

RESEARCH ARTICLE

Molecular mechanisms governing circulating immune cell heterogeneity across different species revealed by single-cell sequencing

Zhibin Li¹ | Chengcheng Sun^{2,3} | Fei Wang^{2,4,5} | Xiran Wang^{6,7} | Jiacheng Zhu^{2,3} | Lihua Luo^{2,3} | Xiangning Ding^{2,3} | Yanan Zhang⁸ | Peiwen Ding^{2,3} | Haoyu Wang^{2,3} | Mingyi Pu² | Yuejiao Li² | Shiyu Wang^{2,3} | Qiuyu Qin² | Yanan Wei² | Jian Sun^{6,7} | Xiangdong Wang^{9,10} | Yonglun Luo^{2,4,5,11} | Dongsheng Chen² | Wei Qiu¹

¹ Department of Neurology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

² BGI-Shenzhen, Shenzhen, China

³ College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

⁴ Department of Biomedicine, Aarhus University, Aarhus, Denmark

⁵ Lars Bolund Institute of Regenerative Medicine, Qingdao-Europe Advanced Institute for Life Sciences, BGI-Qingdao, BGI-Shenzhen, Qingdao, China

⁶ National Risk Assessment Laboratory for Antimicrobial Resistance of Animal Original Bacteria, South China Agricultural University, Guangzhou, China

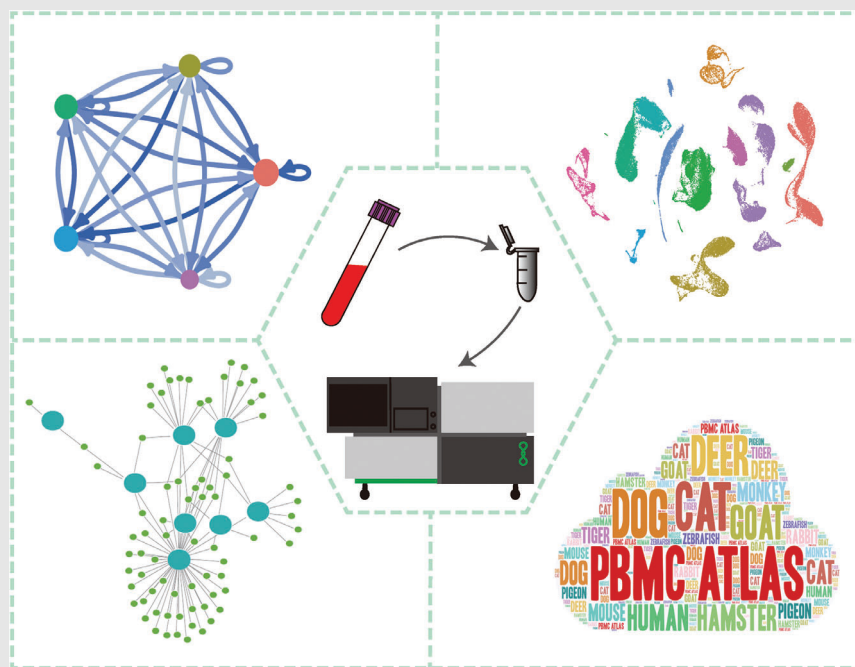
⁷ Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, China

⁸ Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen, China

⁹ Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Shanghai, China

¹⁰ Fudan University Shanghai Medical College, Shanghai, China

Graphical Abstract



Our study compared the single-cell transcriptomes of immune cells from 12 species. The results revealed the heterogeneity and compositions of circulating immune cells. We explored the conserved and divergent cellular crosstalks and genetic regulatory networks among vertebrate immune cells.

¹¹ Steno Diabetes Center Aarhus, Aarhus University Hospital, Aarhus, Denmark

Correspondence

Wei Qiu, Department of Neurology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China.

Email: qiuwei@mail.sysu.edu.cn

Dongsheng Chen, BGI-Shenzhen, Shenzhen, China.

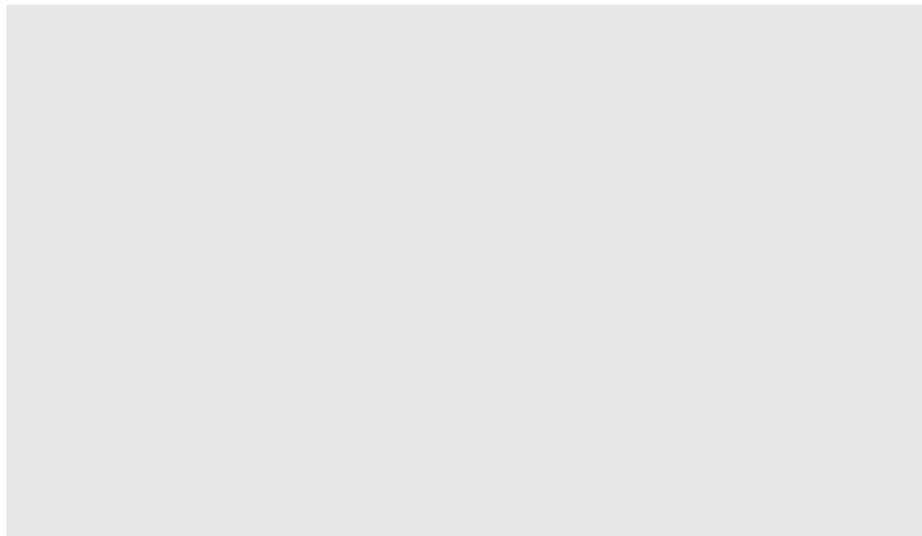
Email: chendongsheng@genomics.cn

Yonglun Luo, Department of Biomedicine, Aarhus University, Aarhus 8000, Denmark.

Email: alun@biomed.au.dk

Xiangdong Wang, Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Shanghai, China.

Email: xdwang@fucgb.com



RESEARCH ARTICLE

Molecular mechanisms governing circulating immune cell heterogeneity across different species revealed by single-cell sequencing

Zhibin Li¹ | Chengcheng Sun^{2,3} | Fei Wang^{2,4,5} | Xiran Wang^{6,7} | Jiacheng Zhu^{2,3} |
Lihua Luo^{2,3} | Xiangning Ding^{2,3} | Yanan Zhang⁸ | Peiwen Ding^{2,3} |
Haoyu Wang^{2,3} | Mingyi Pu² | Yuejiao Li² | Shiyu Wang^{2,3} | Qiuyu Qin² |
Yanan Wei² | Jian Sun^{6,7} | Xiangdong Wang^{9,10} | Yonglun Luo^{2,4,5,11} |
Dongsheng Chen² | Wei Qiu¹

¹ Department of Neurology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

² BGI-Shenzhen, Shenzhen, China

³ College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

⁴ Department of Biomedicine, Aarhus University, Aarhus, Denmark

⁵ Lars Bolund Institute of Regenerative Medicine, Qingdao-Europe Advanced Institute for Life Sciences, BGI-Qingdao, BGI-Shenzhen, Qingdao, China

⁶ National Risk Assessment Laboratory for Antimicrobial Resistance of Animal Original Bacteria, South China Agricultural University, Guangzhou, China

⁷ Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, China

⁸ Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen, China

⁹ Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Shanghai, China

¹⁰ Fudan University Shanghai Medical College, Shanghai, China

¹¹ Steno Diabetes Center Aarhus, Aarhus University Hospital, Aarhus, Denmark

Correspondence

Wei Qiu, Department of Neurology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China.

Email: qiuwei@mail.sysu.edu.cn

Dongsheng Chen, BGI-Shenzhen, Shenzhen, China.

Email: chendongsheng@genomics.cn

Yonglun Luo, Department of Biomedicine, Aarhus University, Aarhus 8000, Denmark.

Email: alun@biomed.au.dk

Xiangdong Wang, Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Shanghai, China.

Email: xdwang@fucb.com

Abstract

Background: Immune cells play important roles in mediating immune response and host defense against invading pathogens. However, insights into the molecular mechanisms governing circulating immune cell diversity among multiple species are limited.

Methods: In this study, we compared the single-cell transcriptomes of immune cells from 12 species. Distinct molecular profiles were characterized for different immune cell types, including T cells, B cells, natural killer cells, monocytes, and dendritic cells.

Results: Our data revealed the heterogeneity and compositions of circulating immune cells among 12 different species. Additionally, we explored the conserved and divergent cellular crosstalks and genetic regulatory networks among

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

Zhibin Li, Chengcheng Sun, Fei Wang, Xiran Wang contributed equally to this work.

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: #81771300, #81971140; Norman Bethune Foundation, Grant/Award Number: 2020009; Natural Science Foundation of Guangdong Province, Grant/Award Number: #2020A1515010053

vertebrate immune cells. Notably, the ligand and receptor pair VIM-CD44 was highly conserved among the immune cells.

Conclusions: This study is the first to provide a comprehensive analysis of the cross-species single-cell transcriptome atlas for peripheral blood mononuclear cells (PBMCs). This research should advance our understanding of the cellular taxonomy and fundamental functions of PBMCs, with important implications in evolutionary biology, developmental biology, and immune system disorders.

KEYWORDS

cross-species, peripheral blood mononuclear cells, single-cell RNA sequencing

1 | INTRODUCTION

Peripheral blood mononuclear cells (PBMCs) are derived from myeloid and lymphoid hematopoietic systems and are mainly comprised of circulating multi-functional immune cell types, such as lymphocytes, monocytes, and dendritic cells (DCs). As the supervisor and executor of body defense, PBMCs play important roles in mediating innate and adaptive immune responses, maintaining immune homeostasis, and reflecting the real-time cellular and humoral immune state of the whole body. As such, PBMCs are widely used in the fields of immunology,^{1,2} infectious diseases,^{3,4} cancer,⁵ vaccine development,⁶ transplantation⁷ and high-throughput screening for therapeutic antibodies.⁸ As a commonly used *ex vivo* cellular model in immunological function studies, PBMCs also play a vital role in immunological research and immunotherapy, and have been used to predict diagnostic biomarkers and discover potential immunotherapy targets.^{9,10}

Traditional RNA sequencing (RNA-seq) provides the ability to measure average gene expression of the entire transcriptome from bulk cells, which hides the potential cellular heterogeneity.¹¹ Single-cell RNA-sequencing (scRNA-seq) offers an unbiased approach to deconvolve the heterogeneity of immune cells and profile cell breadth (cell number) and depth (gene number per cell).¹²

Due to the complexity of PBMCs, it is difficult to study the function of individual immune cells. However, advances in scRNA-seq allow comprehensive analysis of the immune system at the single-cell resolution. scRNA-seq can capture gene expression of individual immune cell types, identify new immune cell populations, reveal pathogenic immune cell subsets and transcriptional modules related to pathogenesis, and evaluate immunotherapy efficacy and response.¹³ Furthermore, differential gene expression and intercellular interactions among immune cell types and samples can be evaluated.¹⁴ In addition to humans, many studies have been con-

ducted on scRNA-seq of PBMCs in mouse models, providing novel understanding of immune system in healthy and disease conditions.^{15–17} These studies not only identified immune cell types, their interactions, and regulatory molecular mechanisms, but also identified potential targets for immune-related disease therapy.^{9,10} In this study, we compared the cellular taxonomy of PBMCs in 12 species, revealing the conserved and divergent patterns of cellular crosstalk and genetic regulatory networks among multiple species.

2 | RESULTS

2.1 | Single-cell transcriptomic profiles of PBMCs

To enable cross-species comparison of the molecular mechanisms governing immune cell heterogeneity and compositions, we first isolated fresh PBMCs from individuals of seven species: cat (*Felis catus*), dog (*Canis lupus familiaris*), rabbit (*Oryctolagus cuniculus domesticus*), hamster (*Mesocricetus auratus*), deer (*Cervus nippon*), goat (*Capra aegagrus hircus*), and pigeon (*Columba livia domestica*). Using scRNA-seq and filtering out doublets and low-quality cells (see Section 5), we obtained high-quality scRNA-seq data of 50 478 cells. We next integrated five publicly available PBMC scRNA-seq datasets: that is, human,¹⁸ tiger¹⁹ monkey,²⁰ mouse,²¹ and zebrafish²² (Supporting information 1a), from which single-cell transcriptome data of 27 479 cells were obtained for further analyses (Figure 1A).

We first humanized the homologous genes for all non-human species and performed unsupervised clustering using the top variable genes. We identified five main types of immune cells, that is, T cells, B cells, natural killer (NK) cells, monocytes, and DCs based on the specific expression of cell-type marker genes (Figure 1B and Figure S1, S2, Supporting information 1b). Cell-type identity was also

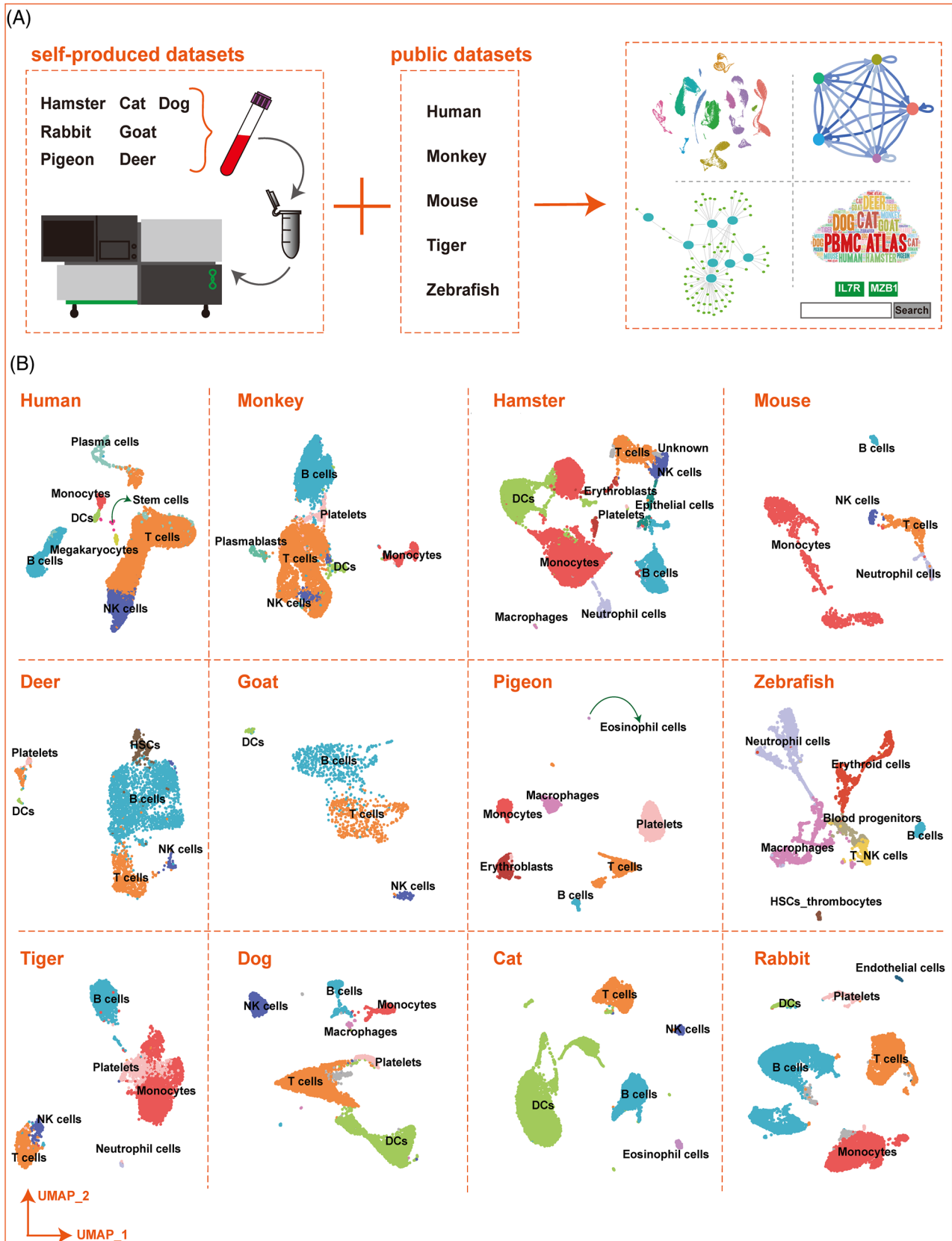


FIGURE 1 Single-cell transcriptome atlas of PBMCs for 12 species. (A) Illustration of the overall project design. (B) UMAP plots showing single cell transcriptome atlas of 12 species. Dots with colours represent different cell types which were indicated above

confirmed using gene ontology (GO) functional enrichment analysis of differentially expressed genes (DEGs) (Supporting information 2 and 3). For example, in the cat, B cells were annotated based on the high expression of markers *CD19*, *IRF8*, and *MS4A1*; T cells were characterized by the specific expression of *TCF7* and *TAPBPL*; NK cells were identified by enrichment of *CCL5*, *GZMA*, and *KLRF1*; and DCs were characterized by enrichment of *TREMI*, *TKT*, and *SOD2* (Figure 2A,B). In the dog, B cells were annotated based on the high expression of markers *CD19*, *BLNK*, *FCRLA*, and *MS4A1*; T cells were characterized by the specific expression of *CD4*, *SOD2*, *ITM2B*, and *SELL*; NK cells were identified via enrichment of *GZMB*; DCs were confirmed by enrichment of *MAFB* and *RETN*; and monocytes were identified by the expression of *IL-7R*, *TCF7*, and *CCR7* (Figure 2C,D). The specific expression patterns of these molecules successfully identified the different cell types and provided a molecular basis for exploring the physiological functions of the respective immune systems (Figure S3). In both cat and dog, functional analysis of DEGs in the five main immune cell types indicated that they were primarily related to regulation of innate immune response, regulation of immune effector process, neutrophil activation involved in immune response, immune response-activating signal transduction, and immune response-activating cell surface receptor signaling pathway (Figure S3).

2.2 | Conservation of PBMC connectomes

To identify potential cellular interactions, we constructed a ligand receptor mediated communication network of the above five immune cell types for each species using the Connectome R package²³ (Figure 3A, Supporting information 4). Based on the interaction network data, we further analysed the relationship between ligand–receptor pairs in the different cell types. For example, CD4OLG was identified as a ligand of T cells, and was able to interact with its receptors CD40, TRAF3, and ITGB2 on B cells (Figure 3B). For target analysis of network centrality, these communication pairs were generally classified into seven signalling modalities (i.e., tumour necrosis factor (TNF), NOTCH, matrix glycoproteins, intracellular trafficking, interleukins, complement, and uncategorized) (Figure).

We further identified connectivity among immune cells. In total, 214 pairs of cell–cell connections were conserved among the five main immune cells in the 12 species (Supporting information 4 m). Among them, there were 19 ligand–receptor pairs showing pan-conserved immune cell interactions. In addition, ligands related to the *VIM* gene appeared most frequently, and receptors related to the *CD44* gene appeared most frequently, with most belong-

ing to the uncategorized signalling modality (Figure 3D, Supporting information 4 n). In addition to immune cell crosstalks, we explored the conservation of cell–cell crosstalks in different species. Several interactions between immune cells in PBMCs, such as the VIM ligand and CD44 receptor, were commonly expressed in the DCs and T cells of human, monkey, hamster, dog, cat, rabbit, and goat. The CD40LG and CD40, CD40LG and TRAF3, CD40LG and ITGB, were common interactions pairs in T cells–B cells among human, monkey, hamster, mouse, tiger, dog, and cat. The GNAI2–C5AR1 ligand–receptor pair was expressed in the monocytes of six species (human, monkey, hamster, mouse, rabbit, and dog), and categorized as complement signalling modality. Additionally, the PSAP–SORT1 ligand–receptor pair was expressed in the DCs of five species (human, hamster, cat, rabbit, and dog) and belonged to the intracellular trafficking signalling modality (Figure 3E, Figures S4,5).

2.3 | Conservation of PBMC regulomes

To explore the regulatory mechanisms underlying the immune system development in the light of evolution, the PBMC genetic regulatory networks were predicted for the 12 species (Figure 4A, Supporting information 5). Subsequently, we analysed the regulatory network of the five immune cell types. A variety of TF–target interactions conserved in at least four species were identified (293 in T cells, 324 in B cells, 108 in DCs, 94 in NK cells, and 194 in monocytes) (Figure 4B, Supporting information 6–10). Enrichment of GO terms for predicted target genes indicated that the regulatory functions of these TFs were closely related to immune response processes (Figures 4C and S6). Many regulatory circuits were highly conserved in each cell type among the multiple species (Figure 4C). Specifically, in B cells, the interactions between *MEF2C* and its targets (*LYN*, *ABRACL*, *ARID5B*, *CANX*, *CCT5*, *HSP90BI*, *PAX5*, *PLEK*, *PNISR*, *RALGPS2*, *SNW1*, *VIM*, and *PAN3*) were highly conserved. The interactions between *PAX5* and its target genes (*LYN* and *VCP*) were conserved, and were mainly related to the differentiation, activation, proliferation, and receptor signalling pathway of B cells. In T cells, the regulatory relationships between *NCOR1* and its target genes (*PTPRC*, *ARHGAP15*, *IQGAP1*, *RPL27*, *SCAF11*, and *YTHDC1*), *TCF7* and its target gene (*LEF1*), and *FOS* and its target genes (*PTPN6*, *TMBIM6*, *RPL27*, *HERPUDI*, and *VIM*) were conserved, and were mainly involved in the differentiation, activation, proliferation, and receptor signalling pathway of T cells. In DCs, *CCDC88A* and its target genes (*PSMA3*, *GOLGA4*, *LGALS1*, *PDIA3*, and *SWAP70*) and *PLEK* and its target genes (*ANXA1* and *SGK1*) were mainly related to antigen processing and presentation. In

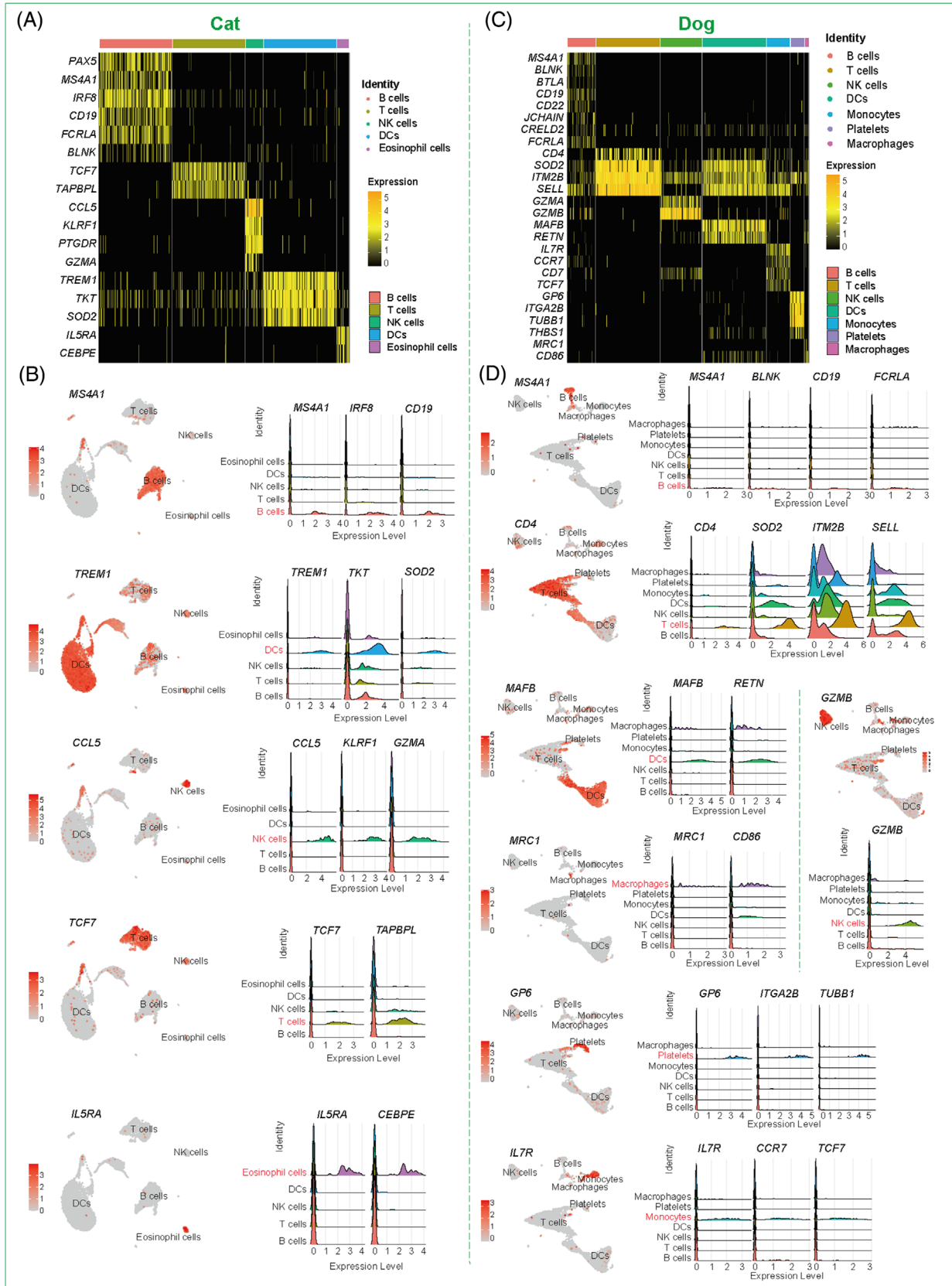


FIGURE 2 The cellular heterogeneity and compositions of PBMCs in cat and dog. (A,C) Heatmap plots showing a series of genes expressions in various annotated cell types of cat (A) and dog (C). The colours indicated the percentage of cells that showed gene expression levels within each cell type. (B,D) Feature plots and ridge plots showing the gene expression pattern of cat (B) and dog (D) respectively in different cell types

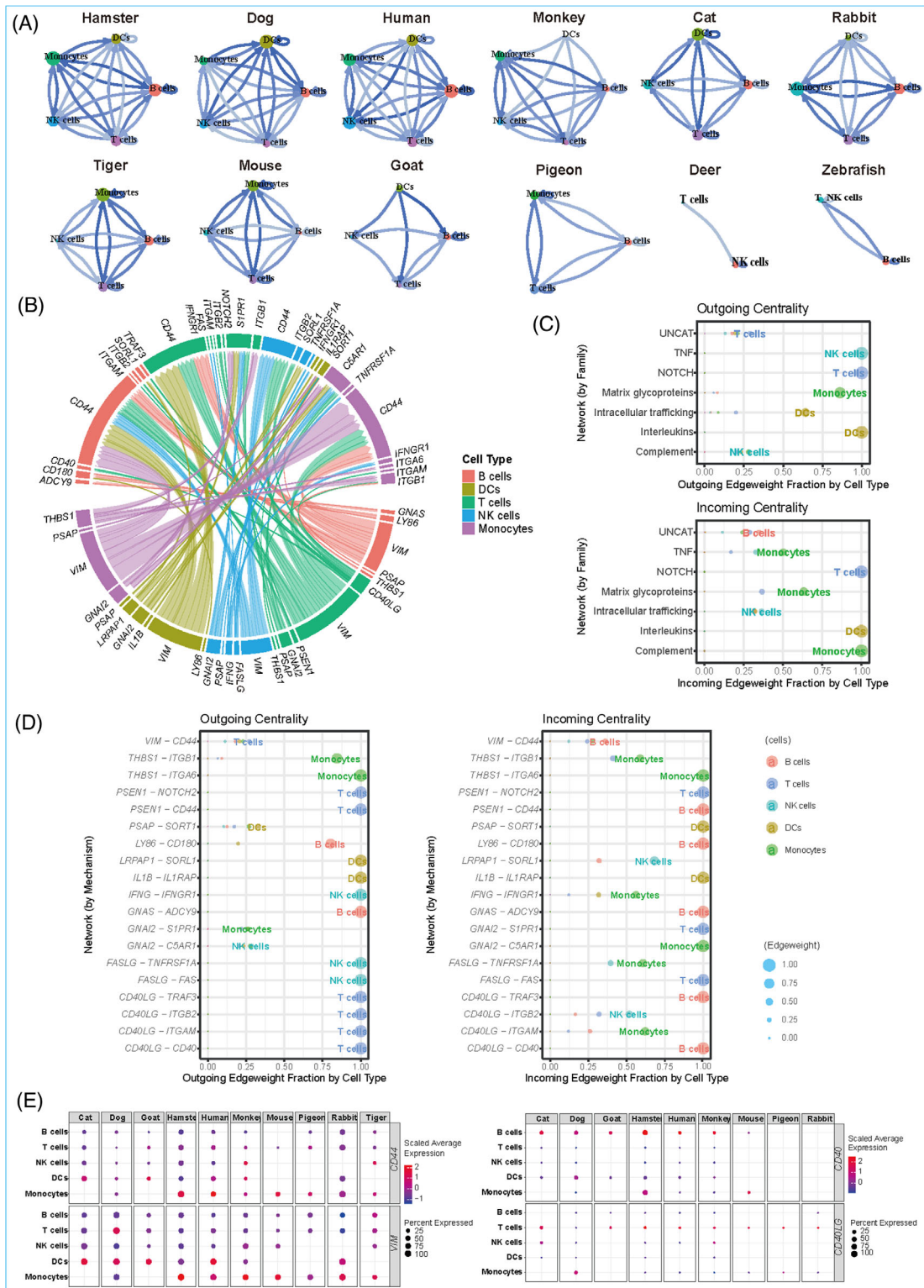


FIGURE 3 Cross species conserved PBMCs cellular connectomes. (A) Communication network of receptor–ligand pairs between five immune cell types including B cells, T cells, NK cells, DCs, and monocytes. Cell types were represented by coloured node, of which the size was proportional to the sum of receptor–ligand pairs between this node and all other nodes. The edge colour was proportional to the number of receptor–ligand pairs between two connected nodes. (B) Circos plot of cross species conserved connectome. Receptors and ligands were displayed near the upper and lower half circle respectively. (C) Centrality analysis of the conserved connectomes grouped by modes (signalling families). In the centrality plot, the outgoing means sending and incoming means receiving, which refers to quantitative metrics of how 'connected' a given edge is to other edge. (D) Centrality analysis of the conserved connectomes grouped by mechanism (ligand–receptor pair). (E) Dot plots showing the co-expression of two ligand–receptor gene pairs (*VIM–CD44*, *CD40LG–CD40*) in five immune cells of indicated species

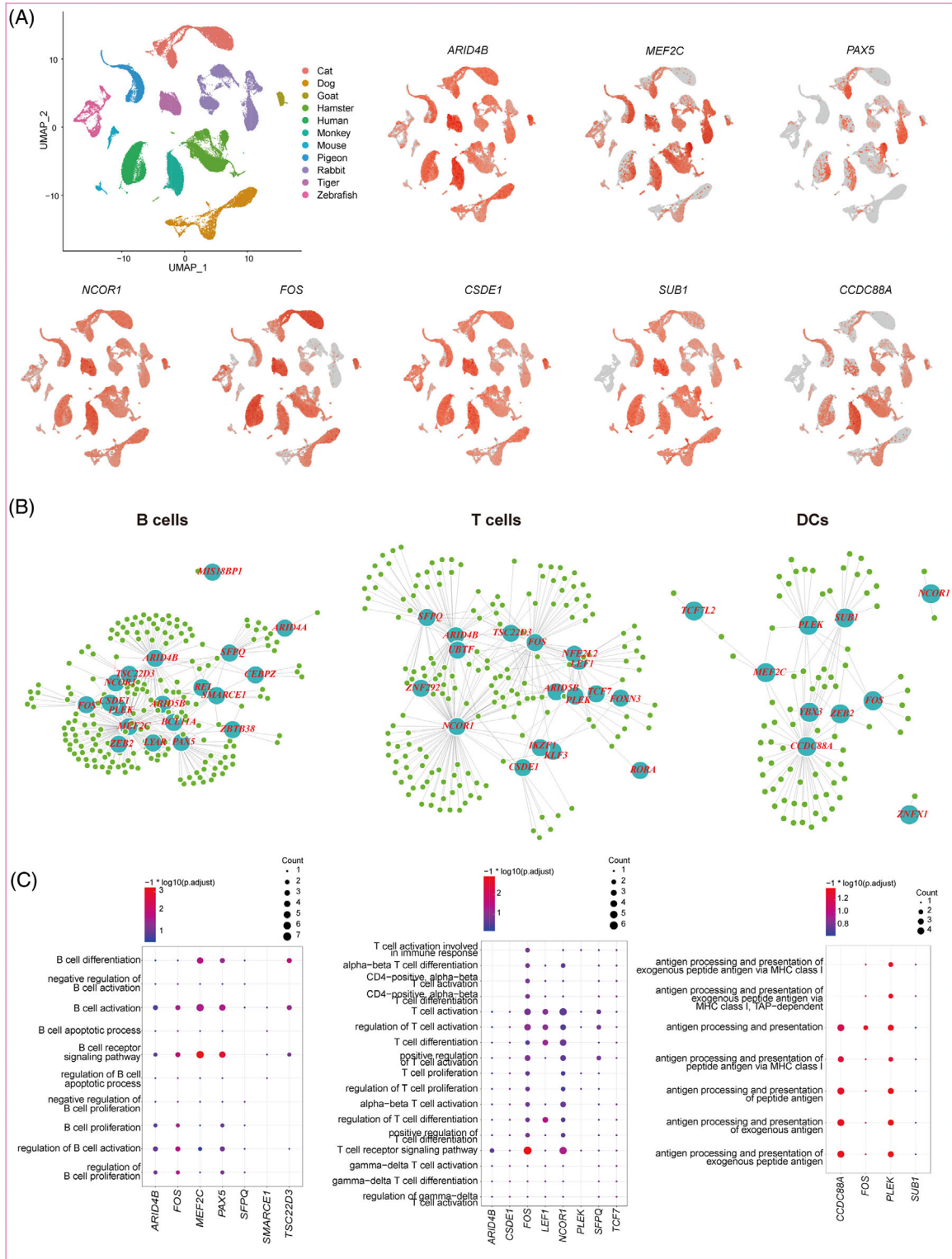


FIGURE 4 Cross species conservation of regulomes in PBMCs. (A) Feature plots showing high expressions of a series of TFs cross species. A big dataset was merged by 12 species datasets and the different colours indicated different species. (B) Conserved genetic regulatory networks in B cells, T cells and DCs. TF–target interactions with a weight (linkList) value ≥ 0.01 of specific cell type in all species was performed, and the frequency of each pair in each species was counted as SpeciesNumber to evaluate its conservation level. If the SpeciesNumber was ≥ 4 , then the TF–target interactions were considered as conserved. Light blue nodes represent regulators, green nodes represent corresponding target genes. Edge width is proportional to weight of regulation, and node size is proportional to the number of target genes of regulator. (C) GO term enrichment related to cellular functions of predicted target genes in B cells, T cells and DCs. Dot colour represents significant level of enrichment analysis and dot size is proportional to the count of target genes classified in GO terms

NK cells, the interactions between *SUB1* and its target gene (*NDUFA8*) were conserved (Figure S6, Supporting information 9). In monocytes, the interactions between *PLEK* and its target genes (*B2M*, *CTLA*, and *LGALS3*) were conserved (Figure S6, Supporting information 10).

2.4 | Integrated online platform for versatile data exploration

To share our data resources, we developed a website named PBMCatlas (<http://120.79.46.200:81/Pandora/PBMC.html>), allowing users to analyze cell populations of interest and quickly visualize the expression patterns of important genes (Figure 5). This function can assist users to clarify if the input gene is specifically expressed in certain cell types. Thus, PBMCatlas offers a user-friendly platform to explore our dataset and flexibly access our analysis results, which should facilitate future-related research.

3 | DISCUSSION

As the most often used cell model in immunological studies, PBMCs can reflect the dynamic changes in circulating innate and adaptive immune systems and play a vital role in immunotherapy. Here, we used scRNA-seq analysis to elucidate the heterogeneity of PBMCs, and investigated the conserved cell–cell communications and genetic regulatory networks of major immune cell types across 12 different species.

3.1 | The cellular heterogeneity and compositions of PBMCs

We surveyed the PBMC atlas by scRNA-seq for comprehensive analysis. In total, 77 957 cells were derived from the 12 species. After cell clustering analysis, we identified five major immune cell types for each species, including NK cells, B cells, DCs, monocytes, and T cells. The divergences of immune cells across species were also investigated in the previous study.²⁴

3.2 | The conserved cellular crosstalks between immune cells

3.2.1 | VIM–CD44 in DCs–T cells

We analysed cell–cell interactions from the classical immune response and focused on DCs–T cells and T cells–B cells, which are involved in antigen presentation processing and adaptive immune response. The interactions between *VIM* and *CD44* in DCs–T cells were iden-

tified among 214 pairs of conserved cell interactions in the 12 species. VIM–CD44 is a highly conserved cell–cell crosstalk pair, especially among ligand–receptor pairs in DCs–T cells. *VIM* is expressed in lymphocytes and can interact with other proteins for intercellular signal transduction and can be released as an antigen component of pathogen infection,^{25,26} with bacterial and viral pathogens able to attach to this protein on the host cell surface.²⁷ *CD44* is up-regulated in activated lymphocytes and is involved in various cellular functions, including activation, recirculation, and homing of T-lymphocytes (T cells are activated, and IL-2 production is elevated under CD44 stimulation), as well as hematopoiesis, inflammation, and response to bacterial infection.²⁸ In our study, the VIM–CD44 interaction is highly conserved among multiple species. VIM on DCs could cooperate with CD44 on T cells, which promotes the antigen presentation and activation of autoreactive T cells. Additionally, the binding of VIM and CD44 activates T cells, triggers activation of a series of possible effector genes, and activates signalling pathways transduction,²⁹ while how this signal is transduced and participates in immune responses and host defense requires further study in immune diseases.

3.2.2 | CD40LG–CD40 in T cells–B cells

CD40LG–CD40 is among one of the conserved ligand–receptor pairs between T cells and B cells. CD40LG is predominantly expressed in activated CD4⁺ T cells and binds to its ligand (CD40) on the surface of B cells, thereby influencing B cell function.^{30,31} This pair also costimulates T-cell proliferation and cytokine production,³² and enhances the expression of IL-4 and IL-10³². *CD40LG* deficiency is a severe primary immunodeficiency caused by mutations in the *CD40L* gene, which can lead to T-cell impairment, B-cell defects, and susceptibility to opportunistic pathogens.³³ Mutation of the *CD40* gene can result in type 3 hyper-IgM immunodeficiency, characterized by an inability to undergo isotype switching, an inability to mount an antibody-specific immune response, and a lack of germinal centre formation.³⁴ Clinical trials have evaluated novel therapeutic approaches targeting the CD40–CD40LG pathway based on T cell–B cell interactions for autoimmune diseases.³⁵

3.3 | The conserved regulatory networks in different immune cells

3.3.1 | CCDC88A and its target genes in DCs

Lastly, we investigated the conserved regulatory networks in different immune cells among the 12 species. The

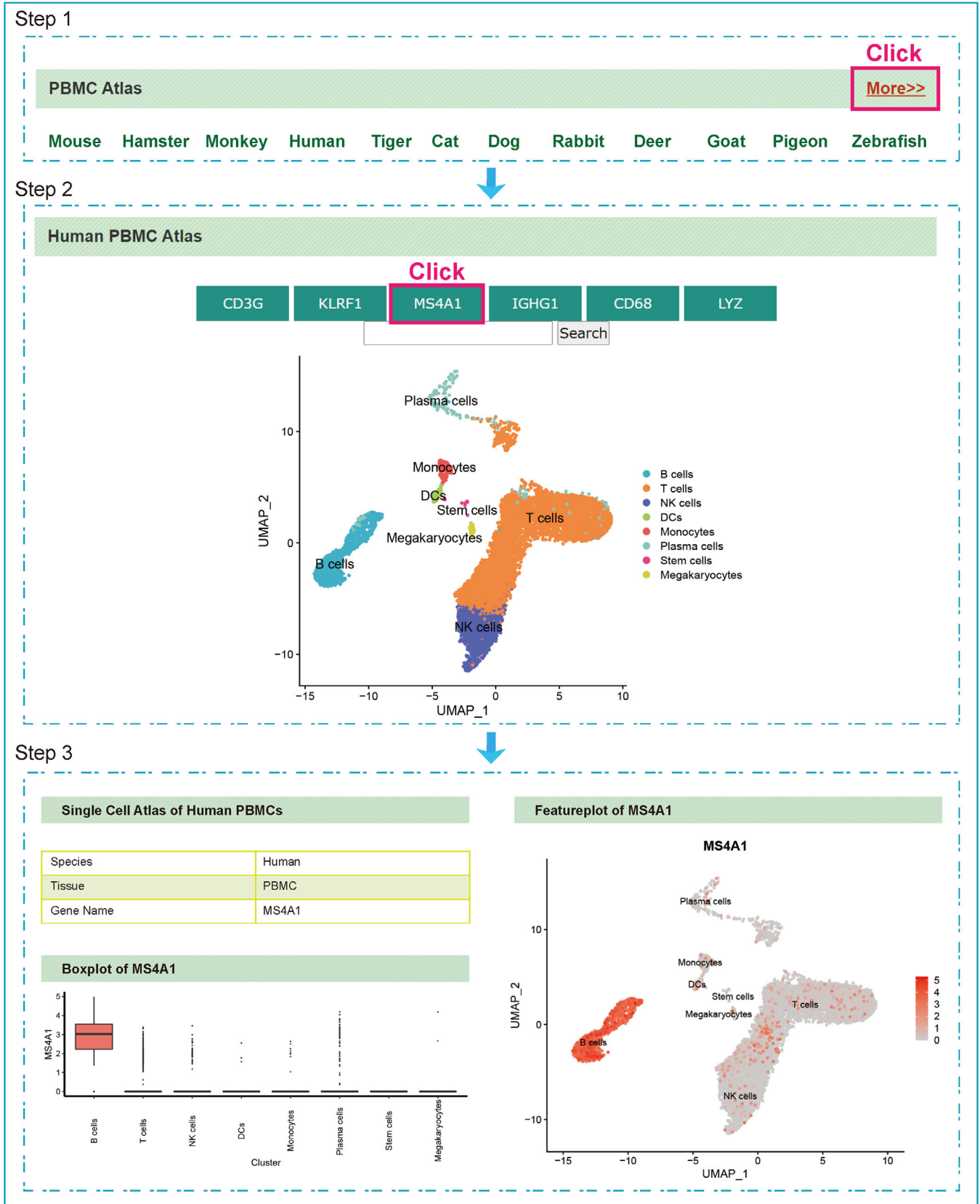


FIGURE 5 Cross species PBMC atlas website. Based on the scRNA-seq data of 12 species, a cross-species PBMC atlas website was generated. It can be used to search for information about gene expression in different species in the form of box plots and feature plots. Additionally, it can be utilized to investigate the conserved genes expression between species at the single-cell level

coiled-coil domain containing 88A (*CCDC88A*) gene encoded protein is a kind of coiled-coil domain containing Girdin family proteins that are activated by Akt and necessary for cytoskeleton remodelling and cell migration.^{36–38} *CCDC88A* accumulates in cell protrusions and contributes to the formation of membrane protrusions and cell migration and invasion,^{39,40} in line with the structural characteristics and functional demands of DCs. In addition, *PSMA3* a target gene of *CCDC88A*, which encodes proteasome, is essential for the generation of a subset of major histocompatibility complex (MHC) class I-presented antigenic peptides, as well as for the maturation of DCs.⁴¹ The guanine nucleotide-binding (G) protein α subunit ($G\alpha$)-interacting vesicle-associated protein (GIV), protein of *CCDC88A* is most highly expressed in DCs and macrophages and participates in the inhibition of proinflammatory signalling via Toll-like receptors (TLRs).⁴² These developmental and functional characteristics of DCs may provide new pathways to restore immune tolerance and inhibit self-antigen presentation processing.

3.3.2 | NCOR1, TCF7 and their target genes in T cells

NCOR1 plays an important role in controlling positive and negative selection of thymocytes during T-cell development.⁴³ Furthermore, NCOR1 is considered a novel regulator of immune tolerance and immune cell development,⁴⁴ and shapes the transcriptional landscape, influences the direction of CD4⁺ T cell differentiation, and controls Th1/Th17 effector functions.⁴⁵ In addition, its target gene *PTPRC* (encoded protein tyrosine phosphatase receptor type C), also known as CD45, is essential for T-cell antigen receptor-mediated activation, and its downstream regulatory imbalance can result in autoimmunity.^{46,47} *TCF7* is predominantly expressed in T cells and plays a critical role in T-cell development.^{48,49} Its encoded protein, T cell factor-1 (TCF-1), belongs to the T-cell factor/lymphoid enhancer-binding factor family. TCF-1 is highly expressed in naive CD8⁺ T cells but is down-regulated after differentiation into effector CD8⁺ T cells, and is necessary for the formation of central CD8⁺ T cell memory in response to infection.^{50–52} Silencing of *Tcf1* facilitates effector CD8⁺ T cell differentiation,⁵¹ and knockout of *TCF7* in mice results in impaired T-lymphocyte differentiation.⁵³ In addition, its target gene *LEF1*, which encodes lymphoid enhancer binding factor 1, can bind to functionally important sites in the T-cell receptor- α enhancer. This is critical for the maturation and development of IL17A-producing T cells, with its imbalance downstream potentially resulting in autoimmunity.^{47,54}

3.3.3 | MEF2C, PAX5 and their target genes in B cells

MEF2C binds the active regulatory region to the V(D)J gene in mouse B cell progenitors and human B lymphoblasts, which is essential for lymphatic fate determination.^{55,56} MEF2C has a highly conserved MADS box and MEF2 domain, which contribute to B cell homeostasis.^{57,58} In addition, MEF2C and early B cell factor-1 together form a co-regulator, which targets and regulates a subset of B cell-specific genes.⁵⁵ Various animal models also show that *MEF2C* is important in myeloid leukaemia. Mutations in *MEF2C* are often found in patients with B cell lymphoma, and these mutations are involved in the pathogenesis of abnormal B cell proliferation.^{57,59,60} *LYN* is the target gene of MEF2C, participating in the regulation of B cell differentiation, proliferation, survival, and apoptosis, and plays an important role in maintaining immune self-tolerance.⁶¹ It also acts downstream of B cell receptors via the down-regulation of signalling pathways. As another crucial regulatory gene of B cells, paired box protein 5 (*PAX5*) is necessary for the differentiation of lymphoid progenitor cells into B lymphocyte lineage.⁶² *PAX5* regulates transcriptional reprogramming processes by restricting uncommitted progenitors to the B cell pathway, promoting V(H)–DJ(H) recombination, inducing B-cell receptor signalling, and facilitating development to the mature B-cell stage.^{62,63} However, *PAX5* inhibition is not necessary for stable plasma cell development and antibody secretion, even though it is essential for immunoglobulin G (IgG) production and long-lived plasma cell increase.⁶⁴ Thus, the role of *PAX5* in plasma cell differentiation needs to be further investigated. *LYN* is also a target gene of *PAX5*, and studies show that both are related to B cell development.^{65,66} Collectively, we identified highly conserved regulomes in the PBMCs of different animal species. Identifying conserved key genes and exploring their functions in multiple species will help improve our understanding of the development, maturation, proliferation, activation, differentiation of immune cells.

In this study, we produced the comprehensive PBMC atlas of 12 species, which holds significance for immunological research. We systematically studied the gene expression profiles and molecular characteristics of each cell type and compared them across species at single-cell resolution. We also identified key genes and highly conserved cell–cell interactions that play important roles in regulating development and immune response. The PBMC atlas website was constructed, which provides an accessible approach to explore different species datasets. These results provide a systematic resource for understanding

immune cell diversity as well as insights into the molecular mechanisms governing conservation of PBMCs across species.

4 | LIMITATIONS OF THE STUDY

A few limitations of this study were also addressed, which should be investigated by further study. First, the conservation of homologous genes in non-human species was calculated to identify distinct cell types of PBMCs in different species. However, the non-traditional species (e.g., deer, rabbit) showed low ratios of homologous genes to humans, resulting in lower diversity in cell types of PBMCs for these species. Second, when comparing cell-type specific gene expression across species, the effects of age, sex, and physiological conditions were not considered. Nevertheless, major immune cell types were identified in all these species, and cellular heterogeneity and compositions of PBMCs were characterized. The connectome analysis may be influenced by potential species differences in ligand–receptor interactions or homolog conversion. Last, the aim of this study was to focus on the single-cell atlas of PBMCs from multiple species, even though mRNA and protein expression in immune cells can show discrepancies when using different methods.^{67–69} In addition, the PBMCs datasets for the 12 species were obtained from different platforms, and differences caused by sample processing, scRNA-seq technical bias and batch effects could impact cell capture and cell-type classifications, resulting in technical differences. Importantly, despite these limitations, this study provides a key resource of PBMCs for understanding the immune cell-type diversity as well as insights into the developmental and evolutionary biology of the circulating immune system across species.

5 | MATERIAL AND METHODS

5.1 | Ethics statement

The collection and experimental processing of all samples used in this article were strictly carried out according to the “Guidelines on the Ethical Treatment of Experimental Animals” established by the Ministry of Science and Technology, China. The Institutional Review Board on Ethics Committee of BGI reviewed and approved this study (NOS. BGI-IRB A20008, BGI-IRB A20008 T1).

5.2 | Blood samples collection

Blood samples were obtained from seven animals, including: cat (*Felis catus*), dog (*Canis lupus familiaris*), hamster

(*Mesocricetus auratus*), goat (*Capra aegagrus hircus*), rabbit (*Oryctolagus cuniculus domesticus*), pigeon (*Columba livia domestica*), and deer (*Cervus nippon*). The blood samples of deer with related genome assembly were provided by the Institute of Special Animal and Plant Sciences (ISAPS) of the Chinese Academy of Agricultural Sciences. Other bloods samples were obtained from farmers’ markets with permission from the BGI Ethics Committee. The gene expression matrices for five other species (human,¹⁸ tiger,¹⁹ monkey,²⁰ mouse,²¹ zebrafish²²) were obtained from published datasets.

5.3 | Peripheral blood mononuclear cells processing

All operations were performed under sterile conditions. Peripheral blood samples (3 mL) were collected into an EDTA anticoagulant tube, gently reversed 4–6 times, and fully mixed, and placed at room temperature. Whole blood was diluted with 3 mL of phosphate-buffered saline (PBS) and transferred into a 15 mL centrifuge tube. After this, 6 mL of Histopaque-1077 (Cat. No. 10771-6X100ml) was slowly added into the 15 mL centrifuge tube, followed by density gradient centrifugation methods to collect peripheral blood mononuclear cells (PBMCs).

5.4 | Single-cell RNA-seq library construction and sequencing

The goat, deer and pigeon PBMCs underwent library construction using DNBelab C Series Single-Cell Library Prep Set (MGI) and sequenced with BGISEQ-500 in China National GeneBank (CNGB). Cat, dog, rabbit and hamster PBMCs underwent library construction using a 10X Chromium Next GEM Single cell 3’ Reagent Kits v3.1 following the guidelines provided by the manufacturer and sequenced using NOVAseq 6000 sequencing platform of Illumina.

5.5 | Cross-species homologous gene conversion

We referred to the previous methodological section for cross-species homologous gene transfer and single-cell RNA-seq data processing.¹⁹

5.6 | Single-cell RNA-seq data preprocessing

Genome information used for read alignment was downloaded from the NCBI Assembly (Supporting information

1b). The raw data was processed using Cell Ranger v3.0.2 (10X Genomics) and Seurat.^{71,72}

5.7 | Cell-type annotation

For annotation of the self-produced datasets, our annotation used several classic cell-type markers from the Cell-Marker database.⁷³ The four published datasets used cell-type markers from corresponding published article. The annotation results are presented in Figure 1B using UMAP plots.

5.8 | Differentially expressed genes (DEGs) and Gene Ontology (GO) term enrichment analysis

All DEGs for each cell type were identified using the FindAllMarkers function in Seurat (Supporting information 2). The hypergeometric test implemented in the clusterProfiler⁷⁴ package with the compareCluster function (ENTREZID~celltype, fun = “enrichGO”, “org.Hs.eg.db”, p-value cutoff = 0.05) was used to carry out GO term enrichment analysis (Supporting information 3).

5.9 | Cellular communication analysis

We applied the Connectome (<https://github.com/msraredon/Connectome>) R package²³ for cellular communication analysis. All ligands and receptors data was downloaded from the FANTOM5 database.^{75,76} First, the five major immune cell types (B cells, T cells, NK cells, DCs, monocytes) were extracted from the annotated datasets of the 12 species for further analysis. The connectome networks were then constructed according to the expression of ligands and receptors.

5.10 | Transcription Factor (TF)-target interaction analysis

We applied the GENIE3⁷⁷ R package for TF-target interaction analysis using data from 11 species (cat, dog, goat, hamster, human, monkey, mouse, pigeon, rabbit, tiger, and zebrafish). The human’s TF list was downloaded from animalTFDB3.0.⁷⁸ The igraph R package⁷⁹ was used to visualize representative regulatory TFs networks.

ACKNOWLEDGEMENTS

We thank the China National GeneBank for producing the sequencing data. This work was supported by the

National Natural Science Foundation of China (#81771300, #81971140), Norman Bethune Foundation (2020009) and Natural Science Foundation of Guangdong Province (#2020A1515010053).

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

- Hritzoy Ahye MK, Golding A. Cytoplasmic FOXO1 identifies a novel disease-activity associated B cell phenotype in SLE. *Lupus Sci Med*. 2018;5:e000296.
- Bijl M, Horst G, Limburg PC, Kallenberg CG. Fas expression on peripheral blood lymphocytes in systemic lupus erythematosus (SLE): relation to lymphocyte activation and disease activity. *Lupus*. 2001;10:866-872.
- Wilk AJ, Rustagi A, Zhao NQ, et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat Med*. 2020;26:1070-1076.
- Jones D, Como CN, Jing L, et al. Varicella zoster virus productively infects human peripheral blood mononuclear cells to modulate expression of immunoinhibitory proteins and blocking PD-L1 enhances virus-specific CD8+ T cell effector function. *PLoS Pathog*. 2019;15:e1007650.
- Gros A, Parkhurst MR, Tran E, et al. Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nat Med*. 2016;22:433-438.
- Lipsit SWL, Wilkinson J, Scruten E, et al. Kinome profiling of peripheral blood mononuclear cells collected prior to vaccination reveals biomarkers and potential mechanisms of vaccine unresponsiveness in pigs. *Sci Rep*. 2020;10:11546.
- Zhang M, Huang B. The multi-differentiation potential of peripheral blood mononuclear cells. *Stem Cell Res Ther*. 2012;3:48.
- Gong C, Ni Z, Yao C, et al. A high-throughput assay for screening of natural products that enhanced tumoricidal activity of NK cells. *Biol Proced Online*. 2015;17:12.
- Mitsialis V, Wall S, Liu P, et al. Single-cell analyses of colon and blood reveal distinct immune cell signatures of ulcerative colitis and Crohn’s disease. *Gastroenterology*. 2020;159:591-608.
- Kamphorst AO, Pillai RN, Yang S, et al. Proliferation of PD-1+ CD8 T cells in peripheral blood after PD-1-targeted therapy in lung cancer patients. *Proc Natl Acad Sci U S A*. 2017;114:4993-4998.
- Ziegenhain C, Vieth B, Parekh S, et al. Comparative analysis of single-cell RNA sequencing methods. *Mol Cell*. 2017;65:631-643.
- Papalexio E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol*. 2018;18:35-45.
- Stubbington MJT, Rozenblatt-Rosen O, Regev A, Teichmann SA. Single-cell transcriptomics to explore the immune system in health and disease. *Science*. 2017;358:58-63.
- Wu P, Chen D, Ding W, et al. The trans-omics landscape of COVID-19. *Nat Commun*. 2021;12:4543.
- Smith SL, Kennedy PR, Stacey KB, et al. Diversity of peripheral blood human NK cells identified by single-cell RNA sequencing. *Blood Adv*. 2020;4:1388-1406.
- Goel RR, Wang X, O’neil LJ, et al. Interferon lambda promotes immune dysregulation and tissue inflammation in

- TLR7-induced lupus. *Proc Natl Acad Sci U S A*. 2020;117:5409-5419.
17. Villani A-C, Satija R, Reynolds G, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science*. 2017;356:4573.
 18. Zhu L, Yang P, Zhao Y, et al. Single-cell sequencing of peripheral mononuclear cells reveals distinct immune response landscapes of COVID-19 and influenza patients. *Immunity*. 2020;53:685-696.e3.
 19. Chen D, Sun J, Zhu J, et al. Single cell atlas for 11 non-model mammals, reptiles and birds. *Nat Commun*. 2021;12(1):7083.
 20. Kotliar D, Lin AE, Logue J, et al. Single-cell profiling of ebola virus disease in vivo reveals viral and host dynamics. *Cell*. 2020;183:1383-1401.
 21. Winkler ES, Shrihari S, Hykes BL, et al. The intestinal microbiome restricts alphavirus infection and dissemination through a bile acid-type I IFN signaling axis. *Cell*. 2020;182:901-918.
 22. Tang Q, Iyer S, Lobbardi R, et al. Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. *J Exp Med*. 2017;214:2875-2887.
 23. Raredon MSB, et al. Connectome: computation and visualization of cell-cell signaling topologies in single-cell systems data. *bioRxiv*. 2021.
 24. Geirsdottir L, David E, Keren-Shaul H, et al. Cross-species single-cell analysis reveals divergence of the primate microglia program. *Cell*. 2019;179:1609-1622.
 25. Hsu P-C, Liao Ya-F, Lin C-Li, Lin W-H, Liu G-Y, Hung H-C. Vimentin is involved in peptidylarginine deiminase 2-induced apoptosis of activated Jurkat cells. *Mol Cells*. 2014;37:426-434.
 26. Moisan E, Girard D. Cell surface expression of intermediate filament proteins vimentin and lamin B1 in human neutrophil spontaneous apoptosis. *J Leukocyte Biol*. 2006;79:489-498.
 27. Su L, Pan P, Yan P, et al. Role of vimentin in modulating immune cell apoptosis and inflammatory responses in sepsis. *Sci Rep*. 2019;9:5747.
 28. Funaro A, Spagnoli GC, Momo M, Knapp W, Malavasi F. Stimulation of T cells via CD44 requires leukocyte-function-associated antigen interactions and interleukin-2 production. *Human Immunol*. 1994;40:267-278.
 29. Bourguignon LYW, Singleton PA, Diedrich F. Hyaluronan-CD44 interaction with Rac1-dependent protein kinase N-gamma promotes phospholipase Cgamma1 activation, Ca(2+) signaling, and cortactin-cytoskeleton function leading to keratinocyte adhesion and differentiation. *J Biol Chem*. 2004;279:29654-29669.
 30. Takada YK, Yu J, Shimoda M, Takada Y. Integrin binding to the trimeric interface of CD40L plays a critical role in CD40/CD40L signaling. *J Immunol*. 2019;203:1383-1391.
 31. Graf D, Korthäuer U, Mages HW, Senger G, Kroczeck RA. Cloning of TRAP, a ligand for CD40 on human T cells. *Eur J Immunol*. 1992;22:3191-3194.
 32. Blotta MH, Marshall JD, DeKruyff RH, Umetsu DT. Cross-linking of the CD40 ligand on human CD4+ T lymphocytes generates a costimulatory signal that up-regulates IL-4 synthesis. *J Immunol*. 1996;156:3133-3140.
 33. França TT, Barreiros LA, Al-Ramadi BK, Ochs HD, Cabral-Marques O, Condino-Neto A. CD40 ligand deficiency: treatment strategies and novel therapeutic perspectives. *Expert Rev Clin Immunol*. 2019;15:529-540.
 34. Karaca NE, Forveille M, Aksu G, Durandy A, Kutukculer N. Hyper-immunoglobulin M syndrome type 3 with normal CD40 cell surface expression. *Scand J Immunol*. 2012;76:21-25.
 35. Karnell JL, Rieder SA, Ettinger R, Kolbeck R. Targeting the CD40-CD40L pathway in autoimmune diseases: humoral immunity and beyond. *Adv Drug Deliv Rev*. 2019;141:92-103.
 36. Lin C, Ear J, Pavlova Y, et al. Tyrosine phosphorylation of the Galpha-interacting protein GIV promotes activation of phosphoinositide 3-kinase during cell migration. *Sci Signal*. 2011;4:ra64.
 37. Ghosh P, Beas AO, Bornheimer SJ, et al. A G{alpha}i-GIV molecular complex binds epidermal growth factor receptor and determines whether cells migrate or proliferate. *Mol Biol Cell*. 2010;21:2338-2354.
 38. Garcia-Marcos M, Ghosh P, Farquhar MG. GIV is a nonreceptor GEF for G alpha i with a unique motif that regulates Akt signaling. *Proc Natl Acad Sci U S A*. 2009;106:3178-3183.
 39. Nechipurenko I, Lavrentyeva S, Sengupta P. GRDN-1/Girdin regulates dendrite morphogenesis and cilium position in two specialized sensory neuron types in *C. elegans*. *Dev Biol*. 2021;472:38-51.
 40. Tanouchi A, Taniuchi K, Furihata M, et al. CCDC88A, a prognostic factor for human pancreatic cancers, promotes the motility and invasiveness of pancreatic cancer cells. *J Exp Clin Cancer Res*. 2016;35:190.
 41. Matsunaga T, Ishida T, Takekawa M, Nishimura S, Adachi M, Imai K. Analysis of gene expression during maturation of immature dendritic cells derived from peripheral blood monocytes. *Scand J Immunol*. 2002;56:593-601.
 42. Swanson L, Katkar GD, Tam J, et al. TLR4 signaling and macrophage inflammatory responses are dampened by GIV/Girdin. *Proc Natl Acad Sci U S A*. 2020;117:26895-26906.
 43. Müller L, Hainberger D, Stolz V, Ellmeier W. NCOR1: a new player on the field of T cell development. *J Leukocyte Biol*. 2018;104:1061-1068.
 44. Lin A, Qiu Z, Zhang J, Luo P. Effect of NCOR1 mutations on immune microenvironment and efficacy of immune checkpoint inhibitors in patient with bladder cancer. *Front Immunol*. 2021;12:630773.
 45. Hainberger D, Stolz V, Zhu Ci, et al. NCOR1 orchestrates transcriptional landscapes and effector functions of CD4(+) T cells. *Front Immunol*. 2020;11:579.
 46. Al Barashdi MA, Ali A, McMullin MF, Mills K. Protein tyrosine phosphatase receptor type C (PTPRC or CD45). *J Clin Pathol*. 2021;74:548-552.
 47. Courtney AH, Shvets AA, Lu W, et al. CD45 functions as a signaling gatekeeper in T cells. *Sci Signal*. 2019;12:604.
 48. Harly C, Kenney D, Wang Y, et al. A shared regulatory element controls the initiation of Tcf7 expression during early T cell and innate lymphoid cell developments. *Front Immunol*. 2020;11:470.
 49. Raghu D, Xue H-H, Mielke LA. Control of lymphocyte fate, infection, and tumor immunity by TCF-1. *Trends Immunol*. 2019;40:1149-1162.
 50. Pais Ferreira D, Silva JG, Wyss T, et al. Central memory CD8(+) T cells derive from stem-like Tcf7(hi) effector cells in the absence of cytotoxic differentiation. *Immunity*. 2020;53:985-1000.
 51. Danilo M, Chennupati V, Silva JG, Siebert S, Held W. Suppression of Tcf1 by inflammatory cytokines facilitates effector CD8 T cell differentiation. *Cell Rep*. 2018;22:2107-2117.

52. Cui Lu, Guan Y, Qu Z, et al. WNT signaling determines tumorigenicity and function of ESC-derived retinal progenitors. *J Clin Invest*. 2013;123:1647-1661.
53. Schilham MW, Wilson A, Moerer P, Benaissa-Trouw BJ, Cumano A, Clevers HC. Critical involvement of Tcf-1 in expansion of thymocytes. *J Immunol*. 1998;161:3984-3991.
54. Jesse S, Koenig A, Ellenrieder V, Menke A. Lef-1 isoforms regulate different target genes and reduce cellular adhesion. *Int J Cancer*. 2010;126:1109-1120.
55. Wang W, Org T, Montel-Hagen A, et al. MEF2C protects bone marrow B-lymphoid progenitors during stress haematopoiesis. *Nat Commun*. 2016;7:12376.
56. Wilker PR, Kohyama M, Sandau MM, et al. Transcription factor Mef2c is required for B cell proliferation and survival after antigen receptor stimulation. *Nat Immunol*. 2008;9:603-612.
57. Dong C, Yang X-Z, Zhang C-Y, et al. Myocyte enhancer factor 2C and its directly-interacting proteins: a review. *Progr Biophysics Mol Biol*. 2017;126:22-30.
58. Canté-Barrett K, Pieters R, Meijerink JPP. Myocyte enhancer factor 2C in hematopoiesis and leukemia. *Oncogene*. 2014;33:403-410.
59. Jingjing Z, Lei M, Jie Z, et al. A novel MEF2C mutation in lymphoid neoplasm diffuse large B-cell lymphoma promotes tumorigenesis by increasing c-JUN expression. *Naunyn-Schmiedeberg's Arch Pharmacol* 2020;393:1549-1558.
60. Nagel S, Meyer C, Quentmeier H, Kaufmann M, Drexler HG, Macleod RAF. MEF2C is activated by multiple mechanisms in a subset of T-acute lymphoblastic leukemia cell lines. *Leukemia*. 2008;22:600-607.
61. Zhang X, Mei D, Zhang L, Wei W. Src family protein kinase controls the fate of B cells in autoimmune diseases. *Inflammation*. 2021;44:423-433.
62. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nat Immunol*. 2007;8:463-470.
63. Medvedovic J, Ebert A, Tagoh H, Busslinger M. Pax5: a master regulator of B cell development and leukemogenesis. *Adv Immunol*. 2011;111:179-206.
64. Liu GJ, Jaritz M, Wöhner M, et al. Repression of the B cell identity factor Pax5 is not required for plasma cell development. *J Exp Med*. 2020;217:e20200147.
65. Tang W-Y, Zhang Y-H, Zhang Yi-S, et al. Abnormal thymic B cell activation and impaired T cell differentiation in pristane-induced lupus mice. *Immunol Lett*. 2021;231:49-60.
66. Wu L, Kong L, Yang Y, et al. Effects of cell differentiation on the phagocytic activities of IgM(+) B cells in a teleost fish. *Front Immunol*. 2019;10:2225.
67. Reimegård J, Tarbier M, Danielsson M, et al. A combined approach for single-cell mRNA and intracellular protein expression analysis. *Commun Biol*. 2021;4:624.
68. Zhou Z, Ye C, Wang J, Zhang NR. Surface protein imputation from single cell transcriptomes by deep neural networks. *Nat Commun*. 2020;11:651.
69. Li J, Zhang Yi, Yang C, Rong R. Discrepant mRNA and protein expression in immune cells. *Curr Genomics*. 2020;21:560-563.
70. Emms DM, Kelly S. Phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20:238.
71. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411-420.
72. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell*. 2019;177:1888-1902.
73. Zhang X, Lan Y, Xu J, et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucl Acids Res*. 2019;47:D721-D728.
74. Yu G, Wang Li-G, Han Y, He Q-Yu. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic*. 2012;16:284-287.
75. Forrest AR, Kawaji H, Rehli M, et al. A promoter-level mammalian expression atlas. *Nature*. 2014;507:462-470.
76. Ramilowski JA, Goldberg T, Harshbarger J, et al. A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat Commun*. 2015;6:7866.
77. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE*. 2010;5:e12776.
78. Hu H, Miao Ya-Ru, Jia L-H, Yu Q-Y, Zhang Q, Guo An-Y. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucl Acids Res*. 2019;47:D33-D38.
79. team T I C igraph. Network analysis software; 2019.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Li Z, Sun C, Wang F, et al. Molecular mechanisms governing circulating immune cell heterogeneity across different species revealed by single-cell sequencing. *Clin Transl Med*. 2022;12:e689. <https://doi.org/10.1002/ctm2.689>