Database of synovial T cell repertoire of rheumatoid arthritis patients identifies cross-reactive potential against pathogens including unencountered SARS-CoV-2

Rheumatoid arthritis (RA) is a chronic autoimmune disease of complex aetiology in which adaptive immune responses mediated by T cells play a major role in driving synovial inflammation and joint damage.[1] Before developing effector functions, naïve T cells are conventionally understood to require an initial antigen-specific stimulus transduced through their T cell receptor (TCR). TCR-mediated recognition is also required for targeted cell lysis by cytotoxic T cells. Previous studies have sought to characterise the functional roles of various T cell populations in RA, and have demonstrated patterns of oligoclonal expansion among them.[2] It has also been shown that some T cell clones may be shared among multiple patients as a public repertoire.[3] However, the antigen specificities of synovial T cells remain poorly studied outside of a few epitopes, and open resources of TCRs and repertoires associated with RA are lacking.

To accelerate research efforts into the TCR repertoires and T cell antigen specificities in RA, we report here synovial repertoires mined using a repertoire reconstruction algorithm from two large-cohort RNA sequencing datasets totaling 196 RA patients, as well as additional controls from peripheral blood and other joint disorders (figure 1A, online supplemental methods). The first, GSE89408, had been generated to identify transcriptome signatures in the synovium of patients with RA associated with disease progression (online supplemental table S1).[4] The second, PEAC, focused on early, treatment-naïve patients and identify transcriptome signatures contributing to establishment of disease in synovial tissue and peripheral blood (online supplemental table S2).[5] By exploring these cohorts in the hitherto unnavigated immune repertoire dimension, we built the first openly accessible online database of high-frequency TCR sequences found in the synovium of patients with RA (http://repertoire.life). This reference database is an ongoing effort to collect and compile human TCR repertoires of patients with RA, and can be searched with TCR repertoires identified in future work to better identify public sequences and contextualise their antigen specificity.

Initial analysis of this database demonstrated that an increased repertoire diversity among synovial (but not peripheral blood) high-frequency clonotypes was associated with higher disease activity (Disease Activity Score-28, (DAS28)) and erythrocyte sedimentation rate (ESR) (figure 1B,C, online supplemental figure S1, online supplemental table S3), highlighting the significant correlations between synovial oligoclonal expansion and clinical indices. Through TCR clustering analyses, we observed that synovial TCR motifs from patients with RA were largely distinct from those found in other joint diseases (online supplemental figure S2), in a manner consistent with their overall transcriptome differences (online supplemental figure S3), but that these motifs could be shared in aggregate among patients with lymphoid pathotype (online supplemental figure S4). Consistently, we also observed that several of these motifs were clonally expanded and enriched in patients with highly active disease as assessed through multiple indices(figure 1D, online supplemental figures S5,S6), further indicating their potential involvement in RA progression. Through comparison with paired peripheral blood samples from the same patients, we observed that a portion of synovial clones were shared with peripheral blood independently of pathotype and disease activity levels (online supplemental figures S7,S8), indicating that clonal migration may generally occur in RA. Collectively, these results show that recovered synovial TCRs correlate with clinical measures of RA disease activity, and suggest that they may potentially originate from circulation.

Beyond observing associations between RA TCR repertoire and disease activity, we further sought to annotate the antigen specificities of these clonotypes using databases of validated antigen-specific sequences. Intriguingly, we observed that many of the clonotypes expanded in RA were highly similar to pathogen antigen-specific sequences (figure 1E, online supplemental figures S9–S11). These include commonly encountered pathogens (such as cytomegalovirus (CMV), Epstein-Barr virus (EBV), influenza), as previously described[6] (figure 1F). At the same time, some clones matched pathogens that these patients may have never encountered (such as dengue virus (DENV) and yellow fever virus (YFV)), even when absolute sequence matching criteria were used (online supplemental figure S12). More surprisingly, we observed a number of TCRs matching annotated sequences specific for multiple peptides originating from SARS-CoV-2 (figure 1G). These include three dominant epitopes shared among betacoronaviruses (online supplemental figure S13). Since these patients had been sequenced years prior to its emergence, these matches between T cell clones found in RA synovium and SARS-CoV-2 antigen-specific clones may potentially originate from cross-reactivity, possibly with uncharacterised viruses and/or autoantigens.

To confirm the presence of these SARS-CoV-2 antigen-specific matching sequences identified from recovered repertoires, we then performed additional analyses using TCRseq, which has higher sensitivity in capturing rarer clones (such as naïve cells). When we searched a pre-pandemic TCRseq dataset of peripheral blood samples of 65 patients with RA and 20 controls,[7] we observed that although annotated SARS-CoV-2 reactive clonotypes could be found in both controls and patients with RA, the frequency of these clonotypes was much higher in patients with RA, with an additional peak not seen in controls (online supplemental figure S14). At the same time, we also performed retrospective analysis on 8 TCRseq libraries of FACS-sorted CD8+T cells derived from synovial fluid of 4 patients with RA we had sequenced in 2018. These libraries included both CD57+ effector cells and an overall pool of CD57- cells (online supplemental figure S15, online supplemental table S4). Consistent with expectations, searching these samples against the reconstructed repertoires of the PEAC and GSE89408 cohorts demonstrated that a substantial number of clonotypes recovered in these patients fell into the same specificity groups as our TCRseq sequences (figure 1H,I, online supplemental figure S16). When we then searched these TCRseq libraries against antigen-specificity databases, we could once again observe that sequences annotated against SARS-CoV-2 spike antigen were found to be in the same specificity clusters (figure 1J).

Of note, TCRβ sequence sharing/similarity alone is not an absolute determinant of TCR specificity, with other factors such as TCRα structure, costimulatory molecules, and HLA-restriction/effective antigen-presentation also providing key contributions. However, these results provide a crucial necessary condition for the possibility that some patients with RA may naturally harbour TCR clones sensitive to SARS-CoV-2 antigens, particularly among CD8+T cells. These clones may then be predisposed to preferentially expand and/or adopt effector functions if the pathogen is subsequently encountered,

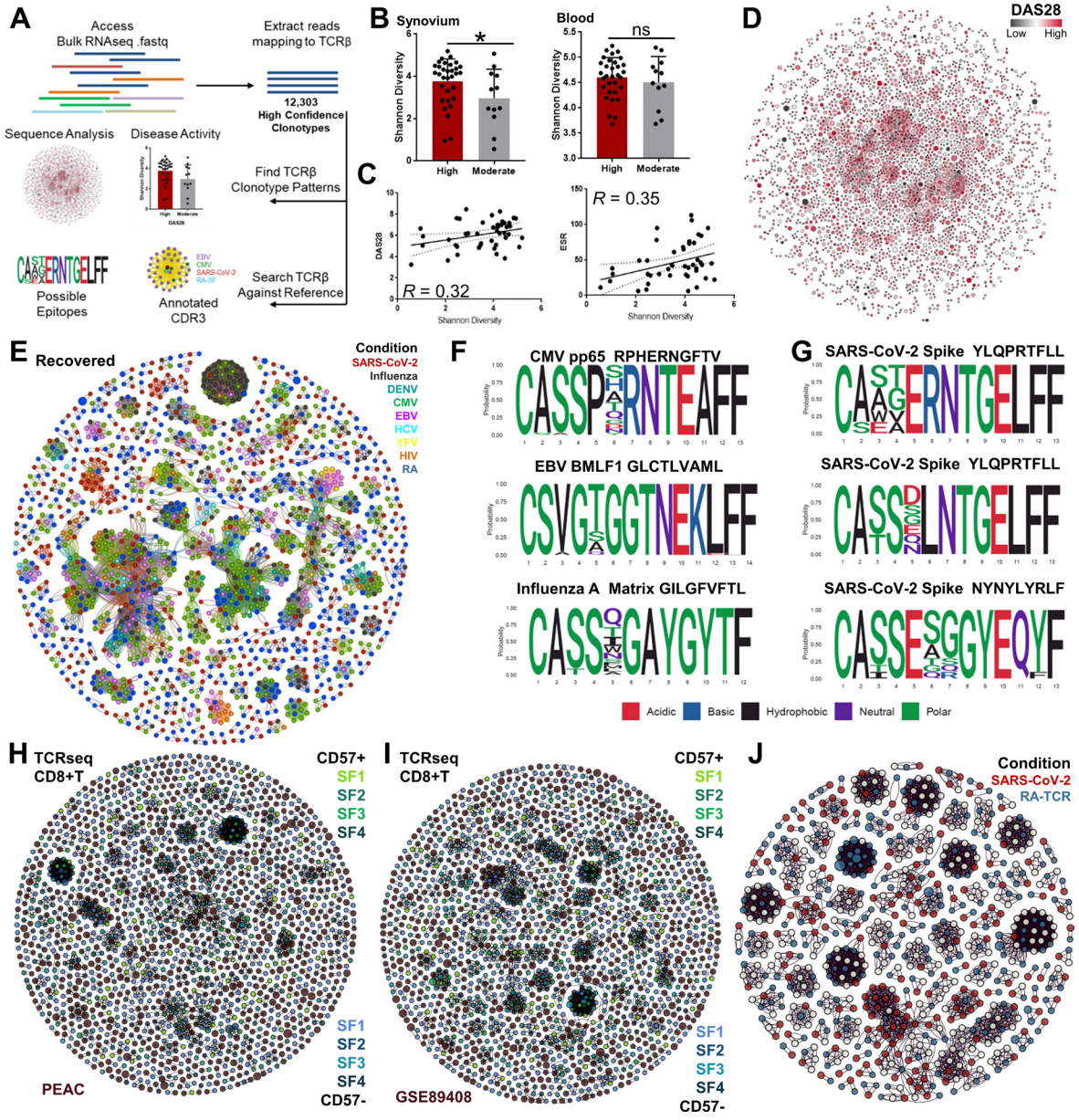

Figure 1 Extracted TCR repertoire analysis of patients with rheumatoid arthritis (RA) captures signatures of clonal expansion and pathogen antigen cross-reactivity. (A) Schematic overview of the analysis pipeline for extracting and processing RA T cell receptor (TCR) repertoire information. Raw RNA sequencing data of patients with RA from different cohorts was collected and aligned using TRUSTV.4 to extract TCR sequences. TCR sequences supported by multiple reads were retained and used to construct an RA-associated TCR repertoire database. Further analyses were performed using this database to identify associations with disease activity, as well as potentially enriched motifs. Annotation of these motifs demonstrates a potential for cross-reactivity between synovial TCR sequences and pathogen antigen-specific TCRs. (B) Shannon diversity of TCRβ clonotypes was higher in patients with high disease activity (DAS28 > 5.1) in synovial tissue, but not significantly different in peripheral blood. (C) Both overall DAS28 and ESR showed positive correlation with increasing synovial TCRβ diversity. (D) Network analyses of RA-associated TCRβ clusters according to CDR3 sequence reveals the presence of shared expanded clonotypes among patients with high disease activity, and demonstrates that many of these sequences are highly similar. Each node represents one CDR3 sequence, with edges connecting nodes assigned to the same specificity group, and larger specificity groups are densely connected together. Larger nodes correspond to higher-frequency clones. (E) Network visualisation of RA-associated TCRβ CDR3 clusters annotated against a combined database of TCRβ sequences of known antigen specificity, indicates possible associations with SARS-CoV-2 antigens as well as other pathogens. (F) Logo plots showing consensus TCRβ CDR3 motifs detected in RA synovial tissue and with known antigen specificity in three commonly encountered pathogens, labeled in order of pathogen, protein and peptide. (G) Logo plot as in (F) showing consensus TCRβ CDR3 motifs detected in RA synovial tissue but matching SARS-CoV-2 antigens, a pathogen not present at the time the sequencing studies were conducted. (H–I) Network analysis of TCR sequences recovered from focused TCRseq of 4 patients with RA (sorted into CD8+CD57+ and CD8+CD57- populations) clustered against the TCRs obtained from repertoire reconstruction of the RA synovial samples in the PEAC (H) and GSE89408 (I) cohorts. An abundance of clustered TCRs can be observed, indicating that sequences recovered from repertoire reconstruction may also be shared. (J) Network analysis of clustering primary TCRseq sequences (as in H–I) with annotated antigen-specific sequences. As in (E), associations between a number of annotated SARS-CoV-2 sequences and RA synovial sequences can be observed. *p-value <0.05. DAS28, Disease Activity Score 28; ESR, erythrocyte sedimentation rate.

and potentially contribute to the observed clinical associations between arthritis/autoimmunity and SARS-CoV-2 infection.[8] While the existence of antigen-specific memory clones in unexposed controls has been previously regarded as controversial, recent studies in the context of SARS-CoV-2 have confirmed that pre-existing antigen-specific CD4+memory clones can be found in uninfected patients,[9] and were intriguingly found to correspond with severe outcomes. Furthermore, these repertoires may be reshaped following vaccination or infection.[10] As such, we speculate that vaccination against SARS-CoV-2 may also reshape CD8+T cell clonal dynamics in the synovium and peripheral circulation of patients with RA and help to dampen the feedback between SARS-CoV-2 infection and autoimmunity. Additional exploration of possible pathogen antigen cross-reactivity in RA is needed to clarify the role of SARS-CoV-2 and other infections diseases in influencing disease progression.

Zihan Zheng ,[1,2] **Ling Chang,**[1] **Jie Mu,**[3] **Qingshan Ni,**[3] **Zhong Bing,**[4] **Qing-Hua Zou,**[4] **Ying Wan,**[3] **Yuzhang Wu,**[1] **Jingyi Li,**[4] **Liyun Zou**[1]

[1]Institute of Immunology, Army Medical University, Chongqing, Chongqing, China
[2]Department of Autoimmune Diseases, Chongqing International Institute for Immunology, Chongqing, Chongqing, China
[3]Biomedical Analysis Center, Army Medical University, Chongqing, Chongqing, China
[4]Department of Rheumatology and Immunology, First Affiliated Hospital (Southwest Hospital) of Army Medical University, Chongqing, Chongqing, China

Correspondence to Professor Liyun Zou, Institute of Immunology, Army Medical University, Chongqing, People's Republic of China; zouliyun_2012@163.comDr Jingyi Li; lijingyi_tmmu@foxmail.comProfessor Yuzhang Wu; wuyuzhang@iiicq.vipProfessor Ying Wan; wanying516@foxmail.com

Correction notice This article has been corrected since it published Online First. The funding statement has been amended.

Handling editor Josef S Smolen

Contributors ZZ, LC, JL and LZ designed and coordinated this study. ZZ, LC, JM and QN performed the primary bioinformatics analyses and built the online database. ZB and Q-HZ, performed supporting analyses. LZ, YWa and YWu supervised the analyses. ZZ, YWa, YWu, JL and LZ wrote the manuscript. All authors were involved in critical review, editing, revision and approval of the final manuscript.

Funding This work was funded by grants to LZ (National Natural Science Foundation of China No. 81971546 and No. 82171787, and the Chongqing International Institute for Immunology No. 2020YJC06), JL (National Natural Science Foundation of China No. 81971537), and WYu (Emergency Key Program of Guangzhou Laboratory No. EKPG21-30-3).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Ethics Committee of the First Affiliated Hospital (Southwest Hospital) of Army Medical University as KY202213. Patients gave informed consent prior to participation.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/ard-2022-222763).

ZZ and LC contributed equally.

ZZ and LC are joint first authors.

YW, YW, JL and LZ are joint senior authors.

To cite Zheng Z, Chang L, Mu J, *et al*. *Ann Rheum Dis* 2023;**82**:438–440.

Received 6 May 2022
Accepted 8 October 2022
Published Online First 19 October 2022

Ann Rheum Dis 2023;**82**:438–440. doi:10.1136/annrheumdis-2022-222763

ORCID iD
Zihan Zheng http://orcid.org/0000-0003-1486-9586

REFERENCES

1 Smolen JS, Aletaha D, Barton A, *et al*. Rheumatoid arthritis. *Nat Rev Dis Primers* 2018;4:18001.
2 Weyand CM, Goronzy JJ. The immunology of rheumatoid arthritis. *Nat Immunol* 2021;22:10–18.
3 Liu X, Zhang W, Zhao M, *et al*. T cell receptor β repertoires as novel diagnostic markers for systemic lupus erythematosus and rheumatoid arthritis. *Ann Rheum Dis* 2019;78:1070–8.
4 Lliso-Ribera G, Humby F, Lewis M, *et al*. Synovial tissue signatures enhance clinical classification and prognostic/treatment response algorithms in early inflammatory arthritis and predict requirement for subsequent biological therapy: results from the pathobiology of early arthritis cohort (PEAC). *Ann Rheum Dis* 2019;78:1642–52.
5 Guo Y, Walsh AM, Fearon U, *et al*. CD40L-dependent pathway is active at various stages of rheumatoid arthritis disease progression. *J Immunol* 2017;198:4490–501.
6 Davignon J-L, Combe B, Cantagrel A. Cytomegalovirus infection: friend or foe in rheumatoid arthritis? *Arthritis Res Ther* 2021;23:16.
7 Savola P, Kelkka T, Rajala HL, *et al*. Somatic mutations in clonally expanded cytotoxic T lymphocytes in patients with newly diagnosed rheumatoid arthritis. *Nat Commun* 2017;8:15869.
8 Redwan EM, Alghamdi MF, El-Aziz TMA, *et al*. The mechanism behind flaring/triggering of autoimmunity disorders associated with COVID-19. *Autoimmun Rev* 2021;20:102909.
9 Bacher P, Rosati E, Esser D, *et al*. Low-avidity $CD4^+$ T Cell responses to SARS-CoV-2 in unexposed individuals and humans with severe COVID-19. *Immunity* 2020;53:1258–71.
10 Minervina AA, Pogorelyy MV, Kirk AM, *et al*. SARS-CoV-2 antigen exposure history shapes phenotypes and specificity of memory $CD8^+$ T cells. *Nat Immunol* 2022;23:781–90.