



OPEN

Dynamic analysis of peripheral blood TCR β -chain CDR3 repertoire in occupational medicamentosa-like dermatitis due to trichloroethylene

Dafeng Lin¹✉, Dianpeng Wang¹, Peimao Li¹, Xiangli Yang¹, Wei Liu², Lu Huang³, Zhimin Zhang¹, Yanfang Zhang¹, Wen Zhang¹, Naixing Zhang¹✉, Ming Zhang¹ & Xianqing Huang¹

Previously, we had cross-sectionally explored the characteristics of T cell receptor (TCR) repertoires from occupational medicamentosa-like dermatitis due to trichloroethylene (OMDT) patients, now we further analyzed the dynamic features of OMDT TCR repertoires. Peripheral blood TCR β -chain complementarity-determining region 3 (CDR3) genes were detected with the high throughput sequencing in 24 OMDT cases in their acute, chronic and recovery stages, respectively, and in 24 trichloroethylene-exposed healthy controls. The TCR repertoire diversity, TRBV/TRBD/TRBJ gene usage and combination, frequencies of CDR3 nucleotide (nt) and amino acid (aa) sequences in the cases in different stages and in the controls were analyzed. TRBV6-4 and TRBV7-9 frequencies significantly differed between the cases and controls (both $P < 6.1 \times 10^{-4}$). TRBV6-4 combination with TRBJ2-1, TRBJ2-2, TRBJ2-3, and TRBJ2-6, and TRBV7-9 combination with TRBJ2-1 were associated with the stage by OMDT severity (all $P < 0.001$). Ten CDR3-nt and 7 CDR3-aa sequences in TRBV7-9-TRBJ2-1 combination and 1 CDR3-nt and 1 CDR3-aa sequences in TRBV6-4-TRBJ2-1 combination were identified as associated with the severity of OMDT (all $P < 0.001$). We revealed further how TCR repertoires vary with the severity in the development of OMDT, and severity-related TCRs may provide important therapeutic targets for OMDT in clinical practice.

Trichloroethylene (TCE) is known as an environmental contaminant and an industrial toxicant mainly used as organic solvent and degreasing agent, mostly in some Asian countries like China, Japan and Korea over the last few decades^{1,2}. Exposure to TCE was reported to cause a notorious occupational disease named occupational medicamentosa-like dermatitis due to trichloroethylene (OMDT). The manifestation of OMDT includes various degrees of generalized skin lesions, fever, hepatitis, and lymphadenopathy. It resembles serious drug hypersensitivities referred to as drug reaction with eosinophilia and systemic symptoms (DRESS) or drug-induced hypersensitivity syndrome (DIHS)². Epidemiological studies predicted that the incidence of OMDT was about 1%², but more than 7% of the patients died of secondary hepatic encephalopathy, severe infection, and multiple organ failure³. Survived patients usually had a very long clinical course, and the symptoms could relapse after cessation of the treatment.

The clinical features of OMDT include a latency of 2–5 weeks, no dose–response relationship, skin lesions not confined to the contact area, the relatively effective glucocorticoid therapy, and rapid recurrence due to re-exposure to trichloroethylene^{2,4}. Based on the above features, OMDT is generally regarded as a T cell-mediated, Type IV hypersensitivity reaction. Other pathological evidence includes that elevated levels of T cells were identified in peripheral blood and affected skin tissues of the patients⁵. Therefore, T cells may play a pivotal role

¹Medical Laboratory, Shenzhen Prevention and Treatment Center for Occupational Diseases, 2019 Buxin Rd., Luohu district, Shenzhen 518020, China. ²Key Laboratory of Modern Toxicology of Shenzhen, Medical Key Laboratory of Guangdong Province, Medical Key Laboratory of Health Toxicology of Shenzhen, Shenzhen Center for Disease Control and Prevention, Shenzhen 518055, China. ³Fuyong Prevention and Health Care Center, Bao'an District, Shenzhen 518103, China. ✉email: david1385@foxmail.com; zhangnx@wjw.sz.gov.cn

Variables	OMDT cases (n = 24)	Controls (n = 24)	P
Age (year, M (MAD))	24.0 (8.2)	28.0 (7.4)	0.261 ^a
Sex (male, %)	17 (70.8)	19 (79.2)	0.739 ^b
Race (Han, %)	20 (83.3)	22 (91.7)	0.666 ^c
Smoking (yes, %)	9 (37.5)	7 (29.2)	0.760 ^b
Drinking (yes, %)	1 (4.2)	2 (8.3)	1.000 ^c
Trichloroethylene exposure concentration T_{WA} (mg/m ³ , M (MAD))	15.2 (8.2)	1.6 (1.2)	<0.001 ^a
Trichloroethylene exposure period (day, M (MAD))	31.5 (6.7)	220.5 (60.0)	<0.001 ^a

Table 1. Baseline characteristics of the study subjects. OMDT, occupational medicamentosa-like dermatitis due to trichloroethylene; M, median; MAD, median absolute deviation; TWA, time weighted average. ^aMann–Whitney U test. ^b χ^2 -test. ^cFisher’s exact test.

in the pathogenesis of OMDT. T cell receptor (TCR) is a membrane protein of T cell responsible for recognizing antigens and transmitting signals to downstream molecules⁶. Because TCR is the key to initiate T cell reactions, studying TCR in OMDT may help to clarify the underlying pathogenetic mechanism and provide basis for more effective prevention and treatment measures.

Each T cell is known to generally encode one single, unique TCR. To ward off a wide variety of pathogens, a vast array of differential TCRs in the human adaptive immune system constitute a TCR repertoire⁷. Most of TCRs in the repertoire are composed of an α -chain and a β -chain, and each of the chains contains a variable region (V region) which has three hypervariable regions, namely complementarity-determining region (CDR) 1, CDR2, and CDR3. The most variable is the CDR3 which directly determines the antigen-binding specificity of the TCR⁷. Rearrangement of variable (V), diversity (D) and joining (J) segments occurs in the CDR3 gene during the maturity process of the lymphocyte⁷. Notably, β -chain CDR3 is the mostly studied because β -chain is a better indicator to distinguish different T cell clones from another.

Previously, we conducted a case–control study to explore the characteristics of TCR repertoires in OMDT by the high-throughput sequencing (HTS) of the TCR β -chain CDR3⁸. In the previous, cross-sectional study, we found 5 TCR β -chain V gene (TRBV) and 1 TCR β -chain J gene (TRBJ) usage, and 48 TRBV-TRBJ combinations that might be related to OMDT. We also identified 17 CDR3 nucleotide (nt) sequences and 16 CDR3 amino acid (aa) sequences of which the corresponding T cell clones might be specific to OMDT and participate in OMDT development⁸. However, the human TCR repertoire is quite sophisticated, and it actually dynamically shifts with the disease status. One cross-sectional exploration with limited sample size is difficult to comprehensively reveal the characteristics of TCR repertoires in OMDT. Therefore, we now increased the sample size and carried out a longitudinal study to further investigate the dynamic features of TCR repertoires in the course of OMDT.

Results

Baseline characteristics of the study subjects and blood test indices. Characteristics of the cases at admission and of the controls are listed in Table 1. The cases had an average age of 24.0 years, comparable to that of the controls (28.0 years). Seventy percent of the cases were male, and 83.3% were Han Chinese, also comparable to the percentage of male (79.2%) and of Han (91.7%) in the controls, respectively. Smoking and drinking habits were matched between the cases and controls. The cases were exposed to a higher average trichloroethylene concentration than the controls (15.2 vs. 1.6 mg/m³, $P = 2.7 \times 10^{-6}$), and their average trichloroethylene exposure period was significantly shorter (31.5 vs. 220.5 days, $P = 2.9 \times 10^{-9}$).

The averages of blood test indices in the controls in Table 2 were all within their respective clinical references. The OMDT cases in acute stage showed significantly higher white blood cell (WBC) count, neutrophil count, alanine aminotransferase (ALT) activity, and aspartate aminotransferase (AST) activity, and lower red blood cell (RBC) count, hemoglobin concentration, total protein (TP) concentration, and albumin concentration than the controls (all $P < 0.05$); the cases in acute stage also showed higher ALT and AST activities than in chronic and recovery stages (all $P < 0.05$). The cases in chronic stage had higher WBC, lymphocyte, and neutrophil counts, ALT and AST activities, and lower RBC count, hemoglobin, TP, and albumin concentrations than the controls (all $P < 0.05$), the cases in chronic stage also had higher ALT and AST activities than in recovery stage (both $P < 0.01$). The cases in recovery stage showed lower RBC count, hemoglobin, TP, and albumin concentrations than the controls (all $P < 0.05$), while the other indices were comparable to their counterparts in the controls.

TCR repertoire diversity. The HTS yielded an average of 16,313,850 raw reads per sample. After filtration of low quality reads and adapter contamination, we obtained an average of 16,068,894 clean reads per sample. Each sample has high clean read Q20 and Q30, indicating good quality of the sequencing data. The overview of sequence statistics is shown in Supplementary Table S1 online.

We used 6 indices to measure the TCR repertoire diversity. None of the 6 indices, including resampled diversity, extrapolate diversity, chaoE, chao1, Shannon–Wiener index, and inverse Simpson index, was found significantly different between the cases and the controls, or between different stages of the cases (see Supplementary Fig. S1 online).

TRBV/TRBD/TRBJ usage and combination. After the HTS, we identified 66 TRBV subtypes, 2 TRBD subtypes, and 14 TRBJ subtypes in total. Frequency of each TRBV/TRBD/TRBJ subtype was plotted accord-

Variables	OMDT cases (n = 24)			Controls (n = 24)
	Acute stage	Chronic stage	Recovery stage	
White blood cell count ($\times 10^9/L$, M (MAD)) ^a	11.75 (5.72)*	13.74 (5.81)** [#]	7.81 (3.97)	6.92 (2.09)
Lymphocyte count ($\times 10^9/L$, M (MAD)) ^a	2.74 (1.88)	3.24 (0.95)*	2.42 (0.90)	2.37 (0.67)
Neutrophil count ($\times 10^9/L$, M (MAD)) ^a	7.06 (3.31)*	8.79 (2.90)** [#]	4.66 (3.23)	4.20 (1.70)
Red blood cell count ($\times 10^9/L$, M (MAD)) ^a	4.86 (0.46)*	4.29 (0.71)**	4.53 (0.47)*	5.43 (0.48)
Hemoglobin concentration (g/L, mean (SD)) ^b	134.61 (20.59)**	129.33 (17.82)**	133.15 (18.93)**	155.12 (16.20)
Alanine aminotransferase activity (U/L, M (MAD)) ^a	503.00 (434.40)** [#] [§]	181.00 (171.98)** [#]	25.00 (19.27)	23.50 (14.83)
Aspartate aminotransferase activity (U/L, M (MAD)) ^a	199.00 (166.05)** [#] [§]	60.00 (51.15)** [#]	21.00 (5.93)	23.50 (7.41)
Total protein concentration (g/L, M (MAD)) ^a	59.40 (6.82)**	58.45 (5.56)**	61.05 (4.08)**	76.50 (3.71)
Albumin concentration (g/L, M (MAD)) ^a	34.00 (4.60)** [#]	36.30 (3.71)**	38.25 (2.59)**	47.50 (3.71)

Table 2. Comparison of blood test indices between the groups. OMDT, occupational medicamentosa-like dermatitis due to trichloroethylene; M, median; MAD, median absolute deviation; SD, standard deviation. ^aMann–Whitney U test with Holm’s adjustment. ^bANOVA with Tukey’s HSD test. Compared with controls, * $P < 0.05$, ** $P < 0.01$; compared with cases in recovery stage, # $P < 0.05$, ## $P < 0.01$; compared with cases in chronic stage, § $P < 0.01$.

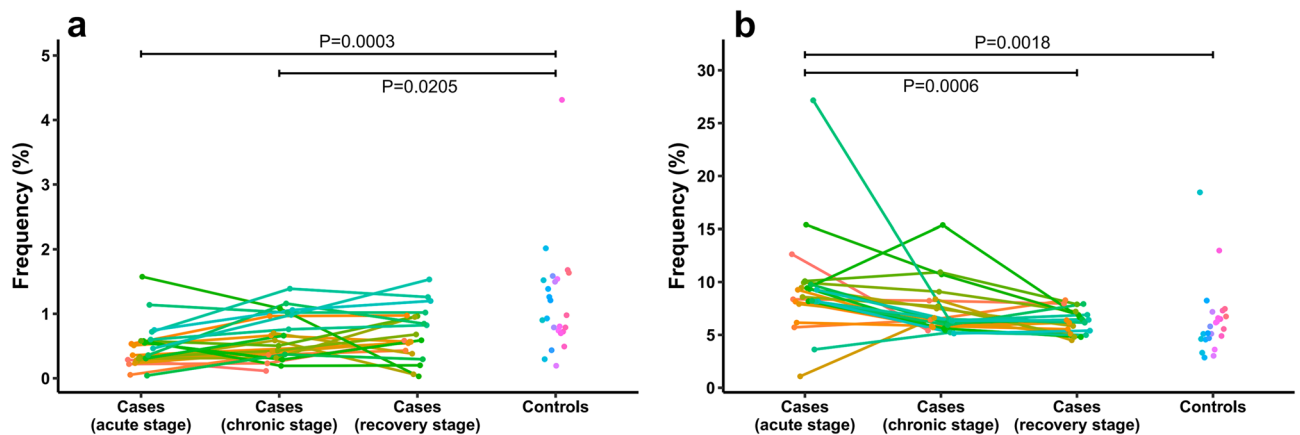


Figure 1. TRBV frequencies of the OMDT cases in acute, chronic, and recovery stages and of the controls. (a) TRBV6-4. (b) TRBV7-9. Dots of the same color represent TRBV frequencies of the same subject, the connecting lines indicate varying trends, and P values were obtained by the Man-Whitney U test with Holm’s adjustment.

ing to case in 3 stages and control, and the averages of frequency were compared between the 4 groups using the Kruskal–Wallis test with adjusted $P = 6.1 \times 10^{-4}$. We found that only TRBV6-4 and TRBV7-9 usage were significantly different between the 4 groups ($P = 2.1 \times 10^{-4}$ and $P = 1.5 \times 10^{-4}$, respectively). Further between-group analysis of TRBV6-4 and TRBV7-9 subtypes by the Man-Whitney U test with Holm’s adjustment revealed that TRBV6-4 frequencies of the cases in acute and chronic stages were significantly lower than the controls ($P = 3.3 \times 10^{-4}$ and $P = 0.020$, respectively), and TRBV7-9 frequencies of the cases in acute stage were significantly higher than the cases in recovery stage and the controls ($P = 6.3 \times 10^{-4}$ and $P = 0.002$, respectively) (Fig. 1). Fold changes of TRBV6-4 and TRBV7-9 frequencies in the OMDT cases relative to the controls were listed in Supplementary Table S2 online.

TRBV-TRBJ combination is an important source of CDR3 sequence diversification. TRBV6-4/TRBV7-9 and 14 TRBJ combinations were compared between the 4 groups using the Kruskal–Wallis test with adjusted $P = 0.001$. We found that only TRBV6-4-TRBJ2-1, TRBV6-4-TRBJ2-2, TRBV6-4-TRBJ2-3, TRBV6-4-TRBJ2-6, and TRBV7-9-TRBJ2-1 combinations were significantly different between the 4 groups ($P = 3.1 \times 10^{-4}$, $P = 4.1 \times 10^{-5}$, $P = 2.6 \times 10^{-6}$, $P = 5.0 \times 10^{-7}$, and $P = 4.2 \times 10^{-4}$, respectively). We further conducted between-group analysis of the 5 TRBV-TRBJ combinations by the Man-Whitney U test with Holm’s adjustment. As shown in Fig. 2, TRBV6-4-TRBJ2-1 combination frequencies of the cases in acute, chronic, and recovery stages were significantly lower than the controls ($P = 3.3 \times 10^{-4}$, $P = 0.020$, and $P = 0.020$, respectively). TRBV6-4-TRBJ2-2 frequencies of the cases in acute and chronic stages were significantly lower than the controls ($P = 1.1 \times 10^{-4}$ and $P = 0.004$, respectively). TRBV6-4-TRBJ2-3 frequencies of the cases in acute, chronic, and recovery stages were significantly lower than the controls ($P = 4.0 \times 10^{-6}$, $P = 0.001$, and $P = 0.042$, respectively), and TRBV6-4-TRBJ2-3 frequencies of the cases in acute stage were significantly lower than the cases in recovery stage ($P = 0.046$). TRBV6-4-TRBJ2-6 frequencies of the cases in acute and chronic stages were significantly lower than the controls ($P = 1.6 \times 10^{-5}$ and $P = 3.3 \times 10^{-4}$, respectively) and the cases in recovery stage ($P = 0.002$ and $P = 0.038$,

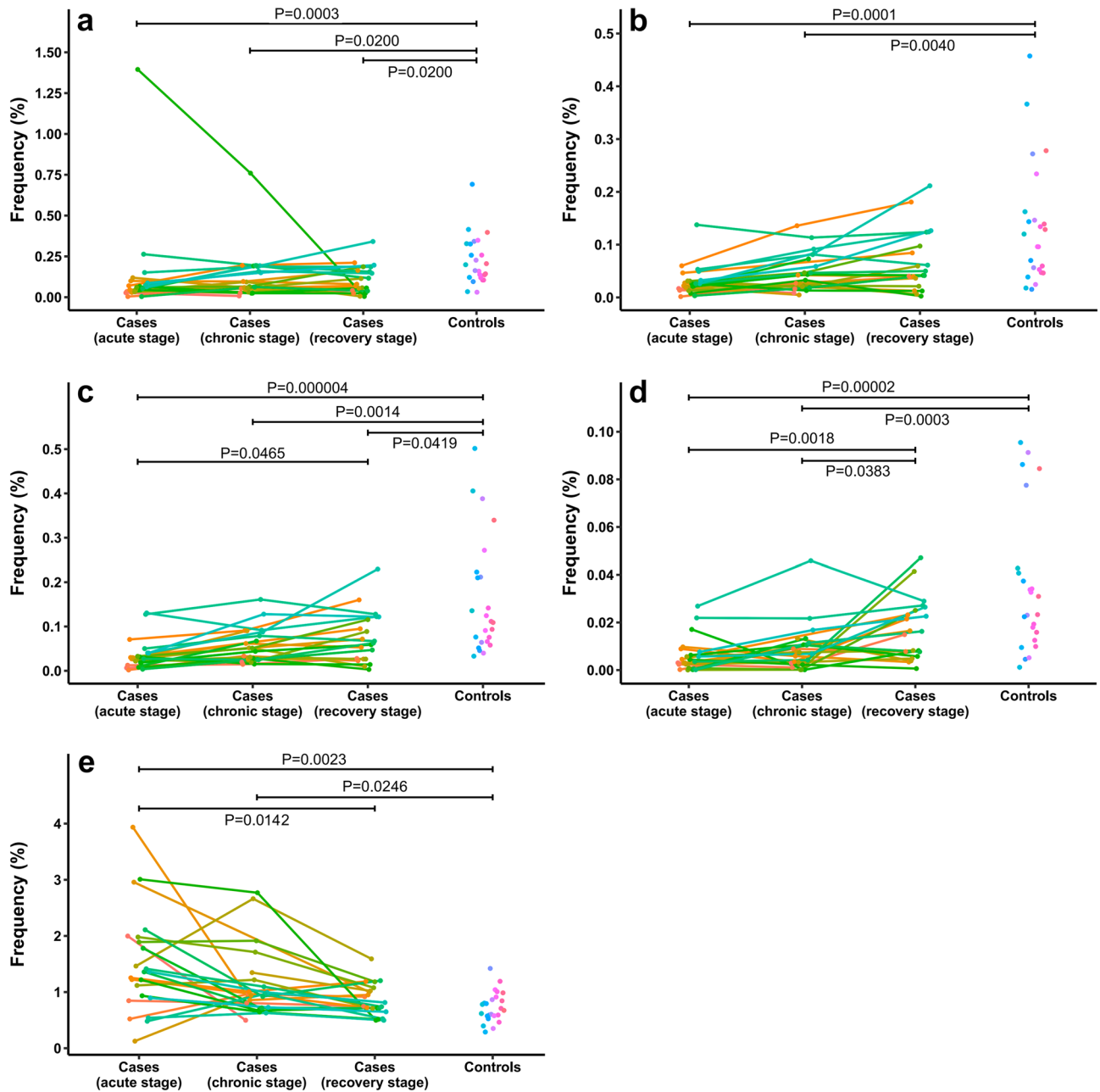


Figure 2. TRBV-TRBJ combination frequencies of the OMDT cases in acute, chronic, and recovery stages and of the controls. (a) TRBV6-4-TRBJ2-1. (b) TRBV6-4-TRBJ2-2. (c) TRBV6-4-TRBJ2-3. (d) TRBV6-4-TRBJ2-6. (e) TRBV7-9-TRBJ2-1. Dots of the same color represent TRBV-TRBJ combination frequencies of the same subject, the connecting lines indicate varying trends, and *P* values were obtained by the Man-Whitney U test with Holm's adjustment.

respectively). TRBV7-9-TRBJ2-1 frequencies of the cases in acute stage were significantly higher than the cases in recovery stage and the controls ($P=0.014$ and $P=0.002$, respectively), and TRBV7-9-TRBJ2-1 frequencies of the cases in chronic stage were also significantly higher than the controls ($P=0.025$). Fold changes of the 5 TRBV-TRBJ combination frequencies in the OMDT cases relative to the controls were listed in Supplementary Table S2 online.

CDR3 nucleotide and amino acid sequences. For each of the 5 significant TRBV-TRBJ combinations, there included too many CDR3-nt or CDR3-aa sequences. For example, the TRBV7-9-TRBJ2-1 combination included 205,117 CDR3-nt sequences and 174,284 CDR3-aa sequences, and most of them were of low frequency. Since the most expanded clonotypes are more likely to show association with OMDT, we only selected CDR3-nt or CDR3-aa sequences with a frequency more than 0.05% from the 5 TRBV-TRBJ combinations. Eventually, there were 8 CDR3-nt and 8 CDR3-aa sequences selected from TRBV6-4-TRBJ2-1 combination, 2 CDR3-nt

Sequences	Percentage of frequency (%) ^a	Subjects	Estimate (SE)	t	P
cdr3-nt					
TRBV7-9-TRBJ2-1					
TGTGCCAGCAGGCCAACGGGGGGCCTGCGGAATGAGCAGTTCTTC	0.407	Case 3, 13, 19	- 0.266 (0.003)	- 91.8	1.2 × 10 ⁻⁴
TGTGCCAGCAGCCCACGTGGGTATGAGCAGTTCTTC	2.600	Case 5, 7, 10	- 1.732 (0.002)	- 870.8	1.3 × 10 ⁻⁶
TGTGCCAGCAGCCCCGGGGTATGAGCAGTTCTTC	0.149	Case 5, 7, 10	- 0.099 (0.000)	- 730.8	1.9 × 10 ⁻⁶
TGTGCCAGCAGCCCGAGAGGATATGAGCAGTTCTTC	0.117	Case 5, 7, 10	- 0.078 (0.000)	- 1350.8	5.5 × 10 ⁻⁷
TGTGCCAGCAGCCCCTTACTAGCGGGGCGTACAATGAGCAGTTCTTC	0.053	Case 65, 66, 68	0.035 (0.000)	5429.0	1.2 × 10 ⁻⁴
TGTGCCAGCAGCTTAAGGGAAAACTCCTACAATGAGCAGTTCTTC	0.101	Case 5, 11, 12	0.067 (0.000)	1094.8	8.3 × 10 ⁻⁷
TGTGCCAGCAGCGAAACTAGCGGGGGCCCTACAATGAGCAGTTCTTC	0.121	Case 7, 9, 13, 61	0.081 (0.000)	2954.0	1.2 × 10 ⁻⁷
TGTGCCAGCAGCGGGCGACTTTCTCTGGGGAATGAGCAGTTCTTC	0.059	Case 7, 9, 13, 61	0.039 (0.000)	3603.0	7.7 × 10 ⁻⁸
TGTGCCAGCAGCTTCGTTTAGCGGGGCTGGAGCTCCACAATGAGCAGTTCTTC	0.236	Case 7, 9, 13, 61	0.154 (0.001)	110.9	1.6 × 10 ⁻⁶
TGTGCTAGCAGCGAGGGCAAGAGCAGTTCTTC	0.063	Case 62, 63 64, 68	0.042 (0.000)	434.1	5.3 × 10 ⁻⁶
TRBV6-4-TRBJ2-1					
TGTGCCAGTCTTCTTCGGTTAGCTCCTACAATGAGCAGTTCTTC	0.865	Case 66,93,95,96 Control 104,117,118,120,121,123,124	- 0.045 (0.007)	- 6.8	8.0 × 10 ⁻⁵
cdr3-aa					
TRBV7-9-TRBJ2-1					
CASRPTGGLRNEQFF	0.407	Case 3, 13, 19	- 0.266 (0.003)	- 91.8	1.2 × 10 ⁻⁴
CASSPLTSGAYNEQFF	0.053	Case 65, 66, 68	0.035 (0.000)	5429.0	1.2 × 10 ⁻⁴
CASLRENSYNEQFF	0.101	Case 5,11,12,13,19	0.067 (0.000)	647.9	3.4 × 10 ⁻¹¹
CASSETSGGAYNEQFF	0.121	Case 7, 9, 13, 61	0.081 (0.000)	2954.0	1.2 × 10 ⁻⁷
CASSGGLSLGNEQFF	0.059	Case 7, 9, 13, 61	0.039 (0.000)	3603.0	7.7 × 10 ⁻⁸
CASSFV*RGWSSYNEQFF	0.236	Case 7, 9, 13, 61	0.154 (0.002)	102.6	2.0 × 10 ⁻⁶
CASSAGQEQFF	0.067	Case 62, 63 64, 68	0.044 (0.000)	441.2	5.1 × 10 ⁻⁶
TRBV6-4-TRBJ2-1					
CASSSSVSSYNEQFF	0.865	Case 66,93,95,96 Control 104,117,118,120,121,123,124	- 0.045 (0.007)	- 6.8	8.0 × 10 ⁻⁵

Table 3. OMDT severity-associated TCR β -chain CDR3 sequences. OMDT, occupational medicamentosa-like dermatitis due to trichloroethylene; SE, standard error. Linear regression of sequence frequencies with decreasing OMDT severity from acute to recovery stage. ^aPercentage of frequency of the sequence in the total frequency of corresponding TRBV-TRBJ combination.

and 2 CDR3-aa sequences from TRBV6-4-TRBJ2-2 combination, 6 CDR3-nt and 6 CDR3-aa sequences from TRBV6-4-TRBJ2-3 combination, 1 CDR3-nt and 1 CDR3-aa sequences from TRBV6-4-TRBJ2-6 combination, and 98 CDR3-nt and 96 CDR3-aa sequences from TRBV7-9-TRBJ2-1 combination.

After comparison of frequencies between the groups, we obtained 10 CDR3-nt and 7 CDR3-aa sequences in TRBV7-9-TRBJ2-1 combination, and 1 CDR3-nt and 1 CDR3-aa sequences in TRBV6-4-TRBJ2-1 combination, reaching the statistically significant levels. Their linear trends with decreasing OMDT severity from acute to recovery stage were also analyzed and listed in Table 3. To be noted, the 10 CDR3-nt and 7 CDR3-aa sequences in TRBV7-9-TRBJ2-1 combination were not found in the controls.

Discussion

OMDT has been regarded as a T cell-mediated, Type IV hypersensitivity reaction^{2,4}, so T cells may play a pivotal role in the development of OMDT. TCR is the key to initiate T cell reactions, therefore, studying TCR is of great significance for the treatment and prevention of OMDT. Previously, we had explored cross-sectionally the characteristics of TCR repertoires from OMDT patients⁸; the dynamic features of OMDT TCR repertoires were further analyzed in the present study.

It is known that OMDT is a kind of hypersensitivity disease, the occurrence of OMDT is not related to the exposure dose of trichloroethylene². Although the average trichloroethylene exposure concentration was higher in the cases, there were cases exposed to trichloroethylene concentration much lower than the controls. All the controls had been exposed to trichloroethylene for more than half a year, and their average trichloroethylene exposure period was significantly longer than the cases.

The staging of OMDT was based on the main clinical manifestation and auxiliary laboratory examination of the cases⁹. It is noticeable that most of the laboratory indices in the cases in acute stage significantly deviated

from those in the controls, and generally returned in recovery stage to the reference levels as in the controls (Table 2), suggesting a decreasing severity of OMDT from acute to recovery stage.

Dynamic analysis of OMDT TCR repertoires started with the repertoire diversity. As reported previously in the pilot cross-sectional study⁸, we found no significant difference in CDR3 repertoire diversity measured by 6 indices between the cases and the controls, although decreased diversity has been observed in similar diseases^{10–12}. Remarkably, the 6 indices all showed an increasing trend with the decreasing severity from acute to recovery stage in the cases, however, the differences between the stages were not statistically significant.

We next examined TRBV/TRBD/TRBJ usage and combination, and found that only TRBV6-4 and TRBV7-9 usage were significantly associated with OMDT, and 5 of their combinations with TRBJ were further identified to be OMDT severity-related. In the previous cross-sectional study⁸, we also identified that TRBV6-4 frequencies of the cases were significantly lower than the controls ($P < 0.05$), and TRBV7-9 frequencies were higher in the cases, but the difference with the controls was not statistically significant. It seems that the time point of repertoire detection is quite critical, because we found in this longitudinal study that only the cases in acute stage showed significantly higher TRBV7-9 frequencies than the controls, suggesting the importance of dynamic analysis in studies of TCR repertoire. Moreover, decreased frequencies of TRBV6-4-TRBJ2-1, TRBV6-4-TRBJ2-2, TRBV6-4-TRBJ2-3, and TRBV6-4-TRBJ2-6 combinations in the cases in the previous cross-sectional study were also validated in the present study, and we further observed their restoring trends from acute to recovery stage by the dynamic analysis. Other significant TRBV usage such as TRBV7-6 and TRBV4-1, and TRBV-TRBJ combinations such as TRBV7-6-TRBJ2-5 and TRBV4-1-TRBJ2-5 from the previous study were not re-identified, because some of the prior positive results were possibly obtained by chance due to the cross-sectional design, and some of them still need to be further validated with larger sample size.

In the next step, we further analyzed part of CDR3-nt and CDR3-aa sequences with relatively high frequency from the 5 significant TRBV-TRBJ combinations. We finally identified 10 CDR3-nt sequences and 7 CDR3-aa sequences in TRBV7-9-TRBJ2-1 combination and 1 CDR3-nt sequence and 1 CDR3-aa sequence in TRBV6-4-TRBJ2-1 combination, of which the frequencies significantly varied with the severity of OMDT. Because these sequences were not commonly shared between the cases, none of them was previously found in the cross-sectional study⁸.

We observed opposite trends in the analysis of TRBV/TRBD/TRBJ usage and combination and CDR3 sequence frequency with varying severity of OMDT from acute to recovery stage. The decreasing trends, such as in the CDR3-nt sequence 'TGTGCCAGCAGGCCAACGGGGGGCCTGCGGAATGAGCAGTTCTTC' and CDR3-aa sequence 'CASRPTGGLRNEQFF', indicated that the corresponding T cells might be effector cells. They could rapidly expand when OMDT occurred, and gradually fell back to normal levels as the disease was cured. On the other hand, the increasing trends, such as in the CDR3-nt sequence 'TGTGCCAGCAGCCCCTTGACTAGCGGGGCGTACAATGAGCAGTTCTTC' and CDR3-aa sequence 'CASSPLTSGAYNEQFF', indicated that the corresponding T cells might be regulatory cells (Tregs), which are as important as effector cells in the pathology of hypersensitivity disease^{13,14}. The assumption was based on the animal experiment report that TCE caused imbalance of T helper cell subsets in BALB/c mice¹⁵, and more importantly, on the clinical evidence that Treg ratios significantly decreased in OMDT in the acute stage¹⁶. Those Tregs could be first inhibited when OMDT occurred, and gradually restored to normal levels with the recovery from the disease.

Additionally, we analyzed the potential public T cell clonotypes associated with OMDT by comparing Jaccard similarity indices between the cases in different stages and the controls. As shown in Supplementary Fig. S2 online, Jaccard similarity indices calculated based on CDR3-aa sequences in the cases in acute stage and in recovery stage were significantly higher than in the controls (both $P < 0.01$), but Jaccard similarity indices calculated based on CDR3-nt sequences were not significantly different between the cases and controls. The results suggested that the situation of public T cell clonotypes was similar to the previous study⁸, and that the possibility of public T cell clonotypes associated with OMDT was relatively small.

Susceptibility of OMDT had been linked to *HLA-B*13:01* genotype in the previous literature^{17,18}, so we tried to screen the TCR β CDR3 repertoire under the same *HLA-B*13:01* background in the previous case-control study⁸. Unexpectedly, different TCR clonotypes were found specific to OMDT cases carrying the same *HLA-B*13:01* allele, and there were also patients carrying other alleles such as *HLA-B*08:01*, *HLA-B*44:03*, *HLA-B*46:01*, and *HLA-B*07:02*, etc., which suggest that different mechanisms of immune recognition might occur in OMDT. Considering the complexity of OMDT mechanism, we focused only on the dynamic change of TCR repertoires with the disease severity in the present study, leaving the mechanisms of immune recognition in OMDT to be further discussed in the future studies.

Taken together, the present study, being an important sequel to the previous cross-sectional study, further explored in depth the dynamic characteristics of OMDT TCR repertoires. Not only did it validate the association of some previously found TRBV and TRBV-TRBJ combinations with OMDT, but also it discovered how TCR repertoires vary with the severity in the course of OMDT. Moreover, identification of specific, severity-related TCRs may provide important therapeutic targets as well as monitoring indicators for OMDT in clinical practice. Because of the dynamic features of TCR repertoires, the time point of repertoire detection is vitally important. Limited by the samples, we only detected TCR repertoires at one time point in each of the stages for each case, which hindered us from outlining the panorama of the whole OMDT TCR repertoire. Decreased repertoire diversity, highly expanded T cell clonotypes, more severity-associated TCRs, and other new features of TCR repertoires might be further uncovered by multiple time-point longitudinal studies in the future.

Methods

Cases and controls. Twenty-four OMDT patients, who were successively admitted to Shenzhen Prevention and Treatment Center for Occupational Diseases (SPTCOD) from January 2014 to December 2018, were recruited as cases. A panel of at least three occupational physicians diagnosed their diseases based on the diagnostic criteria GBZ 185–2006 issued by Ministry of Health of the People's Republic of China¹⁹. The clinical courses of the cases all ended before June 2019. Twenty-four healthy workers, who had been exposed to trichloroethylene during their work for more than half a year, were randomly recruited as controls when they received health check in SPTCOD from January 2017 to June 2019.

We collected detailed information on demographic characteristics, occupational history and medical conditions from the 24 OMDT cases and 24 controls. We also recorded clinical data of the cases during their hospitalization including main clinical manifestation and auxiliary examination results. The clinical courses of the cases were divided into 3 stages, namely acute stage, chronic stage (or early period of chronic stage), and recovery stage (or later period of chronic stage), according to the clustering of the clinical manifestation and laboratory examination indices during the hospitalization, which has been previously reported⁹.

Written informed consent for study participation was obtained from all subjects or their guardians. The use of peripheral blood samples for further studies beyond routine laboratory examination was approved by the Ethics Committee of SPTCOD (No. LL-201802). This study abides by the Helsinki Declaration on ethical principles for medical research involving human subjects.

Blood samples collection and preparation. Peripheral blood samples of the cases were collected after their routine laboratory examination during the hospitalization, and stored at -80°C for later use. One sample in each of the 3 stages of each case was selected for DNA extraction and the deep sequencing. A few samples were missing in the stages, thus there were 23, 22 and 21 samples selected in the acute, chronic and recovery stages, respectively. The average time period for samples collection was 26.35 ± 43.94 days between the acute and chronic stages, and 81.78 ± 67.77 days between the chronic and recovery stages. Blood samples of the 24 controls were all collected during their health check.

Genomic DNA was extracted from about 2.0 mL of each blood sample using a QIAamp DNA Blood Midi Kit (QIAGEN, Shanghai, China) following the manufacturer's instructions. DNA concentration was quantified using a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) and DNA integrity was tested by the agarose gel electrophoresis.

Multiplex-PCR amplification of the TCR β CDR3 region. According to the IMGT unique numbering for V-DOMAIN²⁰, the TCR β CDR3 region starts with the second conserved cysteine (position 104) encoded by the Vb gene segment and ends with the conserved phenylalanine or tryptophan (position 118) encoded by the Jb gene segment. A multiplex-PCR system was designed to amplify rearranged TCR β CDR3 regions from genomic DNA using a set of forward and reverse primers to generate the template library for Genome Analyzer. The primers and PCR conditions had been reported before^{21,22}. After the amplification and agarose gel electrophoresis selection, the PCR products were purified and quantitated by determining the average molecule length using the Agilent 2100 bioanalyzer instrument (Agilent DNA 1000 Reagents) and by real-time quantitative PCR (QPCR) (TaqMan Probe).

High-throughput sequencing and data analysis. As previously reported⁸, the libraries obtained from the last step were amplified with cBot to generate the cluster on the flow cell, and the amplified flow cell was pair-end sequenced using a HiSeq4000 instrument, with a read length of 100 as the most frequently used sequencing strategy. The acquisition of raw data of sequences, basic data processing and sequence alignment were as previously described by Bolotin et al.²³. Thereafter, the V, D, J gene segments and CDR3 regions were identified, and nucleotides were translated into amino acids⁸.

The sequencing data were sorted and calculated into the following indices: (1) repertoire diversity indices including resampled diversity, extrapolate diversity, chaoE, chao1, Shannon–Wiener index, and inverse Simpson index, (2) TRBV/TRBD/TRBJ usage and combination, (3) CDR3-nt and CDR3-aa sequence frequency. Resampled diversity and extrapolate diversity are clonal diversity under the same amount of effective data. Resampled diversity uses the lowest amount of data, while extrapolate diversity uses the maximum amount of data²⁴. ChaoE is the clone diversity obtained after normalizing the amount of clone data, and chao1, also known as chao estimate, is the predicted clone diversity under the assumption that the amount of data is completely saturated²⁴. Shannon–Wiener index and inverse Simpson index are both commonly used measures of diversity, which have been introduced previously⁸.

Statistical analysis. The statistical analyses and graphing were conducted with R v. 3.6.3²⁵. Numerical variables were presented as mean with standard deviation (*SD*) if they were normally distributed, otherwise as median (*M*) with median absolute deviation (*MAD*). Repertoire diversity indices, frequencies of TRBV, TRBD, TRBJ, and TRBV-TRBJ combination, and frequencies of CDR3-nt and CDR3-aa sequences were compared between different stages of the cases or between the cases in different stages and the controls using the Mann–Whitney U test with Holm's adjustment. The trend of CDR3-nt or CDR3-aa sequence frequencies with decreasing OMDT severity from acute to recovery stage was analyzed by the linear regression model. Categorical variables such as the percentage of male and of Han were compared between the cases and controls using χ^2 -test or Fisher's exact test accordingly. All tests were two-sided and $P < 0.05$ was considered statistically significant unless it was adjusted in multiple comparisons.

Ethics approval. All procedures performed in the study involving human participants were in accordance with the ethical standards of the Ethics Committee of Shenzhen Prevention and Treatment Center for Occupational Diseases and with the 1975 Declaration of Helsinki and its later amendments or comparable ethical standards. The study was approved by the Ethics Committee of Shenzhen Prevention and Treatment Center for Occupational Diseases (No. LL-201802).

Consent to participate. Informed consent was obtained from all individual participants included in the study or their legal guardians.

Data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Received: 11 December 2020; Accepted: 27 April 2021

Published online: 11 May 2021

References

- Xu, X. *et al.* Severe hypersensitivity dermatitis and liver dysfunction induced by occupational exposure to trichloroethylene. *Ind. Health* **47**, 107–112 (2009).
- Kamijima, M., Hisanaga, N., Wang, H. & Nakajima, T. Occupational trichloroethylene exposure as a cause of idiosyncratic generalized skin disorders and accompanying hepatitis similar to drug hypersensitivities. *Int. Arch. Occup. Environ. Health* **80**, 357–370 (2007).
- Hua, M. *et al.* Analysis of causes of death and intervention measures of occupational medicamentose-like dermatitis due to trichloroethylene. *Chin. J. Ind. Hyg. Occup. Dis.* **28**, 54–55 (2010) (in Chinese).
- Zhang, J. X., Li, N., Wang, H., Shen, T. & Zhu, Q. X. The immune response in trichloroethylene hypersensitivity syndrome: A review. *Toxicol. Ind. Health* **33**, 876–883 (2017).
- Chen, J. Z., Liang, S. H., Wu, Q. F., Li, L. Y. & Li, Z. G. Study on the clinical skin pathology of medicamentose-like dermatitis induced by trichloroethylene. *Chin. Occup. Med.* **35**, 128–130 (2008) (in Chinese).
- White, K. D., Chung, W. H., Hung, S. I., Mallal, S. & Phillips, E. J. Evolving models of the immunopathogenesis of T cell-mediated drug allergy: The role of host, pathogens, and drug response. *J. Allergy Clin. Immunol.* **136**, 219–234 (2015) (quiz 235).
- Sui, W. *et al.* Composition and variation analysis of the TCR beta-chain CDR3 repertoire in systemic lupus erythematosus using high-throughput sequencing. *Mol. Immunol.* **67**, 455–464 (2015).
- Lin, D. *et al.* A pilot study to assess peripheral blood TCR β -chain CDR3 repertoire in occupational medicamentosa-like dermatitis due to trichloroethylene using high-throughput sequencing. *Environ. Toxicol. Pharmacol.* **71**, 103211 (2019).
- Lin, D., Xie, J. W., Huang, L., Zhang, Y. F. & Huang, X. Q. Staging of occupational medicamentosa-like dermatitis due to trichloroethylene. *Chin. J. Ind. Hyg. Occup. Dis.* **38**, 450–455 (2020) (in Chinese).
- Xiong, H. *et al.* Comprehensive assessment of T cell receptor β repertoire in Stevens-Johnson syndrome/toxic epidermal necrolysis patients using high-throughput sequencing. *Mol. Immunol.* **106**, 170–177 (2019).
- Thapa, D. R. *et al.* Longitudinal analysis of peripheral blood T cell receptor diversity in patients with systemic lupus erythematosus by next-generation sequencing. *Arthritis Res. Ther.* **17**, 132 (2015).
- Liu, S. *et al.* Comprehensive assessment of peripheral blood TCRbeta repertoire in infectious mononucleosis and chronic active EBV infection patients. *Ann. Hematol.* **96**, 665–680 (2017).
- Bacchetta, R., Gambineri, E. & Roncarolo, M. G. Role of regulatory T cells and FOXP3 in human diseases. *J. Allergy Clin. Immunol.* **120**, 227–235 (2007) (quiz 236–237).
- Akdis, M., Blaser, K. & Akdis, C. A. T regulatory cells in allergy: Novel concepts in the pathogenesis, prevention, and treatment of allergic diseases. *J. Allergy Clin. Immunol.* **116**, 961–968 (2005) (quiz 969).
- Li, S. *et al.* Trichloroethylene alters Th1/Th2/Th17/Treg paradigm in mice: A novel mechanism for chemically induced autoimmunity. *Int. J. Toxicol.* **37**, 155–163 (2018).
- Deng, L., Lu, Y., Deng, L., Wei, S. & Liang, S. Role of regulatory T cells and regulatory B cells in pathogenesis of occupational medicamentosa-like dermatitis due to trichloroethylene. *Occup. Health Emerg. Res.* **38**, 7–10 (2020) (in Chinese).
- Li, H. *et al.* HLA-B*1301 as a biomarker for genetic susceptibility to hypersensitivity dermatitis induced by trichloroethylene among workers in China. *Environ. Health Perspect.* **115**, 1553–1556 (2007).
- Wang, D. *et al.* Joint association of carrying HLA-B*13:01 gene and human herpesvirus-6 with occupational trichloroethylene hypersensitivity syndrome. *Int. Arch. Occup. Environ. Health* **92**, 395–401 (2019).
- Ministry of Health of the People's Republic of China. *Diagnostic Criteria of Occupational Medicamentose-Like Dermatitis Due to Trichloroethylene*. (National Standards for Occupational Health, 2007).
- Lefranc, M. P. *et al.* IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* **27**, 55–77 (2003).
- Zhang, W. *et al.* IMonitor: A robust pipeline for TCR and BCR repertoire analysis. *Genetics* **201**, 459–472 (2015).
- Liu, X. *et al.* Systematic comparative evaluation of methods for investigating the TCR β repertoire. *PLoS ONE* **11**, e0152464 (2016).
- Bolotin, D. A. *et al.* MiXCR: Software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).
- Shugay, M. *et al.* VDJtools: Unifying post-analysis of T cell receptor repertoires. *PLoS Comput. Biol.* **11**, e1004503 (2015).
- R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria) <https://www.R-project.org/> (2020).

Acknowledgements

This work was supported by [Science and Technology Planning Project of Shenzhen Municipality] (Grant numbers [JCYJ20180306170306255], [JCYJ20160429090813380], and [JCYJ20190808174815278]) and [National Natural Science Foundation of China] (Grant number [81872666]). The authors would like to thank colleagues from Medical Laboratory, Shenzhen Prevention and Treatment Center for Occupational Diseases, who had been working so hard to help accomplish this work, and to acknowledge Beijing Genomics Institute (BGI) company for the high throughput sequencing service.

Author contributions

D.L.: Conceptualization, Formal analysis, Writing—Original Draft, Supervision, Funding acquisition. D.W.: Methodology, Software, Investigation, Funding acquisition. P.L.: Investigation, Resources. X.Y.: Methodology, Investigation. W.L.: Investigation, Formal analysis. L.H.: Investigation, Visualization. Z.Z.: Investigation, Resources. Y.Z.: Resources, Writing—Review & Editing. W.Z.: Data Curation. N.Z.: Project administration. M.Z.: Writing—Review & Editing. X.H.: Project administration, Funding acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-89431-w>.

Correspondence and requests for materials should be addressed to D.L. or N.Z.

Reprints and permissions information is available at www.nature.com/reprints.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021