

# Integrated analysis of B-cell and T-cell receptors by high-throughput sequencing reveals conserved repertoires in IgA nephropathy

MINGLIN OU<sup>1</sup>, FENGPING ZHENG<sup>1</sup>, XINZHOU ZHANG<sup>1</sup>, SONG LIU<sup>1</sup>,  
DONGE TANG<sup>1</sup>, PENG ZHU<sup>2</sup>, JINGJUN QIU<sup>2</sup> and YONG DAI<sup>1</sup>

<sup>1</sup>Clinical Medical Research Center of Second Clinical Medical College, Jinan University, Shenzhen People's Hospital, Shenzhen, Guangdong 518020; <sup>2</sup>Lab Center of Shenzhen Pingshan People's Hospital, Shenzhen, Guangdong 518118, P.R. China

Received October 9, 2017; Accepted February 1, 2018

DOI: 10.3892/mmr.2018.8793

**Abstract.** Immunoglobulin A nephropathy (IgAN) is a type of glomerular disorder associated with immune dysregulation, and understanding B-/T-cell receptors (BCRs/TCRs) may be valuable for the development of specific immunotherapeutic interventions. In the present study, B and T cells were isolated from IgAN patients and healthy controls, and the composition of the BCR/TCR complementarity-determining region (CDR)3 was analyzed by multiplex polymerase chain reaction, high-throughput sequencing and bioinformatics. The present results revealed that the BCR/TCR CDR3 clones were expressed at very low frequencies, and the composition of clone types in patients with IgAN was skewed; the majority of clones were unique, and only 12 BCR and 228 TCR CDR3 clones were public ones, of which 16 were expressed at a significantly higher frequency in patients with IgAN ( $P < 0.001$ ). There were also certain conserved amino acid residues between unique clones or groups, and the residues GMDV, EQY and EQF were recurring only in the IgAN group. In addition, some VDJ gene recombinations indicated great variation between groups, including 4 high-frequency VDJ gene recombinations in the IgAN patients ( $P < 0.001$ ). Immune repertoires provide novel information, and conserved BCR/TCR CDR3 clones and VDJ gene recombinations with great variation may be potential therapeutic targets for IgAN patients.

## Introduction

Immunoglobulin A nephropathy (IgAN) is a type of glomerular disorder, which is an important cause of kidney failure (1). Patients usually present macroscopic hematuria during a gastrointestinal or upper respiratory infection (2). However, the diagnosis of IgAN requires kidney biopsy, which may lead to various complications (3,4). Although IgAN is increasingly considered to be a systemic immune complex deposition disorder, the kidney is most commonly affected, and damage may frequently recur even subsequent to kidney transplantation (5). Deposited immune complexes contain a number of classes of antibody, although IgA1 is the most frequent and is often the dominant or co-dominant class. Certain autoreactive IgA1 antibodies that are specific for the hinge region of IgA1 have additionally been observed in serum from patients with IgAN (6). Although the trigger of autoantibodies and formation of immune complexes has remained unclear until now, certain studies have indicated that susceptibility alleles in the antigen region of leukocytes and adaptive immunity may serve key roles in the pathogenesis of IgAN (7,8).

Human adaptive immunity is mediated by leukocytes, including B and T cells, which have been demonstrated to be selectively activated by interactions between antigens and their receptors expressed on the surface of lymphocytes (9). Therefore, understanding the spectrum of antigen receptors present on B and T cells, known as the immune repertoire, may aid in identifying biomarkers and the development of therapeutics for IgAN (10). In the antigen-binding site of B-cell receptors (BCRs)/T-cell receptors (TCRs), specific recognition of countless diverse peptide-major histocompatibility complexes are determined by three complementarity-determining regions (CDRs): CDR1, CDR2 and CDR3. Studies have indicated that CDR1 and CDR2 are relatively conserved and are encoded by different germline DNA sequences; however, CDR3 is encoded by genomic loci that undergo somatic recombination of VDJ gene segments and are assembled into highly diverse junctional residues (11,12). Therefore, the DNA sequences encoding CDR3 in the BCR/TCR could provide structural

*Correspondence to:* Dr Yong Dai, Clinical Medical Research Center of Second Clinical Medical College, Jinan University, Shenzhen People's Hospital, 1017 Dongmen North Road, Shenzhen, Guangdong 518020, P.R. China  
E-mail: daiyong22@aliyun.com

Dr Jingjun Qiu, Lab Center of Shenzhen Pingshan People's Hospital, 19 Renmin Road, Shenzhen, Guangdong 518118, P.R. China  
E-mail: 851874308@qq.com

**Key words:** immunoglobulin A nephropathy, immunology, B cell receptor, T cell receptor, high throughput sequencing

information about primary receptor antigenic binding sites. In the literature, few studies have focused on the diversity of the TCR at the nucleotide sequence level in IgAN. In a study by Wu *et al* (13), 14 patients with IgAN were divided into a stable and a progressive group, and TCR CDR3 sequences were cloned and analyzed by the dideoxy chain determination method. The results revealed that certain conserved amino acids in the TCR CDR3 contributed to the recognition of a particular antigen or set of antigens. However, as the human immune repertoire includes B and T cells, TCR CDR3 is only part of the information. In another study, TCR diversity in IgAN was assessed using traditional techniques, included targeted cloning and Sanger sequencing, which are low-throughput methods and allowed for only a descriptive assessment of 20 TCR V $\beta$  families (14). In the present study, a comprehensive analysis of BCR/TCR CDR3 diversity was attempted, in order to provide novel information about IgAN by multiplex PCR, high-throughput sequencing and bioinformatics.

## Materials and methods

**Study subjects.** Written informed consent was obtained from each participant. The study was approved by the Ethics Committee of Shenzhen People's Hospital and abided by the ethical principles of the Helsinki Declaration of 1975, as revised in 2000. Peripheral blood samples with fresh EDTA-K2 anticoagulant were obtained from eight South Chinese patients with IgAN and six South Chinese healthy controls in Shenzhen People's Hospital, China. Samples were collected at the time of the initial diagnosis of IgAN and prior to drug treatments. The renal biopsies were performed under direct ultrasound guidance with automated biopsy needles; the sections of biopsy tissues were stained by periodic acid-Schiff stain at 20°C for 10 min, and the mesangial IgA deposition was confirmed by immunofluorescence using fluorescein isothiocyanate-labelled IgA antibodies (Dako; Agilent Technologies, Inc., Santa Clara, CA, USA) (15). The diagnosis of IgAN was confirmed by clinical and biopsy findings (Fig. 1), which were according to the Lee glomerular grading system (16). Of the 8 patients with IgAN, 5 were female and 3 were male, and the mean age was 34.6 $\pm$ 6.63 years (range, 26-45 years; Table I). The six healthy controls were healthy volunteers, 2 men and 4 women, and the mean age was 34.0 $\pm$ 6.3 years (range, 27-45 years), and they were matched to the patients in terms of sex, age and ethnicity.

**Cell sub-population isolation and DNA extraction.** Peripheral blood mononuclear cells (PBMCs) were isolated from each peripheral blood sample using LymphoPrep (Axis Shield Diagnostics Ltd., Dundee, UK), according to the manufacturer's protocol. For the isolation of B cells, the non-B cells of each PBMC sample were depleted using MicroBeads (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany) using a MACS Separator (Miltenyi Biotec GmbH). The T cells were isolated from each PBMC sample using CD3 MicroBeads (Miltenyi Biotec GmbH), according to the manufacturer's protocol. DNA was extracted from the isolated B and T cells using a QIAamp DNA Mini kit (Qiagen GmbH, Hilden, Germany), according to the manufacturer's protocol.

**Preparation of BCR/TCR libraries by multiplex-polymerase chain reaction (PCR).** The human DNA sequences of BCR/TCR CDR3, which are the most diverse DNA sequences, were obtained from IMGT ([www.imgt.org](http://www.imgt.org)). The relatively conserved regions in the BCR/TCR DNA sequences were designed as the primer combining sites, and amplification by multiplex-PCR was performed according to previous studies, with minor modifications (17,18). Briefly, equal amounts of DNA taken from each sample were used to construct a library, with primers covering the majority of the V, D and J gene families, including 16 primers to amplify BCR CDR3 (Table II), and 44 primers to amplify TCR CDR3 (Table III). The multiplex-PCR was performed using a QIAGEN Multiplex PCR kit. For the amplification of BCR CDR3, the PCR mixture was amplified under the following conditions: 95°C (15 min) followed by 25 cycles of 94°C (15 sec), 60°C (3 min) and 72°C (10 min). For the amplification of TCR CDR3, the PCR mixture was amplified under the following conditions: 95°C (15 min), followed by 30 cycles of 94°C (15 sec), 59°C (30 sec) and 72°C (11 min). The PCR products were purified using a QIA quick PCR Purification kit (Qiagen GmbH) and AMPure XP beads (Beckman Coulter, Inc., Brea, CA, USA). The sequencing index was added to each PCR product sample, and the quantity of each library was analyzed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.) quantitative PCR using the KAPA Library Quantification kit, following manufacturer's protocol (Kapa Biosystems, Inc., Wilmington, MA, USA), as previously described (19).

**Deep sequencing and data analysis.** The libraries were sequenced using Hiseq 2000 following the manufacturer's protocol (Illumina, Inc., San Diego, CA, USA). In order to control the sequencing quality, internal controls consisting of defined DNA fragments were used in the sequence detection. The defined DNA fragments ranged between 170 and 800 bp, and enabled the estimation of error rates (E). The quality scores of sequencing (sQ), which were one important criterion for read filtering, ranged from 0 to 40, and the relationship between sQ and E was calculated by the following formula:

$$\begin{aligned} \text{sQ} &= -10 \times (\log E / 1 - E) / \log 10 \\ E &= Y / 1 + Y \\ Y &= \text{sQ} / (e^{-10 \times \log 10}) \end{aligned}$$

In order to obtain clean immune sequences, the adapters and the reads with sQ < 15 or N bases > 5% in the raw data generated by Hiseq 2,000, were filtered, as previously described (17). The clean data were expected to have sequence lengths of > 60 nucleotides; and the pair-end read pairs, which had identity-aligning tail regions or sequences with 90% matched bases ( $\geq 10$  bp overlapping), were merged into a single-contig sequence using COPE v1.1.3 and FqMerger (both BGI, Shenzhen, China). Following low-quality read filtering, the clean data were used in further alignments with the online tool IMGT/HighV-QUEST (20). The web portal of IMGT® and the international ImmunoGeneTics information system® ([www.imgt.org](http://www.imgt.org)) was used to analyze the nucleotide sequence rearrangement of BCRs/TCRs obtained from high-throughput sequencing. The output data included each distinct DNA

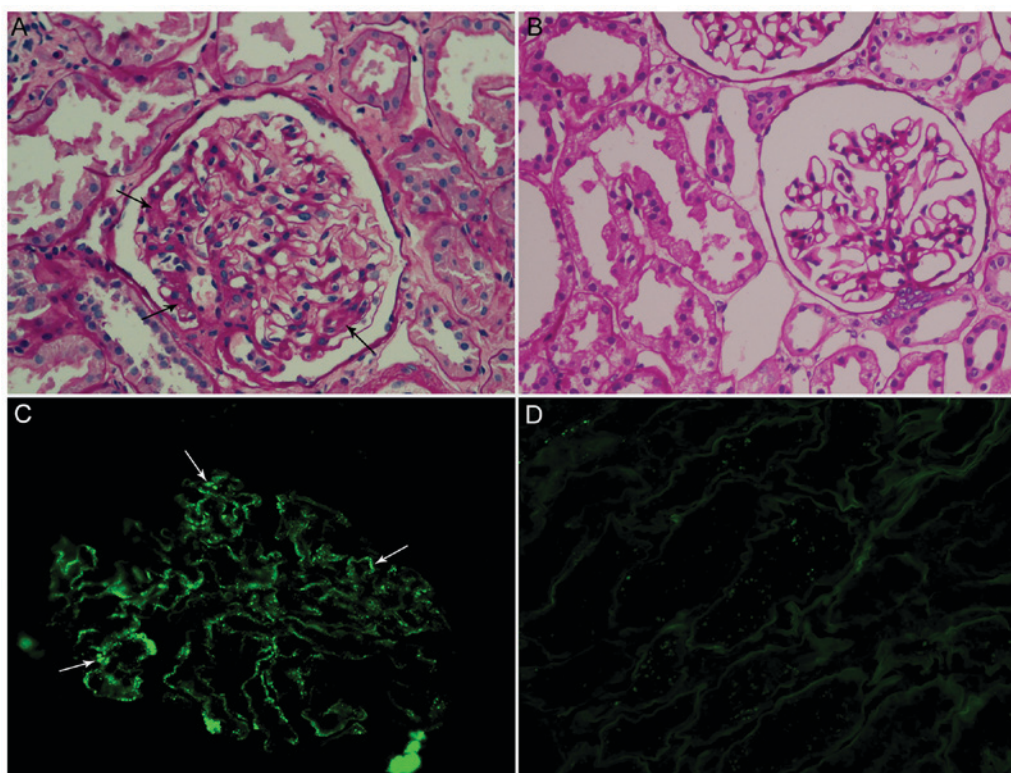


Figure 1. Variation in the sections of renal biopsy (magnification, x400). (A) The IgAN section with mesangial proliferative change stained with PAS (black arrow). (B) The normal adjacent section stained with PAS. (C) The IgAN section with IgA deposition stained by immunofluorescence (white arrow). (D) The normal adjacent section stained by immunofluorescence. IgAN, immunoglobulin A nephropathy; PAS, periodic acid-Schiff.

sequence of BCR/TCR, IMGT clonotypes (AA), VDJ gene assignment and distribution of CDR3 length. Statistical analysis was performed using SPSS Statistics 17.0 (SPSS, Inc., Chicago, IL, USA), and the comparisons of frequency estimates between groups were performed by  $\chi^2$  testes. The residue alignment was performed by MEGA7.0 software (www.megasoftware.net) (21).

## Results

**Summary of BCR/TCR sequencing.** For the BCR libraries, 466,667 and 408,506 reads from the IgAN and the control group, respectively, were obtained. Following filtering out of unknown sequences, 462,372 and 402,373 clean immune sequences were identified, of which 349,990 and 321,803 were CDR3 sequences. 34,078 unique CDR3 nucleotide sequences and 24,049 unique CDR3 amino acid clonotypes were detected in the IgAN group, and 38,191 unique CDR3 nucleotide sequences and 26,919 unique CDR3 amino acid clonotypes were detected in the control group. For the TCR libraries, 467,542 and 491,499 reads from the IgAN and the control group, respectively, were obtained. There were 464,053 and 487,974 clean immune sequences, of which 344,679 and 362,658 were CDR3 sequences; and 46,735 unique CDR3 nucleotide sequences and 41,573 unique CDR3 amino acid clonotypes were identified in the IgAN group, and 27,272 unique CDR3 nucleotide sequences and 22,918 unique CDR3 amino acid clonotypes were identified in the control group. The numbers of unique CDR3 nucleotide sequences and unique CDR3 amino acid clonotypes indicated that 29.43 and

11.05% nucleotide sequences would encode identical amino acid clonotypes in the BCR and TCR, respectively.

**Characteristics of the BCR/TCR CDR3 length distribution.** The length distribution of the BCR/TCR CDR3 loop is an important index for the determination of B/T cell repertoire diversity. In the present study, the average BCR CDR3 lengths did not differ significantly between the IgAN ( $17.01 \pm 4.27$ ) and the control group ( $17.25 \pm 4.15$ ;  $P > 0.05$ ). The average TCR CDR3 lengths did not differ significantly between the IgAN ( $12.56 \pm 2.56$ ) and the control group ( $12.42 \pm 2.20$ ;  $P > 0.05$ ). The length distribution of the BCR/TCR CDR3 loops in the IgAN group and the control group were observed to fit well to a Gaussian distribution curve ( $R^2 \approx 1.0$ ; Fig. 2). In Gaussian distribution curves, the  $R^2$  index ranges between 0 and 1, corresponding to worst and best fit.

**BCR/TCR CDR3 with highly expanded clones.** The expression level of each unique BCR/TCR CDR3 amino acid sequence was evaluated by the unique BCR/TCR CDR3 DNA sequence frequency following alignment. Highly expanded clones (HECs) were defined as those with a frequency  $> 0.01\%$  (22). For the BCR CDR3, a total of 3,133 clones were identified as HECs, with a median HEC ratio of 0.0192664% in the IgAN group; whereas, in the control group, 2,338 clones were HECs, with a median HEC ratio of 0.0277151%. A  $\chi^2$  test indicated that there were significantly more BCR CDR3 HECs in the IgAN compared with the control group ( $P < 0.05$ ; data not shown). The BCR CDR3 HECs in each group were highly diverse, and all of the HECs in the IgAN group were unique clones. For the TCR



Table I. Clinical data of patients with immunoglobulin A nephropathy.

Case no.	Age, years	Lee's grading	SCR, $\mu$ mol/l	BUN, mmol/l	Proteinuria, g/24 h	URBC, cells/ $\mu$ l
1	27	III	62	4.10	0.795	150
2	44	III	122	7.48	1.148	250
3	36	III	70	3.50	0.704	150
4	36	III	129	4.35	1.036	250
5	26	IV	127	7.67	2.739	250
6	34	IV	102	4.30	0.840	150
7	31	III	69	3.99	0.730	250
8	43	IV	116	5.19	1.872	150

BUN, blood urea nitrogen; SCR, serum creatinine; URBC, urinary red blood cell.

Table II. Primers for the amplification of BCR CDR3.

Primer	Forward/reverse	Sequence (5'→3')
IGHV1-18	Forward	CAGACGTGTGCTCTTCCGATCTAGAGAGTCACCATGACCACAGAC
IGHV1-2/1-46	Forward	CAGACGTGTGCTCTTCCGATCTAGAGAGTCACCAKKACCAGGGAC
IGHV1-24	Forward	CAGACGTGTGCTCTTCCGATCTAGAGAGTCACCATGACCGAGGAC
IGHV1-3/1-45	Forward	CAGACGTGTGCTCTTCCGATCTAGAGAGTCACCATTACYAGGGAC
IGHV1-69/1-f	Forward	CAGACGTGTGCTCTTCCGATCTAGAGAGTCACGATWACCRCGGAC
IGHV1-8	Forward	CAGACGTGTGCTCTTCCGATCTAGAGAGTCACCATGACCAGGAAC
IGH2-70/26/5	Forward	CAGACGTGTGCTCTTCCGATCTAGACCAGGCTCACCATYWCCAAGG
IGHV3	Forward	CAGACGTGTGCTCTTCCGATCTAGGGCCGATTACCATCTCMAG
IGH4	Forward	CAGACGTGTGCTCTTCCGATCTAGCGAGTCACCATRTCMGTAGAC
IGHV5-51	Forward	CAGACGTGTGCTCTTCCGATCTAGCAGCCGACAAGTCCATCAGC
IGHV6-1	Forward	CAGACGTGTGCTCTTCCGATCTAGAGTCGAATAACCATCAACCCAG
IGHV7-NEW	Forward	CAGACGTGTGCTCTTCCGATCTAGGACGGTTTGTCTTCTCCTTG
HIGHJ-Rev1	Reverse	CTACACGACGCTCTTCCGATCTCTGAGGAGACRGTGACCAGGGTG
HIGHJ-Rev2	Reverse	CTACACGACGCTCTTCCGATCTCTGAAGAGACGGTGACCATTGTC
HIGHJ-Rev3	Reverse	CTACACGACGCTCTTCCGATCTCTGAGGAGACGGTGACCAGGGT
HIGHJ-Rev4	Reverse	CTACACGACGCTCTTCCGATCTTGAGGAGACGGTGACCGTGGTC

CDR3, a total of 1,804 clones were identified as HECs, with a median HEC ratio of 0.0133458% in the IgAN group; and in the control group, 2277 clones were identified as HECs, with a median HEC ratio of 0.0228866%. The HECs of the TCR CDR3 in each group were highly diverse, with 98.7804878% of TCR HECs in the IgAN group and 98.3311375% of TCR HECs in the control group observed to be unique clones. Furthermore, when the frequency distribution of the BCR/TCR repertoire was plotted, a great variation between the two groups of samples was identified, and the frequency distribution was markedly more skewed in the IgAN group (Fig. 3).

#### Conserved BCR/TCR CDR3 in the IgAN and control groups.

The public clones in the IgAN and control groups were considered to be an important aspect of the conserved clones; a total of 12 BCR CDR3 clones and 228 TCR CDR3 clones were identified as public ones, of which 22 in the IgAN group were HECs. The  $\chi^2$  test indicated that a total of 16 public clones, including one clone of BCR CDR3 and 15 clones of TCR CDR3,

were expressed with a significantly higher frequency in IgAN patients compared with the control group ( $P < 0.001$ ; Fig. 4). In order to identify whether there were any conserved BCR/TCR CDR3 residues among the unique clones, the top 30 highly expressed BCR/TCR CDR3 clones were aligned, and it was demonstrated that the amino acid residues of BCR/TCR CDR3 may not be encoded randomly; there were several conserved BCR/TCR CDR3 residues, including GMDV, EQY and EQF, that were recurrent  $>3$  times only in the IgAN group (Fig. 5).

**VDJ gene recombination of the BCR/TCR CDR3.** Diversity of the BCR/TCR CDR3 is generated by the process of VDJ genetic recombination, in which certain genes in the V, D and J gene family are selected. In order to avoid distortion by dominant clones, which may be expanded as a consequence of an immune response, the composition of unique CDR3 sequences was analyzed and their copies disregarded. The composition of the BCR was additionally analyzed by the gene sub-segment usage of IGHV, IGHD, IGHJ; and TCR

Table III. Primers for the amplification of TCR CDR3.

Primer	Forward/reverse	Sequence (5'→3')
TRBV2F-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGATTTCACTCTGAAGATCCGGTCCAC
TRBV9F-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCCTGACTTGCCTCTGAACTAAACCT
TRBV14F-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGGAGGGACGTATTCTACTCTGAAGG
TRBV15F-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGTTCTTGACATCCGCTCACCAGG
TRBV19F-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGTCCTTTCTCTCACTGTGACATCGG
TRBV3-1-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGAAACAGTTCCAAATCGMTTCTCAC
TRBV4-1/2/3-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCAAGTCGCTTCTCACCTGAATG
TRBV5-1-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGCCAGTTCTCTAACTCTCGCTCT
TRBV5-4/5/6/8-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGTCAGGTCGCCAGTTCCCTAAYTAT
TRBV6-1/2/3/5/8-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCAATGGCTACAATGTCTCYAGAT
TRBV6-4-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGTGATGGTTATAGTGTCTCCAGAG
TRBV6-9-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCGATGGCTACAATGTATCCAGAT
TRBV6-6-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGAATGGCTACAACGTCTCCAGAT
TRBV7-2/4/6/7/8-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGGGATCCGTCTCCACTCTGAMGAT
TRBV7-3-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGGGATCCGTCTCTACTCTGAAGAT
TRBV7-9-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGGGATCTTTCTCCACCTTGGAGAT
TRBV10-1-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCCTCACTCTGGAGTCTGCTGCC
TRBV10-2/3-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCCTCACTCTGGAGTCMGCTACC
TRBV11-1/2/3-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGCAGAGAGGCTCAAAGGAGTAGACT
TRBV12-3/4-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGATCGATTCTCAGCTAAGATGCCT
TRBV12-5-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGATCGATTCTCAGCAGAGATGCCT
TRBV13-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGTCGATTCTCAGCTCAACAGTTC
TRBV18-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGTAGATGAGTCAGGAATGCCAAAG
TRBV20-1-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGAACCATGCAAGCCTGACCTT
TRBV24-1-F2-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCTCCCTGTCCCTAGAGTCTGCCAT
TRBV25-1F-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGCCCTCACATACCTCTCAGTACCTC
TRBV27/28-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGGGAGATGTTCTTGARGGGTACA
TRBV29-1-F4-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGAACTCTGACTGTGAGCAACATGAG
TRBV16-F2-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCTGTAGCCTTGAGATCCAGGCTACGA
TRBV30-F5-IND	Forward	CAGACGTGTGCTCTTCCGATCTAGCAGATCAGCTCTGAGGTGCCCA
TRBJ1.1-R2-P1	Reverse	CTACACGACGCTCTTCCGATCTCTTACCTACAACGTGTGAGTCTGGTG
TRBJ1.2R-P1	Reverse	CTACACGACGCTCTTCCGATCTCTTACCTACAACGGTTAACCTGGTC
TRBJ1.3R-P1	Reverse	CTACACGACGCTCTTCCGATCTCTTACCTACAACAGTGAGCCAACTT
TRBJ1.4R-P1	Reverse	CTACACGACGCTCTTCCGATCTCATACCCAAGACAGAGAGCTGGGTTC
TRBJ1.5R-P1	Reverse	CTACACGACGCTCTTCCGATCTCTTACCTAGGATGGAGAGTCGAGTC
TRBJ1.6R-P1	Reverse	CTACACGACGCTCTTCCGATCTCATACCTGTACAGTGAGCCTG
TRBJ2.1R-P1	Reverse	CTACACGACGCTCTTCCGATCTCCTTCTTACCTAGCACGGTGA
TRBJ2.2R-P1	Reverse	CTACACGACGCTCTTCCGATCTCTTACCCAGTACGGTCAGCCT
TRBJ2.3R-P1	Reverse	CTACACGACGCTCTTCCGATCTCCGCTTACCGAGCACTGTCAG
TRBJ2.4R-P1	Reverse	CTACACGACGCTCTTCCGATCTCCAGCTTACCCAGCACTGAGA
TRBJ2.5-R2-P1	Reverse	CTACACGACGCTCTTCCGATCTCGAGCACCAGGAGCCGCGT
TRBJ2.6R-P1	Reverse	CTACACGACGCTCTTCCGATCTCTCGCCAGCACGGTCAGCCT
TRBJ2.7-R2-P1	Reverse	CTACACGACGCTCTTCCGATCTCTTACCTGTGACCGTGAGCCTG
P1	Reverse	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCT TCCGATCT

by the gene sub-segment usage of TRBV, TRBD and TRBJ. The recombination of genes from the V, D and J gene family in the BCR/TCR CDR3 appeared to be conserved, and it was demonstrated that 73 genes in the IGHV subfamily,

30 genes in the IGHD subfamily, six genes in the IGHI subfamily, 53 genes in the TRBV subfamily, two genes in the TRBD subfamily and 12 genes in the TRBJ subfamily were commonly used in both groups (Fig. 6). However, the VDJ

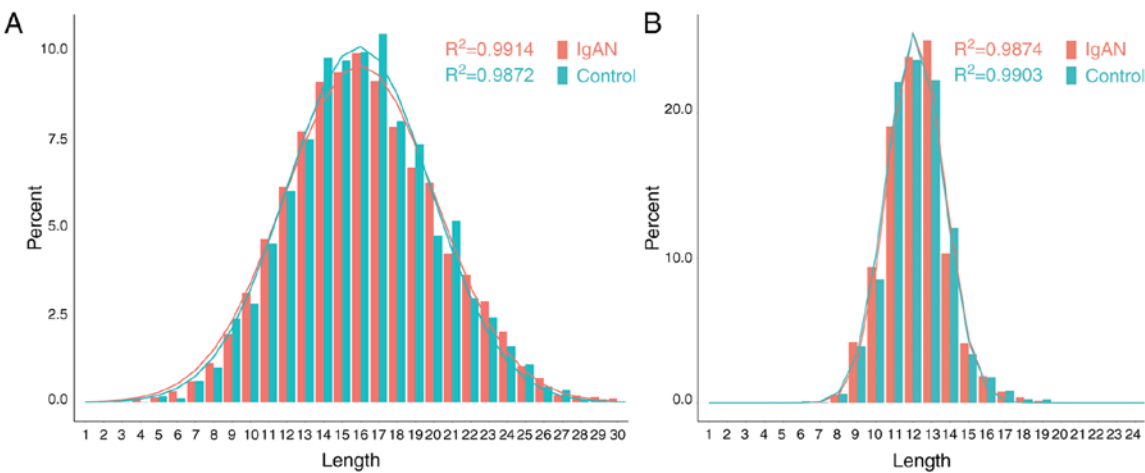


Figure 2. Length distribution of the BCR/TCR CDR3 with a Gaussian distribution curve. (A) BCR CDR3. (B) TCR CDR3. BCR, B-cell receptors; IgAN, immunoglobulin A nephropathy; TCRs, T-cell receptors; CDR3, complementarity-determining region 3.

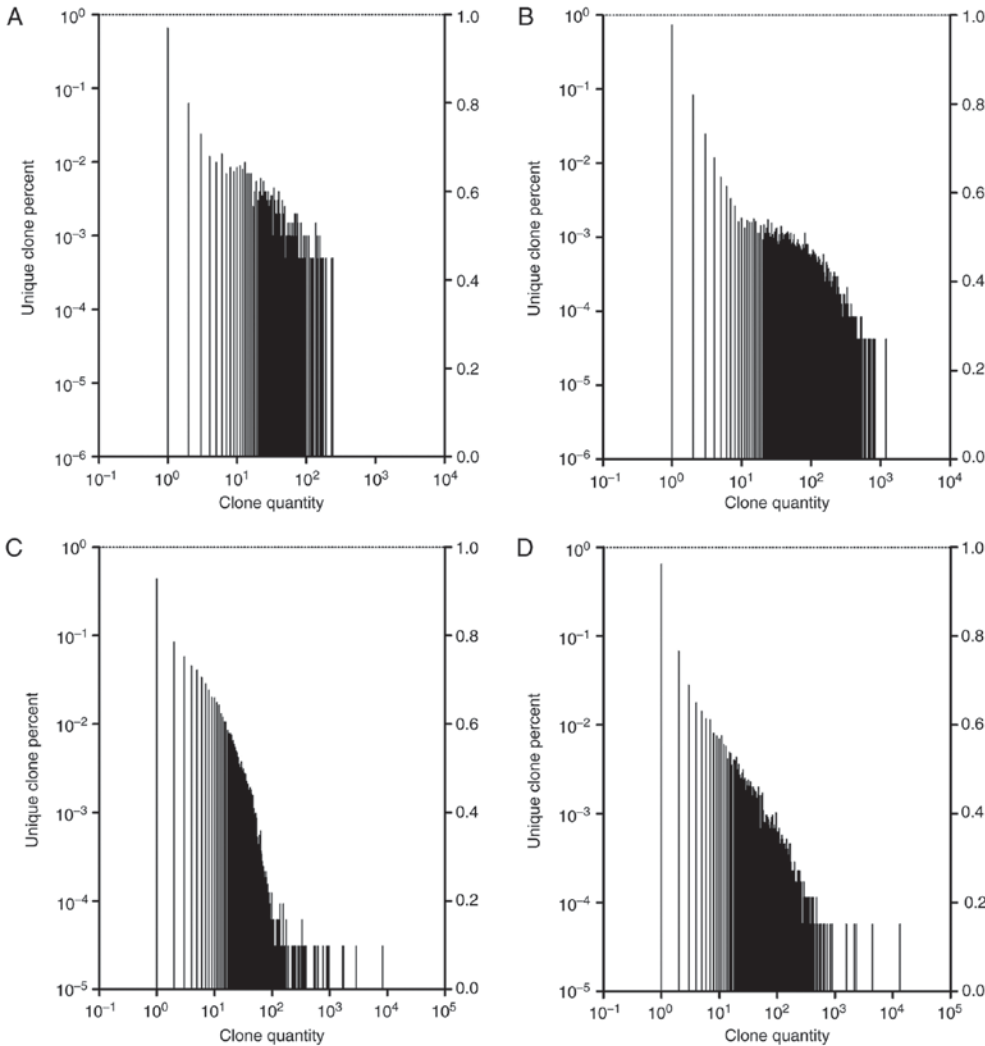


Figure 3. Frequency distribution of BCR/TCR unique clones. (A) BCR CDR3 of the IgAN group. (B) BCR CDR3 of the control group. (C) TCR CDR3 of the IgAN group. (D) TCR CDR3 of the control group. BCR, B-cell receptors; IgAN, immunoglobulin A nephropathy; TCRs, T-cell receptors; CDR3, complementarity-determining region 3.

recombination of BCR/TCR CDR3 exhibited great variation between the two groups of samples, with the  $\chi^2$  test indicating that four of the seven high frequency VDJ gene recombinations

(frequency >1.0%), including IGHV3-11/IGHD3-10/IGHJ6, IGHV3-30/IGHD3-10/IGHJ6, TRBV6-6/TRBJ2-7/TRBD2 and TRBV12-3/TRBJ1-1/TRBD1, were present with a

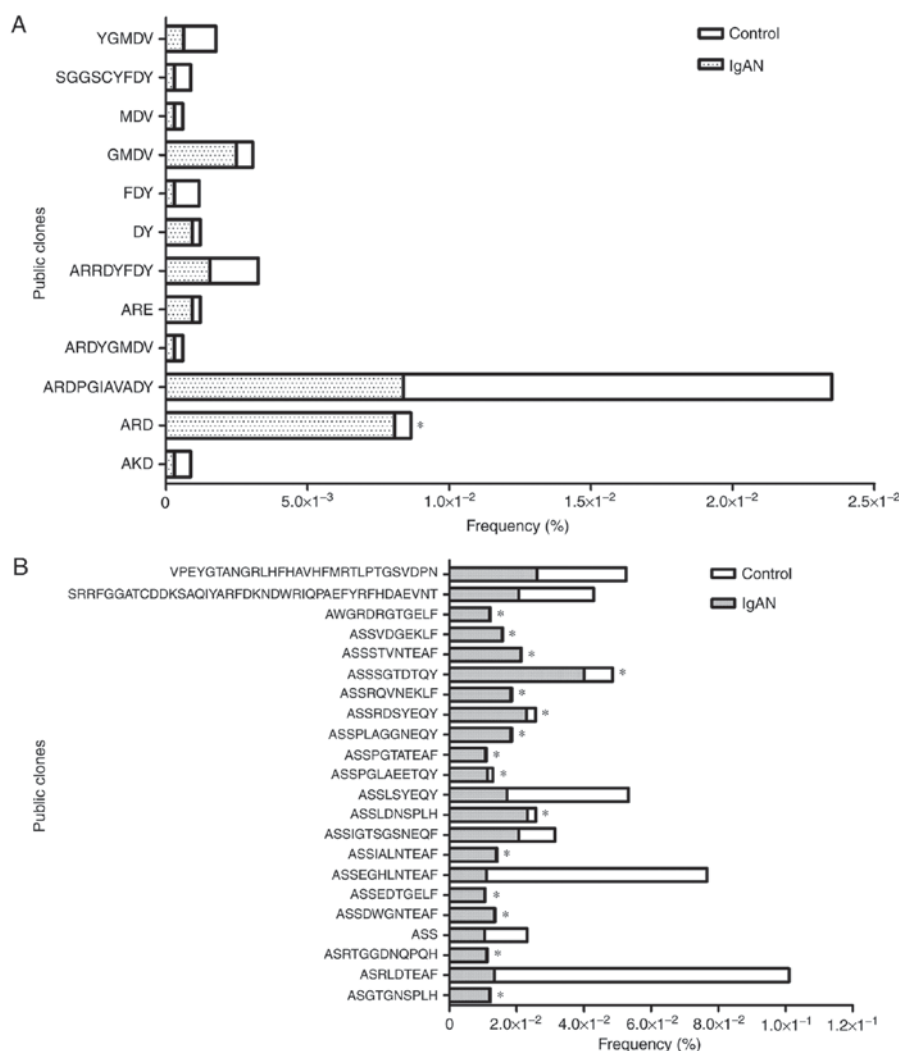


Figure 4. Public clones in the two groups of samples. (A) The public BCR CDR3 clones. (B) The public clones of the highly expanded TCR CDR3. \* $P < 0.001$  vs. control. BCR, B-cell receptors; IgAN, immunoglobulin A nephropathy; TCRs, T-cell receptors; CDR3, complementarity-determining region 3.

significantly higher frequency in the IgAN compared with the control group ( $P < 0.001$ ; data not shown).

## Discussion

B and T cells serve important roles in the human immune system, and understanding their receptor characteristics in human autoimmune diseases may help the development of specific immunotherapeutic interventions for diseases in the clinic (23). IgAN is widely defined as a human autoimmune disease, and certain studies have indicated that a cytokine-mediated switch of naive B cells to antibody-secreting cells or Toll-like receptor ligation pathways in T-cells may be important in the pathogenesis of IgAN (6,24). In the clinical setting, certain classical immunosuppressive drugs that exhibit pleiotropic effects, including cyclophosphamide and azathioprine, have been used for the treatment of IgAN (25,26). With the development of next-generation sequencing, it is possible to analyze the human immune repertoires in different physiological and pathological conditions with a high depth, which was not achievable previously (27,28).

In the present study, a large number of BCR/TCR CDR3 sequences at the DNA level were resolved, providing a large

quantity of novel information regarding immune repertoires in IgAN. With regard to the common index for the determination of B/T cell repertoire diversity, the length distribution of the BCR/TCR CDR3 loop appeared to fit well to a Gaussian distribution curve in the IgAN and control groups. This indicated that there may be no notable alterations in length distribution compared with certain typical autoimmune disorders (18). As another common index, the HECs of BCRs and TCRs exhibited certain differences; for example, the number of BCR HECs was increased in the IgAN compared with the control group, whereas the TCR HECs did not exhibit the same trend. This result indicated that there were more BCR CDR3 sequences with higher frequencies in the disease condition of IgAN. As IgAN is a disease characterized by mesangial IgA deposits, a higher frequency of BCR HECs may contribute to a higher frequency of class-switching of prime naive B cells to IgA<sup>+</sup> antibody-secreting cells following mucosal infection (6).

BCR/TCR CDR3 peptides are believed to be among the most complex and diverse amino acid sequences in humans (29). Studies have indicated that BCR/TCR CDR3 is encoded by somatic rearrangement between IGHV/IGHD/IGHJ (TRBV/TRBD/TRBJ) gene segments, and that receptor diversity may be augmented by non-template nucleotide trimming



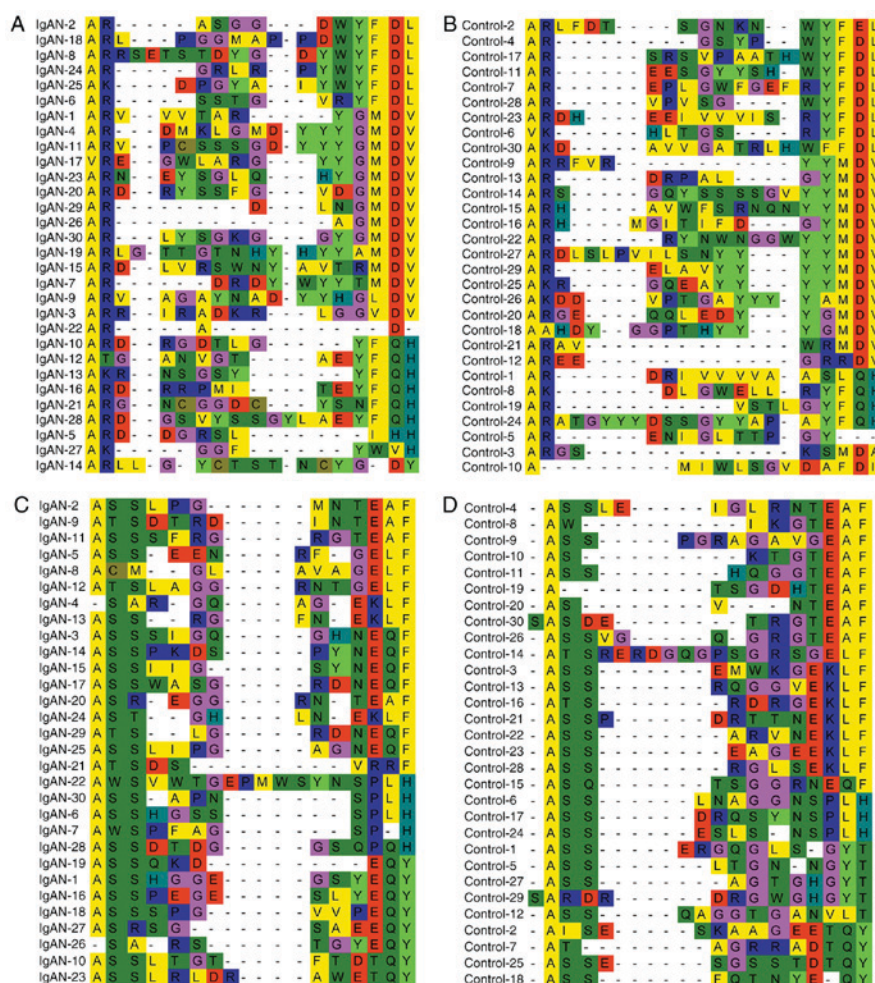


Figure 5. Top 30 most highly expressed BCR/TCR CDR3 clones with conserved residues, different colors indicated different kinds of residues. (A) BCR CDR3 of the IgAN group. (B) BCR CDR3 of the control group. (C) TCR CDR3 of the IgAN group. (D) TCR CDR3 of the control group. BCR, B-cell receptors; IgAN, immunoglobulin A nephropathy; TCRs, T-cell receptors; CDR3, complementarity-determining region 3.

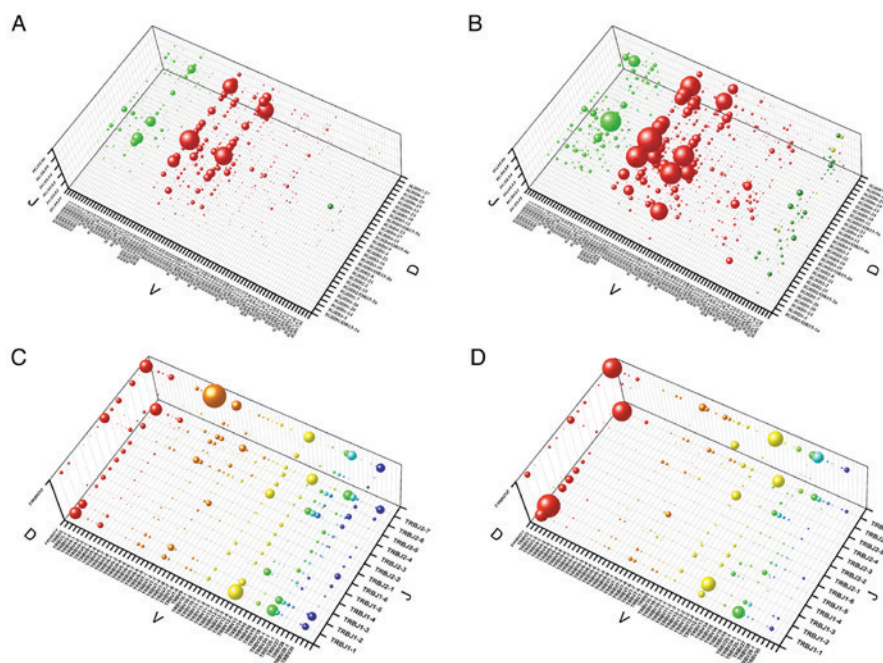


Figure 6. Relative VDJ gene recombination frequency of the BCR/TCR CDR3 in the two groups of samples. (A) BCR CDR3 of the IgAN group. (B) BCR CDR3 of the control group. (C) TCR CDR3 of the IgAN group. (D) TCR CDR3 of the control group. BCR, B-cell receptors; IgAN, immunoglobulin A nephropathy; TCRs, T-cell receptors; complementarity-determining region 3.



or addition at different gene segment junction sites (30,31). Therefore, the original hypervariable sequences of CDR3 may be encoded randomly and are able to recognize millions of antigens (32). Notably, public clones of BCR/TCR CDR3 sequences in the two groups of samples were identified, including 12 BCR and 228 TCR CDR3 clones, certain of which were HECs. Furthermore, certain of the BCR/TCR CDR3 residues appeared to be conserved in different clones. These findings indicated that the BCR/TCR CDR3 may not be produced randomly. Public (conserved) B/T cells are known to be key factors affecting human health (33). Notably, there were public sequences with significantly different frequencies in the IgAN group. It is possible that these public clones or residues with significantly higher frequency may be the initiators of IgAN, and may represent potential targets for immunotherapeutic intervention.

Regarding the use of IGHV/IGHD/IGHJ in encoding the BCR CDR3, there were certain common characteristics in the two groups of samples, including the most highly used gene segment, IGHV3-11/IGHD3-10/IGHJ6; this indicated that gene segment usage of BCR CDR3 may be conserved. Regarding the use of TRBV/TRBD/TRBJ in the TCR CDR3, there were certain common gene combinations in the two groups, while the most highly used gene segment in the IgAN group, TRBV6-6/TRBJ2-7/TRBD2, had a lower frequency in the control group. Although it may be difficult to fully understand the different degrees of BCR/TCR (V/D/J) repertoire usage, fully describing the gene segment usages of BCR/TCR CDR3 may be valuable for understanding the background of IgAN.

In conclusion, the present study attempted to profile the entire B- and T-cell repertoires in patients with IgAN by high-throughput sequencing for the first time, to the best of our knowledge. The global repertoires contain a wealth of conserved characteristics, especially the conserved BCR/TCR CDR3 sequences and the usage of gene segments, which may provide novel information for the treatment of IgAN.

### Acknowledgements

Not applicable.

### Funding

Funding was received from Science and Technology Plan of Shenzhen (grant nos. JCYJ20160422164313440 and JCYJ20150403093323885).

### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

### Author's contributions

YD and JQ carried out the study design, MO participated in experiment and data analysis, and drafted the manuscript. FZ and XZ participated in sample collection and clinical data analysis. SL, DT and PZ participated in experiment and analysis. All authors read and approved the final manuscript.

### Ethics approval and consent to participate

The study was approved by the Ethics Committee of Shenzhen People's Hospital. Written informed consent was obtained from all participants.

### Consent for publication

Written informed consent for the publication of any associated data and accompanying images was obtained from each participant.

### Competing interests

The authors declare that they have no competing interests.

### References

1. Berthelot L, Papista C, Maciel TT, Biarnes-Pelicot M, Tissandie E, Wang PH, Tamouza H, Jamin A, Bex-Coudrat J, Gestin A, *et al*: Transglutaminase is essential for IgA nephropathy development acting through IgA receptors. *J Exp Med* 209: 793-806, 2012.
2. Riispere Ž, Kuudeberg A, Seppet E, Sepp K, Ilmoja M, Luman M, Kõlvald K, Auerbach A and Ots-Rosenberg M: Significance of clinical and morphological prognostic risk factors in IgA nephropathy: Follow-up study of comparison patient groups with and without renoprotection. *BMC Nephrol* 18: 89, 2017.
3. Fiorentino M, Bolignano D, Tesar V, Pisano A, Van Biesen W, D'Arrigo G, Tripepi G and Gesualdo L; ERA-EDTA Immunonephrology Working Group: Renal biopsy in 2015-from epidemiology to evidence-based indications. *Am J Nephrol* 43: 1-19, 2016.
4. Le WB and Liu ZH: Prognostic indicators and treatment of IgA Nephropathy in China. *Pathogen Treat IgA Nephropat*: pp127-139, 2016.
5. Wyatt RJ and Julian BA: IgA Nephropathy. *N Engl J Med* 368: 2402-2414, 2013.
6. Boyd JK, Cheung CK, Molyneux K, Feehally J and Barratt J: An update on the pathogenesis and treatment of IgA nephropathy. *Kidney Int* 81: 833-843, 2012.
7. Gharavi AG, Kiryluk K, Choi M, Li Y, Hou P, Xie J, Sanna-Cherchi S, Men CJ, Julian BA, Wyatt RJ, *et al*: Genome-wide association study identifies susceptibility loci for IgA nephropathy. *Nat Genet* 43: 321-327, 2011.
8. Suzuki H, Kiryluk K, Novak J, Moldoveanu Z, Herr AB, Renfrow MB, Wyatt RJ, Scolari F, Mestecky J, Gharavi AG and Julian BA: The pathophysiology of IgA nephropathy. *J Am Soc Nephrol* 22: 1795-1803, 2011.
9. Mekori YA, Hershko AY, Frossi B, Mion F and Pucillo CE: Integrating innate and adaptive immune cells: Mast cells as crossroads between regulatory and effector B and T cells. *Eur J Pharmacol* 778: 84-89, 2016.
10. Han Y, Li H, Guan Y and Huang J: Immune repertoire: A potential biomarker and therapeutic for hepatocellular carcinoma. *Cancer Lett* 379: 206-212, 2016.
11. Hou X, Lu C, Chen S, Xie Q, Cui G, Chen J, Chen Z, Wu Z, Ding Y, Ye P, *et al*: High throughput sequencing of T cell antigen receptors reveals a conserved TCR repertoire. *Medicine (Baltimore)* 95: e2839, 2016.
12. Madi A, Shifrut E, Reich-Zeliger S, Gal H, Best K, Ndifon W, Chain B, Cohen IR and Friedman N: T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity. *Genome Res* 24: 1603-1612, 2014.
13. Wu H, Zhang GY, Clarkson AR and Knight JF: Conserved T-cell receptor beta chain CDR3 sequences in IgA nephropathy biopsies. *Kidney Int* 55: 109-119, 1999.
14. O'Connell AE, Volpi S, Dobbs K, Fiorini C, Tsitsikov E, de Boer H, Barlan IB, Despotovic JM, Espinosa-Rosales FJ, Hanson IC, *et al*: Next generation sequencing reveals skewing of the T and B cell receptor repertoires in patients with wiskott-Aldrich syndrome. *Front Immunol* 5: 340, 2014.
15. Suzuki K, Honda K, Tanabe K, Toma H, Nihei H and Yamaguchi Y: Incidence of latent mesangial IgA deposition in renal allograft donors in Japan. *Kidney Int* 63: 2286-2294, 2003.

16. Lee HS, Lee MS, Lee SM, Lee SY, Lee ES, Lee EY, Park SY, Han JS, Kim S and Lee JS: Histological grading of IgA nephropathy predicting renal outcome: Revisiting H. S. Lee's glomerular grading system. *Nephrol Dial Transplant* 20: 342-348, 2005.
17. Hou X, Chen J, Lu C, Wang L and Wu W: The conserved T cell receptor repertoire observed in patients with systemic lupus erythematosus. *Int J Clin Exp Med* 10: 2053-2065, 2017.
18. Liu S, Hou XL, Sui WG, Lu QJ, Hu YL and Dai Y: Direct measurement of B-cell receptor repertoire's composition and variation in systemic lupus erythematosus. *Genes Immun* 18: 22-27, 2017.
19. Guda K, Veigl ML, Varadan V, Nosrati A, Ravi L, Lutterbaugh J, Beard L, Willson JK, Sedwick WD, Wang ZJ, *et al*: Novel recurrently mutated genes in African American colon cancers. *Proc Natl Acad Sci USA* 112: 1149-1154, 2015.
20. Meyer EH, Hsu AR, Liliental J, Löhr A, Florek M, Zehnder JL, Strober S, Lavori P, Miklos DB, Johnson DS and Negrin RS: A distinct evolution of the T-cell repertoire categorizes treatment refractory gastrointestinal acute graft-versus-host disease. *Blood* 121: 4955-4962, 2013.
21. Kumar S, Stecher G and Tamura K: MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33: 1870-1874, 2016.
22. Xu H, Wang C, Liu X, Zhao Y and Zhang W: Characterization of the T cell receptor repertoires during pregnancy. *Biomed Res* 28: 1161-1166, 2017.
23. Borroto A, Reyes-Garau D, Jiménez MA, Carrasco E, Moreno B, Martínez-Pasamar S, Cortés JR, Perona A, Abia D, Blanco S, *et al*: First-in-class inhibitor of the T cell receptor for the treatment of autoimmune diseases. *Sci Transl Med* 8: 370ra184, 2016.
24. Renfrow MB and Novak J: What insights can proteomics give us into IgA nephropathy (Berger's disease)? *Expert Rev Proteomics* 14: 645-647, 2017.
25. Rasche FM, Keller F, Rasche WG, Schiekofer S, Boldt A, Sack U and Fahnert J: Why, when and how should immunosuppressive therapy considered in patients with immunoglobulin A nephropathy. *Clin Exp Immunol* 186: 115-133, 2016.
26. Kim MJ, Schaub S, Molyneux K, Koller MT, Stampf S and Barratt J: Effect of immunosuppressive drugs on the changes of serum galactose-deficient IgA1 in patients with IgA nephropathy. *PLoS One* 11: e0166830, 2016.
27. Callan CG Jr, Mora T and Walczak AM: Repertoire sequencing and the statistical ensemble approach to adaptive immunity. *Curr Opin Syst Biol* 1: 44-47, 2017.
28. Hou XL, Wang L, Ding YL, Xie Q and Diao HY: Current status and recent advances of next generation sequencing techniques in immunological repertoire. *Genes Immun* 17: 153-164, 2016.
29. Berry R and Call ME: Modular activating receptors in innate and adaptive immunity. *Biochemistry* 56: 1383-1402, 2017.
30. Krell PF, Reuther S, Fischer U, Keller T, Weber S, Gombert M, Schuster FR, Asang C, Stepensky P, Strahm B, *et al*: Next-generation-sequencing-spectratyping reveals public T-cell receptor repertoires in pediatric very severe aplastic anemia and identifies a  $\beta$  chain CDR3 sequence associated with hepatitis-induced pathogenesis. *Haematologica* 98: 1388-1396, 2013.
31. Aouinti S, Malouche D, Giudicelli V, Kossida S and Lefranc MP: IMGT/HighV-QUEST statistical significance of IMGT clonotype (AA) diversity per gene for standardized comparisons of next generation sequencing immunoprofiles of immunoglobulins and T cell receptors. *PLoS One* 10: e0142353, 2015.
32. Saada R, Weinberger M, Shahaf G and Mehr R: Models for antigen receptor gene rearrangement: CDR3 length. *Immunol Cell Biol* 85: 323-332, 2007.
33. Kitaura K, Shini T, Matsutani T and Suzuki R: A new high-throughput sequencing method for determining diversity and similarity of T cell receptor (TCR)  $\alpha$  and  $\beta$  repertoires and identifying potential new invariant TCR  $\alpha$  chains. *BMC Immunol* 17: 38, 2016.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.