SCIENTIFIC REP☉RTS

Quantitative and qualitative characterization of expanded CD4$^+$ T cell clones in rheumatoid arthritis patients

Kazuyoshi Ishigaki[1], Hirofumi Shoda[1], Yuta Kochi[2], Tetsuro Yasui[3], Yuho Kadono[3], Sakae Tanaka[3], Keishi Fujio[1] & Kazuhiko Yamamoto[1]

Rheumatoid arthritis (RA) is an autoimmune destructive arthritis associated with CD4$^+$ T cell-mediated immunity. Although expanded CD4$^+$ T cell clones (ECs) has already been confirmed, the detailed characteristics of ECs have not been elucidated in RA. Using combination of a single-cell analysis and next-generation sequencing (NGS) in TCR repertoire analysis, we here revealed the detailed nature of ECs by examining peripheral blood (PB) from 5 RA patients and synovium from 1 RA patient. When we intensively investigated the single-cell transcriptome of the most expanded clones in memory CD4$^+$ T cells (memory-mECs) in RA-PB, senescence-related transcripts were up-regulated, indicating circulating ECs were constantly stimulated. Tracking of the transcriptome shift within the same memory-mECs between PB and the synovium revealed the augmentations in senescence-related gene expression and the up-regulation of synovium-homing chemokine receptors in the synovium. Our in-depth characterization of ECs in RA successfully demonstrated the presence of the specific immunological selection pressure, which determines the phenotype of ECs. Moreover, transcriptome tracking added novel aspects to the underlying sequential immune processes. Our approach may provide new insights into the pathophysiology of RA.

Rheumatoid arthritis (RA) is an autoimmune disease that is characterized by systemic chronic synovitis and bone erosion. Although various types of innate and adaptive immune cells have been shown to coordinately contribute to the pathophysiology of RA, evidence obtained from human[1–3] and mouse studies[4–7] has suggested the central roles of CD4$^+$ T cells. Clonal expansion is a unique feature of the adaptive immune system, which makes it efficient and potent for antigen-specific reactions. A detailed characterization of clonal expansion is informative for assessing abnormalities in autoimmune diseases, especially antigen-specific autoimmunity. T cell receptor (TCR) repertoire analysis is the standard approach used to assess the clonal expansion status and many studies have focused on expanded CD4$^+$ T cell clones (ECs) in RA[8–10]. However, only limited fraction of CD4$^+$ T cells were analyzed in previous articles and it was methodologically impossible to characterize the phenotypes of ECs.

Complementarity-determining region 3 (CDR3) of TCR has theoretically 10^{15} diversity[11] and can be used as a unique marker of each T cell clone. Historically, single strand conformation polymorphism (SSCP) and Sanger's sequencing have been used for TCR analyses[9,12,13]. Recent advances in next-generation sequencing (NGS) has enabled us to perform quantitative analyses of the TCR repertoire with large amounts of TCR sequencing data. Although the NGS TCR repertoire analysis is becoming a widely accepted method[14–16], it can be affected by PCR bias and has a higher error rate than that of

[1]Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan. [2]Laboratory for Autoimmune Diseases, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan. [3]Department of Orthopaedic Surgery, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan. Correspondence and requests for materials should be addressed to K.F. (email: kfujio-tky@umin.ac.jp)
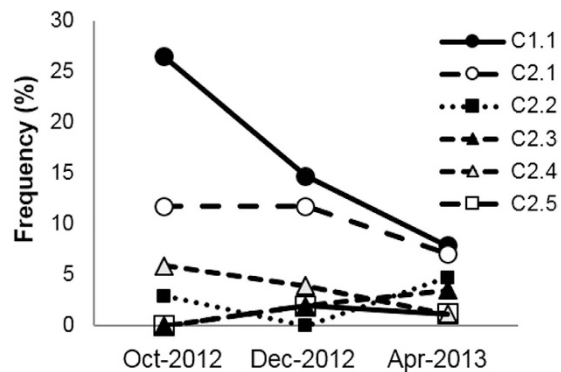
Figure 1. Major memory-ECs were repeatedly detected in RA-PB. Single-cell analysis of PB memory CD4+ T cells from RA1 and RA2 was repeatedly performed with 3-month intervals and the TCR repertoire was analyzed. Each dot and line indicates one clone. Ci.j is the clone ID indicating the j-th expanded major clone in RAi.

classical Sanger's sequencing[17]. Sequencing error is a critical issue especially in TCR sequencing. Random V-D-J recombination makes the diversity of CDR3 so enormous that it is difficult to distinguish sequencing errors from rare clones. Although there have been some studies on error correction algorithms for NGS TCR repertoire analysis[18], careful applications are required because inaccurate algorithm might distort the results of TCR repertoire analysis. On the other hand, TCR repertoire analysis based on single-cell Sanger's sequencing has complementary features to NGS: negligible PCR bias and a low error rate. Therefore, the accuracy and the validity of error correction algorithms for NGS TCR repertoire analysis have to be tested using reference data obtained by single-cell Sanger's sequencing. However, few studies to date have directly compared NGS and single-cell TCR repertoire analyses.

Due to the lack of reliable markers of clonal expansion, single-cell transcriptome analysis is the only method that allows ECs to be characterized because TCR sequencing and gene expression analysis have to be investigated simultaneously and analyzed comprehensively in a single-cell resolution. The extremely small amount of total RNA (approximately 10 pg total RNA/single cells) has been the main obstacle to single-cell gene expression analysis. However, recent improvements in experimental methodology have made single-cell transcriptome analysis feasible by efficient amplifications of single-cell cDNA and the suppression of PCR byproducts[19–21].

The combination of a comprehensive TCR repertoire analysis and single-cell transcriptome analysis now allows ECs to be characterized quantitatively and qualitatively. Moreover, single-cell transcriptome analysis also enables us to track the gene expression changes of the same clones in different sites. In this way, through in-depth characterization of ECs, we might be able to detect pathogenic clones and unveil the new aspects of pathophysiology of RA.

In the present study, we began our experiments by validating the NGS TCR repertoire analysis using a single-cell analysis. The TCR repertoire of different subsets of CD4+ T cells in peripheral blood (PB) from 5 RA patients and 5 healthy controls was then analyzed by NGS. We finally performed a single-cell transcriptome analysis of the most expanded CD4+ T cell clones (mECs) in PB and the synovium and tracked the shift in gene expression profiles between them.

Results

ECs were continuously detected in long-standing RA. The existence of ECs has been reported previously in both healthy individuals and RA patients[22,23]. We first examined the persistence of ECs over several months. In two RA patients with stable disease and fixed treatments, the TCR repertoire analysis of PB memory CD4+ T cells was repeatedly performed with 3-month intervals by a single-cell analysis, which is the gold standard method. Since we sorted a maximum of 102 single cells per one sample, we defined ECs in the single-cell analysis as CD4+ T cell clones observed more than once. We detected several identical memory-phenotype ECs (memory-ECs) repeatedly from each patient and confirmed the persistent oligoclonal expansion of major memory-ECs in RA-PB (Fig. 1). Detailed information on the main memory-ECs detected in this pipeline was listed in Table S3. We showed that memory-ECs continuously occupied a significant percentage of memory CD4+ T cells and these results suggested the existence of chronic selection pressure.

Naive and Memory CD4+ T cells had a distinct TCR repertoire. In order to characterize the TCR repertoire of CD4+ T cells in RA, we sorted naive and memory CD4+ T cells from RA-PB and healthy control (HC) PB and performed the NGS TCR repertoire analysis. We first verified the robustness of our NGS TCR repertoire analysis platform by single-cell analysis and FACS analysis (Fig. S1 and S2). We were able to identify ECs by high-frequency CDR3 sequences in the NGS TCR repertoire analysis. Since
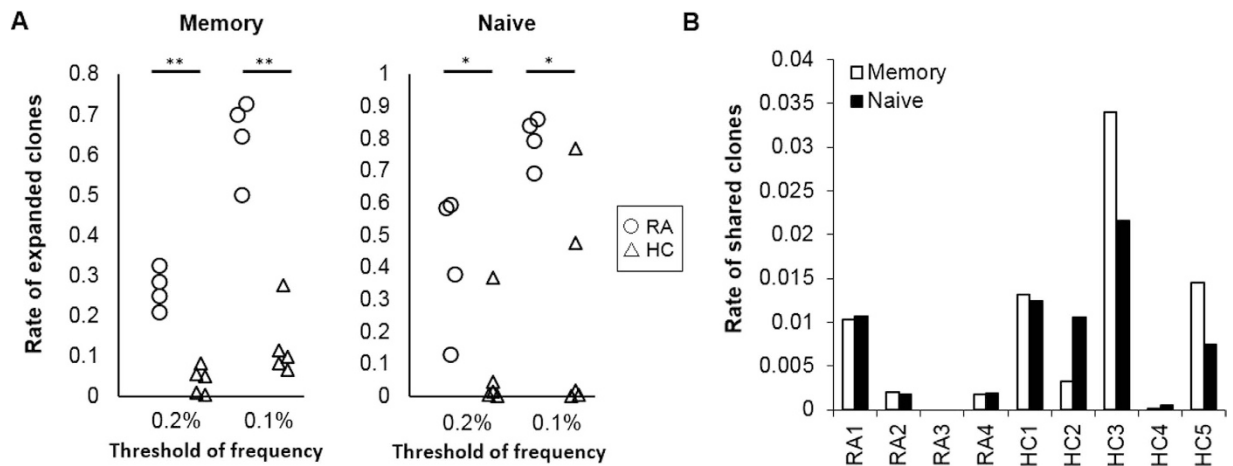
Figure 2. Both of naive and memory CD4+ T cells in RA-PB had many ECs. Naïve and memory CD4+ T cells were sorted from the PB of 4 RA patients and 5 healthy controls and the NGS TCR repertoire analysis was performed. (**A**) Rate of expanded clones (clones with more than 0.2% or 0.1% of total number of reads with functional CDR3 sequences). The cumulative read count of ECs was divided by the total number of reads with functional CDR3 sequences. (**B**) Rate of clones that were shared by both the naive and memory CD4+ T cell subset. The cumulative read count of shared clones was divided by the total number of reads with functional CDR3 sequences. *p < 0.05, **p < 0.005.

there is currently no general definition of ECs in the NGS TCR repertoire analysis, we prepared two thresholds of frequencies for ECs (Fig. 2A, 0.2% and 0.1%). We observed significantly more ECs in not only memory, but also naive CD4+ T cells in RA-PB than in healthy control PB. When the similarities of sequences between memory and naive CD4+ T cells were examined, we found that the majority of clones in each subset were unique and not shared by the other subset both in RA and HC (Fig. 2B). These results revealed marked differences in the TCR repertoire between naive and memory CD4+ T cells.

The non-Th1/Th17/Tfh subset in RA-PB contained the majority of expanded and synovium-infiltrating clones. Memory CD4+ T cells are functionally heterogeneous and can be classified into subpopulations by characteristic cytokine productions, transcription factor networks and epigenomic changes[24–27]. Among the subpopulations of memory CD4+ T cells, Th1, Th17 and Tfh have been reported to have critical roles in the pathogenesis of RA and other autoimmune diseases[28–32]. In order to further characterize the TCR repertoire of memory CD4+ T cells in RA, we sorted Th1, Th17, Tfh, and non-Th1/Th17/Tfh subsets (see details in methods) and performed the NGS TCR repertoire analysis. Non-Th1/Th17/Tfh subsets contained Th2 and other subsets which are not well characterized in the current understanding of human immune system.

We initially focused on the most expanded CD4+ T cell clones (mECs) that persistently expanded with the highest frequency as shown in Fig. 1 (C1.1 for RA1 and C2.1 for RA2). Both of these were the most frequently observed in the non-Th1/Th17/Tfh subsets (Fig. 3A). The distributions of target clones (Fig. 3A–C) were calculated based on the frequency assessed by the TCR repertoire analysis and the size of each subset assessed by FACS analysis (see details in the figure legend). We then examined the distribution of ECs (more than 0.2%) in the 4 subsets of memory CD4+ T cells and found that the majority of ECs were also detected in the non-Th1/Th17/Tfh subsets (Fig. 3B). We tracked synovial tissue-infiltrating CD4+ T cells and examined their localization in PB. A comparison of the PB and synovial tissue TCR repertoires from RA1 revealed that the non-Th1/Th17/Tfh subsets contained the majority of synovial tissue-infiltrating CD4+ T cells (Fig. 3C). The distinct clonality of ECs between naïve and memory CD4+ T cells and evident skewing of memory ECs toward the non-Th1/Th17/Tfh subsets collectively suggested that the expansion of ECs reflected a specific differentiation process rather than a non-specific generalized activation process.

Gene expression profiles of mECs in RA-PB. In order to further characterize the memory ECs in RA, we next performed single-cell transcriptome analysis of the most expanded CD4+ T cell clones (mECs) and non-expanded CD4+ T cell clones (NECs) in the peripheral blood of 2 patients, RA1 and RA2 (Fig. 4). We defined mECs and NECs as the CD4+ T cell clones detected in the single-cell analysis of each sample with the highest frequency and only once, respectively. We focused on "the most expanded" clones for this purpose because multiple single cells are required for the reliable characterization of the gene expression profile of each clone. We also confined our target to memory-phenotype CD4+ T cell clones (memory-mEC and memory-NECs). By comparing memory-mECs with memory-NECs from the
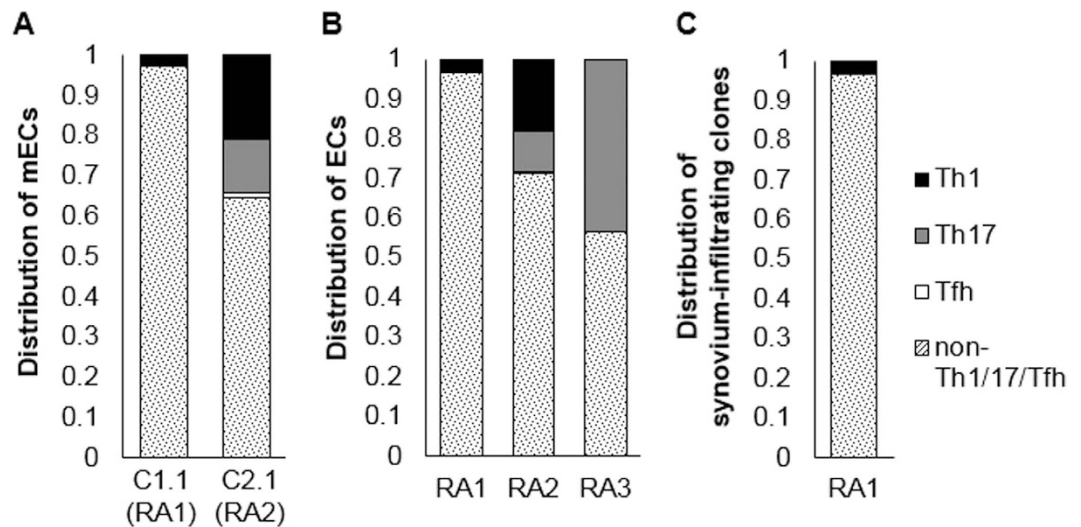
Figure 3. The non-Th1/Th17/Tfh subset in RA-PB contained the majority of expanded and synovium-infiltrating clones. The Th1, Th17, Tfh and non-Th1/Th17/Tfh subsets were sorted from 3 RA-PB and the NGS TCR repertoire analysis was performed. In order to assess the distribution of target clones within different subsets of CD4$^+$ T cells, frequencies assessed by the TCR repertoire analysis (the total number of reads of target clones divided by the total number of reads with functional CDR3 sequences) were multiplied by the frequency of each CD4$^+$ T cell subset within all CD4$^+$ T cells assessed by FACS in order to correct for the subset size. (**A**) The distribution of the mECs in RA1 and RA2 (corresponding to C1.1, C2.1 in Fig. 1). (**B**) The distribution of ECs (clones with more than 0.2% of total number of reads with functional CDR3 sequences). (**C**) The distribution (within PB) of clones that were also detected in the synovium of the same patient.

same patient, we could extract the "clonal expansion" molecular signatures, without being influenced by inter-individual variability.

In order to capture the whole view of the gene expression characteristics of the memory-mECs, we first divided our single-cell cDNA samples into 4 groups: memory-mECs (8 cells) and memory-NECs (65 cells) in RA1 PB and memory-mECs (6 cells) and memory-NECs (48 cells) in RA2 PB. Due to a limitation in the cDNA sample volume, every single-cell cDNA of each group was mixed and quantitative PCR (qPCR) was performed for each pooled sample.

The down-regulation of CD28 and up-regulation of KLRG1 and GZMB were observed in memory-mECs, and this was similar to the phenotype of senescent CD4$^+$ T cells[33]. T cell senescence is one of the characteristic findings in the immunological abnormalities of RA[34,35]. Memory-mECs also had a gene expression profile that was similar to Th1 (up-regulation of TBX21 and IFNG). The down-regulation of RORC, IL17, and Foxp3 suggested that memory-mECs in RA-PB did not belong to the Th17 or Treg subsets.

Gene expression profiles of mECs in the RA synovium. We next performed single-cell transcriptome analysis of synovium of RA1 (Fig. 5). Three single-cell cDNA samples were randomly selected from both synovium memory-mECs and memory-NECs, and analyzed using single-cell RNA-Seq. An average of 1.46 million reads was mapped per sample. In order to assess differentially expressed gene sets, we performed a gene set enrichment analysis[36] and found that the expression of the gene sets of TNF targets were significantly higher in synovium memory-mECs than in synovium memory-NECs (FDR = 0.029. Heatmap was shown in Fig. 5A).

When gene expression was compared between the pooled cDNA of synovium memory-mECs (4 cells) and memory-NECs (15 cells), the significant up-regulation of CXCR4 was observed in memory-mECs (Fig. 5B). CXCR4 is an important chemokine receptor for homing to synovial tissues and was previously shown to be associated with the pathogenicity of arthritis[37,38]. The down-regulation of RORC and Foxp3 also suggested that memory-mECs in the synovium did not belong to the Th17 or Treg subsets, as in RA-PB. We collected immune senescence marker genes[39,40] from RNA-seq data and summarized them in a heatmap (Fig. 5A). The heatmap and qPCR data collectively suggested that synovium memory-mECs also possessed a senescent CD4$^+$ T cell-like gene expression profile (up-regulation of KLRG1, NCAM1, GZMB, and PRF1 and down-regulation of CD27, CD28, and PDCD1).

Gene expression tracking of mECs. The combined use of single-cell TCR repertoire analysis and single-cell transcriptome analysis enabled us to track the gene expression shift within the same clones by localization. Since memory-mECs in the RA1 PB and the RA1 synovium were identical clones
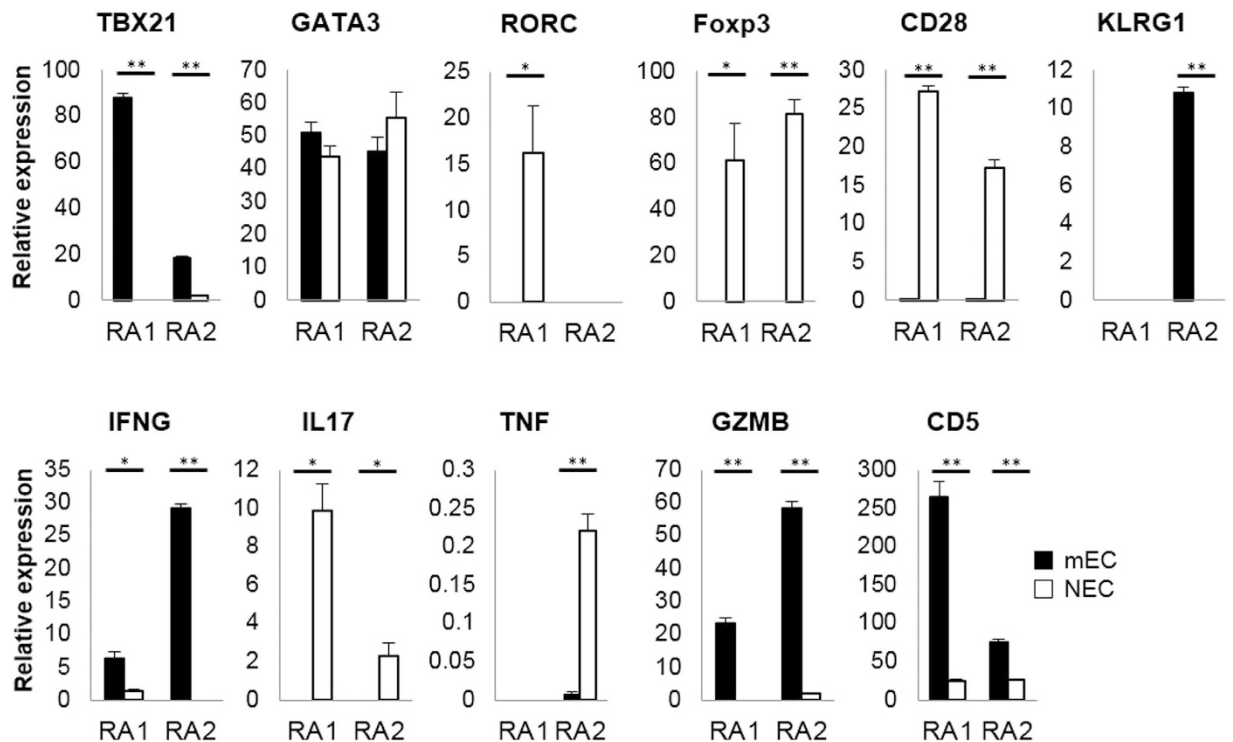
Figure 4. Gene expression profiles of PB memory-mECs. cDNAs of PB memory-mECs (corresponding to C1.1 for RA1 and C2.1 for RA2 as shown in Fig. 1) and memory-NECs obtained by the single-cell analysis were mixed (4 pooled samples in total), and qPCR was performed. ACTB was used as an internal control gene. *p < 0.05, **p < 0.005.
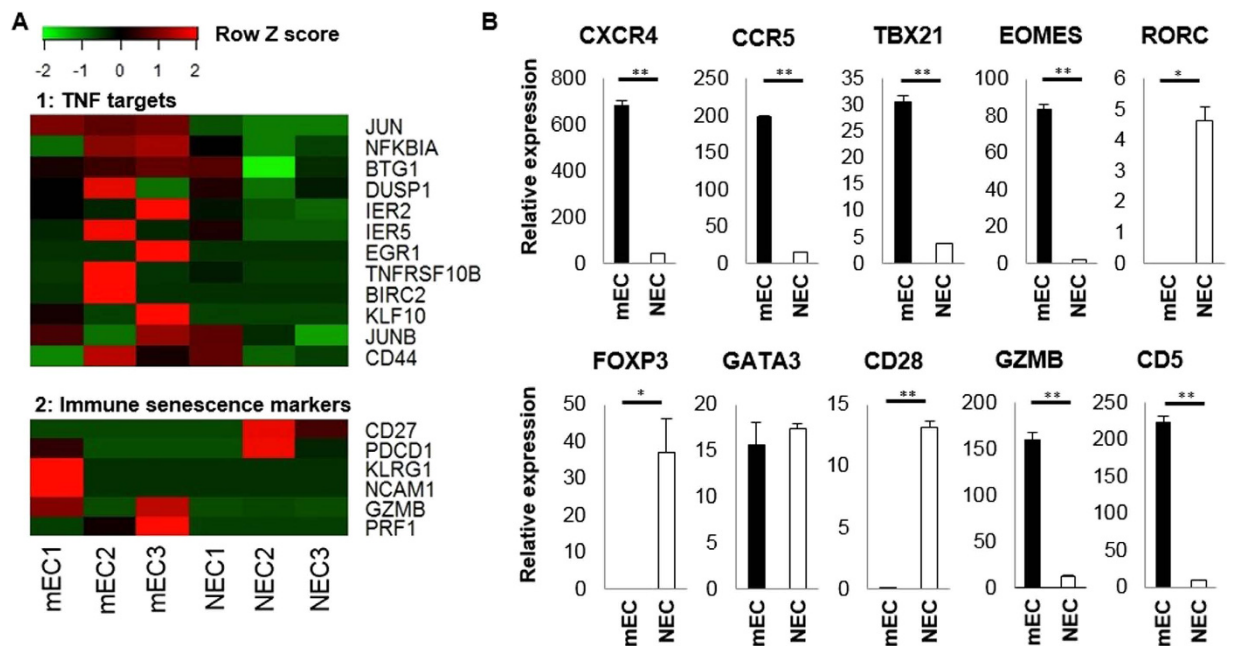

Figure 5. Gene expression profiles of synovium memory-mECs. Three single-cell cDNA samples were randomly selected from both memory-mECs and memory-NECs in the synovium, and analyzed by single-cell RNA-Seq. (**A**) A heatmap of TNF target genes and immune senescent markers. The TNF targets genes were shown to be significantly up-regulated in memory-mECs by the gene set enrichment analysis (see details in the text and supplementary methods). (**B**) cDNAs of synovium memory-mECs and memory-NECs obtained by the single-cell analysis were mixed (2 pooled samples in total), and qPCR was performed. ACTB was used as an internal control gene. *p < 0.05, **p < 0.005.
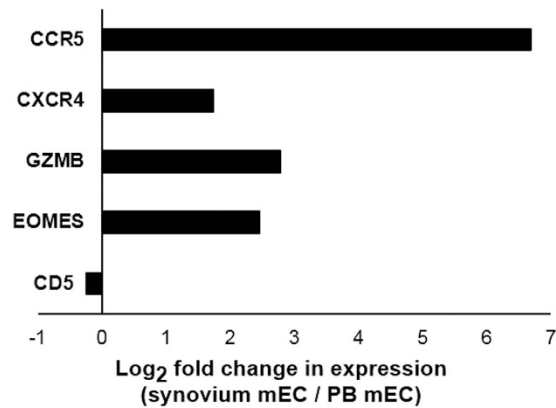
Figure 6. Tracking of the gene expression shift in memory-mECs in RA1 between PB and the synovium. Memory-mECs in the RA1 PB and RA1 synovium were identical clones (corresponding to C1.1 in Fig. 1). cDNAs of RA1 memory-mECs in PB and the synovium obtained by the single-cell analysis were mixed (2 pooled samples in total), and qPCR was performed. ACTB was used as an internal control gene. The log-transformed ratio (gene expression value in the synovium / gene expression value in PB) was shown.

(corresponding to C1.1 in Fig. 1), we examined the phenotypic changes of memory-mECs caused by infiltration into the synovium (Fig. 6). Chemokine receptors, which are important for infiltration into the synovium (CXCR4 and CCR5), and effector molecules (GZMB) were up-regulated in the synovium. On the other hand, CD5, a marker of autoreactivity[41], was stably expressed.

Discussion

In the present study, we proposed an original and rational strategy for the detailed characterization of each EC quantitatively (level of clonal expansion) and qualitatively (infiltration into target organs, cell surface markers, and gene expression status) by combining a comprehensive TCR repertoire analysis and single-cell transcriptome analysis. Although the clonal size and senescence features of ECs in RA-PB have already been reported[23,33], the mode of proliferation of ECs has not yet been clarified. ECs may expand non-specifically among various kinds of T cell subpopulations including naïve and memory T cell subsets. We succeeded in adding novel aspects to the dynamic processes underlying memory-ECs in RA. The persistently high level of expansion of memory-ECs and apparent skewing of their phenotype to distinct subsets clearly demonstrated that ECs were constantly exposed to specific immunological selection pressure. Moreover, tracking of the gene expression profiles of memory-mECs suggested that they underwent sequential immunological modifications. Memory-mECs in PB acquire an immune senescence phenotype by chronic selection pressure and up-regulate the expression of chemokine receptors for homing to the synovium. Memory-mECs are exposed to TNF in the synovium and augment pathogenic activity. We speculated that this sequential phenomenon is not a unique event specific to memory-mECs, but a generalized process for other autoreactive and joint-homing memory CD4+ T cells.

Most of the results obtained in the present study are consistent with previous findings. Senescent CD4+ T cells were previously reported to exhibit high levels of clonal expansion and play pathogenic roles in RA through cytotoxic molecules such as perforin and granzymes[33,40,42,43]. TNF signaling appeared to function as a booster of senescence in synovium memory-mECs because TNF alpha was shown to augment the senescent phenotype in CD4+ T cells[44,45]. EOMES is one of the molecular signatures of senescence, although this finding came from CD8+ T cells[46]. The upregulation of this transcript in RA-PB memory-mECs and more prominently in synovium memory-mECs also appeared to be associated with the senescent phenotype.

Unlike Th1, Th17 and Tfh subsets, non-Th1/Th17/Tfh subset (Th2 and other subsets which are not well characterized) has not been considered and investigated as a pathogenic population in RA. Our findings suggested there might be a novel pathogenic subpopulation in this subset. Although the majority of mECs were detected in the non-Th1/Th17/Tfh subset, some of them were also detected in Th1 and they had the transcriptomes similar to those of the Th1 subset. These results can be explained by inter-subset plasticity or the instability of CXCR3 expression because the discrimination of non-Th1/Th17/Tfh and Th1 was solely dependent on CXCR3 in this study, according to a standard strategy[47].

The critical question of what the cognate antigens of these ECs are remains. The strong expression of CD5 in memory-mECs, which has been associated with autoreactivity[41], indicated that they were auto-reactive clones. Recent studies identified citrullinated vimentin, aggrecan, fibrinogen and BiP as the candidate auto-antigens of CD4+ T cells in RA based on the autoantibody repertoire and other factors[48–51]. The MHC class II tetramer assay showed that the frequency of CD4+ T cells specific to citrullinated epitopes was approximately 0.1–10 cells per 1,000,000 T cells without *in vitro* antigen stimulation[2,29]. The clonal expansion size of ECs was more than one thousand times larger than that of citrullinated

auto-antigen-specific CD4[+] T cells. Therefore, we speculated that the ECs detected in the present study recognized autoantigens other than the above antigens.

Current treatment of RA suppresses inflammation nonspecifically and the risk of adverse events such as infection is inevitable. Although data is limited, several autoantigens have already been identified and there have been several reports about the tolerance-inducing antigen-specific strategies[52–57]. By expanding our knowledge of autoantigens in RA and tolerance-inducing mechanism in human, we may be able to develop antigen-specific therapies which have fewer side effects and possibly become alternative treatment.

The analyzed sample size was evidently not sufficient to generalize our results. Nevertheless, we consider our combined analysis of PB and the synovium in one patient to have reflected at least some common RA features because RA1 was a prototypical seropositive patient with shared-epitope HLA-DRB1 alleles (HLA-DRB1*04:05/0101, Table S1).

Information regarding the physiological or pathogenic roles of ECs is currently limited. The unbiased and robust nature of our strategy should shed light on these roles and may ultimately be able to identify a new aspect of RA pathophysiology.

Methods

Patients and controls.   PB was obtained from 5 RA patients visiting outpatient clinics of the Department of Allergy and Rheumatology at The University of Tokyo Hospital and 5 healthy donors. Synovial tissue was collected from 1 active RA patient (harboring HLA-DRB1*04:05, positive for both RF and anti-CCP antibodies) during elbow synovectomy at the Department of Orthopaedic Surgery and Spinal Surgery. Detailed patient information is summarized in Table S1. RA was diagnosed according to the American College of Rheumatology/European League Against Rheumatism 2010 criteria[58]. This study was approved by the Ethical Committee of The University of Tokyo. A written form of informed consent was obtained from all patients and healthy donors. The methods were carried out in accordance with the approved guidelines.

Cell isolation.   PB mononuclear cells were separated by Ficoll-Hypaque density-gradient centrifugation[59]. Synovial lymphoid cells were also enriched by Ficoll-Hypaque from cell suspensions made by passing synovial tissue through a stainless steel sieve.

Flow cytometry (FACS).   We analyzed and sorted all samples for single-cell and NGS analyses by MoFlo XDP (Beckman Coulter)[60,61]. Cell staining was performed by CXCR5-FITC (clone: RF8B2, BD Pharmingen), CXCR3-PE (clone: 1C6, BD Pharmingen), CD4-PerCP-Cy5.5 (clone: OKT4, Biolegend), CD3-PE-Cy7 (clone: VCHT1, Biolegend), CCR6-APC (clone: 11A9, BD Pharmingen), and CD45RO-APC-Cy7 (clone: UCHL1, BD Pharmingen). The definitions of each subset in this study were as follows: CD3[+]CD4[+]CD45RO[+] for memory CD4[+] T cells, CD3[+]CD4[+]CD45RO[-] for naive CD4[+] T cells, CD3[+]CD4[+]CD45RO[+]CXCR5[+] for follicular helper T cells (Tfh), CD3[+]CD4[+]CD45RO[+]CXCR5[−]CXCR3[+]CCR6[−] for Th1, CD3[+]CD4[+]CD45RO[+]CXCR5[−]CXCR3[−]CCR6[+] for Th17, and CD3[+]CD4[+]CD45RO[+]CXCR5[−]CXCR3[−]CCR6[−] for non-Th1/Th17/Tfh. We followed standard immunophenotyping approach of human immune cells based on chemokine receptors, which is now becoming widely accepted[47]. The total count of sorted cells for the NGS analysis was listed in Table S2.

Single-cell analysis.   Single-cell whole transcriptome amplification was performed as an initial process for every single-cell analysis. We followed the original Kurimoto method[19]. Pre-amplified cDNA was used for the TCR repertoire and gene expression analysis. We focused on the beta chains of TCR because they generally have a more strict allelic exclusion than alpha chains. Detailed methods are described in the supplementary methods.

NGS TCR repertoire analysis.   We also focused on the beta chains of TCR in the NGS TCR repertoire analysis. We modified the 5′-RACE-based PCR strategy for the TCR region described previously[62]. Mapping to the TCR region, identification of the TRBV gene, extraction of CDR3 sequences, and error corrections were performed by MiTCR with default settings[18]. Detailed methods are described in the supplementary methods.

Statistical analysis.   Comparisons of numerical data between samples were analyzed by the unpaired Student's t-test. $p < 0.05$ was considered significant. Spearman's correlation coefficient was used to indicate the relationship between two ordered sets of numbers and was calculated by R (version3.0.2).

References

1. Okada, Y. *et al.* Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat Genet* **44,** 511–516, doi: 10.1038/ng.2231 (2012).
2. Scally, S. W. *et al.* A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med* **210,** 2569–2582, doi: 10.1084/jem.20131241 (2013).
3. Kremer, J. M. *et al.* Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. *N Engl J Med* **349,** 1907–1915, doi: 10.1056/NEJMoa035075 (2003).

4. Kadowaki, K. M., Matsuno, H., Tsuji, H. & Tunru, I. CD4+ T cells from collagen-induced arthritic mice are essential to transfer arthritis into severe combined immunodeficient mice. *Clin Exp Immunol* **97,** 212–218 (1994).

5. Jang, E. *et al.* Prevention of spontaneous arthritis by inhibiting homeostatic expansion of autoreactive CD4+ T cells in the K/BxN mouse model. *Arthritis Rheum* **54,** 492–498, doi: 10.1002/art.21567 (2006).

6. Sakaguchi, N. *et al.* Altered thymic T-cell selection due to a mutation of the ZAP-70 gene causes autoimmune arthritis in mice. *Nature* **426,** 454–460, doi: 10.1038/nature02119 (2003).

7. Hirota, K. *et al.* Preferential recruitment of CCR6-expressing Th17 cells to inflamed joints via CCL20 in rheumatoid arthritis and its animal model. *J Exp Med* **204,** 2803–2812, doi: 10.1084/jem.20071397 (2007).

8. Goronzy, J. J. *et al.* Dominant clonotypes in the repertoire of peripheral CD4+ T cells in rheumatoid arthritis. *J Clin Invest* **94,** 2068–2076, doi: 10.1172/JCI117561 (1994).

9. Yamamoto, K. *et al.* Accumulation of multiple T cell clonotypes in the synovial lesions of patients with rheumatoid arthritis revealed by a novel clonality analysis. *Int Immunol* **4,** 1219–1223 (1992).

10. Wagner, U. *et al.* Clonally expanded CD4+CD28null T cells in rheumatoid arthritis use distinct combinations of T cell receptor BV and BJ elements. *Eur J Immunol* **33,** 79–84, doi: 10.1002/immu.200390010 (2003).

11. Davis, M. M. & Bjorkman, P. J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334,** 395–402, doi: 10.1038/334395a0 (1988).

12. Fujio, K. *et al.* Gene therapy of arthritis with TCR isolated from the inflamed paw. *J Immunol* **177,** 8140–8147 (2006).

13. Tahara, H. *et al.* Reconstitution of CD8+ T cells by retroviral transfer of the TCR alpha beta-chain genes isolated from a clonally expanded P815-infiltrating lymphocyte. *J Immunol* **171,** 2154–2160 (2003).

14. Klinger, M. *et al.* Combining next-generation sequencing and immune assays: a novel method for identification of antigen-specific T cells. *PLoS One* **8,** e74231, doi: 10.1371/journal.pone.0074231 (2013).

15. Warren, E. H., Matsen, F. A. & Chou, J. High-throughput sequencing of B- and T-lymphocyte antigen receptors in hematology. *Blood* **122,** 19–22, doi: 10.1182/blood-2013-03-453142 (2013).

16. Sherwood, A. M. *et al.* Tumor-infiltrating lymphocytes in colorectal tumors display a diversity of T cell receptor sequences that differ from the T cells in adjacent mucosal tissue. *Cancer Immunol Immunother* **62,** 1453–1461, doi: 10.1007/s00262-013-1446-2 (2013).

17. Bolotin, D. A. *et al.* Next generation sequencing for TCR repertoire profiling: platform-specific features and correction algorithms. *Eur J Immunol* **42,** 3073–3083, doi: 10.1002/eji.201242517 (2012).

18. Bolotin, D. A. *et al.* MiTCR: software for T-cell receptor sequencing data analysis. *Nat Methods* **10,** 813–814, doi: 10.1038/nmeth.2555 (2013).

19. Kurimoto, K., Yabuta, Y., Ohinata, Y. & Saitou, M. Global single-cell cDNA amplification to provide a template for representative high-density oligonucleotide microarray analysis. *Nat Protoc* **2,** 739–752, doi: 10.1038/nprot.2007.79 (2007).

20. Yan, L. *et al.* Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* **20,** 1131–1139, doi: 10.1038/nsmb.2660 (2013).

21. Sasagawa, Y. *et al.* Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol* **14,** R31, doi: 10.1186/gb-2013-14-4-r31 (2013).

22. Klarenbeek, P. L. *et al.* Human T-cell memory consists mainly of unexpanded clones. *Immunol Lett* **133,** 42–48, doi: 10.1016/j.imlet.2010.06.011 (2010).

23. Klarenbeek, P. L. *et al.* Inflamed target tissue provides a specific niche for highly expanded T-cell clones in early human autoimmune disease. *Ann Rheum Dis* **71,** 1088–1093, doi: 10.1136/annrheumdis-2011-200612 (2012).

24. Vahedi, G. *et al.* Super-enhancers delineate disease-associated regulatory nodes in T cells. *Nature* **520,** 558–562, doi: 10.1038/nature14154 (2015).

25. Chang, J. T., Wherry, E. J. & Goldrath, A. W. Molecular regulation of effector and memory T cell differentiation. *Nat Immunol* **15,** 1104–1115, doi: 10.1038/ni.3031 (2014).

26. Mosmann, T. R. & Coffman, R. L. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annu Rev Immunol* **7,** 145–173, doi: 10.1146/annurev.iy.07.040189.001045 (1989).

27. Ivanov, I. I. *et al.* The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells. *Cell* **126,** 1121–1133, doi: 10.1016/j.cell.2006.07.035 (2006).

28. Yamada, H. *et al.* Th1 but not Th17 cells predominate in the joints of patients with rheumatoid arthritis. *Ann Rheum Dis* **67,** 1299–1304, doi: 10.1136/ard.2007.080341 (2008).

29. James, E. A. *et al.* Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. *Arthritis Rheumatol* **66,** 1712–1722, doi: 10.1002/art.38637 (2014).

30. Leipe, J. *et al.* Role of Th17 cells in human autoimmune arthritis. *Arthritis Rheum* **62,** 2876–2885, doi: 10.1002/art.27622 (2010).

31. van Hamburg, J. P. *et al.* Th17 cells, but not Th1 cells, from patients with early rheumatoid arthritis are potent inducers of matrix metalloproteinases and proinflammatory cytokines upon synovial fibroblast interaction, including autocrine interleukin-17A production. *Arthritis Rheum* **63,** 73–83, doi: 10.1002/art.30093 (2011).

32. Ma, C. S. & Deenick, E. K. Human T follicular helper (Tfh) cells and disease. *Immunol Cell Biol* **92,** 64–71, doi: 10.1038/icb.2013.55 (2014).

33. Goronzy, J. J. & Weyand, C. M. Aging, autoimmunity and arthritis: T-cell senescence and contraction of T-cell repertoire diversity - catalysts of autoimmunity and chronic inflammation. *Arthritis Res Ther* **5,** 225–234, doi: 10.1186/ar974 (2003).

34. Weyand, C. M., Fulbright, J. W. & Goronzy, J. J. Immunosenescence, autoimmunity, and rheumatoid arthritis. *Exp Gerontol* **38,** 833–841 (2003).

35. Weyand, C. M., Yang, Z. & Goronzy, J. J. T-cell aging in rheumatoid arthritis. *Curr Opin Rheumatol* **26,** 93–100, doi: 10.1097/BOR.0000000000000011 (2014).

36. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102,** 15545–15550, doi: 10.1073/pnas.0506580102 (2005).

37. Buckley, C. D. *et al.* Persistent induction of the chemokine receptor CXCR4 by TGF-beta 1 on synovial T cells contributes to their accumulation within the rheumatoid synovium. *J Immunol* **165,** 3423–3429 (2000).

38. Tamamura, H. *et al.* Identification of a CXCR4 antagonist, a T140 analog, as an anti-rheumatoid arthritis agent. *FEBS Lett* **569,** 99–104, doi: 10.1016/j.febslet.2004.05.056 (2004).

39. Larbi, A. & Fulop, T. From "truly naïve" to "exhausted senescent" T cells: when markers predict functionality. *Cytometry A* **85,** 25–35, doi: 10.1002/cyto.a.22351 (2014).

40. Broux, B., Markovic-Plese, S., Stinissen, P. & Hellings, N. Pathogenic features of CD4+CD28– T cells in immune disorders. *Trends Mol Med* **18,** 446–453, doi: 10.1016/j.molmed.2012.06.003 (2012).

41. Mandl, J. N., Monteiro, J. P., Vrisekoop, N. & Germain, R. N. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* **38,** 263–274, doi: 10.1016/j.immuni.2012.09.011 (2013).

42. Namekawa, T., Wagner, U. G., Goronzy, J. J. & Weyand, C. M. Functional subsets of CD4 T cells in rheumatoid synovitis. *Arthritis Rheum* **41,** 2108–2116, doi: 10.1002/1529-0131(199812)41:12<2108::AID-ART5>3.0.CO;2-Q (1998).

43. Fasth, A. E., Björkström, N. K., Anthoni, M., Malmberg, K. J. & Malmström, V. Activating NK-cell receptors co-stimulate CD4(+)CD28(−) T cells in patients with rheumatoid arthritis. *Eur J Immunol* **40,** 378–387, doi: 10.1002/eji.200939399 (2010).

44. Bryl, E., Vallejo, A. N., Weyand, C. M. & Goronzy, J. J. Down-regulation of CD28 expression by TNF-alpha. *J Immunol* **167,** 3231–3238 (2001).

45. Bryl, E. *et al.* Modulation of CD28 expression with anti-tumor necrosis factor alpha therapy in rheumatoid arthritis. *Arthritis Rheum* **52,** 2996–3003, doi: 10.1002/art.21353 (2005).

46. Wherry, E. J. *et al.* Molecular signature of CD8+ T cell exhaustion during chronic viral infection. *Immunity* **27,** 670–684, doi: 10.1016/j.immuni.2007.09.006 (2007).

47. Maecker, H. T., McCoy, J. P. & Nussenblatt, R. Standardizing immunophenotyping for the Human Immunology Project. *Nat Rev Immunol* **12,** 191–200, doi: 10.1038/nri3158 (2012).

48. Snir, O. *et al.* Identification and functional characterization of T cells reactive to citrullinated vimentin in HLA-DRB1*0401-positive humanized mice and rheumatoid arthritis patients. *Arthritis Rheum* **63,** 2873–2883, doi: 10.1002/art.30445 (2011).

49. Hill, J. A. *et al.* Arthritis induced by posttranslationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice. *J Exp Med* **205,** 967–979, doi: 10.1084/jem.20072051 (2008).

50. Hill, J. A. *et al.* Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J Immunol* **171,** 538–541 (2003).

51. von Delwig, A., Locke, J., Robinson, J. H. & Ng, W. F. Response of Th17 cells to a citrullinated arthritogenic aggrecan peptide in patients with rheumatoid arthritis. *Arthritis Rheum* **62,** 143–149, doi: 10.1002/art.25064 (2010).

52. Thomas, R. Dendritic cells and the promise of antigen-specific therapy in rheumatoid arthritis. *Arthritis Res Ther* **15,** 204, doi: 10.1186/ar4130 (2013).

53. Farooq, S. M., Kumar, A. & Ashour, H. M. Eye-mediated immune tolerance to Type II collagen in arthritis-prone strains of mice. *J Cell Mol Med* **18,** 2512–2518, doi: 10.1111/jcmm.12376 (2014).

54. Farooq, S. M. & Ashour, H. M. Type II collagen induces peripheral tolerance in BALB/c mice via the generation of CD8+ T regulatory cells. *PLoS One* **7,** e48635, doi: 10.1371/journal.pone.0048635 (2012).

55. Chen, L., Bao, B., Wang, N., Xie, J. & Wu, W. Oral Administration of Shark Type II Collagen Suppresses Complete Freund's Adjuvant-Induced Rheumatoid Arthritis in Rats. *Pharmaceuticals (Basel)* **5,** 339–352, doi: 10.3390/ph5040339 (2012).

56. Yoshinari, O. *et al.* Water-soluble undenatured type II collagen ameliorates collagen-induced arthritis in mice. *J Med Food* **16,** 1039–1045, doi: 10.1089/jmf.2013.2911 (2013).

57. Shoda, H. *et al.* Autoantigen BiP-Derived HLA-DR4 Epitopes Differentially Recognized by Effector and Regulatory T Cells in Rheumatoid Arthritis. *Arthritis Rheumatol* **67,** 1171–1181, doi: 10.1002/art.39054 (2015).

58. Aletaha, D. *et al.* 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* **62,** 2569–2581, doi: 10.1002/art.27584 (2010).

59. English, D. & Andersen, B. R. Single-step separation of red blood cells. Granulocytes and mononuclear leukocytes on discontinuous density gradients of Ficoll-Hypaque. *J Immunol Methods* **5,** 249–252 (1974).

60. Bhattacharya, P. *et al.* A novel pancreatic β-cell targeting bispecific-antibody (BsAb) can prevent the development of type 1 diabetes in NOD mice. *Clin Immunol* **153,** 187–198, doi: 10.1016/j.clim.2014.04.014 (2014).

61. Chen, Z. *et al.* The novel role of IL-7 ligation to IL-7 receptor in myeloid cells of rheumatoid arthritis and collagen-induced arthritis. *J Immunol* **190,** 5256–5266, doi: 10.4049/jimmunol.1201675 (2013).

62. Sun, X. *et al.* Unbiased analysis of TCRα/β chains at the single-cell level in human CD8+ T-cell subsets. *PLoS One* **7,** e40386, doi: 10.1371/journal.pone.0040386 (2012).

Acknowledgments

Author Contributions

K.I., H.S. and K.F. experimental conception, design of the study, analysis and interpretation of data and drafting the article. Y.Kochi, T.Y., Y.Kadono, S.T. and K.Y. analysis and interpretation of data. All authors have given their final approval for the manuscript to be published as presented.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: K.Y. received financial support or fees from AbbVie, Astellas, BMS, Daiichi-Sankyo, MitsubishiTanabe, Pfizer, Sanofi, Santen, Takeda, Teijin., Boehringer Ingelheim, Chugai, Eisai, Ono, Taisho Toyama,UCB., ImmunoFuture, Asahi Kasei, and Janssen. K.F. received financial support or fees from Astellas, BMS, Daiichi-Sankyo, MitsubishiTanabe, Pfizer, Santen, Takeda, Chugai, Eisai, Taisho Toyama and UCB., and Janssen. All other authors declare no competing financial interests.

How to cite this article: Ishigaki, K. *et al.* Quantitative and qualitative characterization of expanded CD4+ T cell clones in rheumatoid arthritis patients. *Sci. Rep.* **5,** 12937; doi: 10.1038/srep12937 (2015).