OPEN ACCESS

TRANSLATIONAL SCIENCE

Rheumatoid arthritis, systemic lupus erythematosus and primary Sjögren's syndrome shared megakaryocyte expansion in peripheral blood

Yukai Wang ,[1] Xuezhen Xie,[1] Chengpeng Zhang,[2] Miaotong Su,[2] Sini Gao,[2] Jing Wang,[2] Changhao Lu,[2] Qisheng Lin,[1] Jianqun Lin,[1] Marco Matucci-Cerinic ,[3] Daniel E Furst,[4] Guohong Zhang [2]

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2021-220066).

[1]Department of Rheumatology and Immunology, Shantou Central Hospital, Shantou, China
[2]Department of Pathology, Provincial Key Laboratory of Infectious Diseases and Molecular Immunopathology, Shantou University Medical College, Shantou, China
[3]Department of Internal Medicine, University of Florence, Firenze, Italy
[4]Rheumatology, University of California Los Angeles, Los Angeles, California, USA

Correspondence to
Dr Guohong Zhang, Department of Pathology, Provincial Key Laboratory of Infectious Diseases and Molecular Immunopathology, Shantou University Medical College, Shantou, China; g_ghzhang@stu.edu.cn

YW and XX contributed equally.

Received 1 February 2021
Accepted 15 August 2021
Published Online First 30 August 2021

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Wang Y, Xie X, Zhang C, *et al*. *Ann Rheum Dis* 2022;**81**:379–385.

ABSTRACT

Objectives Rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and primary Sjögren's syndrome (pSS) share many clinical manifestations and serological features. The aim of this study was to identify the common transcriptional profiling and composition of immune cells in peripheral blood in these autoimmune diseases (ADs).

Methods We analysed bulk RNA-seq data for enrichment of biological processes, transcription factors (TFs) and deconvolution-based immune cell types from peripheral blood mononuclear cells (PBMCs) in 119 treatment-naive patients (41 RA, 38 pSS, 28 SLE and 12 polyautoimmunity) and 20 healthy controls. The single-cell RNA-seq (scRNA-seq) and flow cytometry had been performed to further define the immune cell subsets on PBMCs.

Results Similar transcriptional profiles and common gene expression signatures associated with nucleosome assembly and haemostasis were identified across RA, SLE, pSS and polyautoimmunity. Distinct TF ensembles and gene regulatory network were mainly enriched in haematopoiesis. The upregulated cell-lineage-specific TFs *PBX1*, *GATA1*, *TAL1* and *GFI1B* demonstrated a strong gene expression signature of megakaryocyte (MK) expansion. Gene expression-based cell type enrichment revealed elevated MK composition, specifically, $CD41b^{+}CD42b^{+}$ and $CD41b^{+}CD61^{+}$ MKs were expanded, further confirmed by flow cytometry in these ADs. In scRNA-seq data, MKs were defined by TFs *PBX1/GATA1/TAL1* and pre-T-cell antigen receptor gene, *PTCRA*. Cellular heterogeneity and a distinct immune subpopulation with functional enrichment of antigen presentation were observed in MKs.

Conclusions The identification of MK expansion provided new insights into the peripheral immune cell atlas across RA, SLE, pSS and polyautoimmunity. Aberrant regulation of the MK expansion might contribute to the pathogenesis of these ADs.

Key messages

What is already known about this subject?

► Rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and primary Sjögren's syndrome (pSS) share many clinical and serological features. The peripheral blood mononuclear cells (PBMCs) are the common origin for immune cells infiltrating specific targeted organs in these autoimmune diseases (ADs). However, the initial immune cells regulated by core transcription factors (TFs) in the PBMC remain unknown.

What does this study add?

► This study uncovers common gene expression signatures in platelet activation, consisting of increased megakaryocyte (MK) composition with upregulated expression of cell-lineage specific TFs *PBX1*, *GATA1*, *TAL1* and *GFI1B* in PBMC across RA, SLE, pSS and polyautoimmunity. In peripheral blood, MKs are a heterogeneous cell population that comprises a subpopulation with distinct immune characteristics in these ADs. Speculatively, this subpopulation of immunologically active MKs might be an initial stimulus for T-cell initiation of these diseases.

How might this impact on clinical practice?

► Our results elucidate MK expansion for the atlas of peripheral immune cells across RA, pSS and SLE and support the hypothesis that regulatory events in MK expansion might act as pivotal components, which could be an entry point toward developing targeted treatment for patients with these ADs.

INTRODUCTION

Rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and primary Sjögren's syndrome (pSS) are common autoimmune disease (ADs) in women, which preferentially affect specific target organs. Indeed, these ADs share several clinical manifestations, serological profiles and immunological characteristics. Furthermore, the co-occurrence of these ADs within a single patient (polyautoimmunity) and within members of a nuclear family (familial autoimmunity) indicate that they have common aetiological components, including genetic and epigenetic factors and sex hormones. The genetic variants in the T-cell receptor (TCR) pathway and *TNFAIP3*[1 2] and DNA methylation signatures had been previously uncovered across RA, SLE and pSS.[3] However, it is still unknown which antigen initially primes autoimmune T cells.

Beyond affected organs, peripheral blood represents the main highway for the immune system for RA, SLE and pSS. Peripheral blood mononuclear cells (PBMCs) in this context are the immune cells which initiate the autoimmune inflammatory process directed against target organs. Therefore, the gene expression signatures of PBMC could shed light on the molecular features of the immune cells in the targeted organs in these ADs. Shared type I interferon (IFN)-stimulated genes were identified via meta-analysis of PBMC transcriptomes across RA, SLE and pSS.[4] However, PBMC comprises several cell types and each cellular subtype expresses a unique set of genes; thus, cell-specific signatures may further define immune cell composition for AD pathogenesis. On the other hand, underlying immune responses are the developmental trajectories that determine immune cells' fates.[5] The transcription factor (TF) network controlling cell lineage commitment in the bone marrow could determine the landscape for immune cell expansion in the peripheral blood in pathogenesis of these ADs.

Herein, we combined bulk RNA-seq and single-cell RNA-seq (scRNA-seq) of gene expression signatures, immune cell subsets and TF networks to identify potential common mechanisms in the immunopathogenesis of SLE, RA and pSS.

Patients and methods

Subjects

PBMCs were obtained from 42 RA, 41 pSS, 28 SLE and 12 polyautoimmunity patients and 21 gender-matched healthy controls (online supplemental table S1). Patients met the 2002 American-European Consensus Group for pSS, the 2012 Systemic Lupus Collaborating Clinics for SLE and the 2010 ACR/EULAR for RA, respectively. Polyautoimmunity was defined as patients with two ADs, RA/pSS or SLE/pSS.[6] Fully informed consent was obtained from all participants prior to sample collection. For more details about the study design, experimental and bioinformatic methods, see online supplemental methods (online supplemental figure S1).

RESULTS

RA, pSS and SLE shared common gene expression signatures enriched in coagulation and nucleosome assembly

Using bulk RNA-seq data, we initially assessed clustering of RA, SLE, pSS and polyautoimmunity by principal component analysis that demonstrated that these ADs were similar (figure 1A). Compared with healthy controls, these ADs had similar transcriptional profiles, sharing differentially expressed genes ((DEGs) figure 1B–C and online supplemental figures S2–6 and table S2), including type I IFN-stimulated gene *IFI27* plus chemokine receptors *CXCR1* and *CXCR2*.[4] Indeed, 446 common upregulated and 165 downregulated genes overlapped across these ADs (figure 1D and online supplemental table S2). Among the upregulated genes, the major biological processes that were enriched were related to nucleosome assembly and coagulation cascades (figure 1E). The most impacted pathway was the 'SLE' pathway (figure 1F). To further illustrate this point, haemostasis and megakaryocyte (MK) differentiation had been identified via gene ontology (GO) term networks (online supplemental figure S6A). Protein–protein interactions were demonstrated among histone genes H2A and H2B, including *H2AC11*, *H2AC13*, *H2BC11* and *H2BC12*, which were consistent with the GO term of nucleosome assembly (online supplemental figure S6B). Gene set enrichment analysis showed significant enrichment of platelet activation in these ADs as well (online supplemental figure S6C). Collectively, transcriptional profiling suggested potential regulation of MK/platelet-related processes emerged as the gene expression signatures in these ADs.

Common TFs highlighted MK expansion responding to the gene expression signatures

We next sought to identify TFs linked with gene enrichment involved in biological processes in ADs. Transcriptional regulatory networks indicated *GATA1* as the top-ranked regulator by enrichment analysis of upregulated genes (figure 2A). Only 17 common upregulated and 8 downregulated TFs were identified (figure 2B and online supplemental figure S6G) and were mainly enriched in embryonic haematopoiesis and granulocyte differentiation (figure 2C–D). We further determined correlations among TFs, reasoning that the distinct TF ensembles could be correlated with expression. The correlated expression pattern was comprised of: *GRHL1, MEIS1, THRB, PBX1, GATA1, TAL1, GFI1B* and *E2F1* (figure 2E). Focusing on TF function, estradiol promotes haematopoietic stem cell (HSC) division by enrichment of cell cycle genes, harbouring a binding motif for the TF *E2F1*.[7] In addition, oestrogen receptor (ER) interacted with *MEIS1*, *THRB* and *GRHL1*;[8] consequently, interaction of *MEIS1* and *PBX1* acts upstream of *GATA1* to regulate primitive haematopoiesis and induce lineage commitment toward a MK-erythroid progenitor cell.[9,10] Thus, *GRHL1, MEIS1, THRB* and *PBX1* formed a compound linking oestrogen to haematopoietic and MK-erythroid commitment. We mapped TFs *GATA1*, *TAL1* and *GFI1B* to their source immune cell lineage according to the order of expression of haematopoietic transcriptional networks[11] and identified a strong gene expression signature in MK expansion (figure 2F). The upregulated expression of *MEIS1, PBX1*, *GATA1, TAL1* and *GFI1B* in ADs was validated by real-time quantitative PCR (online supplemental figure S7A).

We also observed downregulated TFs, including *EGR1*, *EGR2*, *EGR3* and *CEBPE*. Egr2 and Egr3 have long been regarded as negative regulators of T-cell activation.[12] *CEBPE* is expressed in a stage-specific manner during myeloid differentiation and is an essential TF for granulocytic differentiation.[13] Therefore, the TF network highlighted MK expansion responding to the gene expression signatures.

Immune cell composition further supported the MK expansion

To dissect the MK in the PBMC, reasoning that the gene signature of MK was enriched in the bulk RNA-seq data, we integrated three central algorithms of deconvolution: xCell, CIBERSORT and ABIS (figure 2G and online supplemental figure S7B). The xCell results demonstrated that MKs and erythrocytes were positively enriched ($p<0.001$), while neutrophils, eosinophils and basophils were negatively enriched in ADs versus healthy individuals ($p<0.001$). Consistent with the results of xCell, the ABIS results identified decreased absolute deconvolution values of low-density neutrophil and basophil in ADs ($p<0.001$). Furthermore, the CIBERSORT results confirmed decreased neutrophils in ADs ($p<0.001$, figure 2G). We identified well-known MK marker genes, including *PPBP*, *PF4*, *GNG11* and *GP9* (*CD42b*), which were upregulated in ADs (online supplemental figure S7C). Thus, in accordance with gene expression signatures and TF regulation networks, the composition of immune cells demonstrated MK expansion in PBMC from ADs. To validate cellular composition, we analysed the MKs by flow cytometry. The percentage of $CD41b^{+}CD42b^{+}$ and $CD41b^{+}CD61^{+}$ MKs was significantly elevated in ADs compared with healthy controls (figure 2H and online supplemental figure S8A). Thus,

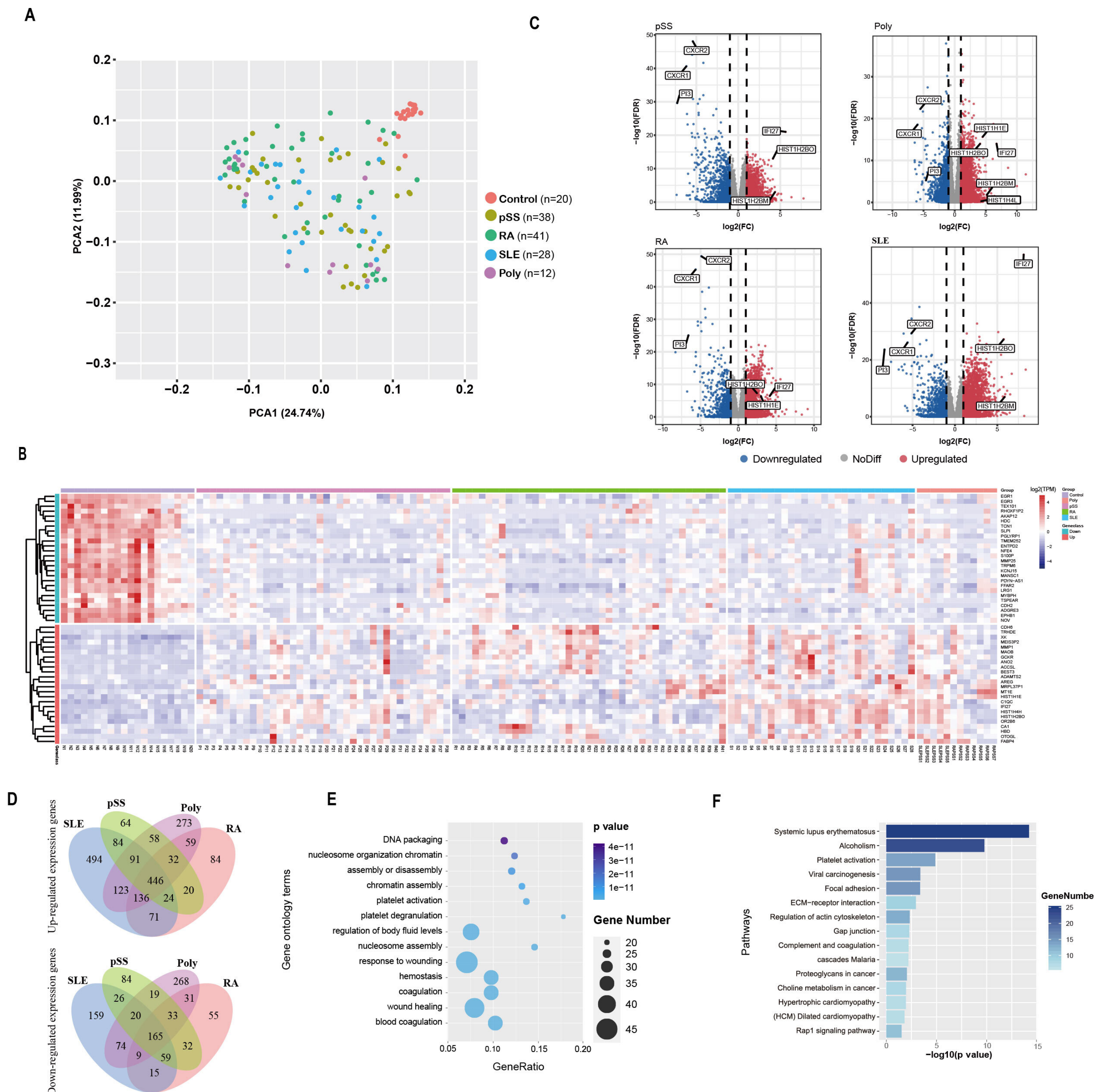

Figure 1 Shared transcriptional profiling and platelet activation across RA, SLE and pSS. (A) Principal component analysis (PCA) of gene expression profiles for PBMCs from RA, SLE, pSS and polyautoimmunity, indicating absence of a clear differentiation among these ADs. Each point is assigned a location to illustrate potential clusters of neighbouring samples, which contain similar gene expression patterns. (B) Heatmap illustrating the top differentially expressed genes (DEGs) across RA, SLE, pSS and polyautoimmunity. (C) Volcano plots showing DEGs across RA, SLE, pSS and polyautoimmunity, in which some representative genes were highlighted. (D) Venn diagram showing 446 upregulated (top panel) and 165 downregulated genes (bottom panel) in common across RA, SLE, pSS and polyautoimmunity. (E) Gene ontology term enrichment of biological processes for common 446 upregulated genes showing nucleosome assembly and platelet degranulation. (F) Kyoto Encyclopedia of Genes and Genomes pathway enrichment highlighted 'systemic lupus erythematosus' and platelet activation pathways. Controls denote healthy controls. ADs, autoimmune diseases; FC, fold change; poly, polyautoimmunity; PBMCs, peripheral blood mononuclear cells; pSS, primary Sjögren's syndrome; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus.

these findings were consistent with our original gene expression signatures indicating the MK expansion.

MK and cellular heterogeneity identified by scRNA-seq

To map the MKs in the immune cell population of PBMC, we initially combined 57486 individual cells from pSS (n=3), SLE (n=3, datasets obtained from Mistry *et al* 2019; GSE139360),[14] patients with RA (n=1) and a healthy control (n=1). MKs were identified by type-specific markers of *PPBP*, *PF4* and *PTCRA* (figure 3A–C) and TFs, *PBX1, MEIS1, GATA1* and *TAL1* (online supplemental figure S8B-E and table S3). GO analysis of MKs further indicated enrichment of upregulated genes related to

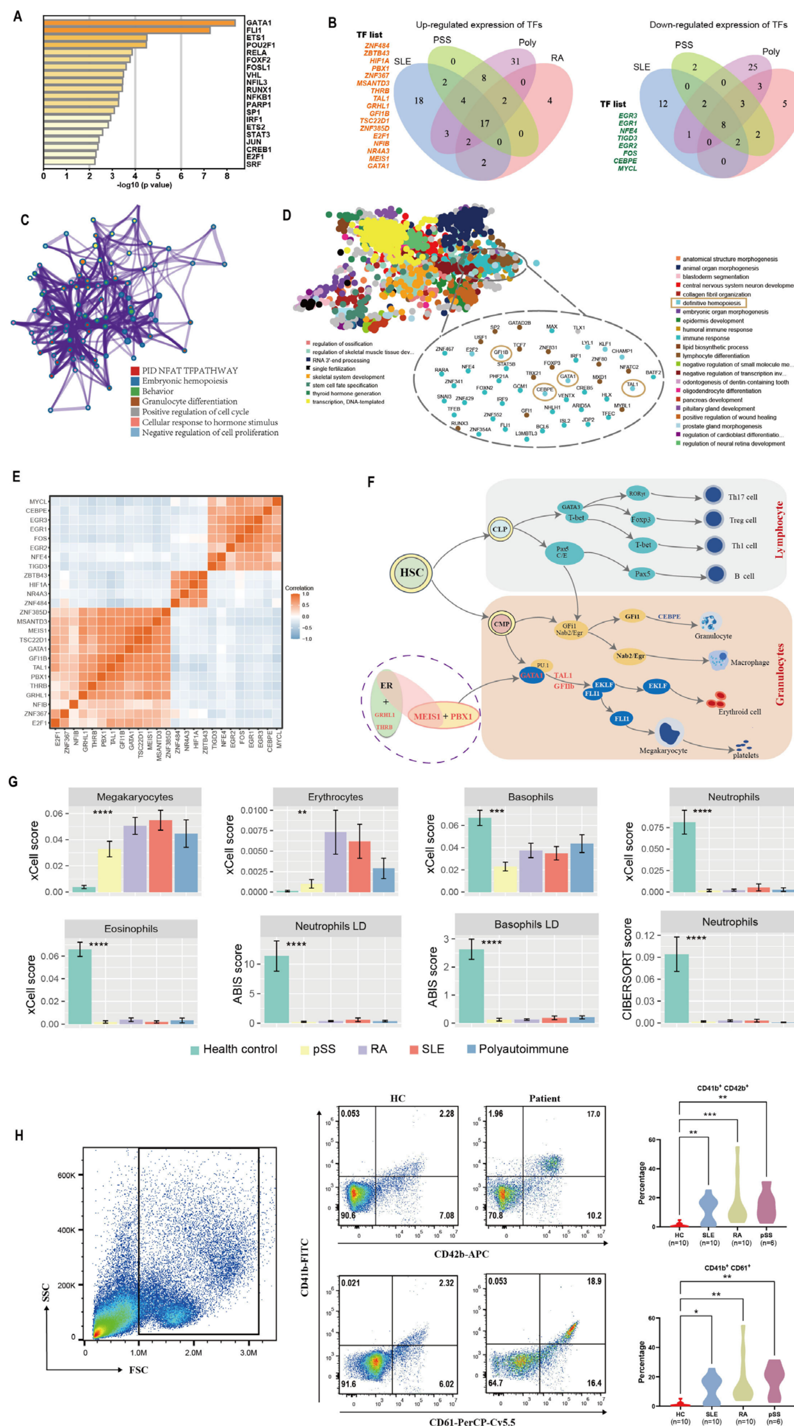

Figure 2 Core transcription factors presented the megakaryocyte (MK) expansion. (A) Enrichment of regulator by Transcriptional Regulatory Relationships Unraveled by Sentence-based Text mining (TRRUST) showing *GATA1* as the top-ranking TF. (B) Venn diagram showing 25 common TFs across RA, SLE, pSS and polyautoimmunity. Left panel, upregulated expression of TFs, right panel, downregulated expression of TFs. (C) Gene ontology term for 25 common TFs significantly enriched in biological process of haematopoiesis. (D) TF enrichment had been performed by ChIP-X Enrichment Analysis 3 (ChEA3), which offers associations among involved TFs. TFs that are covered by the ChEA3 database, including *GATA1*, *TAL1*, *GFI1B* and *CEBPE*, are significantly related to definitive haematopoiesis. (E) TF correlation heatmap generated by the upregulated coexpression of TFs. Red colour indicates correlation. (F) TFs defining, showing MK expansion. Oestrogen interacted with *MEIS1*, *THRB* and *GRHL1* and *MEIS1* and *PBX1* act upstream of *GATA1* to regulate primitive haematopoiesis with *TAL1* and *GFI1B* to determine MK lineage. TF in red means upregulated expression, while in blue means downregulated expression. (G) Immune cell composition generated by xCell-inferred, ABIS-inferred and CIBERSORT-inferred enrichment score of cell types across ADs and healthy controls. *p<0.05; **p<0.01; ***p<0.001 by Kruskal-Wallis test. (H) Flow cytometry and its quantification of MKs from PBMC. Representative fluorescence-activated cell sorting plots for the identification of MKs. After gating for MKs by forward versus side scatter (FSC vs SSC), MKs were characterised as $CD41b^{+}CD42b^{+}$ and $CD41b^{+}CD61^{+}$. ***p<0.001 by Mann-Whitney U test. ADs, autoimmune diseases; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; ER, oestrogen receptor; HC, healthy controls; HSCs, haematopoietic stem cells; LD, low-density; PBMC, peripheral blood mononuclear cell; poly: polyautoimmunity; pSS, primary Sjögren's syndrome; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; TF, transcription factor.

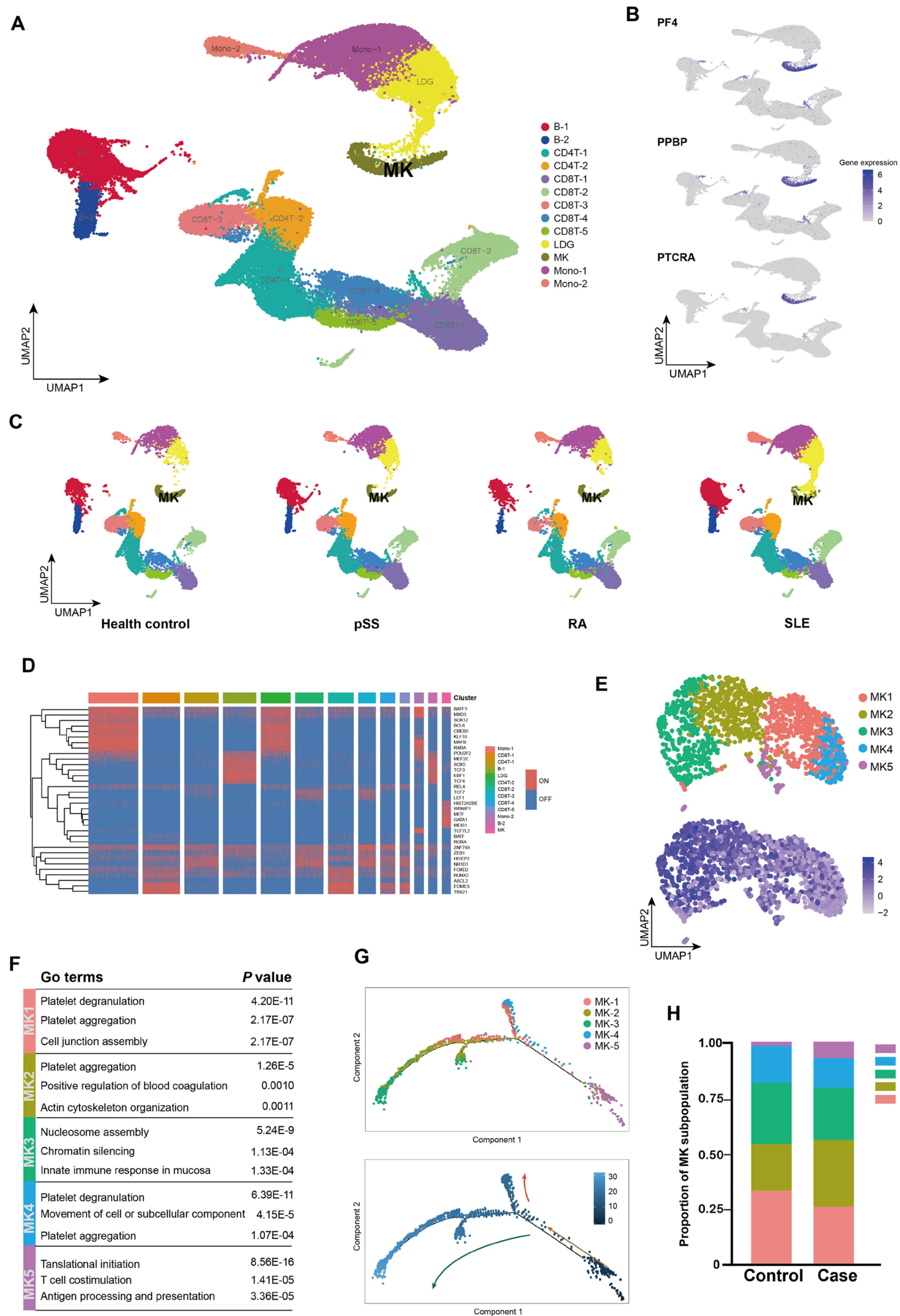

Figure 3 Cell type confirmed the megakaryocyte (MK) and subset with high expression of *PTCRA*. (A) Uniform manifold approximation and projection (UMAP) embedding of the entire dataset coloured by orthogonally generated clusters labelled by cell type annotation. Thirteen putative cell clusters were identified from all profiled samples (n=57 486 cells). (B) The expression of cell-lineage marker genes for MK, including *PF4*, *PPBP* and *PTCRA*. (C) UMAP embedding split by ADs and healthy control highlighted the MK cluster. (D) Heatmap showing the activity of top five cell type-specific transcription factors (rows) in each cell type (columns) as identified by single-cell regulatory network inference and clustering. Term of ON indicates activity exceeds a regulon-specific area under the curve threshold. ON, active; OFF, inactive. (E) Five putative MK subpopulations were identified and subpopulations with high expression of *PTCRA* were highlighted (bottom). (F) The representative terms of gene ontology (GO) (biological processes) term enriched in each MK subpopulation. (G) Monocle pseudotime trajectory analysis of MKs, indicating two developmental directions. MK subpopulations along the branching trajectories (top). Inferred pseudotime for each cell is shown (bottom). (H) Stacked bar charts showing the percentage of each MK subpopulation among total MKs in ADs and healthy control. ADs, autoimmune diseases; B, B cell; CD4T, $CD4^+$ T cell; CD8T, $CD8^+$ T cell; mono, monocyte; pSS, primary Sjögren's syndrome; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus.

platelet degranulation and aggregation (online supplemental figure S8F). We constructed a gene regulatory network among the TFs predicted by single-cell regulatory network inference and clustering and revealed that *MEIS1* and *GATA1* were active in the MK compared with other immune cell types (figure 3D and online supplemental figure S8G). GO enrichment in biological processes of DEGs of MKs in ADs compared with healthy controls revealed that upregulated genes were associated with translation and translational initiation, while downregulated genes were associated with platelet degranulation and aggregation (online supplemental figure S9A).

To dissect MK heterogeneity, five putative subpopulations (MK1–MK5) of MKs were identified by subclustering (figure 3E–F and online supplemental table S4). We then characterised the gene sets enriched in these MK subpopulations. MK1, MK2 and MK4 mainly showed enrichment of GO terms related to platelet degranulation and aggregation, whereas MK3 exhibited enriched GO terms affecting nucleosome assembly. MK5 highly expressed genes were associated with translational initiation and antigen processing and presentation. The GO term 'translational initiation' was perhaps indicative of a less mature MK population.[15] Cellular trajectory analysis revealed distinct differentiation trajectories underpinning MK heterogeneity with a major bifurcation and highlighted MK5 at the origin of the trajectory (figure 3G and online supplemental figure S9B-E). Furthermore, the proportion of the MK5 subpopulation was increased in ADs, compared with healthy control (figure 3H). Thus, the subpopulation proportion of MKs coincided with the GO terms of DEGs in ADs.

Given the immune characteristics of MKs, we characterised cell communication by ligand–receptor interactions between MK and other immune cell subsets. We identified ligand–receptor interactions between MKs and $CD4^+$ and $CD8^+$ T cells such as a PF4–CXCR3 pair (online supplemental figure S9F). We next focused on the granulocytes; low-density granulocytes were identified by genes highly specific for neutrophils, including *FCGR3B* and *CMTM2*. Biological processes of GO enrichment suggested that upregulated genes in ADs were associated with viral transcription and cellular response to type I interferon (online supplemental figure S10). We also used peripheral TCR repertoire sequencing to find clonotype in RA and healthy control. TCR gene rearrangement and variable gene usage are presented in online supplemental figure S11.

DISCUSSION

Using transcriptomic profiling, we demonstrated common gene expression signatures relating to haemostasis via the regulation of transcriptomes by the TF network, *PBX1*/*GATA1*/*TAL1*/*GFI1B*. This provides novel evidence of MKs expansion in PBMC in treatment-naive RA, SLE and pSS. It is in this context that we examined patients with polyautoimmunity, seeking to elucidate gene expression signatures and TFs common across ADs; through these patients we feel we have supported such an AD-associated pathway.

In this study, bulk RNA-seq was used to find the gene expression signature, TFs and composition of immune cells, while the scRNA-seq was used to identify the cell type of interest in the PBMC. Bulk RNA-seq TF and composition studies confirmed the presence of MKs expansion in these ADs. scRNA-seq data further defined the MK subpopulations. We observed similar transcriptional profiles linking RA, SLE, pSS and polyautoimmunity and noted across-disease upregulated expression of the type I IFN-stimulated gene *IFI27*, as well as downregulated expression of chemokine receptors, *CXCR1* and *CXCR2*. These genes play significant roles in the suppression of megakaryopoiesis.[16] Importantly, the gene expression signatures enriched in haemostasis might explain common immunological characteristics via regulation with transcriptomic reprogramming.

It is noteworthy that dysregulated transcriptomic reprogramming might introduce disturbances in immune homeostasis leading to ADs.[17] Transcriptomic data of bone marrow (BM)-derived haematopoietic stem and progenitor cells from SLE mice showed myeloid skewing, with granulocytic differentiation arrest and a positive correlation with platelet degranulation that indicated expansion of stem cell-like MK-committed cells.[18] In accordance with BM observations, we further detected upregulated expression of MK-lineage TFs *PBX1, GATA1*, *TAL1* and *GFI1B* and downregulated expression of granulocytic-lineage TF *CEBPE*.

Furthermore, via deconvolution of bulk RNA-seq data, we identified increased MKs, accompanied by decreased neutrophils, eosinophils and basophils in these ADs. Thus, we have expanded the previously documented MK expansion in the BM to peripheral blood in ADs. Conventional antigen-presenting cells (APCs) are essential for AD progression, but it is unknown what initially primes autoimmune T cells. MKs express MHC I and II molecules, thus acting as professional APC that enhance Th17 and Th1/Th17 responds to lupus autoantigens.[19,20] We reasoned, preliminary, that elevated levels of MKs enhance their intrinsic antigen presenting function in peripheral blood across RA, SLE and pSS.

In scRNA-seq, we identified MK across patients with pSS, SLE and RA. MK expansion had been previously observed[21,22] and was a critical peripheral source of cytokine storms in COVID-19.[23] We described MKs with highly expressed *PTCRA*, encoding the pre-TCRα chain (pTα). Normally, pTα along with TCRβ and CD3 form the pre-TCR, which are exclusively expressed in immature thymocytes during early T-cell development.[24] *PTCRA* (pTα) is also required for TCR rearrangement for extrathymic T-cell development.[25] Cell surface $PTCRA^+$ MKs had been identified in early human embryonic yolk sacs.[15] Furthermore, a less mature immune MK subpopulation had been found to be enriched in ADs. This subpopulation of MKs presenting immune characteristics with antigen processing and presentation had been previously demonstrated in yolk sac and fetal liver cells.[15] Therefore, we speculated that MKs act as specific endogenous APC, resulting in abnormal TCR arrangements which, in turn, trigger the initial autoimmune T cell for AD pathogenesis.

We also speculate that there might be a connection between sex hormones and megakaryocytopoieses. A predominant role of sex hormones has been suggested as the main cause of sex-biased ADs.[26] Oestrogen stimulates HSC self-renewal, megakaryocytopoiesis and erythropoiesis in females.[7,27] Megakaryopoiesis is dynamic and adaptive to biological needs, termed as 'emergency haematopoiesis' that biases toward the MK lineage.[28] *MEIS1* interacts with ER[8] and *PBX* acts upstream of *GATA1* to regulate primitive haematopoiesis.[9] Oestrogen promotes MK polyploidisation via ERβ-mediated transcription of *GATA1*.[29] Therefore, an upregulated TF network *MEIS1*/*PBX1*/*GATA1*/*TAL1*/*GFI1B* might connect estrogens and MK expansion in RA, SLE and pSS.

To summarise, we have presented evidence for peripheral MK expansion across RA, SLE and pSS. Our discovery provides clues that MK expansion might initially prime autoimmune T cells in the pathogenesis of these ADs.

Correction notice This article has been corrected since it published Online First. Figure 3 has been amended.

Acknowledgements The authors thank the patients for participation in the study.

Contributors GZ and YW designed the study. YW, XX, CZ, MS, JW, SG and CL performed experiments and analysed data. GZ, DEF and MM-C wrote the manuscript. JL and QL helped with sample collection.

Funding This work was supported by the grant from the Natural Science Foundation of Guangdong Province (grant number: 2014A030307003).

Competing interests MM-C is an editorial board member of the *Annals of the Rheumatic Diseases*. The other authors declare no competing interests.

Patient consent for publication Not required.

Ethics approval This study was approved by the Ethics Committee of Shantou Central Hospital. Each participant signed an informed consent before data collection.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All sequencing data have been deposited in public, open access repository of the Genome Sequence Archive for Human (http://bigd.big.ac.cn/gsa-human/) with the accession ID. HRA000911 and HRA000916.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Yukai Wang http://orcid.org/0000-0003-2468-3208
Marco Matucci-Cerinic http://orcid.org/0000-0002-9324-3161
Guohong Zhang http://orcid.org/0000-0002-3856-3111

REFERENCES

1. Wang Y, Chen S, Chen J, *et al*. Germline genetic patterns underlying familial rheumatoid arthritis, systemic lupus erythematosus and primary Sjögren's syndrome highlight T cell-initiated autoimmunity. *Ann Rheum Dis* 2020;79:268–75.
2. Ciccacci C, Latini A, Perricone C, *et al*. TNFAIP3 gene polymorphisms in three common autoimmune diseases: systemic lupus erythematosus, rheumatoid arthritis, and primary Sjogren Syndrome—Association with disease susceptibility and clinical phenotypes in Italian patients. *J Immunol Res* 2019;2019:1–6.
3. Wang X, Lei D, Ding J, *et al*. A DNA-methylated sight on autoimmune inflammation network across RA, pSS, and SLE. *J Immunol Res* 2018;2018:1–13.
4. Toro-Domínguez D, Carmona-Sáez P, Alarcón-Riquelme ME. Shared signatures between rheumatoid arthritis, systemic lupus erythematosus and Sjögren's syndrome uncovered through gene expression meta-analysis. *Arthritis Res Ther* 2014;16:489.
5. Rothenberg EV. Lineage determination in the immune system. *Immunol Rev* 2010;238:5–11.
6. Anaya J-M. The autoimmune tautology. *Arthritis Res Ther* 2010;12:147.
7. Nakada D, Oguro H, Levi BP, *et al*. Oestrogen increases haematopoietic stem-cell self-renewal in females and during pregnancy. *Nature* 2014;505:555–8.
8. Zheng Y, Shao X, Huang Y, *et al*. Role of estrogen receptor in breast cancer cell gene expression. *Mol Med Rep* 2016;13:4046–50.
9. Pillay LM, Forrester AM, Erickson T, *et al*. The Hox cofactors MEIS1 and PBX act upstream of GATA1 to regulate primitive hematopoiesis. *Dev Biol* 2010;340:306–17.
10. Zeddies S, Jansen SBG, di Summa F, *et al*. MEIS1 regulates early erythroid and megakaryocytic cell fate. *Haematologica* 2014;99:1555–64.
11. Zhu J, Emerson SG. Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene* 2002;21:3295–313.
12. Morita K, Okamura T, Inoue M, *et al*. Egr2 and Egr3 in regulatory T cells cooperatively control systemic autoimmunity through Ltbp3-mediated TGF-β3 production. *Proc Natl Acad Sci U S A* 2016;113:E8131–40.
13. Shyamsunder P, Shanmugasundaram M, Mayakonda A, *et al*. Identification of a novel enhancer of CEBPE essential for granulocytic differentiation. *Blood* 2019;133:2507–17.
14. Mistry P, Nakabo S, O'Neil L, *et al*. Transcriptomic, epigenetic, and functional analyses implicate neutrophil diversity in the pathogenesis of systemic lupus erythematosus. *Proc Natl Acad Sci U S A* 2019;116:25222–8.
15. Wang H, He J, Xu C, *et al*. Decoding human megakaryocyte development. *Cell Stem Cell* 2021;28:535–49.
16. Adeli EK, Abolghasemi H, Ebtekar M, *et al*. Effects of CXCR1 and CXCR2 inhibition on expansion and differentiation of umbilical cord blood CD133+ cells into megakaryocyte progenitor cells. *Cytokine* 2011;55:181–7.
17. Wardowska A. The epigenetic face of lupus: focus on antigen-presenting cells. *Int Immunopharmacol* 2020;81:106262.
18. Grigoriou M, Banos A, Filia A, *et al*. Transcriptome reprogramming and myeloid skewing in haematopoietic stem and progenitor cells in systemic lupus erythematosus. *Ann Rheum Dis* 2020;79:242–53.
19. Finkielsztein A, Schlinker AC, Zhang L, *et al*. Human megakaryocyte progenitors derived from hematopoietic stem cells of normal individuals are MHC class II-expressing professional APC that enhance Th17 and Th1/Th17 responses. *Immunol Lett* 2015;163:84–95.
20. Kang H-K, Chiang M-Y, Ecklund D, *et al*. Megakaryocyte progenitors are the main APCS inducing Th17 response to lupus autoantigens and foreign antigens. *J Immunol* 2012;188:5970–80.
21. Stephenson E, Reynolds G, Botting RA, *et al*. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat Med* 2021;27:904–16.
22. Bernardes JP, Mishra N, Tran F, *et al*. Longitudinal multi-omics analyses identify responses of megakaryocytes, erythroid cells, and Plasmablasts as hallmarks of severe COVID-19. *Immunity* 2020;53:1296–314.
23. Ren X, Wen W, Fan X, *et al*. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell* 2021;184:1895–913.
24. Pang SS, Berry R, Chen Z, *et al*. The structural basis for autonomous dimerization of the pre-T-cell antigen receptor. *Nature* 2010;467:844–8.
25. McClory S, Hughes T, Freud AG, *et al*. Evidence for a stepwise program of extrathymic T cell development within the human tonsil. *J Clin Invest* 2012;122:1403–15.
26. Taneja V. Sex hormones determine immune response. *Front Immunol* 2018;9:9.
27. Carreras E, Kovats S. Girl power: estrogen promotes HSC self-renewal. *Cell Stem Cell* 2014;14:137–8.
28. Noetzli LJ, French SL, Machlus KR. New insights into the differentiation of megakaryocytes from hematopoietic progenitors. *Arterioscler Thromb Vasc Biol* 2019;39:1288–300.
29. Du C, Xu Y, Yang K, *et al*. Estrogen promotes megakaryocyte polyploidization via estrogen receptor beta-mediated transcription of GATA1. *Leukemia* 2017;31:945–56.