Single-cell RNA sequencing illuminates the ontogeny, conservation and diversification of cartilaginous and bony fish lymphocytes

Hong-Yan Wang [1,2,12], Jian-Yang Chen [3,4,12], Yanan Li [1,2,5,12], Xianghui Zhang [1,2,12], Xiang Liu [1,2,6], Yifang Lu [1,2], Hang He [3,4], Yubang Li [1,2], Hongxi Chen [3,4], Qun Liu [3,4], Yingyi Huang [1,2], Zhao Jia [7], Shuo Li [1,2], Yangqing Zhang [1,2], Shenglei Han [1,2], Shuhong Jiang [1,2], Mingming Yang [3,4], Yingying Zhang [3,4], Li Zhou [3,4], Fujian Tan [3,4], Qianyue Ji [8], Liang Meng [3,4], Rui Wang [1,2], Yuyan Liu [1,2], Kaiqiang Liu [1,2], Qian Wang [1,2], Inge Seim [9,10], Jun Zou [7], Guangyi Fan [3,4,11], Shanshan Liu [8] ✉ & Changwei Shao [1,2] ✉

Elucidating cellular architecture and cell-type evolution across species is central to understanding immune system function and susceptibility to disease. Adaptive immunity is a shared trait of the common ancestor of cartilaginous and bony fishes. However, evolutionary features of lymphocytes in these two jawed vertebrates remain unclear. Here, we present a single-cell RNA sequencing atlas of immune cells from cartilaginous (white-spotted bamboo shark) and bony (zebrafish and Chinese tongue sole) fishes. Cross-species comparisons show that the same cell types across different species exhibit similar transcriptional profiles. In the bamboo shark, we identify a phagocytic B cell population expressing several pattern recognition receptors, as well as a T cell sub-cluster co-expressing both T and B cell markers. In contrast to a division by function in the bony fishes, we show close linkage and poor functional specialization among lymphocytes in the cartilaginous fish. Our cross-species single-cell comparison presents a resource for uncovering the origin and evolution of the gnathostome immune system.

An increasing number of cell types arose over the evolutionary history of species[1]. Recent studies have provided new insights into the evolutionary history of cell type diversification across organ systems, including the central nervous and reproductive systems[2–4].

The immune system is, similarly, a complex regulatory network characterized by the coordinated functioning of specialized cell types, but little is known about the evolution of cell types within this system.

[1]State Key Laboratory of Mariculture Biobreeding and Sustainable Goods, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, Shandong 266071, China. [2]Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao Marine Science and Technology Center, Qingdao, Shandong 266237, China. [3]BGI Research, Qingdao 266555, China. [4]Qingdao Key Laboratory of Marine Genomics, BGI Research, Qingdao 266555 Shandong, China. [5]College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China. [6]Qingdao-Europe Advanced Institute for Life Sciences, BGI Research, Qingdao 266555, China. [7]Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Ministry of Education, Shanghai Ocean University, Shanghai 201306, China. [8]MGI Tech, Shenzhen 518083, China. [9]Integrative Biology Laboratory, College of Life Sciences, Nanjing Normal University, Nanjing 210023, China. [10]School of Biology and Environmental Science, Queensland University of Technology, Brisbane 4000, Australia. [11]BGI Research, Shenzhen 518083, China. [12]These authors contributed equally: Hong-Yan Wang, Jian-Yang Chen, Yanan Li, Xianghui Zhang. ✉e-mail: liushanshan@mgi-tech.com; shaocw@ysfri.ac.cn

The adaptive immune system is a common feature of all gnathostomes (jawed vertebrates) but is not fully resolved in invertebrates[5,6]. Adaptive immunity consists of two types of immune response: humoral immunity driven by B cells and cellular immunity steered by T cells. Cartilaginous fishes (Chondrichthyes) are the oldest jawed vertebrates reported to possess an adaptive immune system with B cell receptors (BCRs), T cell receptors (TCRs), and a major histocompatibility complex (MHC)[5,7,8]. Comparing cartilaginous fishes to other species, especially teleost fishes, is essential to reveal the ancestral and derived evolutionary features of adaptive immunity[6,9].

Research on the evolution of immune genes in cartilaginous and bony fishes has relied heavily on cloned genes or, more recently, whole-genome sequencing data[10–13]. Cartilaginous fishes possess an immunoglobulin (Ig)-based adaptive immune system and have two conventional immunoglobulin classes (i.e., IgM and IgW), and a specific Ig new antigen receptor (IgNAR)[14–16]. The discovery of phagocytic B cells in teleost bridges innate and adaptive immunity[17] and warrants a more comprehensive analysis of B cell types in cartilaginous and bony fishes. In teleost, αβT cell receptors (αβTCRs) function as mechanosensors of CD4+ and CD8+ T cells, and evolutionarily conserved αβTCRs have been found in cartilaginous fishes[18]. A CD4 molecule with a different structure to bony vertebrates and the absence of some related cytokines in the shark genome suggest a primitive or atypical T-helper-cell (Th-cell) system in sharks[12,19–21]. Genomic analyzes have also revealed a close linkage of immunoglobulin and T cell receptor genes in cartilaginous fishes, suggesting that their antigen receptor genes evolved from a common ancestor[12,22,23]. Another important component of adaptive immunity is an antigen presentation system dominated by MHC molecules. The MHC presents endogenous and exogenous antigens to the cell surface that are further recognized by the BCR/TCR system[24]. MHCI and MHCII are recognized by CD8 and CD4, respectively, during T cell activation. Taken together, while genomic and evolutionary data have provided crucial insights into the evolution of the key adaptive immune genes in fishes, the cell types and their patterns of expression across various species remain unclear, as are the ancestral and derived evolutionary features of adaptive immunity.

Here, we used single-cell transcriptomic sequencing (scRNA-seq) to study the immune cells of a cartilaginous fish (the white-spotted bamboo shark, *Chiloscyllium plagiosum*) and two bony fishes (the zebrafish, *Danio rerio*, and the Chinese tongue sole, *Cynoglossus semilaevis*). Using scRNA-seq profiles, we systematically characterized the cell type architecture of the immune system and gene expression patterns of immune cells in these species. The evolutionary features of the immune cells were revealed through cross-species lineage construction and comparative analysis. Furthermore, subtype analysis of B and T cells in the shark revealed specific immunity characteristics of cartilaginous fishes and provided new insights into the origin and evolution of adaptive immunity.

Results

Immune cell architecture of cartilaginous and bony fish

To reveal the cellular architecture of the cartilaginous and bony fish immune systems, we performed scRNA-seq of the immune organs of the white-spotted bamboo shark, zebrafish, and Chinese tongue sole (Supplementary Fig. 1a, b). The spleen, liver, and kidney (in the case of the white-spotted bamboo shark) or the head kidney (specific to zebrafish and Chinese tongue sole) play crucial roles in immune responses and harbor a wide array of immune cells. We constructed a total of 60 single-cell libraries (Supplementary Data 1). Following stringent quality control, we obtained transcriptome data of 58,866 cells (Supplementary Fig. 1c and Supplementary Data 1).

Next, we integrated the cells from the different immune organs of each species (Fig. 1a–c). Our results showed that organ-specific cell types were mainly tissue parenchymal cells. In contrast, immune cells were present in all organs with only quantitative differences (Supplementary Fig. 1c). Cell types were identified through the analysis of marker gene expression (Fig. 1d–f and Supplementary Data 2), revealing 12, 17, and 16 cell types in bamboo shark, zebrafish, and Chinese tongue sole, respectively. We observed shared cell types among the three fish species, including hematopoietic stem cells (HSCs), B cells, T cells, and erythrocytes. The HSCs were defined by the transcription factor *meis1b* and the stem cell regulatory factor *ahnak*) (*ahnak-like* in bamboo shark). In the bamboo shark, HSCs were also marked by *lmo2*, a crucial factor in hematopoietic development[25]. HSCs were identified by the transcription factors *mef2aa* in zebrafish and *zeb2* in Chinese tongue sole. We also identified hematopoietic progenitor cells (HPCs) by the stem cell-related genes *myb*, *npm1a*, and *csf1rb* in zebrafish, and *npm1a*, *myb*, and *kita* in Chinese tongue sole. B cells were characterized by *blnk*, *btk*, *flt3*, and *igll1-like* in bamboo shark, *cd79a* and *pax5* in the two bony fishes. T cells were defined by *tcf7*, *cd247*, *zap70*, and *ccr7* in the bamboo shark and *zap70* in the bony fishes. We further defined bamboo shark myeloid cells from the expression of myeloid genes such as *mpeg1-like*, *mmp9*, and *fam133b*. In zebrafish, we also identified different myeloid cells, such as M1 macrophages (*cd74a*, *irf8*, and *mpeg1.1*), M2 macrophages (*grn1*, *marco*, and *mfap4*), neutrophils-I (*mmp9*, *mmp13a*, and *adam8a*), neutrophils-II (*adam8a*, *lect2l*, and *cpa5*) and eosinophils cells (*gata2a*, *gyg1a*, and *viml*). In Chinese tongue sole, we identified M1 macrophages (*cd74a*, *irf8*, and *cd74b-like*), M2 macrophages (*marco*, *csf3r*, and *ctss2.1*), neutrophils-I (*mmp9* and *cebpb*), neutrophils-II (*cpa1*, *mmp9*, and *cebpb*) and dendritic cells (*flt3*, *s1pr4*, and *spi1*). Furthermore, we defined erythrocytes by the marker genes including *blvrb*, *cahz-like*, and *rhag-like* in bamboo shark, and *cahz* and *alas2* in two bony fishes. And we defined a cluster of proliferating cells with cell cycle related genes including *top2a*, *ube2c*, and *mki67-like* in bamboo shark, and *pcna* and *mki67* in zebrafish, and *top2a* and *ube2c* in Chinese tongue sole. Other non-immune cells identified were hepatocytes (*fabp10a*), endothelial cells (*kdrl*), acinar cells (*prss1*), and goblet cells (*gata3*, *agr2*, and *cldn7-like*) (Supplementary Fig. 1d–f). We also generated expression profiles of additional marker genes for each cell type in the three species (Supplementary Fig. 2a–c). We analyzed the proportion of immune cell types from the three species, revealing a higher proportion of B and T cells in the bamboo shark (Supplementary Fig. 2d). Our results indicated that the two bony fishes share more marker genes than they do with the cartilaginous bamboo shark. This indicates a high conservation of immune cells in the two bony fish species examined.

Evolution of immune cells in cartilaginous and bony fish

The expression patterns of one-to-one orthologous genes allow the study of immune system evolution[26,27]. In this study, we analyzed the transcriptional similarity of immune cells from three species and the expression of 7337 one-to-one orthologous genes (Fig. 2a and Supplementary Data 3). Same cell types across different species exhibited similar transcriptional profiles (Fig. 2a). For example, all T cells from the three species clustered together and then clustered with B and NK cells (Fig. 2a). This pattern indicates a conserved expression pattern of orthologs in the lymphocyte lineages from cartilaginous to bony fishes. Furthermore, although previous studies have shown that erythrocytes and myeloid cells differentiated from common myeloid progenitors[28], we found a higher similarity between myeloid and lymphoid cells than between erythrocytes and myeloid cells (Fig. 2a). Similar results were observed comparing all immune cells from different organs of the three fish species (Supplementary Fig. 3).

Next, we constructed cell-type phylogenies with single-cell transcriptomes and performed pairwise comparisons (Supplementary Fig. 4a). Generally, we found that the broad immune cell types were in the same branch across species. We also found a one-to-one connection between cell types such as erythrocytes, HSCs, and B cells, demonstrating the higher similarity of these cells (Supplementary
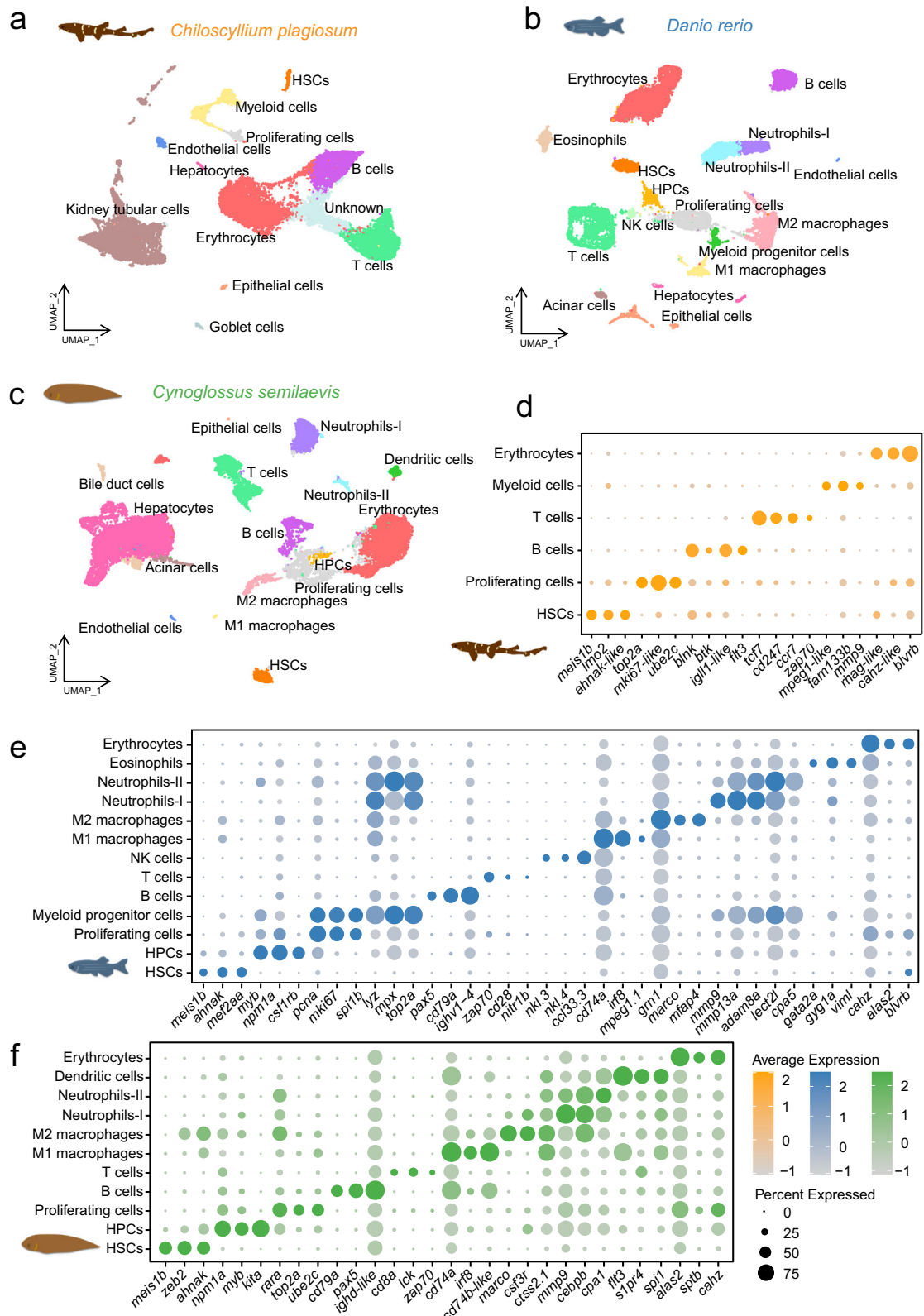
Fig. 1 | Single-cell atlas of immune organs from white-spotted bamboo shark, zebrafish, and Chinese tongue sole. a–c UMAP visualization of all cells. Bamboo shark (**a**), zebrafish (**b**), and Chinese tongue sole (**c**). Each dot represents one cell, with labeled cell types being the predominant cell types in each cluster. **d–f** Dot plots showing the expression of key marker genes (x-axis) of major cell types (y-axis) in bamboo shark (**d**), zebrafish (**e**), and Chinese tongue sole (**f**). The depth of the color from light to dark and the size of the dots represent the average expression from low to high and the percentage of cells expressing the gene.
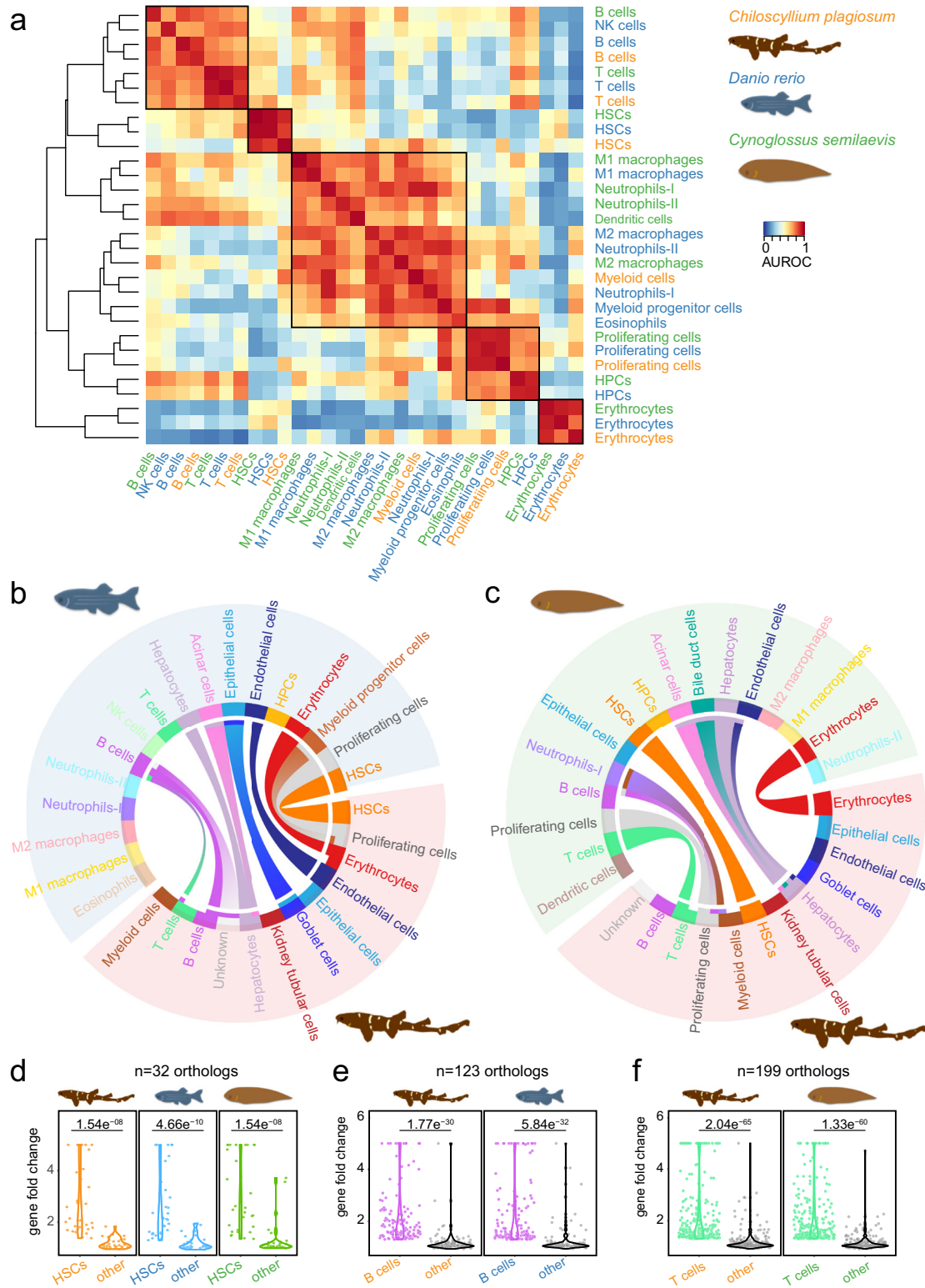
Fig. 2 | Analysis of the cell-type evolution of three species. a Heatmap displays the 'one versus best MetaNeighbour' scores for immune cell types across three distinct species. AUROCs are determined by computing between the two closest neighbors in the test dataset, with the assumption that proximity equates to a higher scoring relationship. **b, c** Cross-species pairwise cell type similarities between bamboo shark and zebrafish (**b**), bamboo shark and Chinese tongue sole (**c**), based on Kullback-Leibler divergence (KLD). The top 5% highest values are depicted as arches linking different cell types, where the breadth of each arch corresponds to the degree of KLD-based similarity. **d**−**f** Violin plot showing the normalized expression of top shared orthologous genes (FC > 1.3, *P*-value < 0.001) in highly similar cell types across species, compared to the average expression of the same genes in all other cell types. Significance was calculated using paired Wilcoxon test between in- and outgroup cell types, Wilcoxon signed rank test adjusted *P*-value < 0.001 for all comparisons. **d** HSCs. **e** B cells in bamboo shark and zebrafish. **f** T cells in bamboo shark and Chinese tongue sole.

Fig. 4a). We proceeded to examine the similarities among cell types across species using the Kullback-Leibler divergence (KLD)[29,30] (Fig. 2b, c, and Supplementary Fig. 4b). We found that some cell types are similar to their counterparts across species (Fig. 2b, c, and Supplementary Fig. 4b). For example, HSCs were highly conserved among the three species, characterized by co-expression of 32 orthologs of stemness maintenance-related genes and transcription factors such as *meis1b*, *gfi1b*, and *tal1* (Fig. 2d, Supplementary Fig. 4c, and Supplementary Data 4). The proliferating cells co-expressed 104 orthologs (e.g., *mki67*, *cdk1*, *pcna*, and *mcm7*) involved in DNA replication and cell cycle regulation (Supplementary Fig. 4c, d and Supplementary Data 4). Interestingly, proliferating cells also expressed epigenetic enzymes such as *dnmt1*, *dek*, and *smarca5*. In addition, erythrocytes were also highly conserved among the three fish species, manifested as the co-expression of 36 orthologs, including the transcription factors *gata2*, *zfpm1*, and *lmo2* and erythrocyte marker genes (e.g., *epor*, *urod*, and *fech*) (Supplementary Fig. 4c, d and Supplementary Data 4).

We found that the species-specific orthologs in the same cell type contribute to a similar function. For example, there were 154, 130, and 146 orthologs specifically expressed in the proliferating cells of bamboo shark, zebrafish, and Chinese tongue sole, respectively (Supplementary Fig. 4c and Supplementary Data 4). Gene enrichment analysis revealed an intersection of GO terms or pathways such as 'DNA Replication', 'Cell cycle', and 'Metabolism of RNA'. For 'DNA Replication', we detected *topbp1*, *orc1*, and *parp1* in bamboo shark proliferating cells; *prim1*, *orc4*, *dscc1*, and *chrac1* in zebrafish; and *pole*, *polq*, and *pold3* in Chinese tongue sole. This indicates a similar functional gene expression pattern of these cell types among the three species. We also found some species-specific characteristics. For example, genes responsible for cartilage morphogenesis (e.g., *chsy1*, *ids*, and *furina*) were specifically expressed in the HSCs of bamboo sharks (Supplementary Data 4).

Bamboo shark immune cell types showed the greatest differences in transcriptomic profiles relative to the two bony fish species based on the number of connections between immune cells of different species, consistent with its more distant phylogenetic position (Fig. 2b, c, and Supplementary Fig. 4b). For example, HPCs, B cells, T cells, neutrophils, and M1 and M2 macrophages showed high similarity between zebrafish and Chinese tongue sole, along with a large number of shared orthologous genes (Supplementary Fig. 4e and Supplementary Data 4). We also observed partial similarity of immune cells between the cartilaginous fish and teleost fishes. A high similarity of B cells between the bamboo shark and zebrafish was characterized by 123 co-expressed orthologs (Fig. 2e). Furthermore, T cells were similar (199 co-expressed orthologs) between the bamboo shark and Chinese tongue sole (Fig. 2f). Surprisingly, T cells from the bamboo shark and B cells from zebrafish shared 136 co-expressed orthologs (Supplementary Fig. 4f). This indicates a close linkage of B and T cells in the bamboo shark, revealing its divergent adaptive immune system.

Conserved cell function with divergent gene sets in B cells of cartilaginous and bony fish

Both cartilaginous and bony fishes possess Ig-based adaptive immune systems[5]. The shared orthologs of DEGs between cartilaginous and bony fish indicate conservation of B cells (Fig. 2e and Supplementary Data 4). To better understand the conservation and diversification of B cells between cartilaginous and bony fish, we compared the B cells from the three species in our dataset. The transcriptional profiles of 7,184 orthologous genes from 900 B cells from each species were integrated and visualized using UMAP (Fig. 3a).

Generally, B cells from the three species studied were evenly distributed in the integration results, indicating a high similarity of B cells among species (Supplementary Fig. 5a). We identified six B cell subsets (Supplementary Fig. 5b and Supplementary Data 5). The B-5 cluster was characterized by the expression of *rag1* and *rag2* and

transcription factors such as *zeb2b* and *rbx1*, indicating that they are B cell precursors (pre-B cells) (Fig. 3b). DNA replication and cell cycle genes, including *lig1*, *pcna*, *mcm7*, *slbp*, and *dnmt1*, were highly expressed in the B-3 cluster, indicating that B-3 are plasmablasts (i.e., plasma cell precursors). The B-6 cluster expressed plasma cells marker genes including *pdia4*, *ppib*, *ostc*, and *ddost*. The B-1 cluster was characterized by a high expression of *mef2cb*, *sf3b1*, *kmt2a*, and *smc5*, indicating that they are immature B cells. Genes including *cd74a*, *rpl37*, and *rps14* characterized the B-2 cluster. Interestingly, the B-4 cluster expressed phagocytosis-related genes such as *mmp9*, *wbp4*, *prdx6*, and *sod1* (Fig. 3b). Although the same cell types were identified in the three species, the proportion of cells in the six clusters differed for each species (Supplementary Fig. 5c). For example, the proportion of B-1 cells was about 25% in the two bony fishes but more than 52% in the cartilaginous fish.

We further analyzed the similarity of B cells (Supplementary Fig. 5d). B-3 were the most similar among the three species, followed by the B-6 sub-cluster (Supplementary Fig. 5d). Analysis of shared DEGs among the three species also showed a higher similarity of B-3 and B-6 (Fig. 3c and Supplementary Data 6). The 217 shared DEGs among the three species included 144 DEGs in B-3 (Fig. 3d) and 32 in B-6 (Fig. 3e). The 144 shared B-3 DEGs in three species involved cell cycle proteins and transcription factors, including *dkc1*, *ybx1*, and *ncl* (Supplementary Fig. 5e).

Previous studies have revealed that cartilaginous fish possess three types of antibodies, IgM, IgW, and IgNAR, different from those in teleost fish. We annotated six Ig heavy chain constant region encoding genes in the bamboo shark genome. *IgNAR* (LOC122543616) was highly expressed in the B cells of the white-spotted bamboo sharks. While *IgNAR* (LOC122543617 and LOC122543620), *IgM* (LOC122552345), and *IgW* (LOC122547843 and LOC122547844) were relatively lowly expressed (Supplementary Fig. 5f and Supplementary Data 7). Interestingly, the ostensible immunoglobin-secreting plasma cell cluster B-6 shared 32 DEGs including protein folding-related genes (e.g., *pdia4*, *ppib*, and *hspa5*) and genes related to protein localization in the endoplasmic reticulum (e.g., *erp44*, *hyou1*, and *derl2*) (Fig. 3e and Supplementary Data 6). Shared high expression of protein glycosylation modification-related genes (e.g., *ost4*, *ddost*, and *rpn2*) was also observed in B-6 (Fig. 3f). Meanwhile, we observed high expression of RNA splicing-related genes *snw1*, *snrpf*, and *thoc5* and mitochondrial translation elongation related genes *mrpl3*, *gfm1*, *oxa1l*, and *mrpl2* in B-6 of bamboo shark, indicating the unusual regulation of plasma cells in cartilaginous fish (Supplementary Fig. 5g and Supplementary Data 6).

We also analyzed 269 teleost-specific (zebrafish and Chinese tongue sole) DEGs in the six B cell sub-clusters (Fig. 3c). Interestingly, the shared DEGs and teleost-specific DEGs of the B-3 cluster showed enrichment for the gene ontology categories 'Metabolism of RNA', 'Proteasome' and cell cycle-related processes. 'Protein processing in endoplasmic reticulum' and 'protein glycosylation' were enriched by the shared DEGs and 'protein localization to endoplasmic reticulum' was enriched by the teleost-specific DEGs of the B-6 cluster. Species-specific DEGs (786 in bamboo shark, 234 in zebrafish, and 1008 in tongue sole) revealed the gene expression diversity in each species (Supplementary Fig. 5g). The enrichment of pathways and ontologies by different gene sets reflect that gene adaptations in one species can be substituted with other genes in different species while performing similar cellular functions (i.e., functional convergence).

Characterization of phagocytosis of B cells in cartilaginous fish

Previous studies have demonstrated that B cells in teleost fish are involved in innate immune functions such as phagocytosis[10]. In agreement, many innate immunity genes were found in the signature genes of B cells in zebrafish and Chinese tongue soles (Supplementary Fig. 6a, b). Notably, the genes specifically expressed in bamboo shark B
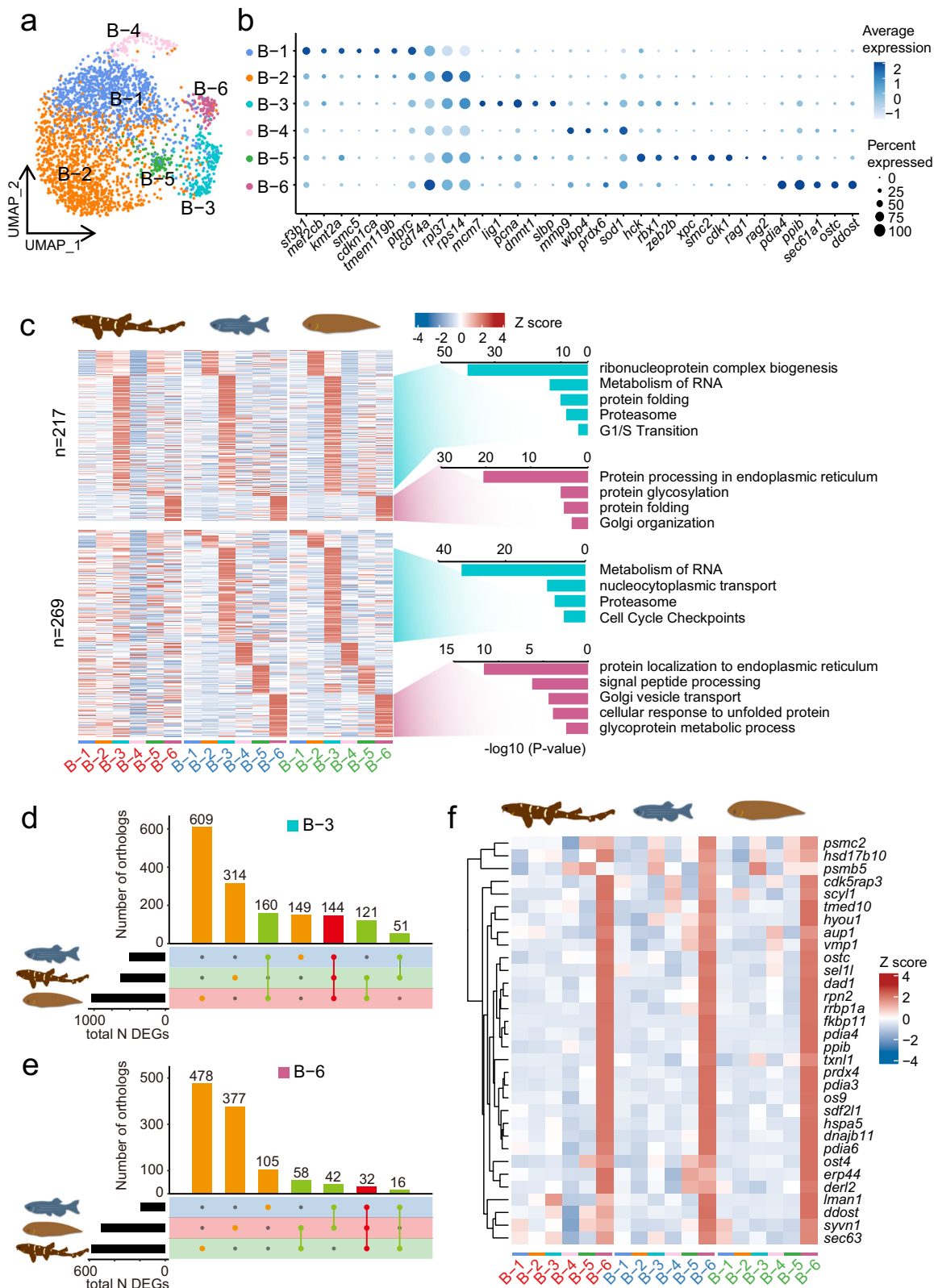
Fig. 3 | Conservation and divergence of B cell sub-types. a UMAP plot showing the clustering of B cell subclusters from three species. **b** Dot plot showing the expression of most significant DEGs in each sub-cluster of B cells. **c** Heatmap showing the conserved and teleost-specific differentially expressed genes in B cell sub-clusters. Enriched Gene Ontology categories for DEGs of B-3 (colored green) and B-6 (colored purple), respectively. The enrichment analysis was generated using Fisher's exact test on the Metasape web server, with Bonferroni correction for multiple hypotheses testing. *P*-values are indicated on the *x*-axis. **d, e** UpSet plots showing shared orthologs of three species in B-3 (**d**) and B-6 (**e**). **f** Heatmap showing the expression of conserved DEGs in B-6 sub-cluster of three species.
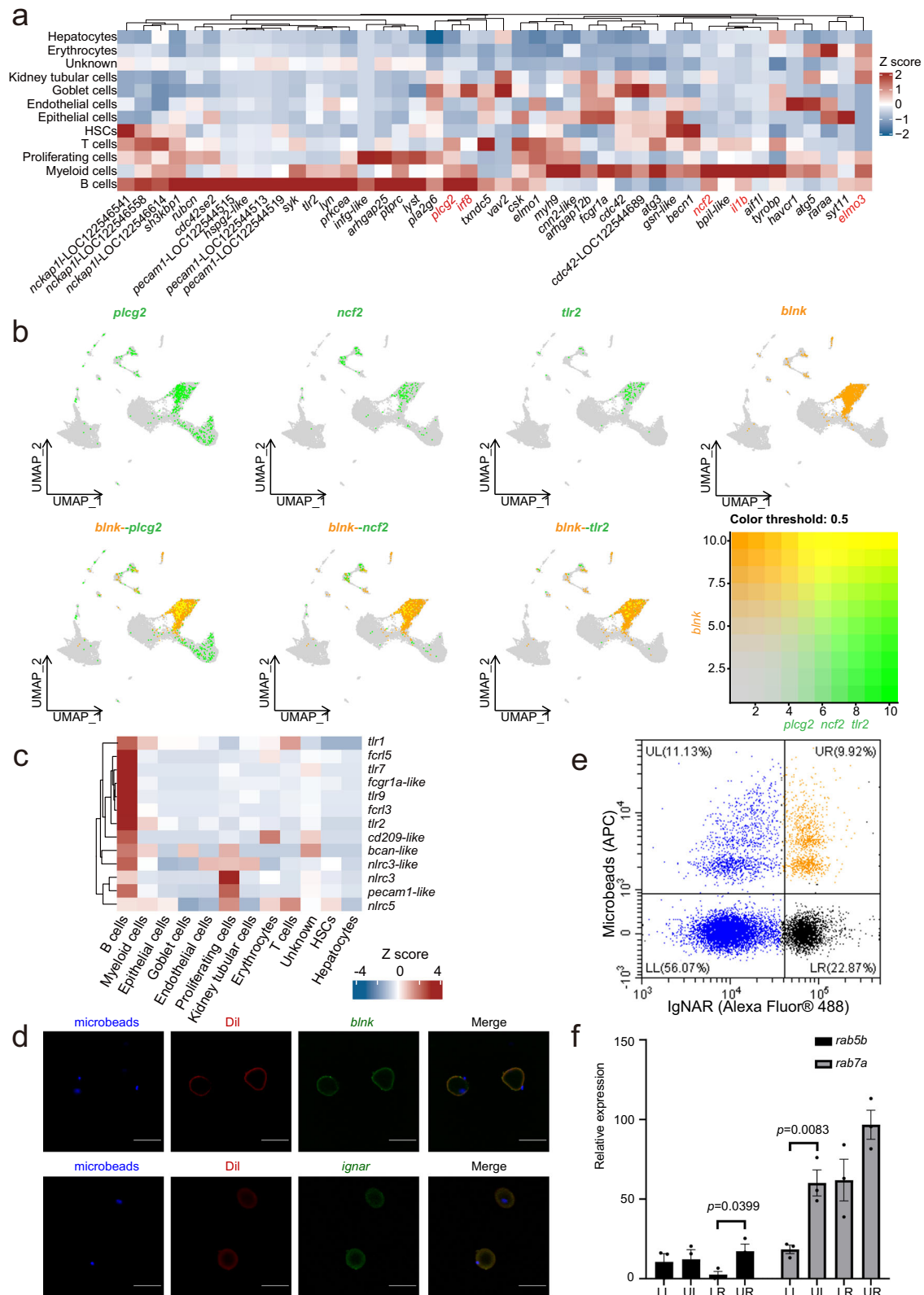
cells are also associated with innate immunity (Supplementary Fig. 6c). Pathway enrichment analysis revealed that 'Fc-gamma receptor (FCGR) dependent phagocytosis' and 'Regulation of actin dynamics for phagocytic cup formation' signaling pathways were enriched in B cell of bamboo sharks (Supplementary Fig. 6d). Furthermore, compared with bony fish, the bamboo shark B cells exhibit upregulation of genes involved in phagocytosis (such as *rhoh*, *met*, *pikfyve*, and *tmem175*) as

well as myeloid cell differentiation (*kmt2a*, *sf3b1*, *ncor1*, *ddx46*, *pdcd2*, etc.) (Supplementary Fig. 7). These results suggest that the potential phagocytic function of B cells present in cartilaginous fishes.

We generated a phagocytosis-related gene set using the GO terms of the zebrafish, mouse, and human[31] (Supplementary Data 8). Many phagocytic genes showed high expression in the B cells of the bamboo shark, including *pecam1*, *nckap1l*, and *rubcn*(Fig. [4]a). Both B cells and

Fig. 4 | Phagocytic function of B cells in bamboo shark. a Heatmap showing the expression of phagocytosis gene set in B cells of bamboo shark. **b** UMAP plots showing co-expression of *blnk* with *plcg2*, *ncf2*, and *tlr2*, respectively, in B cells of bamboo shark. The co-expression color threshold for all genes is based on the same standard of 0.5. **c** Heatmap showing the expression of pattern recognition receptors in B cell. **d** Confocal laser scanning microscopic images of phagocytosis of fluorescence-labeled microbeads by B cells (n = 3 biological replicates), B cells were detected by fluorescence in situ hybridization (FISH) with B cell markers (*blnk* and *ignar*). Cell membranes were stained with Dil. Scale bar = 10 μm. **e** Flow cytometry

identification of phagocytic B cells in bamboo shark spleen cells by the combination of fluorescence-labeled microbeads and anti-IgNAR antibody. LL, lower left: indicates IgNAR- non-phagocytic cells; UL, upper left: indicates IgNAR- phagocytic cells; LR, lower right: indicates IgNAR+ non-phagocytic cells; UR, upper right: indicates IgNAR+ non-phagocytic cells. **f**, RT-qPCR analysis of the expression of phagocytotic gene in FACS-sorted cells in Fig. 4e. n = 3 independent biological replicates. Bars represent mean values +/− SEM. *P*-values were determined using the two-tailed Mann−Whitney test.

myeloid cells expressed *plcg2*, *irf8*, *ncf2*, *il1b*, *elmo3*, etc. The phagocytic genes *plcg2*, *tlr2*, and *ncf2* were co-expressed with B cell marker genes such as *blnk* and *flt3* (Fig. 4b and Supplementary Fig. 8a). Interestingly, bamboo shark B cells highly expressed several pattern recognition receptors. Fc-gamma receptors (e.g., *fcgr1a-like*, *fcrl3*, and *fcrl5*) and toll-like receptors (e.g., *tlr1*, *tlr2*, *tlr7*, and *tlr9*) were highly expressed in B cells (Fig. 4c). Considering the different types of Ig genes, we analyzed the co-expression of Ig genes and phagocytic genes (*plcg2*, *tlr2*, and *ncf2*), and found a high co-expression of *IgNAR* (LOC122543616) with phagocytic genes (Supplementary Fig. 8b–f).

Next, we investigated the phagocytosis of B cells in the bamboo shark by flow cytometry. Forward and side scatter parameters (FSC/SSC) were used to gate lymphocytes. After incubating with microbeads, 3.27% of lymphocyte cells gave a positive result. An increased microsphere-cell ratio increased the proportion of fluorescence-positive lymphocytes to 6.23% (Supplementary Fig. 9a). Next, we observed the phagocytosis of B cells by confocal microscopy. Microspheres with blue fluorescence were found in the cytoplasm of *blnk*-positive or IgNAR-positive cells, indicating that B cells had ingested microbeads (Fig. 4d). Next, we sorted both IgNAR+ phagocytotic, IgNAR- phagocytotic cells, and non-phagocytic cells by using a combination of fluorescence-labeled microbeads and anti-IgNAR antibody by FACS (Fig. 4e and Supplementary Fig. 9b). The increased expression of phagosome maturation genes including *rab5b*, indicated the activation of phagocytosis (Fig. 4f).

We also detected the expression of these pattern recognition receptors in B cells from zebrafish and Chinese tongue soles (Supplementary Fig. 9c, d). Many receptors were expressed in B cells in the two bony fishes. Flow cytometry showed that only about 0.36% of lymphocyte cells in the Chinese tongue sole were fluorescence-positive after microbead incubation (Supplementary Fig. 9e). We also observed the phagocytosis of Chinese tongue sole B cells by confocal microscopy (Supplementary Fig. 9f). These results hint at an enhanced phagocytosis function of cartilaginous fish B cells compared to bony fish.

Identification of T cell subsets in cartilaginous fish

Much effort has been expended to identify CD4 genes and related cytokines in cartilaginous fishes to understand the function of T cell subtypes better[12,21]. We explored the main functions of T cells by KEGG pathway enrichment. DEGs of T cells in the bamboo shark were enriched in pathways related to CD4+ T cell subsets such as 'Th17 cell differentiation' and 'Th1 and Th2 cell differentiation' (Fig. 5a).

We divided T cells from the bamboo shark into six subclusters (Fig. 5b; Supplementary Fig. 10a, b and Supplementary Data 9). We analyzed the expression of TCR genes and found high *tcrb* expression and low *tcrd* expression (Supplementary Fig. 10c, d, and Supplementary Data 7). We then annotated each subset based on DEG expression (Fig. 5c). T-0 cells were characterized by tissue-resident memory T cell-related genes such as *itgae-like* and *btnl2*. The naïve T cell marker genes *lef1*, *ccr7*, and *tcf7* were highly expressed by T-1 cells. T-2 cells primarily expressed transcription factors (*junb-like*[32] and *aiolos-like*[33]), and CD4+ T cells differentiation-related genes (*il12rb2*[34] and *sh2d1a-like*[35]), exhibiting characteristics of the helper T cell subset. Interestingly, T-3 cells highly expressed many genes associated with B cells, including *blnk*

and *igj-like*. The genes *flt3* and *tmem119b*, which were DEGs of bamboo shark B cells (see above), were also expressed in this cluster. Cluster T-4 was characterized as proliferating T cells based on the high expression levels of cell proliferation-related genes (e.g., *pcna*, *myc-like*, and *eif1ad*). T-5 cells specifically expressed T cell activation-related genes such as *myb*, *egr2b* and *tagapb* (Fig. 5c).

Many studies have investigated the presence of canonical CD4 and T cell lineage diversification[36–38]. Our study aimed to elucidate the diversification of the T cell lineage in the white-spotted bamboo sharks using single-cell data. We found that T-2 with helper T cell characteristics is *CD8a* negative (Fig. 5d). We identified the sequence of CD4 gene in white-spotted bamboo shark (Supplementary Fig. 11) based on a previous study[21]. However, the expression of the CD4 gene was not subset specific (Supplementary Fig. 12a), which may be related to the structural differences (stemming from a highly diverged gene locus and protein sequence) of CD4 in cartilaginous and bony fish. Like other species, MHCI genes were expressed by most immune cell types in the bamboo shark, while MHCII genes were specifically expressed in antigen-presenting cells (Supplementary Fig. 12b).

Antigens presented by the MHC are recognized by different T cell subsets, and through the cascade signaling of *lck* and *zap70*, ultimately lead to T cell activation and differentiation[39]. *Lck* was primarily expressed in the T-2 subset and *zap70* was ubiquitously expressed in all T cell sub-clusters, indicating different mechanisms of T cell activation in the bamboo shark (Fig. 5d). In addition, some Th1-related genes were expressed in the T-2 subset, including *ifng-like*, *cxcr3-like*, and *cxcl10-like* (Fig. 5e). However, we failed to detect the expression of other Th2 effectors, including *il12* and *tnfa* in T-2. Furthermore, *ccl11-like* and *ccl20-like*, involved in the IL-17 signaling pathway[40], were also expressed in the T-2 subset (Supplementary Fig. 12c).

Transcription factors are crucial in determining cell fate. We analyzed the expression of transcription factors of bamboo shark T cells (Fig. 5f). Some of the signature transcription factors of CD4+ T cells were expressed (Fig. 5f and Supplementary Fig. 12d). Both the CD4+ T cell-related transcription factor *gata3* and the CD8+ T cell-related transcription factor *runx3* were ubiquitously expressed in all T cell subsets (Fig. 5g). The transcription factor ZBTB7B is essential for the development of the CD4 lineage, and its lack results in the commitment to the CD8 lineage[41]. We confirmed that the *zbtb7b-like* gene in the bamboo shark is intact at its conserved domain; however, the gene was not specifically expressed by T cells (Fig. 5g). In conclusion, we have characterized the subsets of T cells in a cartilaginous fish and further elucidated the function of the T-2 subset in a bamboo shark. Our results indicate that T-2 is a Th-like subpopulation in sharks. Moreover, it may have cell fate determinants that differ from those identified in bony vertebrates.

A T cell subtype co-expressed T and B cell markers in cartilaginous fishes

Unexpectedly, the differentially expressed genes of the bamboo shark T cell subcluster T-3 included some B cell-related genes (Fig. 5c). Genome studies have reported a close linkage between immunoglobulin and T-cell receptor genes[12]. To demonstrate that our result did not occur due to cell admixture, we used co-expression analysis to verify that these genes were expressed by T cells. Many cells in T-3
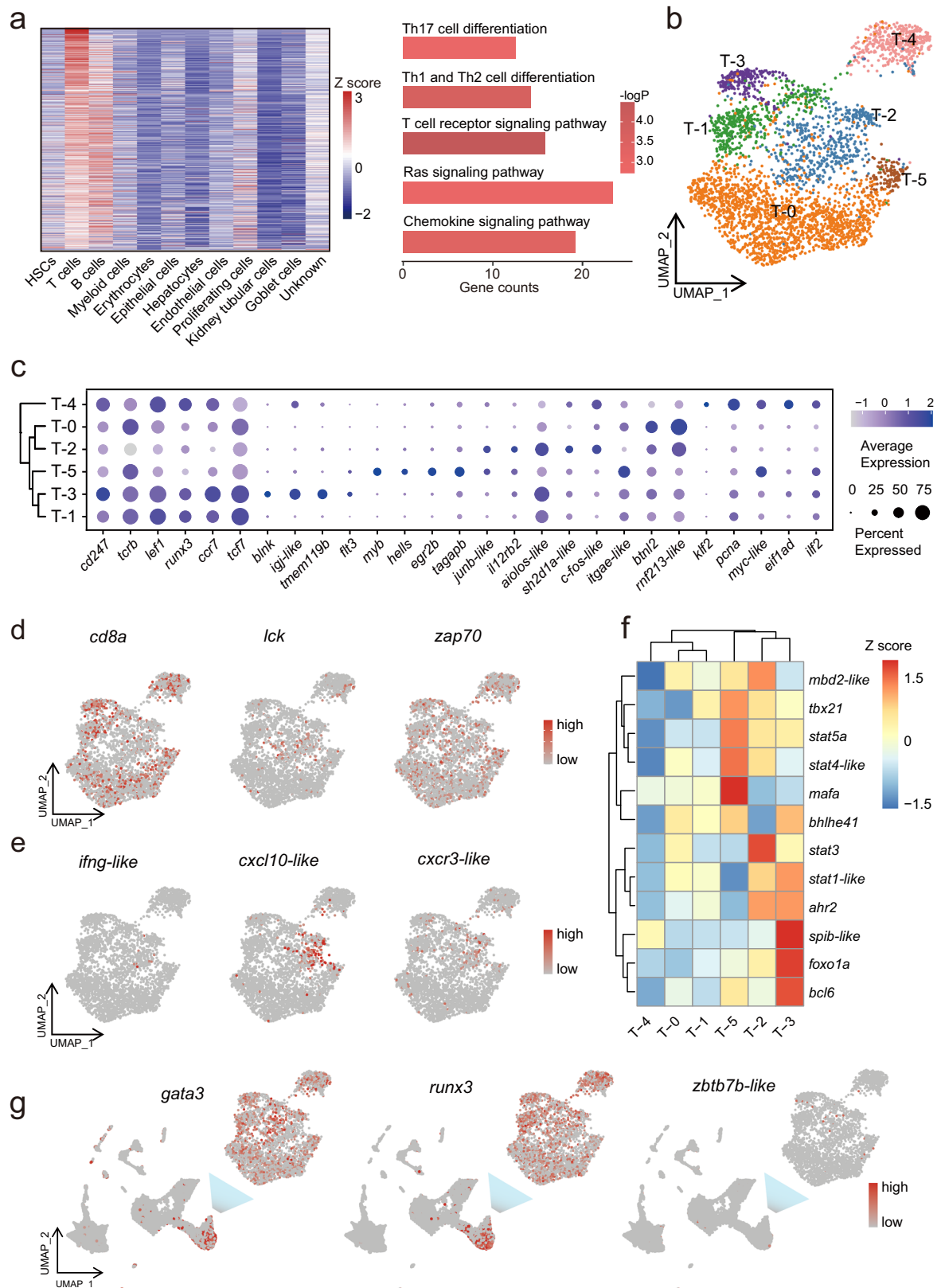
Fig. 5 | Characterization of T cell sub-types in the bamboo shark. a Heatmap showing the expression of all DEGs of T cells in bamboo shark. Gene expression levels utilize a Z score, which depicts variance from the mean, as defined on the color key in the right top corner (left). Histogram showing KEGG pathways enriched by these genes (right). The enrichment analysis was generated using Fisher's exact test, with Benjamini-Hochberg (BH) corrected for multiple hypotheses testing. The P-values is represented by the depth of red. b UMAP plot showing 6 T cell subsets of bamboo shark. c Dot plot showing expression levels of selected signature genes in T cell subsets. Dot size indicates the fraction of expressing cells, colored based on normalized expression levels. d UMAP plot showing T cell activation-related gene expression in different T cell subsets. e UMAP plot showing Th2-related gene expression across T cell subsets. f Heatmap showing CD4+ T cell subset transcription factor expression across T cell subsets. g UMAP plot showing CD8+ T cell and CD4+ T cell key transcription factor expression across T cell subsets and immune cell lineages.
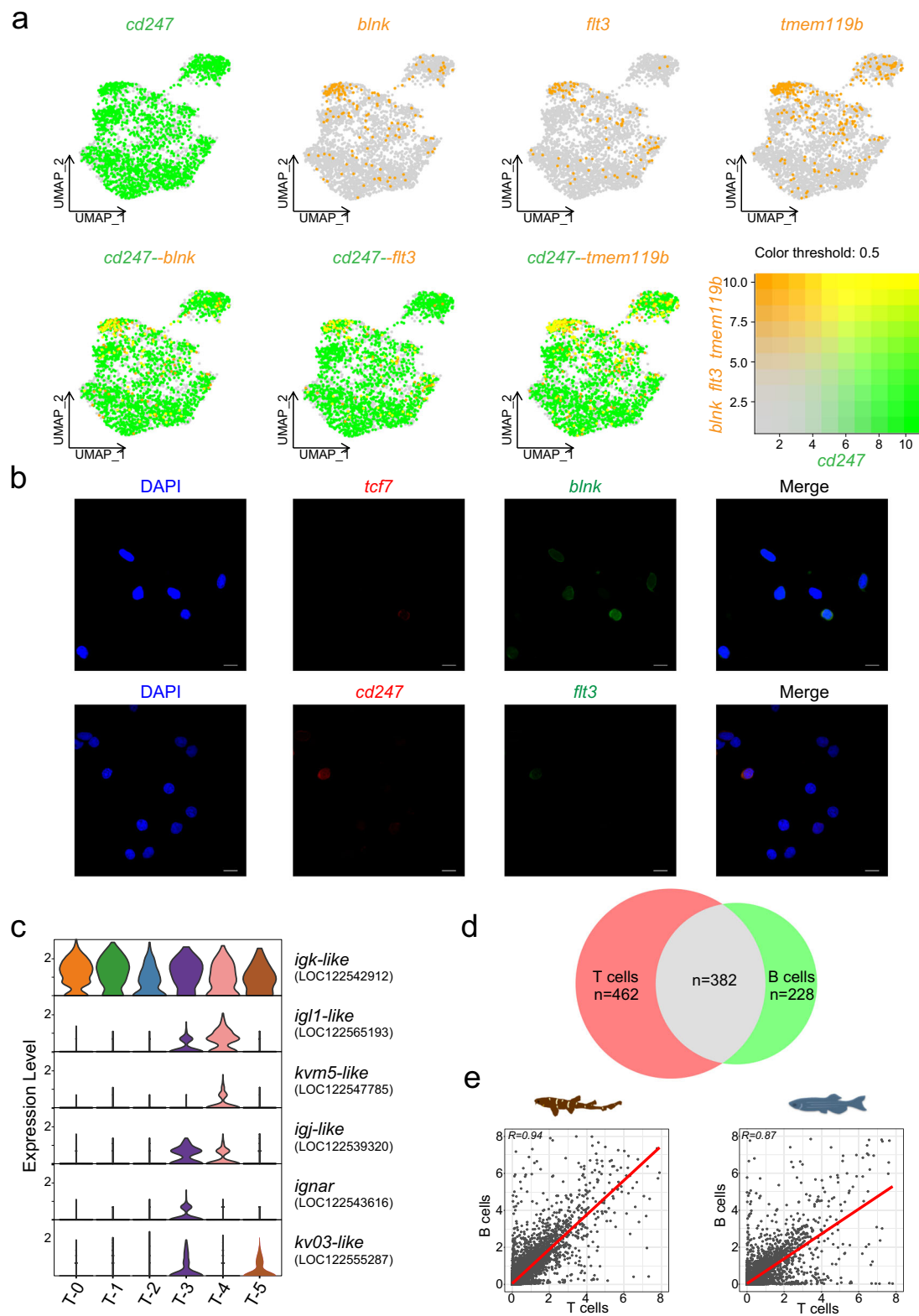
Fig. 6 | **Co-expression of T cell and B cell marker genes in T-3 subsets. a** UMAP plots showing co-expression of *cd247* and *blnk*, *flt3*, *tmem119b*, respectively, in T cell subsets of bamboo shark. The co-expression color threshold for all genes is based on the same standard of 0.5. **b** fluorescence in situ hybridization (FISH) showing co-expression of B cell markers (*blnk* and *flt3*) and T cell markers (*cd247* and *tcf7*). **c** Violin plots showing immunoglobulin gene expression in T cell subsets.

d Venn diagram showing the intersection of all DEGs in T cells and B cells of bamboo shark ($n = 3$ biological replicates), Scale bar = 10 μm. **e** Scatter plots show the correlation of gene expression in T cells and B cells in bamboo shark and zebrafish, respectively. The correlation coefficient *R* is calculated by Pearson correlation test (two-sided).

showed co-expression of *blnk*, *flt3*, and *tmem119b* with the T cell markers *cd247* (Fig. 6a). We verified the expression of *tcf7*, *cd247*, *blnk*, and *flt3* by fluorescence in situ hybridization (FISH), and detected the cells co-expressing *tcf7* and *blnk*, *cd247* and *flt3* (Fig. 6b). Furthermore, we also analyzed the cell populations that co-expressing *tcf7*, *cd247* and IgNAR (Supplementary Fig. 13a). In addition, many immunoglobulin genes were expressed by bamboo shark T cells, especially in the T-3 and T-4 clusters (Fig. 6c). We also found a substantial overlap between bamboo shark T and B cell DEGs (Fig. 6d). The correlation of the gene expression pattern of T and B cells from the bamboo shark was higher than that of zebrafish (Fig. 6e).

To further expand on our results, we analyzed the scRNA-seq dataset of the nurse shark (*Ginglymostoma cirratum*)[42] with the same quality control flow performed in this study. T and B cell markers co-expressed in a specific subset of T cells in nurse sharks (Supplementary Fig. 13b). Meanwhile, there is no subcluster co-expressing T and B cell markers in Chinese tongue sole or zebrafish (Supplementary Fig. 13c, d). This observation suggests that the T-3 subset (co-expressing T-cell and B-cell markers) appears to be specific to cartilaginous fishes. Furthermore, we added additional cross-species analysis. We extracted zebrafish and Chinese tongue sole T cells for clustering and identified several subgroups. Cross-species analysis revealed that there are no T cell subpopulations in zebrafish and Chinese tongue soles with similar transcriptional characteristics to T-3 in sharks (Supplementary Fig. 14). These results support the claim that the T-3 subset is exclusive to cartilaginous fishes. This cell type may represent a unique branch of T cell evolution, conferring specific immune advantages to cartilaginous fishes. However, whether the T-3 is a unique group that distinguishes cartilaginous fish from all bony fishes still needs to be determined. Further expansion of the species range for in-depth comparative analysis is needed.

Discussion

The immune landscape of lamprey has revealed the heterogeneity of lymphocytes between jawless and jawed vertebrates[43]. The cell atlases of a cartilaginous fish and two bony fishes generated here will further advance our understanding of the evolution of immunity in jawed vertebrates. Previous studies of genomes indicated that investigating the immune system of cartilaginous fish is essential for revealing the origin and evolution of adaptive immunity[12,13]. By creating a cellular atlas of a bamboo shark using three organs (liver, spleen, and kidney), we have systematically revealed the composition of the immune system and gene expression profiles in cartilaginous fish. The main cell types and cell type-specific marker genes were conserved in three species. For example, *ahnak* and *meis1b* were specifically expressed in HSCs, and *zap70* were expressed in T cells. While *blnk* is highly expressed in the B cells of all three species, *pax5* is specifically expressed in the B cells of the bony fishes. Furthermore, we also found that *flt3*, an important cytokine receptor involved in hematopoiesis and early B cell development[44], is specifically expressed in the B cells of the cartilaginous fish. The construction of an immunology cell atlas presented in this study reveals the similarity and diversity of immune cell types in cartilaginous and bony fishes, enhancing our knowledge of the molecular mechanisms, cell lineage decisions and regulatory processes within the immune system. Considering the tissue-specific distribution of immune cells, an integrated cell atlas of all the immune-related tissue and a more detailed cell type identification is needed.

In mammals, phagocytosis is generally completed by specialized phagocytes such as monocytes, neutrophils, and macrophages[45]. Moreover, phagocytosis in the literature has gradually evolved from a role only in innate immunity to a role in adaptive immunity[46,47]. It plays a key role in the uptake and destruction of microorganisms and initiating and developing the adaptive immune response[47]. Previous studies have shown that malignant mouse and human B cell lines can be transformed into macrophage-like cells with the ability to phagocytize

large particles[48,49]. Later, it was found that bony fish B cells effectively carry out phagocytosis both in vivo and in vitro. It does so by inducing a 'downstream' degradation pathway by ingesting particles, resulting in the fusion of lysosomes and phagosomes to form phagocytic lysosomes, which then kill microorganisms in cells[10].

We found that cartilaginous fish B cells expressed many genes involved in innate immunity and phagocytosis. We verified that bamboo shark B cells also possess the ability to phagocytize immune microspheres. It is worth noting that several vertebrate B cells can phagocytize particles in a non-BCR-mediated manner but in the same way as traditional phagocyte pattern recognition receptors[46]. Interestingly, we also observed the expression of pattern recognition receptors in bamboo shark B cells. Whether the mode and mechanism by which cartilaginous fish B cells recognize antigen particles are the same as those of higher vertebrates remains to be studied. In conclusion, our exploration of the phagocytic function of B cells in cartilaginous fish illustrates the key role of B cells in connecting innate immunity and adaptive immunity. However, the potential phagocytic functions of B cell perhaps derived from the cooption of a gene regulatory network, and how it differs from the phagocytic function of the innate immune system need to be further explored.

Our single-cell sequencing results provide a new strategy for analyzing T cell subtypes of cartilaginous fishes. T helper 1-related cytokines, including IFN-γ, IL-12 and TNF-α, were found in the genome of the elephant shark (*Callorhinchus milii*), indicating the presence of helper T cell sub-types[12]. We found that MHCII tends to be specifically expressed in myeloid cells and B cells, while MHCI is ubiquitously expressed in all cell types. We propose the existence of an antigen-presenting signaling pathway mediated by MHCII and TCR in cartilaginous fishes. The lack of co-expression of *lck* with CD8 and a broad expression of *zap70* in all T cells suggest a different activation mechanism of TCR in cartilaginous fishes. The absence of CD8 and specific expression of *ifng-like*, *cxcl10-like*, and *cxcr3-like* genes in T-2 suggest that this sub-cluster may be helper T cells. However, the expression of CD4 in the bamboo shark is not specific. Previous studies have found that the cytoplasmic tail of shark CD4 has a CxH motif instead of a CxC motif for (presumably) binding to LCK, which may lead to functional changes[21]. Further research on the shark TCR activation pathway, the T-2 subset, and *CD4* gene function will help us elucidate the characteristics of shark T cells. In addition, the genomes of the Atlantic cod[50] and anglerfish[51] suggest that some bony fishes also lack CD4 and related genes. More research is recommended to reveal the origin and evolution of CD4+ cells by comparing the immune systems of cartilaginous and bony fishes. The limited diversity of T cell subtypes and the non-specific expression of transcription factors in sharks suggest evolutionary divergence in the T cell lineage of cartilaginous fishes.

Jan Klein first proposed the idea that the immune system should be compartmentalized into conserved components when studying its evolution[52]. Our study found that the HSCs are highly conserved in three fish species. This is consistent with the previous studies, which show that the regulatory mechanism of HSCs is highly conserved across species, from zebrafish to mammals[53]. We found that lymphoid lineages, such as B and T cells, have similar connections between cartilaginous and bony fishes (here, two distantly related teleost species). This suggests that adaptive immunity in cartilaginous fish is highly consistent with that in bony fishes. We found that B cells were evenly distributed after integrating data from our three study species, indicating that the subtypes of B cells in cartilaginous fish are the same as those in teleost fish. Among them, B-6 cells of the three species share genes involved in protein folding and transport, such as *pdia3* and *ppib*. This shows that although the types of immunoglobulins differ, antibody synthesis and processing mechanisms are conserved. When comparing all cell types of the three species, proliferating cells were found to have the most homologous genes. Similarly, B-3 (the

proliferating B sub-cluster cells) had the highest similarity among the three species compared to other B cell subsets. These results suggest that genes involved in basic life processes, such as DNA replication and cell division, are more conserved across species. This prompts us to think about the key factors that drive evolution and the most important gene sets responding to the driving factors. The evolution of additional functions that may respond to environmental changes is more important for forming new cell types. Cell-type-specific genomic information (coding and non-coding) plays a pivotal role in defining the unique identity of each cell type. Therefore, more attention should be given to investigating such specific gene sets.

Classically, lymphoid and myeloid cells are differentiated from HSCs[54]. Our research found that, like bony fishes, cartilaginous fish B cells also have a phagocytic function, indicating an ancient connection between the lymphoid and myeloid lineages in cartilaginous fishes. We revealed B and T cell markers in the T-3 cluster, hinting at a close link between T cells and B cells. This result is compatible with previous research showing a close linkage of immunoglobulin and T cell receptor genes in the elephant shark genome[12]. Our results also revealed limited T cell sub-types, supporting a unique cartilaginous fish differentiated T cell state. An increasing number of studies have reported the function of innate-like lymphocytes, including NKT cells, mucosal-associated invariant T cells, B1 cells, marginal zone B cells, and innate lymphoid cells[55,56]. Interestingly, the newly discovered paired immunoglobulin-like receptors-A (PIR-A) mediates the acquired memory of monocytes and macrophages to major MHCI antigens[57]. Therefore, our results add to the hypothesis that the emergence of adaptive immune cells originates from innate immune cells with a gain of a new function. The 'Big Bang' and 'gradual accumulation of changes' hypotheses should be investigated further.

In conclusion, our immune cell atlas of both cartilaginous and bony fishes provides data that supports a comprehensive understanding of immune system architecture and gene expression characteristics. Cell-type evolution lineage and comparative studies across species provide new perspectives on the origins and development of adaptive immunity at the single-cell resolution.

Methods

Ethics and experimental fish

The animal experimental protocol of this study was reviewed and approved by the Experimental Ethics Review Committee of the Yellow Sea Fisheries Research Institute of the Chinese Academy of Fishery Sciences (YSFRI-2023036). The white-spotted bamboo shark (*Chiloscyllium plagiosum*) approximately 25-40 cm in length was purchased from the fish and aquarium market in Xiamen, Fujian, China. The shark was maintained in a controlled environment with a water temperature of $25 \pm 1\,^\circ C$ and a seawater salinity 25‰ $\pm$ 0.5‰. The two-year-old Chinese tongue sole (*Cynoglossus semilaevis*) was obtained from an aquaculture factory in Qingdao, Shandong, China. The Chinese tongue sole was maintained in a controlled environment with a water temperature of $24 \pm 1\,^\circ C$ and a seawater salinity 16‰ $\pm$ 0.5‰. The Tubingen background experimental adult zebrafish (*Danio rerio*) were used and maintained in the circulation water system on a 14:10 h light-dark cycle at 28°C and fed twice daily. The sample information is shown in Supplementary Data 1.

Immune organ collection and cell sample preparation

Prior to immune organ collection, fish were placed in sterile collection chambers for 15 minutes. To obtain a high-quality cell suspension, we meticulously managed the preparation time. First, each fish was anesthetized with a concentration suitable for MS222. After anesthetization, fish were weighed. And briefly, each immune organ (liver, spleen, kidney/head kidney) of these three fish was carefully excised on ice.

For white-spotted bamboo shark and Chinese tongue sole, each organ was dissociated for one single-cell library. The organs were washed twice with DMEM medium (0.2% sea salt added). Other tissues adhered to the immune organs (liver, spleen, kidney/head kidney) were removed with tweezers. Each immune organ tissue was sliced with a blade into tissue homogenate. The tissue was digested in 0.05% trypsin and 0.1 mg/mL collagenase IV at 28 °C for 5 or 10 min (10 min for the liver and bamboo shark kidney, and 5 min for the remaining three organs). The digestion reaction was rapidly terminated by adding fetal bovine serum (FBS) to 10%. The cell suspension, after digestion, was passed through a 40 µm filter to eliminate tissue remnants. We then collected roughly 6 mL of this strained suspension into a 15 mL sterile tube. The tube was then centrifuged for 5 min at 500 g using a horizontal rotor in a low-temperature centrifuge. The cells were then washed twice and resuspended with PBS (0.05% BSA added). For zebrafish, cells were isolated from 3-5 animals for one spleen single-cell library, and about 10 animals for one head kidney single-cell library. The organs were washed with ice-cold PBS, then triturated with needle and syringe and digested in 0.05% Trypsin-EDTA and 0.1 mg/mL collagenase IV at 32 °C for 5 or 10 min (5 min for spleen and head kidney, and 10 min for liver). The digestion reaction was terminated by adding fetal bovine serum (FBS) to 10%. The digested cell suspension was washed with ice-cold 0.05% bovine serum albumin (BSA)/PBS buffer and filtered through a 40 µm filter to obtain single-cell suspensions[58]. All single-cell suspensions were counted utilizing a hemocytometer under microscopic observation, we enumerated the cells and adjusted their concentration to approximately 1000 cells per microliter. The cell viability was measured and recorded using a Trypan Blue Staining kit (Sangon Biotech).

Single-cell RNA libraries construction and sequencing

After each high viability (>90%) single-cell suspension of fish immune organ was prepared, the single-cell RNA-seq library was constructed immediately. For library construction, we used the DNBelab C Series Single Cell RNA Library Preparation Kit (MGI, China) according to the manufacturer's protocol, based on droplet microfluidics technology. In detail, single cells were prepared as droplets, and subjected to cell lysis and mRNA capture using the DNBelab C4 portable single-cell system. Subsequently, the microdroplets containing the captured mRNA were retrieved utilizing an emulsion-breaking recovery system. The mRNA bound to magnetic beads was then converted into cDNA, which underwent 16 cycles of PCR amplification for enrichment. Then, following the manufacturer's protocol, the enriched cDNA products progressed through a series of steps, including fragmentation, end repair, ligation, and an additional 12 cycles of PCR amplification. Following this, the cDNA was subjected to denaturation, circularization, and digestion procedures, ultimately yielding single-stranded DNA libraries. Finally, 10 ng of digested products were obtained and sequenced using the MGISEQ 2000 platform (BGI, China). The detailed information on the sequencing quality of each library was provided in Supplementary Data 1.

Identification of CD4 gene in the transcriptome of white-spotted bamboo shark spleen

RNA-seq datasets from spleen (GenBank SRR18297556-SRR18297559) of bamboo shark (*C. plagiosum*) were retrieved from the Bioproject "*Chiloscyllium plagiosum Raw sequence reads*" (GenBank PRJNA811401). Raw RNA-Seq data underwent rigorous quality control, employing the fastp tool, version 0.23.2, which effectively identified and removed any reads contaminated with adapter sequences or exhibiting substandard quality. The White-spotted bamboo shark transcriptome was de novo assembled from clean data using Trinity (version 2.8.5) to generate a contig file for RNA reads mapping. BLAST was used to align the assembled transcriptome data with the coding

sequences of the *CD4* genes of five shark species in previous studies to obtain candidate CD4 transcripts[21]. For verification of CD4 sequences, total RNA from spleen tissue of bamboo shark was reverse-transcribed into cDNA. CD4 was amplified with primers designed according to contig assemblies (F: GATCTGCAGGCTGCTCCG, R: CTGCTTGA GGATCCTCTCG). PCR reactions for sequence validation were performed with 2 × Rapid Taq Master Mix (Vazyme) according to the manufacturer's guidebook. PCR products were cloned into pEASY T1 Vector with TA Cloning Kit (Transgen Biotech), and sequenced by Ruibiotech Co., Ltd. (Beijing, China).

Bioinformatic processing of single cell RNA raw sequencing data
We used the reference genome and annotation files downloaded from NCBI (*Danio rerio* GCA_000002035.4, *Chiloscyllium plagiosum* GCF_004010195.1, *Cynoglossus semilaevis* GCF_000523025.1). The validated *CD4* gene sequence is appended to the *C. plagiosum* genome file. These reference genome files are in turn used to create custom references using the CellRanger mkref command with default settings and for subsequent analysis. The cell barcodes and UMIs of raw data were embedded into the fastq file through the parsing feature of PISA. Subsequently, the data underwent a quality check using the FastQC (version 0.11.3). We employed STAR (version 2.7.9a) (https://github.com/alexdobin/STAR) to map the reads to the genome[59]. PISA's *sam2bam* and *anno* options were sequentially utilized to transform the barcodes into tag format within the BAM files. Subsequently, the reads within the BAM files were annotated and appended with tags. Sambamba (version 0.7.0) (https://github.com/biod/sambamba/releases/tag/v0.7.0)[60] was used to sort BAM files. The gene-cell UMI count matrix was constructed with PISA. The gene-cell expression matrices were served as processed raw data for further analyzes.

Library integration and clustering analysis
The raw gene expression matrix from the PISA pipeline was processed using the Seurat (version 3.0) (https://satijalab.org/seurat/) R package[61]. First, individual UMI count matrices from the libraries were imported into the R environment through the *Read10X* function, followed by the creation of Seurat objects for each experimental dataset. To obtain high-quality gene readings, we first filter out cells that detect less than 200 genes and 500 transcripts. Based on transcriptional differences between species, cells with excessive gene count and mitochondrial gene expression in each species were also removed (Supplementary Data 1). Subsequently, DoubletFinder (version 2.0) (https://github.com/ddiez/DoubletFinder)[62] was employed to identify doublets, assuming a 4% doublet formation rate for the loaded cells of each library. From the filtered cells, the gene expression matrices were standardization by the function of *SCTransform*. Next, minimize the differences between technical repetitions and use the RunCCA function to combine Seurat objects (each library from the same organ). After single organ integration, the integration of three organ datasets used Harmony (https://github.com/immunogenomics/harmony)[63] for corrected batch effects between organs. Briefly, we performed principal component analysis (PCA) on the integrated objects using *RunPCA* with "npcs = 30". The cell clustering was performed using the *FindNeighbors* and *FindClusters* functions with a resolution of 0.5. The data was visualized by non-linear dimensional reduction using the *RunUMAP* functions (reduction = "pca", dims = 1:30). The differentially expressed genes for each of the identity classes were computed with *FindAllMarkers* function (min.pct=0.25, logfc.threshold=0.25, test.use = "wilcox").

Cross-species orthologous gene pairs construction
To facilitate the integration of cross-species single-cell datasets and the comparison of cell similarities across species, we calculated pairwise orthologous genes for transformation. Orthofinder (version 2.3.7)

was utilized to pinpoint single-copy orthologous genes between a pair of species[64]. Extract single copy genes from the output file of Orthofinder (11,722 *D.rerio-C.semilaevis* pairs, 7865 *C.plagiosum-C.semilaevis* pairs, 8,651 *C.plagiosum-D.rerio* pairs). During subsequent cross-species integration, if there is a 1:1 match between non-zebrafish and zebrafish genes, the non-zebrafish gene name will be converted to zebrafish gene name. And then, in order to compare cell types across species, we have generated paired datasets of orthologous gene expression. In cases where one species had multiple orthologous genes in the other species, we established orthology relationships of one-to-many or many-to-many by duplicating gene entries from a single species. This resulted in 22,255 *D.rerio-C.semilaevis* pairs, 20,370 *C.plagiosum-C.semilaevis* and 20,673 *C.plagiosum-D.rerio* pairs.

Cell type similarity comparison analysis
For cross-species comparisons, we used two methods to measure similarity between cell types. The first is the MetaNeighbor[65]. Similarity analysis based on pseudocellular data (https://github.com/JiaqiLiZju/CelltypeEvolution/tree/master/Pseudocell_analysis). When comparing the overall immune profiles of three species (Fig. 2a), we selected 20 cells, and when comparing B cells across species, we selected 5 cells (Supplementary Fig. 5d). The degree of similarity among cell types was quantified using the average area beneath the receiver operator characteristic curve (AUROC) score. In the cross-organ comparison, after taking the negative log of the AUROC value, the hclust function in the R language was used to construct a hierarchical clustering tree (Supplementary Fig. 3). For cross-species comparison, we used single copy orthologous genes to construct a cell lineage tree in the order of white-spotted bamboo shark, zebrafish, and Chinese tongue sole (Supplementary Fig. 4a). Paired cell types with an AUROC threshold above 0.9 or the highest AUROC value were used to construct the lineage tree (https://github.com/JiaqiLiZju/CelltypeEvolution/tree/master/evolution_tree).

The second alternative approach to assess the similarity of cell types involves the computation of the Kullback-Leibler divergence (https://github.com/sebepedroslab/Stylophora_single_cell_atlas/blob/master/08_Cross_species_comparison.Rmd). We first calculated the normalized gene expression matrix for each dataset (https://github.com/BGI-Qingdao/Lungfish_scRNA_data_analysis/blob/main/6.FC_mtx.r). To compare those species, we generated pairwise datasets of orthologous gene expression with cross-species orthologous gene pairs. This approach permitted the establishment of one-to-many orthologous connections by replicating gene entries from one species in instances where it exhibited a maximum of three orthologous genes with expression in another species. The top 5% (Fig. 2b, c; Supplementary Fig 4b) or 20% (Supplementary Fig. 14) highest values are visually represented as arches connecting cell types in the circos plot, generated using the R package circlize[30]. The width of these arches is proportional to the KLD similarity between the cell types. In the analysis of the derived cell type clusters, we evaluated the expression levels of common orthologous genes (with fold change >1.3) with their expression in other cell types. The significance was calculated using paired Wilcoxon test between in- and outgroup cell types, Wilcoxon signed rank test adjusted *P*-value < 0.001 for all comparisons.

Cross-species clustering and comparison analysis for B cell
For cross-species comparisons, the one-to-one homolog genes shared by three species were retained for data integration of B cells. 900 B cells were randomly selected from the white-spotted bamboo shark, zebrafish, and Chinese tongue sole to ensure the average integration. We integrated the 900 B cells from each species using CCA (Default parameters), and then using *FindClusters* function to divided them into 6 clusters. 6 clusters were identified by specific DEG sets. Positive DEGs in each species were identified using

FindAllMarkers function (min.pct=0.25, logfc.threshold=0.25, test.use = "wilcox"). The conserved DEGs in 6 clusters were obtained by intersecting DEGs of subclusters of each species. Subsequently, we employed the UpSetR package in R to graphically represent the overlapping sets of differentially expressed genes (DEGs) among the three species. Heatmap of conserved and species-specific DEGs for all subclusters were obtained by calculating the average value of each subcluster.

Reclustering of T cells
We parsed out cells in corresponding T cell clusters from the white-spotted bamboo shark and zebrafish data, separated them according to different organs, and re-SCTransform normalized. Use the *FindIntegrationAnchors* function (Set default parameters for zebrafish, set k.filter=43 for bamboo shark) to identify different anchors of Seurat objects and use the found anchors to integrate the T cell data of three organs through the *IntegrateData* function (Set default parameters for zebrafish, Set k.weight=43 for bamboo shark). Finally, 2554 zebrafish T cells and 3356 shark T cells were obtained. These cells were re-clustered and analyzed by UMAP using the first 3000 highly variable genes and PC 1-30. Use the *FindAllMarkers* function to identify positive DEGs in each cluster, with the same parameters as above.

Pathway enrichment analysis
To understand the main functions of white-spotted bamboo shark B and T cells, we screened the differential genes according to the avg_log2FC >0.25 and *P*-value < 0.05 for pathway enrichment analysis. We used two enrichment methods. In B cell section, gene function enrichment analysis was performed using Metascape (https://metascape.org/gp/index.html). The white-spotted bamboo shark genes used for the enrichment analysis were orthologous to zebrafish genes. In T cell part, functional annotation and enrichment of these genes were performed through Kyoto Encyclopedia of Genes and Genomes (KEGG)[66]. Enrichment analysis of the annotations of DEGs obtained from the KEGG database was performed using the enricher function in the R package "clusterProfiler", and the results were visualized through the R package "enrichplot" (pAdjustMethod = "BH", pvalueCutoff = 0.05).

Co-expression analysis
The co-expression UMAP plot and the single gene expression UMAP plot are all displayed using the *FeaturePlot* function in the R package Seurat. The former needs to set the parameter "blend=TRUE".

Multiple sequence alignment of shark CD4 protein
The CD4 gene identified in the *Chiloscyllium plagiosum* was compared with the CD4 protein sequences from five other species (*Chiloscyllium griseum*, *Ginglymostoma cirratum*, *Scyliorhinus canicula*, *Scyliorhinus torazame,* and *Heterodontus zebra*)[21] using the CLUSTALW2 alignment tool[67]. Subsequently, the conservation of regions was visualized using the ESPript3 software[68] to highlight areas of sequence conservation.

Phagocytosis assay
The cell of spleen from white-spotted bamboo sharks was cultured with medium [RPMI-1640 medium with 15 mM Hepes, 10% FBS, 333 mM urea, 250 mM NaCl, 1.78 mM MgCl2, 45 mM trimethylamine oxide, 14.6 mM glucose, 7.01 mM CaCl2]. For the phagocytosis experiment, the cell density of primary cells was adjusted to 5×10^6/mL. After culturing for 4 hours, add fluorescent microsphere beads (final concentration of 5×10^8/mL), and incubate with the cells for 4 hours in the dark. After incubation, the cells were washed 3 times with washing buffer (PBS containing 4.5% D-glucose, 3% BSA, 333 mM urea, 15 mM Hepes), and centrifuged at $500\,g$ for 10 min.

Flow cytometry
Cells were fixed with 4% PFA for 20 min and permeabilized with washing buffer containing 0.1% Triton X-100 for 10 min at room temperature (RT). Then the cells were incubated for 60 min with rabbit anti-IgNAR (1:500, GTX128445, GeneTex, USA) antibody and followed by Goat Anti-Rabbit IgG H&L (Alexa Fluor® 488) (1:2000, ab150077, Abcam, US) to stain B cells. The cells were resuspended in washing buffer and subjected to cytometry analysis with a CytoFLEX flow cytometer (Beckman Coulter, USA). For FACS, the cells were sorted with CytoFLEX SRT cytometer (Beckman Coulter, USA). For detecting the phagocytosis of B cells, compensation for spectral overlap was performed with cells incubated with fluorescent microspheres only, cells stained with anti-IgNAR antibody only, and negative control cells. For detecting the co-expression, compensation for spectral overlap was performed with cells stained with anti-IgNAR antibody only, cells stained with Cy3-tagged probes only, and negative control cells. Further data acquisition and analysis were performed using CytExpert 2.4 software.

RNA Fluorescence in situ hybridization (FISH)
For fluorescence in situ hybridization, the 4% PFA fixed cell samples were cytocentrifuged on Polysine Microscope Adhesion Slides (Solarbio, China) to provides superior cell adhesion. The oligonucleotide probe sequence was designed using Primer 5.0 (Primer-E Ltd., Plymouth, UK) (Supplementary Data 10). The fluorescent probe was synthesized by BGI (China). Fluorescence in situ hybridization was performed using the RNA-FISH Kit (GenePharma, China). Afterwards adhesion slides were stained with DAPI for the nucleus and 1,1'-Dioctadecyl-3,3,3',3'-Tetramethylindocarbocyanine Perchlorate (DiI) (D282, Thermo, USA) for membrane, followed by PBS washed 10 minutes. To determine co-localization of red fluorescent microspheres and B cells. Phagocytized microspheres and fluorescent probes were photographed using a confocal microscope. Images were captured on a widefield Olympus microscope with 63× oil objective. Standard filters for Alexa Fluor 405, 488, and 568 were used in detecting the phagocytosis of B cells.

Fluorescence in situ hybridization coupled to flow cytometry (FISH-flow)
The detection of B and T marker co-expression cells in the bamboo shark was performed by immunostaining of IgNAR and fluorescence in situ hybridization targeting *tcf7*, and *cd247* genes, using FISH-flow, a technique described previously[69]. In brief, cells were fixed with 4% PFA for 20 min and permeabilized with washing buffer (PBS containing 4.5% D-glucose, 3% BSA, 333 mM urea, 15 mM Hepes) containing 0.1% Triton X-100 for 10 min at room temperature. Then the cells were incubated for 60 min with rabbit anti-IgNAR (1:500, GTX128445, GeneTex, USA) antibody and followed by Goat Anti-Rabbit IgG H&L (Alexa Fluor® 488) (1:2000, ab150077, Abcam, US) to stain B cells. The cells were then hybridized with Cy3 labeled probes targeting *tcf7*, and *cd247* genes in hybridization buffer (containing 10% dextran sulfate, 2× saline-sodium citrate, 10% formamide, 1 mg/ml tRNA, and 0.2 mg/ml BSA) at 37 °C overnight. The probe sets are detailed in Supplementary Data 10. The cells were analyzed with CytoFLEX flow cytometer (Beckman Coulter, USA). The gating strategy is shown in Supplementary Fig. 13a.

Quantitative real-time PCR analysis
We performed RNA isolation and reverse transcription using TRIzol reagent (Invitrogen, 15596026), and single cell sequence-specific amplification kit (Vazyme, P621-01) according to the manufacturer's instructions. QuantiFast SYBR Green RT-PCR Kit (Qiangen, 204156) was used for quantitative real-time PCR with the Roche LightCycler® 96 Instrument. The primers have been included in Supplementary Data 10.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The single cell RNA sequencing data and matrix data generated in this study have been deposited in the Genome Sequence Archive (Geromics, Proteomics & Bioinformatics 2021)[70] in National Genomics Data Center (Nucleic Acids Res 2022)[71], China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA017585) [https://ngdc.cncb.ac.cn/gsa/search?searchTerm=CRA017585]. Public scRNA-seq data of nurse shark used in this study are downloaded from the NCBI GEO database (GSE232302)[https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE232302][42].

References

1. Arendt, D. et al. The origin and evolution of cell types. *Nat. Rev. Genet* **17**, 744–757 (2016).
2. Network, B. I. C. C. A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86–102 (2021).
3. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
4. Shami, A. N. et al. Single-cell RNA sequencing of human, macaque, and mouse testes uncovers conserved and divergent features of mammalian spermatogenesis. *Dev. Cell* **54**, 529–547.e512 (2020).
5. Flajnik, M. F. A cold-blooded view of adaptive immunity. *Nat. Rev. Immunol.* **18**, 438–453 (2018).
6. Flajnik, M. F. & Kasahara, M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet* **11**, 47–59 (2010).
7. Boehm, T. Evolution of vertebrate immunity. *Curr. Biol.* **22**, R722–732, (2012).
8. Shimeld, S. M. & Holland, P. W. Vertebrate innovations. *Proc. Natl Acad. Sci. USA* **97**, 4449–4452 (2000).
9. Dornburg, A. & Yoder, J. A. On the relationship between extant innate immune receptors and the evolutionary origins of jawed vertebrate adaptive immunity. *Immunogenetics* **74**, 111–128 (2022).
10. Li, J. et al. B lymphocytes from early vertebrates have potent phagocytic and microbicidal abilities. *Nat. Immunol.* **7**, 1116–1124 (2006).
11. Tan, M. et al. The whale shark genome reveals patterns of vertebrate gene family evolution. *Elife* **10**, https://doi.org/10.7554/eLife.65394 (2021).
12. Venkatesh, B. et al. Elephant shark genome provides unique insights into gnathostome evolution. *Nature* **505**, 174–179 (2014).
13. Zhang, Y. et al. The white-spotted bamboo shark genome reveals chromosome rearrangements and fast-evolving immune genes of cartilaginous fish. *iScience* **23**, 101754 (2020).
14. Litman, G. W., Anderson, M. K. & Rast, J. P. Evolution of antigen binding receptors. *Annu Rev. Immunol.* **17**, 109–147 (1999).
15. Matz, H., Munir, D., Logue, J. & Dooley, H. The immunoglobulins of cartilaginous fishes. *Dev. Comp. Immunol.* **115**, 103873 (2021).
16. Greenberg, A. S. et al. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature* **374**, 168–173 (1995).
17. Wu, L. et al. Recent advances on phagocytic B cells in teleost fish. *Front Immunol.* **11**, 824 (2020).
18. Rast, J. P. et al. alpha, beta, gamma, and delta T cell antigen receptor genes arose early in vertebrate phylogeny. *Immunity* **6**, 1–11 (1997).
19. Flajnik, M. F., Ohta, Y., Namikawa-Yamada, C. & Nonaka, M. Insight into the primordial MHC from studies in ectothermic vertebrates. *Immunol. Rev.* **167**, 59–67 (1999).
20. Kulski, J. K., Shiina, T., Anzai, T., Kohara, S. & Inoko, H. Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunol. Rev.* **190**, 95–122 (2002).
21. Takizawa, F. et al. CD4 and LAG-3 from sharks to humans: related molecules with motifs for opposing functions. *Front. Immunol.* **14**, 1267743 (2023).
22. Deiss, T. C. et al. Ancient use of Ig variable domains contributes significantly to the TCRδ repertoire. *J. Immunol.* **203**, 1265–1275 (2019).
23. Stanfield, R. L., Dooley, H., Flajnik, M. F. & Wilson, I. A. Crystal structure of a shark single-domain antibody V region in complex with lysozyme. *Science* **305**, 1770–1773 (2004).
24. Horton, R. et al. Gene map of the extended human MHC. *Nat. Rev. Genet* **5**, 889–899 (2004).
25. Li, L. et al. Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. *Nat. Immunol.* **12**, 129–136 (2011).
26. Shay, T. et al. Conservation and divergence in the transcriptional programs of the human and mouse immune systems. *Proc. Natl Acad. Sci. USA* **110**, 2946–2951 (2013).
27. Hagai, T. et al. Gene expression variability across cells and species shapes innate immunity. *Nature* **563**, 197–202 (2018).
28. Iwasaki, H. & Akashi, K. Myeloid lineage commitment from the hematopoietic stem cell. *Immunity* **26**, 726–740 (2007).
29. Levy, S. et al. A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell* **184**, 2973–2987.e2918 (2021).
30. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
31. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet* **25**, 25–29 (2000).
32. Hsieh, T. et al. JunB is critical for survival of T helper cells. *Front. Immunol.* **13**, 901030 (2022).
33. Powell, M. D., Read, K. A., Sreekumar, B. K. & Oestreich, K. J. Ikaros zinc finger transcription factors: regulators of cytokine signaling pathways and CD4+ T helper cell differentiation. *Front. Immunol.* **10**, 462704 (2019).
34. Zhao, Z. et al. IL-12Rβ2 promotes the development of CD4+ CD25+ regulatory T cells. *J. Immunol.* **181**, 3870–3876 (2008).
35. McCausland, M. M. et al. SAP regulation of follicular helper CD4 T cell development and humoral immunity is independent of SLAM and Fyn kinase. *J. Immunol.* **178**, 817–828 (2007).
36. Kasheta, M. et al. Identification and characterization of T reg–like cells in zebrafish. *J. Exp. Med.* **214**, 3519–3530 (2017).
37. Wan, F. et al. Characterization of γδ T cells from zebrafish provides insights into their important role in adaptive humoral immunity. *Front. Immunol.* **7**, 675 (2017).
38. Dee, C. T. et al. CD4-transgenic zebrafish reveal tissue-resident Th2- and regulatory T cell–like populations and diverse mononuclear phagocytes. *J. Immunol.* **197**, 3520–3530 (2016).
39. Luckheeram, R. V., Zhou, R., Verma, A. D. & Xia, B. CD4+ T cells: differentiation and functions. *Clin. developmental Immunol.* **2012**, 925135 (2012).
40. Onishi, R. M. & Gaffen, S. L. Interleukin-17 and its target genes: mechanisms of interleukin-17 function in disease. *Immunology* **129**, 311–321 (2010).
41. Wang, L. et al. The zinc finger transcription factor Zbtb7b represses CD8-lineage gene expression in peripheral CD4+ T cells. *Immunity* **29**, 876–887 (2008).
42. Matz, H. et al. Organized B cell sites in cartilaginous fishes reveal the evolutionary foundation of germinal centers. *Cell Rep.* **42**, 112664 (2023).
43. Huang, Y. et al. Discovery of an unconventional lamprey lymphocyte lineage highlights divergent features in vertebrate adaptive

immune system evolution. https://doi.org/10.1038/s41467-024-51763-2. NCOMMS-23-60270C.

44. Busslinger, M. Transcriptional control of early B cell development. *Annu Rev. Immunol.* **22**, 55–79 (2004).

45. Rabinovitch, A., Suarez-Pinzon, W. L., Lapchak, P. H., Meager, A. & Power, R. F. Tumor necrosis factor mediates the protective effect of Freund's adjuvant against autoimmune diabetes in BB rats. *J. Autoimmun.* **8**, 357–366 (1995).

46. Desjardins, M., Houde, M. & Gagnon, E. Phagocytosis: the convoluted way from nutrition to adaptive immunity. *Immunol. Rev.* **207**, 158–165 (2005).

47. Stuart, L. M. & Ezekowitz, R. A. Phagocytosis: elegant complexity. *Immunity* **22**, 539–550 (2005).

48. Bauer, S. R., Holmes, K. L., Morse, H. C. 3rd & Potter, M. Clonal relationship of the lymphoblastic cell line P388 to the macrophage cell line P388D1 as evidenced by immunoglobulin gene rearrangements and expression of cell surface antigens. *J. Immunol.* **136**, 4695–4699 (1986).

49. Hanecak, R., Zovich, D. C., Pattengale, P. K. & Fan, H. Differentiation in vitro of a leukemia virus-induced B-cell lymphoma into macrophages. *Mol. Cell Biol.* **9**, 2264–2268 (1989).

50. Star, B. et al. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207–210 (2011).

51. Swann, J. B., Holland, S. J., Petersen, M., Pietsch, T. W. & Boehm, T. The immunogenetics of sexual parasitism. *Science* **369**, 1608–1615 (2020).

52. Klein, J. & Figueroa, F. Evolution of the major histocompatibility complex. *Crit. Rev. Immunol.* **6**, 295–386 (1986).

53. Rowe, R. G., Mandelbaum, J., Zon, L. I. & Daley, G. Q. Engineering hematopoietic stem cells: lessons from development. *Cell Stem Cell* **18**, 707–720 (2016).

54. Haas, S., Trumpp, A. & Milsom, M. D. Causes and consequences of hematopoietic stem cell heterogeneity. *Cell Stem Cell* **22**, 627–638 (2018).

55. Kang, J. & Malhotra, N. Transcription factor networks directing the development, function, and evolution of innate lymphoid effectors. *Annu Rev. Immunol.* **33**, 505–538 (2015).

56. Vivier, E., van de Pavert, S. A., Cooper, M. D. & Belz, G. T. The evolution of innate lymphoid cells. *Nat. Immunol.* **17**, 790–794 (2016).

57. Dai, H. et al. PIRs mediate innate myeloid cell memory to nonself MHC molecules. *Science* **368**, 1122–1127 (2020).

58. Huang, Y. et al. Single-cell transcriptome landscape of zebrafish liver reveals hepatocytes and immune cell interactions in understanding nonalcoholic fatty liver disease. *Fish Shellfish Immunol* **146**, 109428 (2024).

59. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

60. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).

61. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

62. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst.* **8**, 329–337.e324 (2019).

63. Linderman, G. C., Rachh, M., Hoskins, J. G., Steinerberger, S. & Kluger, Y. Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data. *Nat. methods* **16**, 243–245 (2019).

64. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 1–14 (2019).

65. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat. Commun.* **9**, 1–12 (2018).

66. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27–30 (2000).

67. Larkin, M. A. et al. Clustal W and clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

68. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–324 (2014).

69. Arrigucci, R. et al. FISH-Flow, a protocol for the concurrent detection of mRNA and protein in single cells using fluorescence in situ hybridization and flow cytometry. *Nat. Protoc.* **12**, 1245–1260 (2017).

70. Chen, T. et al. The genome sequence archive family: toward explosive data growth and diverse data types. *Genomics Proteom. Bioinforma.* **19**, 578–583 (2021).

71. Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2022. *Nucleic Acids Res* **50**, D27-d38, https://doi.org/10.1093/nar/gkab951 (2022).

Acknowledgements

Author contributions

C.S., G.F., S.L., and J.Z. conceived and supervised the study. C.S., H.Y.W., J.Y.C., Xianghui Zhang., X.L., Yubang Li., M.Y., Y.Z. Q.J., R.W., and Shanshan Liu coordinated and performed sample collection, single-cell RNA experiments, and sequencing. Yanan Li., H.H., H.C., L.Z., F.T., Y.H., Shuo Li, S.H., and S.J. performed quality control, integration, clustering, and annotation of the scRNA-seq data. H.Y.W., H.H., H.C. Yanan Li. and Yifang Lu conducted the comparative analyzes across species. J.Y.C., H.Y.W., Yangqing Zhang. and J.Z. performed a phagocytosis assay. H.Y.W., J.Y.C., Yanan Li., Xianghui Zhang and Yubang Li. drafted the manuscript with contributions from all other authors. C.S., G.F., S.L., I.S., Q.W., L.M., Q.L., K.L., and Yuyan Liu revised the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51761-4.

Correspondence and requests for materials should be addressed to Shanshan Liu or Changwei Shao.

Peer review information *Nature Communications* thanks Jonathan Rast and Kun Wang for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.