD4<sup>+</sup> T cells sorted with the peptide tetramers with different HLA-DRB1 alleles showed clonal expansion that performed the identical CDR3 sequence and amino-acid sequences' length dominant distribution in TB patients.

Then we found that the clonal expansion of CD4<sup>+</sup> T cells was due to peptide specificity rather than HLA-DR specificity. However, in a few cases, the clonal expansion of CD4<sup>+</sup> T cells to a certain extent is dependent on HLA-DR restriction. Moreover, we genotyped the HLA-DR type of PLF11 in order to further analyze the effect of patient's HLA-DR allele background on CD4<sup>+</sup> T cells in combination with peptide/HLA-DR tetramers constructed with different HLA-DR alleles. The MTB peptide/HLA-DR tetramers with different HLA-DRB1 alleles (DRB1\*0818, DRB1\*150101, and DRB1\*1503) not completely matching PLF11's genetic background

PLF No.	tetramer <sup>A</sup>	Clones	V $\beta$ <sup>B</sup>	CDR3 Region <sup>C</sup>	J $\beta$ <sup>D</sup>	AAN <sup>E</sup>
9	1B	1	TACTTCTGT	GCCAGCAGTCTGGGACAGGGGGCCACGAGCAGTAC	TTCGGG	12
			YFC	ASSPRTGGHEQY	FG	
9	1B	2	TATCTCTGC	AGCGTTGATGGAGGGGGTGGTGGACTGAAGCTTTC	TTTGGA	12
			YLC	SVDGGGWWEAF	FG	
9	9B	2	TACTTCTGT	GCTGCCAGTTCCTCCCAGTCCCTCCAAG	TTCCAC	10
			YFC	AAQFLPSPSK	FH	
9	11B	1	TACTTCTGT	GCCAGCAGTATGAGTTACGAGACCCAGTAC	TTCGGG	10
			YFC	ASSMSYETQY	FG	
10	2B	1	TATCTCTGC	AGCGTTGGCGAGCCGGGACAGAACACCGGGGAGCTGTTT	TTTGGA	13
			YLC	SVGAPGQNTGELF	FG	
10	2B	2	TACTTCTGT	GCCAGCAGTTTATGGGGGAAAGCACTGAAGCTTTC	TTTGGA	12
			YFC	ASSLRGESTEAF	FG	
10	2B	1	TACTTCTGT	GCCATCAGTGAGAGTGGCGGCGAAACGAATGAGCAGTTC	TTCGGG	13
			YFC	AISESGGETNEQF	FG	
10	2	2	TACCTCTGT	GCCAGCAGTTCCTCCAGGGTGGGGAAACACCATATAT	TTTGGA	12
			YLC	ASSSPGWNTIY	FG	
10	AK2B	1	TACCTCTGT	GCCAGCAGTGAATATACAGCCAATCAGCCCCAGCAT	TTTGGT	12
			YLC	ASSEYTANQPQH	FG	
10	AK2B	2	TACCTCTGT	GCCAGCAGCTTAGGAACAGGCGAGCAGTAC	TTCGGG	10
			YLC	ASSLGTGEQY	FG	

**Table 6.** Nucleotide and amino-acid sequences of TCR $\beta$  CDR3 of CD4<sup>+</sup> T cells blocked by CD3 and CD4 polyclonal antibodies in PLFs 9 and 10. <sup>A</sup>Tetramers 1/2/9/11 were E7/HLA-DR tetramers with different HLA-DR alleles. “B” represented blocked by CD3 and CD4 polyclonal antibodies. <sup>B</sup>Nucleotide and amino-acid sequences of V region terminal. <sup>C</sup>Nucleotide and amino-acid sequences of CDR3 region. <sup>D</sup>Joining nucleotide and amino-acid sequences. <sup>E</sup>CDR3 amino-acid sequences’ length.

with HLA-DRB1\*0101 and HLA-DRB1\*0803 were capable of binding CD4<sup>+</sup> T cells with the identical CDR3 amino-acid sequence in PLF11. Although the E7/HLA-DRB1\*0818 and C5/HLA-DRB1\*0818 tetramers shared the same classification of HLA-DRB1\*08 allele with PLF11, the CDR3 sequence of E7 or C5/HLA-DRB1\*08 tetramers bound CD4<sup>+</sup> T cells did not show any specific results. Likewise, E7/HLA-DRB1\*0818 tetramer bound CD4<sup>+</sup> T cells in PLF12 with HLA-DRB1\*1101 and HLA-DRB1\*1302 displayed the consistent CDR3 sequence only at a suitable concentration of the tetramer. All these results provided the novel discovery of antigen presented process between CD4<sup>+</sup> T cells and MHC class II in reference to HLA-DR type. Besides, CD8<sup>+</sup> T cells also play a part in the protective immunity against MTB infections, and several reports showed similar findings that the MTB antigen reactive CD8<sup>+</sup> T cells doing their jobs independent of MHC restriction<sup>31,32</sup>. In other areas of micro-organisms, Roeder et al. even put forward an idea that a new CD8<sup>+</sup> T cell response towards the mutated antigen which caused HIV evade phenomenon could be generated in a population not selected for certain HLA alleles<sup>33</sup>. However, no one has studied the detailed CDR3 sequences as we did.

Tully *et al.* demonstrated that there are individualized skewed spectratyping in the TCR of TB patients<sup>34</sup>. TCR repertoire can reflect the function, status, and composition of CD4<sup>+</sup> T cells. In this study, we detected the spectratyping of  $\alpha$ - and  $\beta$ -chain in three patients, which showed some TCR V $\alpha$  and V $\beta$  gene families of E7-bound CD4<sup>+</sup> T cells were oligoclonal. Some reports evidenced that during the process of immune response to the MTB antigen, certain gene families had become predominant, which restricted the expression of other gene families<sup>35–37</sup>, which is consistent with our findings. Not only is the TCR V $\alpha$  and V $\beta$  gene families of E7-bound CD4<sup>+</sup> T cells were oligoclonal, the CDR3 amino-acid sequences were also identical, which proved powerfully that the CD4<sup>+</sup> T cells against one specific MTB antigen in one single individual were clonal expansion and their CDR3 region plays an important role in immune response against this specific antigen. Moreover, in one of the participants, we also observed the presence of several E7-bound CD4<sup>+</sup> T cells with different amino-acid sequences in the CDR3 region. In addition, E7-bound CD4<sup>+</sup> T cells shared one CDR3 amino-acid sequence with unbound CD4<sup>+</sup> T cells in PLF12. Several possible reasons may explain these observations. First, the different CDR3 sequences likely recognize antigens with the same size and similar spatial structures. Second, the pathogenicity of MTB and the molecular weight of the antigen, as well as the types of Ig, scattered or additional antigen sites, and variability of the MTB antigen composition could contribute to the diversity established. Third, the specificity between E7 and CD4<sup>+</sup> T cells might primarily depend on the flanking sequence of CDR3, which might be depend on a small number of key amino-acids. Fourth, patient-specific factors may also contribute to the degree of CDR3 variation against a specific MTB antigen.

The TCR antigen-binding loops use different conformations in their interaction with each distinct ligand<sup>38</sup>. In our examination, not every TB patients showed the identical CDR3 sequence of E7-bound CD4<sup>+</sup> T cells. Importantly, we also found completely different CDR3 sequences in some of the TB patients, whose architectural properties were consistent or not aroused our interest. In PLF13, the protein tertiary structure of the CD4<sup>+</sup> TCR  $\beta$ -chain CDR3 region showed similar each other, in spite of the completely different amino-acid sequences. These structures provided unique insights into the mechanisms of MTB antigen recognition between CD4<sup>+</sup> T cells and

MHC. E7/HLA-DR tetramers with different HLA-DRB1 alleles were capable of recognizing and binding CDR3 fragments with an identical sequence or with different sequences but a similar structure and function in one individual. In different TB patients, the diverse CDR3 sequences of E7-bound CD4<sup>+</sup> T cells seemed to show mostly similar protein tertiary structures (Fig. S1a). By contrast, the diverse CDR3 sequences of non-peptide tetramers bound CD4<sup>+</sup> T cells displayed almost different protein tertiary structures (Fig. S1b–d). We suggested that CDR3 region combined with an identical peptide, such as E7 or C5, had the same function and a similar complementary groove even if found in different TB patients. We imagine that each non-peptide/HLA-DRB1 tetramers combined with CD4<sup>+</sup> T cells in a nonspecific binding manner. When the conserved flanking regions bind to the side walls of antigen-binding-groove of MHC molecules and no peptide combine with TCR $\alpha\beta$  CDR3 region, the rearrangement of the TCR CDR3 is random and the sequences are messy. Thus, after the release of its specific peptide, the CDR3 region became oddly shaped. Furthermore, we compared the amino-acid properties of E7/C5/non-peptide bound CD4<sup>+</sup> T cells  $\beta$ -chain CDR3 region in PLF13 and found no difference between the three types of CD4<sup>+</sup> T cells, regardless of whether it was hydrophilic or hydrophobic; acidic, alkaline, or neutral. When we blocked the molecule CD3 and CD4 on the surface of CD4<sup>+</sup> T cells, the amino-acid sequences of the CDR3 region in E7-bound CD4<sup>+</sup> T cells were scattered. Thus, we believed that molecular chaperones CD3 and CD4 could assist in TCR recognition of the antigen. Therefore, to minimize the impact on the specific binding between antigen-MHC complex and TCRs by molecular chaperones due to the effect of CD3 and CD4, it is important that the dosage of CD3 and CD4 antibodies need to be carefully determined when used in combination with peptide/MHC tetramer, and special care should be taken in the experimental operation.

In conclusion, our findings reveal that in one individual, the MTB peptide E7/HLA-DR tetramers constructed with different HLA-DRB1 alleles can bind to CD4<sup>+</sup> T cells due to sharing the same CDR3 amino-acid sequence or a similar CDR3 structure and function, which suggests that E7-bound CD4<sup>+</sup> T cells with different HLA-DRB1 alleles undergo clonal expansion in TB patients. These results indicate that the CDR3 region of CD4<sup>+</sup> T cells is a key region for the recognition of a specific MTB antigen. The predominant TCR  $\alpha$ - and  $\beta$ -chain CDR3 sequence exhibit peptide specificity, and there is also certain HLA-DR restriction which may provide a clue of the possible causes and mechanisms of peptide-specific CD4<sup>+</sup> T cell-related presentation against MTB in different TB patients with different HLA-DR alleles.

## Materials and Methods

**Ethics and methods statements.** The recruitment process, delivery and use of clinical samples obtained from TB patients, and the experimental procedures were approved by the Medical Ethics Committee of Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China (Protocol number 2012-28). All of the patients provided written informed consent before enrollment in the study. And all methods and experiments were performed in accordance with the relevant guidelines and regulations.

**Patients and collection of samples.** In the present investigation, patients diagnosed with tuberculosis pleurisy were recruited from the Guangzhou Chest Hospital, Guangzhou City, Guangdong Province, China. The diagnosis of active tuberculosis pleurisy was made based on the following criteria: (1) positive pleural fluid smear or culture results for MTB; (2) detection of active tuberculosis pleurisy lesions by X-ray examination or pleural biopsy sections by MTB antigen-specific immunohistochemistry; and (3) the presence of typical systemic and local symptoms, such as fever, chills, sweating, loss of appetite, fatigue, night sweats, chest distress, chest pain, cough, and short breath or dyspnea. The possibility of a malignant lesion of the lung or lymph node was ruled out. PLF and peripheral blood samples were collected from inpatients that were untreated or in the initial stage of treatment.

**Preparation of MTB peptide or non-peptide/HLA-DR tetramers.** Stable drosophila S2 cell lines expressing soluble biotinylated MTB peptide or non-peptide/HLA-DR monomers were established, and then biotinylated monomers were expressed and purified for building tetramers as described earlier<sup>14,39</sup>. MTB peptide or non-peptide/HLA-DR tetramers were obtained by mixing the corresponding monomers with PE-labeled streptavidin (eBioscience) in an 8:1 molar ratio at room temperature for MACS.

**Obtaining MTB peptide or non-peptide/HLA-DR tetramers bound CD4<sup>+</sup> T cells.** Plural fluid cells samples were isolated by Ficoll-Hypaque (Axis-Shield) density gradient centrifugation and were suspended to a final concentration of  $1 \times 10^6$  cells/mL in complete RPMI 1640 medium (Gibco). CD4<sup>+</sup> T cells were purified from  $1 \times 10^7$  of PFCs in 40  $\mu$ L sorting buffer using 10  $\mu$ L of CD4<sup>+</sup> T-cell Negative Sorting Microbeads (Miltenyi Biotec). When required, 2  $\mu$ L CD3 and 2  $\mu$ L CD4 goat polyclonal antibodies (Santa Cruz Biotechnology) were used to block the action site before sorting by E7/HLA-DR tetramers in total 50  $\mu$ L sorting solution. Then, MTB peptide or non-peptide/HLA-DR tetramers bound CD4<sup>+</sup> T cells were purified from CD4<sup>+</sup> T cells using final concentration 6  $\mu$ g/mL of MTB peptide or non-peptide/HLA-DR tetramers labeled with PE prepared before and Anti-PE MicroBeads (Miltenyi Biotec, 10  $\mu$ L added in 90  $\mu$ L sorting solution). Meanwhile, the cells that were not in combination with Anti-PE magnetic beads namely unbound CD4<sup>+</sup> T cells were also preserved.

**Extraction of RNAs and synthesis of the first cDNAs.** Using TRIzol reagent (Invitrogen), total RNA was extracted from CD4<sup>+</sup> T cells obtained before the experiment. Total RNA was reverse-transcribed into cDNA by a reverse transcription-polymerase chain reaction (RT-PCR) using SuperScript TM III Reverse Transcriptase kit (Invitrogen). Double-stranded cDNA was synthesized by long-distance polymerase chain reaction (LD-PCR) with the primers and reagents provided from the BD SMART PCR cDNA Synthesis Kit (Agilent). The LD-PCR products were stored in  $-20^\circ\text{C}$  and utilized as DNA templates for cloning and sequencing of the CDR3 region and V region of TCR  $\alpha$ - and  $\beta$ -chain by multiple PCR reactions.



**Cloning and sequencing of TCR  $\alpha$ - and  $\beta$ -chain CDR3 region and V family.** TCR CDR3 size analysis of each sample was performed by multiple PCR amplification. We used mixed primer pairs to perform multiple PCR amplification. Each group of mixed primer pairs contains 1 to 5 forward primers and one reverse primer. One clone represents the CDR3 sequence successfully amplified by using one group of mixed primer pair. The reactions for each of the samples were carried out in 20  $\mu$ L mixtures that contained 0.4  $\mu$ L of a sense primer and 0.4  $\mu$ L of an antisense primer, 10  $\mu$ L of a Taq polymerase mix (Genestar), 1  $\mu$ L of DNA template, and 8.2  $\mu$ L of ddH<sub>2</sub>O. The experimental conditions consisted of the following cycles: 95 °C for 5 min, melting at 95 °C for 30 s, primer annealing at 60 °C for 30 s, 72 °C for 1 min (30 cycles), followed by extension at 72 °C for 1 min. Additionally, TCR V size analysis of each sample was performed by PCR amplification. The reaction for each samples were performed in 25  $\mu$ L mixtures that included 1  $\mu$ L of the sense primer and 1  $\mu$ L of the antisense primer, 0.125  $\mu$ L of Go-Taq flexi polymerase (Promega), 3  $\mu$ L of Mg<sup>2+</sup> (Promega), 0.5  $\mu$ L of dNTP (Promega), 1  $\mu$ L of a DNA template, 5  $\mu$ L 5  $\times$  buffer and 12.375  $\mu$ L of ddH<sub>2</sub>O. The experimental program contained the following cycles: 95 °C for 5 min, melting at 95 °C for 50 s, primer annealing at 63 °C for 1 min, 72 °C for 1 min (30 cycles), followed by extension at 72 °C for 5 min. Finally, all of the PCR products were separated on a 1% agarose gel (Biowest), stained with SYBR dye (Probe Bioscience), and purified using gel extraction kits (Omega). The purified DNA fragments were ligated into the pGEM-T easy vector (Promega), and the resulting plasmids were transfected by heat shock into competent *Escherichia coli* DH5h (TransGen Biotech) for propagation. The colonies which were positive for ampicillin resistance were suspended and analyzed by the Beijing Genomics Institute (Beijing, China).

**Genotyping of HLA-DR alleles of selected typical TB patients.** Peripheral blood samples were collected from two typical diagnosed TB patients (PLFs 11 and 12), and genotyping of HLA-DR alleles was performed, by the Beijing Genomics Institute. The HLA sequence-based typing (SBT) approach was adopted to obtain full-length coding sequences, and the results were compared in the IMGT/HLA database to define alleles.

**Statistics.** Data were statistically analyzed by using Lasergene and DNASTar software. We compared the nucleotide sequences in the IMGT/TCR database on the International ImmunoGeneTics information system website<sup>®</sup> (IMGT/V-QUEST, [http://www.imgt.org/IMGT\\_vquest/share/textes/](http://www.imgt.org/IMGT_vquest/share/textes/)). Then, we analyzed their V-(D)-J subgroup type and CDR3 region spectral type of TCR  $\alpha$ - and  $\beta$ -chains, respectively. Raw CDR3 amino-acid sequences were put into the Swiss-Model Workspace (<https://www.swissmodel.expasy.org/interactive>) to predict tertiary protein structure.

## References

- Lin, P. L. *et al.* CD4 T cell depletion exacerbates acute Mycobacterium tuberculosis while reactivation of latent infection is dependent on severity of tissue depletion in cynomolgus macaques. *AIDS Res Hum Retroviruses* **28**, 1693 (2012).
- Sonnenberg, P. *et al.* How soon after infection with HIV does the risk of tuberculosis start to increase? A retrospective cohort study in South African gold miners. *J Infect Dis* **191**, 150 (2005).
- Caruso, A. M. *et al.* Mice deficient in CD4 T cells have only transiently diminished levels of IFN- $\gamma$ , yet succumb to tuberculosis. *J Immunol* **162**, 5407 (1999).
- Lodi, S. *et al.* Risk of tuberculosis following HIV seroconversion in high-income countries. *Thorax* **68**, 207 (2013).
- Mogues, T., Goodrich, M. E., Ryan, L., LaCourse, R. & North, R. J. The relative importance of T cell subsets in immunity and immunopathology of airborne Mycobacterium tuberculosis infection in mice. *J Exp Med* **193**, 271 (2001).
- Bousso, P. Generation of MHC-peptide tetramers: a new opportunity for dissecting T-cell immune responses. *Microbes Infect* **2**, 425 (2000).
- Janeway, C. J. The T cell receptor as a multicomponent signalling machine: CD4/CD8 coreceptors and CD45 in T cell activation. *Annu Rev Immunol* **10**, 645 (1992).
- Reinherz, E. L. alpha $\beta$  TCR-mediated recognition: relevance to tumor-antigen discovery and cancer immunotherapy. *Cancer Immunol Res* **3**, 305 (2015).
- Ibrahim, F. A. *et al.* An Unusual Case of Hepatosplenic alpha $\beta$  T-Cell Lymphoma Presenting with Coombs'-Negative Hemolytic Anemia. *Clin Med Insights Oncol* **9**, 123 (2015).
- Rodriguez-Rodriguez, N. *et al.* Pro-inflammatory self-reactive T cells are found within murine TCR-alpha $\beta$ (+) CD4(-) CD8(-) PD-1(+) cells. *EUR J IMMUNOL* **46**, 1383 (2016).
- Birnbaum, M. E. *et al.* Molecular architecture of the alpha $\beta$  T cell receptor-CD3 complex. *Proc Natl Acad Sci USA* **111**, 17576 (2014).
- Altman, J. D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94 (1996).
- Yang, F. F. *et al.* Monitoring of peptide-specific and gamma interferon-productive T cells in patients with active and convalescent tuberculosis using an enzyme-linked immunosorbent spot assay. *Clin Vaccine Immunol* **19**, 401 (2012).
- Li, Y. *et al.* Use of HLA-DR\*08032/E7 and HLA-DR\*0818/E7 tetramers in tracking of epitope-specific CD4+T cells in active and convalescent tuberculosis patients compared with control donors. *Immunobiology* **216**, 947 (2011).
- Mason, D. A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol Today* **19**, 395 (1998).
- Kudela, P. *et al.* Cross-reactive CD4+T cells against one immunodominant tumor-derived epitope in melanoma patients. *J Immunol* **179**, 7932 (2007).
- D'Orsogna, L. J. *et al.* New tools to monitor the impact of viral infection on the alloreactive T-cell repertoire. *Tissue Antigens* **74**, 290 (2009).
- Caccamo, N. *et al.* Cytokine profile, HLA restriction and TCR sequence analysis of human CD4+T clones specific for an immunodominant epitope of Mycobacterium tuberculosis 16-kDa protein. *Clin Exp Immunol* **133**, 260 (2003).
- Ko, B. K. *et al.* Affinity Maturation of Monoclonal Antibody 1E11 by Targeted Randomization in CDR3 Regions Optimizes Therapeutic Antibody Targeting of HER2-Positive Gastric Cancer. *Plos One* **10**, e134600 (2015).
- Pickman, Y., Dunn-Walters, D. & Mehr, R. BCR CDR3 length distributions differ between blood and spleen and between old and young patients, and TCR distributions can be used to detect myelodysplastic syndrome. *Phys Biol* **10**, 56001 (2013).
- Herzig, C. T., Mailloux, V. L. & Baldwin, C. L. Spectratype analysis of the T cell receptor delta CDR3 region of bovine gammadelta T cells responding to leptospira. *Immunogenetics* **67**, 95 (2015).
- He, X. Y. *et al.* TRAV gene expression in PBMCs and TILs in patients with breast cancer analyzed by a DNA melting curve (FQ-PCR) technique for TCR alpha chain CDR3 spectratyping. *Neoplasma* **59**, 693 (2012).

23. Niemi, H. J., Laakso, S., Salminen, J. T., Arstila, T. P. & Tuulasvaara, A. A normal T cell receptor beta CDR3 length distribution in patients with APECED. *Cell Immunol* **295**, 99 (2015).
24. Campregher, P. V., Srivastava, S. K., Deeg, H. J., Robins, H. S. & Warren, E. H. Abnormalities of the alphabeta T-cell receptor repertoire in advanced myelodysplastic syndrome. *Exp Hematol* **38**, 202 (2010).
25. Joseph, C. G. *et al.* Association of the autoimmune disease scleroderma with an immunologic response to cancer. *Science* **343**, 152 (2014).
26. Wang, C. *et al.* Effects of aging, cytomegalovirus infection, and EBV infection on human B cell repertoires. *J Immunol* **192**, 603 (2014).
27. Zhang, W. *et al.* I-Monitor: A Robust Pipeline for TCR and BCR Repertoire Analysis. *Genetics* **201**, 459 (2015).
28. Chinnasamy, N. *et al.* A TCR targeting the HLA-A\*0201-restricted epitope of MAGE-A3 recognizes multiple epitopes of the MAGE-A antigen superfamily in several types of cancer. *J Immunol* **186**, 685 (2011).
29. Toso, J. F. *et al.* MAGE-1-specific precursor cytotoxic T-lymphocytes present among tumor-infiltrating lymphocytes from a patient with breast cancer: characterization and antigen-specific activation. *Cancer Res* **56**, 16 (1996).
30. Kedzierska, K., La Gruta, N. L., Stambas, J., Turner, S. J. & Doherty, P. C. Tracking phenotypically and functionally distinct T cell subsets via T cell repertoire diversity. *Mol Immunol* **45**, 607 (2008).
31. Axelsson-Robertson, R. *et al.* Extensive major histocompatibility complex class I binding promiscuity for Mycobacterium tuberculosis TB10.4 peptides and immune dominance of human leucocyte antigen (HLA)-B\*0702 and HLA-B\*0801 alleles in TB10.4 CD8 T-cell responses. *Immunology* **129**, 496 (2010).
32. Weichold, F. F. *et al.* Impact of MHC class I alleles on the M. tuberculosis antigen-specific CD8+T-cell response in patients with pulmonary tuberculosis. *Genes Immun* **8**, 334 (2007).
33. Roeder, J. *et al.* Adaptation of CD8 T cell responses to changing HIV-1 sequences in a cohort of HIV-1 infected individuals not selected for a certain HLA allele. *Plos One* **8**, e80045 (2013).
34. Tully, G. *et al.* Highly focused T cell responses in latent human pulmonary Mycobacterium tuberculosis infection. *J Immunol* **174**, 2174 (2005).
35. Zhou, J. W., Kong, C., Luo, J., Cao, J. & Shi, Y. Comparing TCR beta chain variable gene profile skewness between children with tuberculosis and BCG-vaccinated children. *Arch Iran Med* **16**, 104 (2013).
36. Xu, C. *et al.* Diagnostic significance of TCR gene clonal rearrangement analysis in early mycosis fungoides. *Chin J Cancer* **30**, 264 (2011).
37. Kim, S. T. *et al.* Distinctive CD3 heterodimeric ectodomain topologies maximize antigen-triggered activation of alpha beta T cell receptors. *J Immunol* **185**, 2951 (2010).
38. Yin, L. *et al.* A single T cell receptor bound to major histocompatibility complex class I and class II glycoproteins reveals switchable TCR conformers. *Immunity* **35**, 23 (2011).
39. Yang, J., Jaramillo, A., Shi, R., Kwok, W. W. & Mohanakumar, T. *In vivo* biotinylation of the major histocompatibility complex (MHC) class II/peptide complex by coexpression of BirA enzyme for the generation of MHC class II/tetramers. *Hum Immunol* **65**, 692 (2004).

## Acknowledgements

We would like to express our gratitude to the staff of Guangzhou Chest Hospital, Guangzhou city, Guangdong province, China, for providing PLFs of TB patients. We are also thankful to Liangliang Ren (Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China), Na Zhang (Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China), and Ming Gao (Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China) for the plasmid construction. This work was supported by the National Natural Sciences Foundation of China (project 81271779) (to XL), the National Science and Technology Major Project of China (project 2012ZX10004903-004-002) (to XL), the Natural Sciences Foundation of Guangdong Province of China (project 2014A030313725) (to YF), and the Guangzhou Science and Technology Program of China (projects 155700012 and 201604020019) (to XL).

## Author Contributions

X.L. designed the research; Y.F., J.L., L.Z., Y.T., T.C. and Y.S. provided specimens; Y.G., C.W., Y.Y., X.T. and J.W. conducted the experiments and acquired data; Y.G., C.W., Y.Y., X.T., X.H. and K.Z. analyzed the data obtained; Y.G. and X.L. wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-28344-7>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018