Briefings in Bioinformatics, 22(4), 2021, 1–10

https://doi.org/10.1093/bib/bbaa283
Problem Solving Protocol

FunRes: resolving tissue-specific functional cell states based on a cell–cell communication network model

Sascha Jung, Kartikeya Singh and Antonio del Sol

Corresponding author: Antonio del Sol, Computational Biology Group, Luxembourg Centre for Systems Biomedicine (LCSB), Esch-sur-Alzette, L-4362, Luxembourg. Tel.: +352 46 66 44 6982; Email: Antonio.delsol@uni.lu

Abstract

The functional specialization of cell types arises during development and is shaped by cell–cell communication networks determining a distribution of functional cell states that are collectively important for tissue functioning. However, the identification of these tissue-specific functional cell states remains challenging. Although a plethora of computational approaches have been successful in detecting cell types and subtypes, they fail in resolving tissue-specific functional cell states. To address this issue, we present FunRes, a computational method designed for the identification of functional cell states. FunRes relies on scRNA-seq data of a tissue to initially reconstruct the functional cell–cell communication network, which is leveraged for partitioning each cell type into functional cell states. We applied FunRes to 177 cell types in 10 different tissues and demonstrated that the detected states correspond to known functional cell states of various cell types, which cannot be recapitulated by existing computational tools. Finally, we characterize emerging and vanishing functional cell states in aging and disease, and demonstrate their involvement in key tissue functions. Thus, we believe that FunRes will be of great utility in the characterization of the functional landscape of cell types and the identification of dysfunctional cell states in aging and disease.

Key words: computational biology; cell–cell communication; cell states; clustering

Introduction

In multicellular organisms, the functional specification of cell types arises during development and is further shaped by signals from other cells. In particular, the exchange of these signals through receptor–ligand-mediated cell–cell communication networks determines a distribution of different functional cell states that are collectively relevant for tissue functioning [1]. Thus, cell types are composed of a variety of functional cell states that are shaped by their tissue environment. More specifically, in response to different stimuli, cells of the same cell type can exhibit different phenotypes defined by physical, molecular and functional characteristics, called functional cell states. Great efforts have been devoted to characterize tissue-specific functional differences imparted by cell–cell interactions, which has led to the identification of functional cell states in various cell types. For instance, comparison of tissue-resident macrophages in different organs revealed functionally relevant differences in their gene expression programs [2]. While ileal and colonic macrophages showed higher expression of CD74 compared to other tissue-resident macrophages, TGFB2 is exclusively expressed by peritoneal macrophages [2]. Moreover, transplantation of these macrophages to other organs reprograms their expression profile towards tissue-resident cells, which demonstrates that the observed differences are largely

Sascha Jung is a postdoctoral researcher at CIC bioGUNE, Spain, whose main research interest is in systems biology and bioinformatics.
Kartikeya Singh is a PhD student in Systems Biology at the Luxembourg Centre for Systems Biomedicine (LCSB). His research interest is to decipher the immune response to pathogens using computational/bioinformatics tools.
Antonio del Sol is professor of bioinformatics at the Luxembourg Centre for Systems Biomedicine (LCSB) and Ikerbasque professor at CIC bioGUNE. He is interested in the areas of computational biology of stem cells and systems biology.
Submitted: 7 July 2020; **Received (in revised form):** 24 September 2020

imparted by their environment [2]. In addition, previous studies investigated the effect of the environment on the functional specification of stem cells with respect to their activity and the tissue compartment in which they reside. For example, neural stem cells in the subgranular and subventricular zones show subtle phenotypic differences driven by the niche. While subgranular zone stem cells express the transcription factor (TF) HES5, which is induced by Notch signaling, subventricular zone stem cells express Id proteins, which are induced by BMP signaling, to maintain their function [3, 4]. Although these studies have enabled the characterization of certain functional cell states determined by cell–cell interactions in specific cell types, the development of computational methods would greatly aid the systematic identification and characterization of such states.

The identification of cell types and subtypes has been traditionally performed by molecular biology approaches based on the shape and size of cells, while, more recently, these populations have been characterized by their expression of cellular membrane proteins. However, due to the limited number of surface proteins, it is likely that important functional differences between cell states cannot be described by combinations of these proteins alone. Advances in single-cell RNA sequencing (scRNA-seq) technologies have provided an unprecedented view on the cellular heterogeneity within individual cell types. However, the exploitation of these datasets requires computational tools for processing and clustering the data. A plethora of computational tools have been developed in recent years for the unsupervised clustering of scRNA-seq profiles [5–8]. Although these tools have enabled the identification of several novel cell subtypes, they do not specifically consider the effect of cell–cell interactions on tissue-specific functional processes, and therefore fail in resolving the heterogeneity in functional states of different cell types.

In order to address this challenge, we present FunRes, a computational method for the identification of tissue-specific functional cell states that are induced by receptor–ligand-mediated cell–cell interactions. FunRes employs single-cell RNA-seq data to reconstruct the cell–cell communication network among cell types. In particular, FunRes exclusively detects interactions that are functionally relevant to each cell type by warranting the compatibility of the signal with the intracellular signaling and transcriptional network. Next, FunRes leverages this network to identify functional cell states by partitioning cell types based on cell–cell interactions and the functional annotation of downstream target TFs. We applied FunRes to 177 cell types in 10 different tissues and demonstrate that the detected states correspond to known functional cell states of various immune and non-immune cell types, including macrophages, NK and endothelial cells, which cannot be recapitulated by existing computational tools. Moreover, comparison of the identified functional cell states of the same cell types in different tissues showed conserved states carrying out essential cell-type functions as well as the existence of states with unique functionalities. Finally, we assessed the effect of aging and disease on the composition of functional cell states. Although the overall number of functional cell states does not significantly change, certain cell types, such as enterocytes of the large intestine, hepatic sinusoidal epithelial cells and pancreatic alpha cells, display vast differences in their composition of functional states. Moreover, we validate the detected functional cell states in aged and pathologic tissues by providing evidence for their involvement in pathological tissue functions.

In summary, FunRes constitutes the first computational method specifically designed for identifying functional cell states induced by receptor-ligand mediated cell–cell interactions and complements current clustering methodologies that successfully detect cell types and subtypes. We demonstrate that FunRes is applicable to a wide range of tissue conditions and can accurately detect functional cell states. Thus, we believe that FunRes will be of great utility in the characterization of the functional landscape of cell types and the identification of dysfunctional cell states in aging and disease.

Material and Methods

Assembly of cell–cell communication scaffolds, intracellular signaling and TF-gene regulatory interactions

A cell–cell communication scaffold was generated for human and mouse on the basis of a previously published dataset including manually curated, validated and predicted intercellular interactions (Figure 1A) [9]. These interactions have been collected from different databases, i.e. DLRP [10], HPMR [11], IUPHAR [12], HPRD [13] and String DB [14] or manually validated in previous studies. The interactions were further filtered based on UniProt [15] annotations to include only ligands annotated to be ‘Secreted’. Since the original dataset only consists of human data, these interactions were mapped to mouse orthologs using BioMart from Ensembl [16].

The intracellular signaling network is composed of pathway interactions included in Omnipath [17], Reactome [18] and MetaCore from Thomson Reuters. In particular, all pathways from MetaCore were obtained including all signal transduction interactions while discarding transcriptional gene regulatory interactions. MetaCore objects were mapped to gene symbols using the provided mapping table. In case MetaCore objects of the regulator or regulated gene mapped to multiple gene symbols, all possible interactions between the associated gene symbols were generated using a mapping table provided by MetaCore.

Gene regulatory interactions were obtained from MetaCore from Thomson Reuters, a manually curated resource of gene–gene interactions, on 01 April 2019 for human and mouse genes. Only transcriptional regulatory interactions with known effects, i.e. activation or inhibition, were selected by filtering for ‘direct interactions’ with reported effects ‘activation’ or ‘inhibition’. MetaCore objects were again mapped to gene symbols using the provided mapping table. As in case of the intracellular signaling network, if MetaCore objects of the regulator or regulated gene mapped to multiple gene symbols, all possible interactions between the associated gene symbols were generated using an in-house script.

Selection of preserved TFs

To select preserved TFs, FunRes identifies transcription factors that are expressed in at least a user-defined fraction of cells. For the analysis presented in this manuscript, we selected a permissive cutoff of 10% for all samples. For that, FunRes transforms the expression data into a binary format in which TFs with at least one count become ‘1’ while not-expressed TFs become ‘0’ and aggregates the binary data by cell type. Finally, FunRes selects in each subpopulation the TFs that are expressed in the top five-percentile of cells.

Detecting receptors inducing preserved TFs

To detect receptors inducing the expression of the selected TFs, FunRes employs a Markov Chain model of intracellular signaling, called SigHotSpotter, to identify high-probability intermediate molecules (Figure 1A) [19]. In brief, SigHotSpotter uses single-cell RNA-seq data of a subpopulation and the assembled intracellular signaling network to create a state transition matrix representing the traversal of a signal through the network. The state transition matrix represents a finite discrete Markov Chain and is subsequently evolved to create the stationary distribution. The stationary distribution displays the intermediate molecules exhibiting the highest steady state probabilities. Afterwards, SigHotSpotter connects the intermediate molecules to interface TFs, i.e. the first transcription factors in the signal transduction chain, and characterizes the compatibility with their downstream targets. Calculating all shortest paths from high-probability intermediate molecules to downstream preserved TFs defines a compatibility score that classifies each molecule as being active or inactive. Here, a high-probability gene is deemed compatible with its downstream target if the expression of the gene and target TF agrees with the sign of the interaction path, i.e. an even number of inhibitions is an activating path while all other paths are inhibiting. In case of an activation, the intermediate and target genes have to be expressed whereas, in case of an inhibition, the target gene must not be expressed. Next, a gene is compatible, if a significant number of its targets is compatible. Significance is assessed using a hypergeometric test with *P*-value cutoff 0.05. Following the same rationale, receptors are identified that target the compatible high-probability intermediates.

Inference of the cell–cell communication network

The main algorithm consists of four steps. First, preserved TFs and the receptors regulating them are selected in each cell population, as described before. Second, ligands expressed in a user-defined fraction of cells are selected in each cell population. For this study, a fraction of 10% was selected. Third, ligand–receptor interactions are established between two cell populations if (i) the receptor was selected in the first step for the first population, (ii) the ligand was selected in the second step for the second population and (iii) the receptor-ligand interaction is contained in cell–cell communication scaffold. Every interaction between receptor r in subpopulation $p1$ and ligand l in subpopulation $p2$ is augmented with an interaction strength $s_{r,l,p1,p2}$ defined by:

$$s_{r,l,p1,p2} = \left(\frac{1}{\left|\{x_{r,p1} | x_{r,p1} > 0\}\right|} \sum_{\{x_{r,p1} | x_{r,p1} > 0\}} x_{r,p1} \right) \cdot \left(\frac{1}{\left|\{x_{l,p2} | x_{l,p2} > 0\}\right|} \sum_{\{x_{l,p2} | x_{l,p2} > 0\}} x_{l,p2} \right) \quad (1)$$

Informally, the score is the product of average non-zero receptor expression values and average non-zero ligand expression values with respect to the receptor r and its expressing population $p1$ as well as to the ligand l and its expressing population $p2$.

Significance of each interaction is determined by comparing the score of an interaction between two cell types against a background distribution of scores between these cell types based on all interactions in the scaffold. Interactions in the 90th percentile that have at least one significant downstream TF target are retained in the final cell–cell communication network.

Identification of functional cell states

After inferring the cell–cell communication network, FunRes creates an incidence matrix of receptor/downstream TF expression for each cell type (Figure 1B). In particular, each cell in the cell type under study is represented by a binary vector in which each entry corresponds to a receptor/downstream TF pair. If both the receptor and downstream TF are expressed in a cell, the corresponding entry in the vector will be '1' and '0' otherwise. Hierarchical clustering is performed on the resulting incidence matrix using the 'hclust' R function with Euclidean distance and complete linkage. The optimal number of clusters is computed based on the Dunn index, an internal cluster evaluation metric comparing within-cluster with the between-cluster distances. Finally, Gene Ontology [20, 21] enrichment of biological processes is performed using the DOSE R package (Figure 1B) [22]. Terms with false discovery rate lower than 0.01 are considered significant. Clusters having identical functional annotations are subsequently merged. The resulting clusters constitute the identified functional cell states. Importantly, FunRes does not require the proteins within the enriched GO terms to be identical, but only the terms themselves.

Comparison with state-of-the-art clustering methods

We selected Seurat [5, 6], SC3 [7] and SINCERA [8] for assessing their ability to detect functional cell states. All methods were employed in a standard workflow with default parameters in R. In particular, for Seurat, the 2000 most variable genes were detected using the 'FindVariableFeatures' method with selection method 'vst'. Subsequently, the data were scaled ('ScaleData' function) and principal component analysis (PCA) ('RunPCA' function) was performed. Finally, the data were clustered by employing the 'FindNeighbors' and 'FindClusters' functions on the first 10 principal components. SC3 was invoked using the 'sc3' function and the maximum number of clusters was set to 10. SINCERA was invoked without specific pre-processing steps. All analyses were carried out in R v3.6.1.

Results

Identification of functional cell states based on cell–cell communication networks

For the purpose of identifying niche-induced functional cell states, FunRes initially reconstructs the cell–cell communication networks among all cell types in a dataset (Figure 1A). In particular, it integrates transcriptional and signaling networks with extracellular ligand–receptor interactions following a bottom-up approach by first selecting transcription factors (TFs) whose expression is preserved across cells. Next, FunRes identifies receptors that regulate these preserved TFs by employing a previously introduced Markov Chain model that assesses signal transduction probabilities from receptors to their co-expressed TFs [19]. Finally, FunRes finds cognate ligands for these receptors from a previously curated set of ligand–receptor interactions [9]. Based on a previous study about the proportions of functional cell states in heterogeneous T cell populations, we required throughout this study both the ligands and receptors to be expressed in more than 10% of cells of the secreting and receiving cell population, respectively [23].

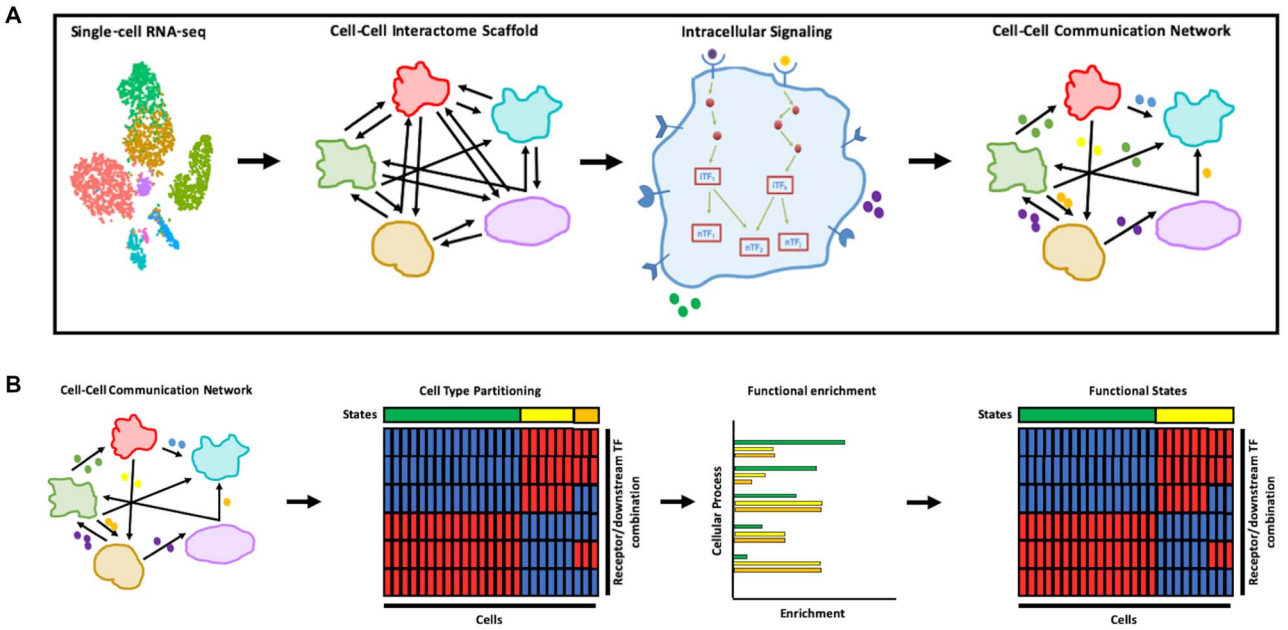

Figure 1. Method overview. (**A**) Based on single-cell RNA-seq data of a tissue, ResFun infers active receptors in each cell population. Based on these receptors, ResFun reconstructs a cell–cell communication network by identifying interactions having a significantly higher score compared to other interactions in the interaction scaffold. (**B**) Given a reconstructed cell–cell communication network of a tissue, FunRes identifies functional cell states for each cell type individually. First, cells are represented as a binary vector of receptor/downstream TF expression. If a receptor and its corresponding downstream TF is expressed in a cell, it will be represented as '1' (red) and '0' (blue) otherwise. An optimal number of clusters is determined, which will be subjected to functional enrichment analysis. If two clusters are enriched in the same processes, they will be merged. The merged clusters define the functional cell states.

After reconstructing the functional cell–cell communication network, FunRes partitions each cell type into different functional cell states based on the identified cell–cell interactions and their downstream TF targets (Figure 1B). More specifically, hierarchical clustering is performed to group cells by their function based on the downstream effect of cell–cell interactions they are participating in. Finally, the optimal number of functional states in each cell type is determined by evaluating cluster consistency. Notably, the optimal number of states could be equal to one, which results in no partitioning of the cell type under consideration, suggesting no niche-induced functional heterogeneity.

Functional evaluation of niche-induced cell states

We employed FunRes to identify niche-induced functional cell states of 177 cell types in 10 mouse tissues from Tabula Muris Senis [24]. On average, each cell type was partitioned into 2.49 functional states (median: 2), with a maximum of nine identified states in enterocytes of the intestinal epithelium and pancreatic alpha cells (Figure 2A). Nevertheless, the vast majority of cell types is partitioned into at most three functional cell states with 16% of cell types remaining unpartitioned.

To assess the performance of our method, we evaluated the function of the identified cell states with known cell state markers compiled from literature and compared the performance of FunRes against current clustering methods. As a result, we were able to collect evidence for multiple functional cell states in the heart, kidney and liver. Firstly, we identified functional cell states of cardiac coronary vascular endothelial cells. As a result, FunRes detected two clusters differentiated by Bmpr2 signaling. While the Bmpr2-positive state corresponds to quiescent coronary vascular endothelial cells, the Bmpr2-negative state consists of activated cells undergoing endothelial mesenchymal transition, which is required for cardiac functioning [25]. In this regard, FosB is activated by Bmpr2 according to FunRes and is downregulated in active vascular endothelial cells [26]. Thus, Bmpr2 and FosB serve as markers for quiescent and active coronary vascular endothelial cells. In order to assess the ability of current clustering methods in identifying functional cell states, we compared the results from FunRes against Seurat [5, 6], SC3 [7] and SINCERA [8]. As a result, SINCERA and SC3 detected 19 and 9 clusters, respectively, none of which could be distinguished using the marker genes of active and quiescent endothelial cells. Similar to FunRes, Seurat detected three clusters. However, no significant difference in the expression of these cell state markers could be detected. Consequently, none of the clusters corresponds to the functional cell states identified by FunRes (Supplementary Figure S1).

Secondly, we assessed the functional cell states of kidney macrophages and identified two states that are differentiated based on Nrp1 signaling. In particular, the first cluster expresses Nrp1 and Il10, all of which are known markers of M2 macrophages, an immunosuppressive cell state necessary for wound healing and cessation of the inflammatory response to pathogens (Figure 2B) [27, 28]. Previous reports showed that activation of Nrp1, which is involved in the cell–cell interactions differentiating the two clusters, induces this functional state, which underscores the accuracy of FunRes in detecting functionally relevant cell–cell interactions. In addition, the other functional cell state is characterized by active Tlr2 and Tlr4 signaling, which results in the polarization into M1 macrophages [29]. Indeed, cells belonging to this cluster express the M1 markers Nfatc1, Tnfsf9 and Il16 [30–32]. Thus, the identified cell states correspond to M1 and M2 polarization of macrophages, respectively. In contrast, current clustering methodologies cannot resemble this finding. Seurat

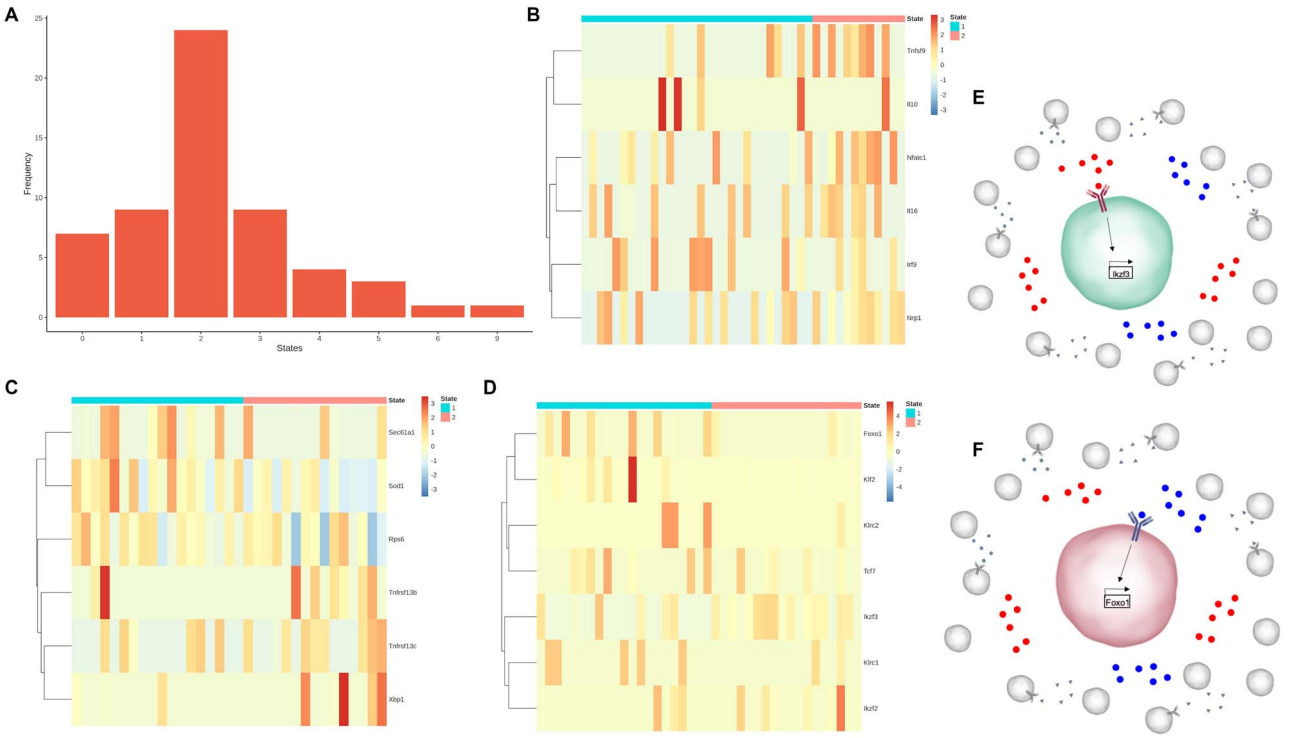

Figure 2. Validation of FunRes and comparison to the state-of-the-art methods. (**A**) Histogram of the number of identified functional cell states in all cell types of all tissues. (**B–D**) Heatmaps of genes important for the function of the identified functional cell states in kidney macrophages (B), liver B cells (C) and liver NK cells (D). The expression values are scaled for each gene individually. (**E** and **F**) Visualization of how different cues from the environment induce downstream marker genes of peripheral (E) and adaptive (F) NK cell states.

was unable to cluster the data due to the failure of PCA with default parameters as the number of kidney macrophage cells is below the number of principal components to be computed. Sincera partitioned the cells into 35 different clusters having on average 1.09 cells whereas SC3 was the only tool providing a reasonable number of clusters. However, these populations do not correspond to M1 and M2 macrophages as known marker genes are expressed across clusters (Supplementary Figure S2).

Thirdly, we investigated the functional states of liver B cells and identified two states that are distinguished by IL4R signaling [33]. Previous studies elucidated the role of IL4R signaling and found that it induces autophagy, a key mechanism for the maintenance of B cell memory against pathogens and synthesis of antigens, such as IgE [34]. Moreover, B cells with active IL4R signaling express the conventional B cell genes Tnfrsf13b and Tnfrsf13c (Figure 2C) [35, 36]. In contrast, the other identified cell state is deficient of IL4R signaling and express the plasma cell markers Sec61a1 (Figure 2C) [37]. Comparison of the cell states identified by FunRes to current clustering methodologies reveals that all of the other methods fail to detect the underlying functional heterogeneity. While Seurat was unable to perform clustering due to the failure of PCA with default parameters as the number of liver B cells is below the number of principal components to be computed, Sincera partitioned the data such that each cell belongs to a different cluster. In addition, SC3 did not detect any functional heterogeneity in liver B cells and leaves the data unpartitioned (Supplementary Figure S3).

Finally, we applied our method to liver NK cells and detected two functional NK cell states defined by differential downstream signaling targets. In particular, cells of the second cluster specifically activate Ikzf3 (Figure 2D), whereas other downstream signaling targets, such as Stat3, Stat1 and Ets1, are commonly activated in both clusters. Indeed, the second cluster is characterized by Ikzf2 and Ikzf3 expression, two markers of peripheral NK cells (Figure 2E) [38]. In contrast, the first cluster shows high expression of the adaptive NK cell markers Klrc1 and Klrc2 (Figure 2F) [39, 40]. Similar to the case of kidney macrophages, Sincera produced an excess of clusters with, on average, only 1.38 cells whereas Seurat produced no clusters due to the failure of PCA with default parameters as the number of liver NK cells is below the number of principal components to be computed. Although SC3 detected 2 clusters like FunRes, they could not be attributed to different functional NK cell states (Supplementary Figure S4).

In summary, FunRes was able to dissect the functional heterogeneity of cell types in various tissues and accurately identified known functional states, while current clustering methods failed to provide satisfactory results. Moreover, comparison of the clusterings identified by each method shows only a moderate agreement between the functional cell states detected by FunRes and the results of Seurat, SC3 and SINCERA (Supplementary Figure S5).

Comparison of functional cell states in different tissues

Next, we set out to characterize the niche-induced functional states of macrophages and B cells in different tissues. As expected, we detected conserved functional states of kidney and limb muscle macrophages with activated Tlr2, Tlr4 and Tnfrsf1a signaling pathways, which constitute the key signaling cascades for the recognition of antigens and the activation of macrophages [41, 42]. However, while kidney macrophages show an Nrp1-positive subpopulation corresponding to M2 polarization, Nrp1 signaling is inactive in all limb muscle

macrophages. However, a subset of cells in the limb muscle display activity of the Il1 decoy receptor Il1r2, which is unique to M2 polarized macrophages. Moreover, these cells display an activated Il6 signaling pathway, which has been previously shown to be an alternative mechanism for M2 macrophage polarization [43]. This demonstrates that M2 macrophages reside in both tissues, but are alternatively activated.

In contrast to macrophages, B cells were detected in the spleen, liver, limb muscle and lung having vastly different functional states that greatly differ between the spleen and liver/limb muscle/lung tissues (Figure 3A). In particular, the identified functional B cell states in the liver and limb muscle are commonly characterized by Il2 receptor signaling activity, which is involved in B cell proliferation and plasma cell differentiation [44, 45]. However, unique functional cell states are acquired through Il4ra and Cxcr4 signaling activity in these tissues. As previously described, Il4ra promotes B cell autophagy required for memory maintenance and antigen presentation [33]. In contrast, Cxcr4 signaling is uniquely activated in the functional state of limb muscle B cells demonstrating the differentiation of these cells into plasma cells [46]. While B cells lose their responsiveness to Cxcl12, the cognate ligand of Cxcr4, during development, they regain their sensitivity upon differentiation into mature B cells [46]. Due to the involvement of Il4ra and Cxcr4 in plasma cell differentiation, we compared the downstream targets of these signaling cascades identified by FunRes and found Foxo1 to be the only common TF. Previous studies already revealed the key role of Foxo1 in plasma cell function and differentiation as an activator of Prdm1, a plasma cell master regulator [47]. Therefore, the identified functional B cell states unique to liver and limb muscle tissue, respectively, induce plasma cell differentiation through distinct signaling pathways. Opposed to liver and limb muscle tissues, the functional B cell states in the lung are characterized by Il2 and interferon gamma signaling, which is the only commonality with B cells from the spleen. Whereas a B cell state expressing Ifngr1 was detected in both spleen and lung tissues, their functions vary due to the presence of co-stimulatory signals. As expected, the identified functional B cell states in the spleen are characterized by active Cd40 and interferon gamma signaling corresponding to a proliferative state [48]. In contrast, no co-stimulatory signals were detected in pulmonary B cell states, which suggests an inhibition of B cell activity [48]. This finding is consistent with the second functional state identified in pulmonary B cells, which shows Il2 receptor signaling activity as well as the absence of interferon gamma signaling, similar to the functional states of limb muscle and liver B cells [44, 45].

Functional heterogeneity underlying aging and disease identifies emerging cell states

Finally, we investigated the influence of different tissue conditions, such as aging and disease, on the functional states of cell types. In order to interrogate how aging affects the functional cell states of macrophages, we compared the identified states in kidney and limb muscle tissue in young and old mice. Based on our analysis, we observed a significant decline in macrophage polarization due to the inactivation of Nrp1 and Il6r signaling pathways. In particular, while a functional state of kidney macrophages exists in old mice that is characterized by the ability to respond to antigens through Tlr2 signaling, the capacity of limb muscle macrophages to antigen response is significantly reduced. Moreover, the heterogeneous activation pattern of macrophages in young tissues has generally declined, which results in more homogeneous, quiescent states. This observation is consistent with previous studies reporting a significant decline of macrophage activation during aging, which is mediated by reduced Tlr signal transduction capacity due to dysfunctional MyD88 [49].

In order to determine whether the loss of niche-induced functional states is a general characteristic of aging tissues, we compared the number of identified functional states for each cell type (Figure 3B). However, no statistically significant difference could be detected (two-sided Wilcoxon-signed-rank test, *P*-value: 0.64). While in more than 85% of cell types, the number of states differed by at most two, three cell types differed in more than six niche-induced functional states. For example, the increase of detected functional states in hepatic sinusoidal endothelial cells is largely driven by the specific activation of the transcription factors Mef2c, Nr3c1 and Jun through distinct signaling cascades of the same receptors. Indeed, these TFs carry out protective functions in hepatic endothelial cells by stimulating cellular survival, modulating the immune response and preventing the breakdown of the extracellular matrix [50–52]. Similarly, an increase in niche-induced functional states can be observed in pancreatic alpha cells due to an elevated selectivity in the activation of Bmpr1a, Itgav and Ddr1. These receptors regulate key functions of alpha cells, such as glucagon secretion [53], even though their combinatorial effect remains elusive. Furthermore, we observe a few of cellular processes unique to young and old tissues, respectively (Figure 3C and D). In contrast to hepatic sinusoidal endothelial and pancreatic alpha cells, enterocytes of the large intestinal epithelium show a marked reduction of heterogeneity with age, which is associated to the impairment of key signaling pathways, including protective signals, such as Lpa and Somatostatin, as well as fat uptake through the sortilin receptor.

Likewise, we investigated the effects of pathologic tissue conditions on the functional states of cell types exemplified in a dataset of human liver cirrhosis [54]. Hepatic cirrhosis is a chronic disease characterized by the progressive formation of permanent scar tissue through fibrosis. Hepatocytes are the main parenchymal cells of the liver, making up 50–60% of the tissue and are involved in key functions, such as lipid synthesis and detoxification. However, their dysregulation in the context of liver fibrosis and the functional states they attain remains elusive. Therefore, we applied FunRes, to healthy and cirrhotic human liver samples for identifying the functional states of hepatocytes. As a result, our method detected two functional states in healthy liver tissue characterized by Sdc1 and Cd74 signaling. Indeed, Sdc1 is necessary for the uptake and degradation of triglyceride-rich lipoproteins in the liver and is therefore involved in a key function of the liver [55]. In addition, Cd74 is a key component of the response to acute liver injury by antigen processing and host defense [56]. In contrast, a great diversity of functional states can be observed in cirrhotic liver tissue. In total, FunRes identified six functional states that are characterized by interferon gamma, leptin receptor, Cxcr4, Tnf, Cd40 and Caveolin-1 signaling (Figure 3E). While two functional states show a highly pro-inflammatory phenotype mediated by interferon gamma, Tnfrsf1a and Cxcr4, three states are characterized by Lepr, Cd40 and Cav1 signaling. The function of leptin receptor in the context of cirrhosis has been subject to previous studies that demonstrated the production of Tnf and Il1 due to leptin signaling activation [57]. Moreover, in a rat

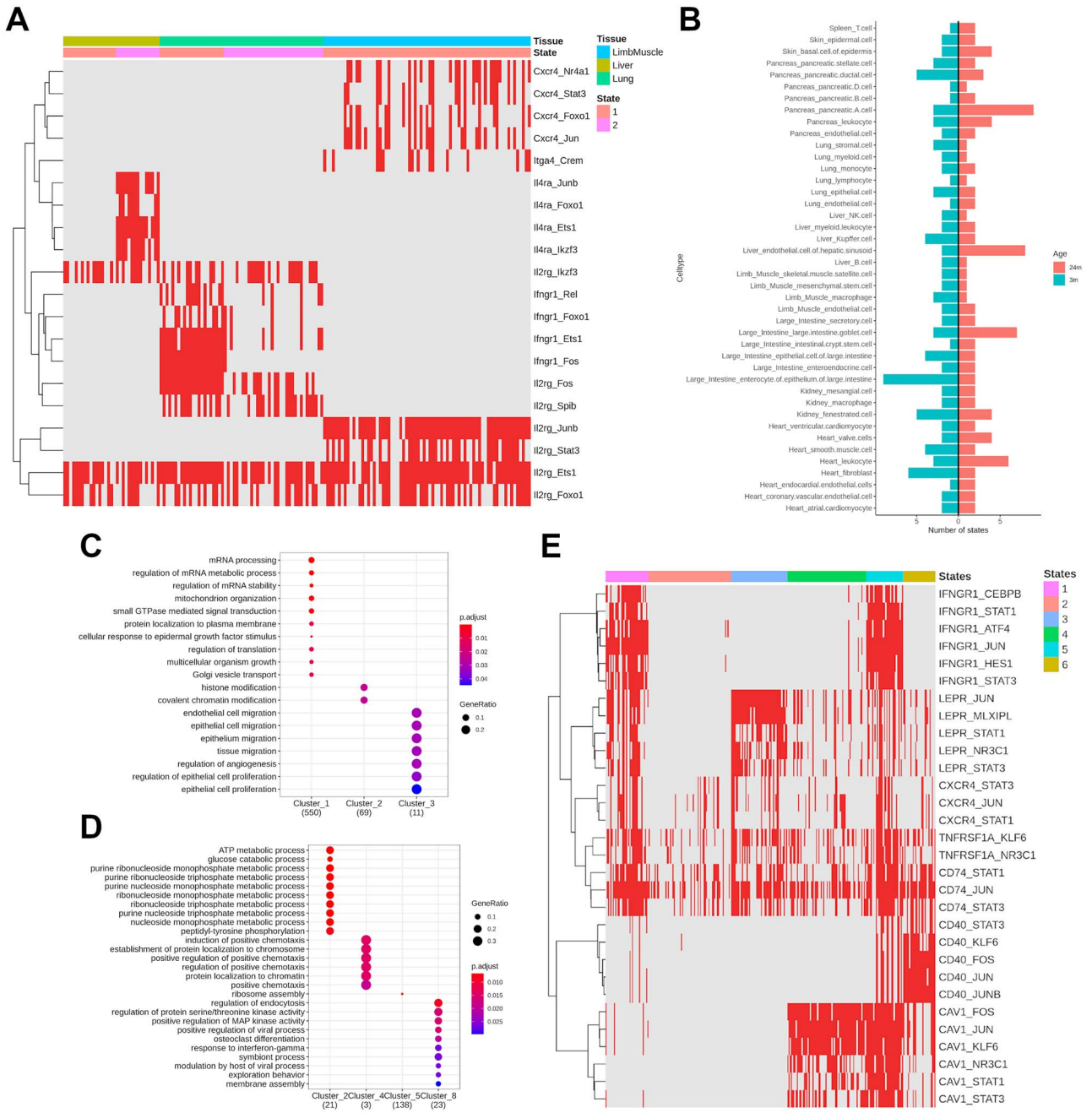

Figure 3. Comparison of functional states across tissues and tissue conditions. (**A**) Heatmap of active signaling cascades induced by cell–cell interactions in B cells of three different tissues. Each row corresponds to a pair of receptor and downstream TF target, whereas columns correspond to cells. Expression of both the receptor and downstream TF target in a cell is depicted in red. Failure to express any of the two genes is depicted in grey. B cells from all tissues share active Il2 signaling cascades targeting common (Ets1, Foxo1) and unique downstream TFs (Stat3, Junb, Spib, Fos). (**B**) Bar chart of identified functional cell states per cell type in young and old tissues. The number of functional states does not significantly change in aging, except for certain cell types, such as pancreatic alpha cells, endothelial cells of the hepatic sinusoid and enterocytes of the large intestine epithelium. (**C** and **D**) Cellular processes unique to pancreatic alpha cells in young and old tissues. The size of the dots represents the fraction of genes in the category that are expressed in the corresponding state, whereas the color indicates significance from blue (less significant) to red (most significant). (**E**) Heatmap of active signaling cascades induced by cell–cell interactions in hepatocytes of cirrhotic liver tissue. Each row corresponds to a pair of receptor and downstream TF target, whereas columns correspond to cells. Expression of both the receptor and downstream TF target in a cell is depicted in red. Six functional cell states were identified that are characterized by Interferon-gamma, CD40, Cav1 and Leptr signaling, respectively.

model of liver cirrhosis, no significant changes in the expression of Lepr have been found, which is consistent with our data. Nevertheless, FunRes only detects significant signaling activity in pathologic livers highlighting the significance of this method. On the contrary, previous studies suggested that the functional hepatocyte state characterized by active Cd40 signaling likely amplifies Fas-dependent apoptosis, thus contributing to the progressive formation of scar tissue in response to injury [58]. Finally, functional states with Caveolin-1 signaling activity regulate endocytosis, energy metabolism and fatty acid

uptake as well as inhibit NOS3, which promotes liver fibrosis [59].

Discussion and Conclusion

In this study, we presented FunRes, a computational method for resolving tissue-specific functional cell states that are determined by receptor–ligand-mediated cell–cell interactions. This method accounts for the effect of cell–cell interactions on specific cellular processes, which have been shown to play a fundamental role in the specification of functional cell states [60–62]. Namely, FunRes identifies functional cell states based on the downstream target genes of cell–cell interactions and the cellular processes they participate in. Hence, our approach is conceptually different from current clustering methods solely relying on differences in the expression of the most variable genes regardless of their cellular function. As a result, unlike other methods, FunRes was able to resolve functional states of various cell types. Importantly, the cell–cell interactions underlying functional cell states were largely supported by previous studies, indicating that FunRes can inform about the cues inducing these states.

The validation of our method was further corroborated by the comparison of functional cell states in different tissues and tissue conditions. In particular, FunRes identified previously reported differences in the functional states of macrophages and B cells, including the appearance of specific functional cell states, such as wound-healing macrophages in the limb muscle. Furthermore, we applied FunRes to several examples of aging and disease to disentangle the composition of functional cell states in these conditions. For instance, in the case of liver cirrhosis, results show the emergence of novel functional hepatocyte states sustained by Caveolin-1 and Leptin signaling that are characterized by dysregulated metabolic processes contributing to pathological tissue functioning.

A limitation of FunRes is that it solely considers receptor–ligand-mediated cell–cell interactions whereas it ignores other ways of cellular communication, such as via exosomes. In addition, this method does not incorporate the effect of environmental factors, such as metabolites, on the functional cell states [63]. In this regard, the method could be extended to overcome some of these limitations.

In summary, FunRes is the first computational method that systematically resolves tissue-specific functional cell states, thereby providing a functional characterization of the identified states. Thus, we believe FunRes will be of great utility in the characterization of the tissue-specific functional landscape and in the identification of dysfunctional cell states underlying aging and disease.

Author Contributions

S.J. implemented the software, conducted the analysis and wrote the manuscript. K.S. implemented the software and conducted the analysis. A.d.S. conceived the idea, supervised the work and wrote the manuscript.

Data Availability

The data underlying this article are available in *figshare*. The weblinks and identifiers of the individual datasets can be found in Supplementary Table S1. The source code of FunRes is deposited in a GitHub repository (https://git-r3lab.uni.lu/kartikeya.singh/funres).

Key Points

- The first method, FunRes is the first method proposed for resolving functional cell states of heterogeneous cell types. As such, FunRes is complementary to existing clustering methods for single-cell RNA-seq data, which have been successfully applied to delineate cell types and subtypes.
- Due to its underlying cell–cell interaction model, FunRes provides insights into the environmental cues inducing the identified functional cell states. Moreover, we showed in several examples that the identified cues have been implicated in the function of the identified states.
- FunRes can be used to compare the functional cell states between different tissues and tissue conditions to reveal differences in their composition. In the presented case studies, FunRes identified known differences between the same cell states in different tissues as well as emerging and vanishing functional states in the context of aging and disease.

Supplementary Data

Supplementary data are available online at *Briefings in Bioinformatics*.

Acknowledgement

We thank Ilya Potapov for fruitful discussions about the cell–cell communication network inference.

Funding

K.S. is supported by the Luxembourg National Research Fund (FNR) under the PRIDE program (Project code: 11012546) within the NextImmune DTU.

Conflict of interest

The authors declare no conflict of interest.

References

1. Gartner ZJ, Prescher JA, Lavis LD. Unraveling cell-to-cell signaling networks with chemical biology. *Nat Chem Biol* 2017;**13**:564–8.
2. Lavin Y, Winter D, Blecher-Gonen R, *et al*. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* 2014;**159**:1312–26.
3. Lugert S, Basak O, Knuckles P, *et al*. Quiescent and active hippocampal neural stem cells with distinct morphologies respond selectively to physiological and pathological stimuli and aging. *Cell Stem Cell* 2010;**6**:445–56.
4. Hsieh J. Orchestrating transcriptional control of adult neurogenesis. *Genes Dev* 2012;**26**:1010–21.
5. Stuart T, Butler A, Hoffman P, *et al*. Comprehensive integration of single-cell data. *Cell* 2019;**177**:1888–1902.e21.

6. Butler A, Hoffman P, Smibert P, *et al*. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;**36**:411–20.
7. Kiselev VY, Kirschner K, Schaub MT, *et al*. SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods* 2017;**14**:483–6.
8. Guo M, Wang H, Potter SS, *et al*. SINCERA: a pipeline for single-cell RNA-Seq profiling analysis. *PLoS Comput Biol* 2015;**11**:e1004575.
9. Ramilowski JA, Goldberg T, Harshbarger J, *et al*. A draft network of ligand–receptor-mediated multicellular signalling in human. *Nat Commun* 2015;**6**:7866.
10. Salwinski L. The database of interacting proteins: 2004 update. *Nucleic Acids Res* 2004;**32**:449D–51.
11. Ben-Shlomo I, Yu Hsu S, Rauch R, *et al*. Signaling receptome: a genomic and evolutionary perspective of plasma membrane receptors involved in signal transduction. *Sci Signal* 2003;**2003**:re9–9.
12. Armstrong JF, Faccenda E, Harding SD, *et al*. The IUPHAR/BPS guide to Pharmacology in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV guide to MALARIA PHARMACOLOGY. *Nucleic Acids Res* 2020;**48**: D1006–21.
13. Keshava Prasad TS, Goel R, Kandasamy K, *et al*. Human protein reference database–2009 update. *Nucleic Acids Res* 2009;**37**:D767–72.
14. Szklarczyk D, Gable AL, Lyon D, *et al*. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;**47**:D607–13.
15. Consortium TU. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 2019;**47**:D506–15.
16. Cunningham F, Achuthan P, Akanni W, *et al*. Ensembl 2019. *Nucleic Acids Res* 2019;**47**:D745–51.
17. Türei D, Korcsmáros T, Saez-Rodriguez J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat Methods* 2016;**13**:966–7.
18. Jassal B, Matthews L, Viteri G, *et al*. The reactome pathway knowledgebase. *Nucleic Acids Res* 2020;**48**:D498–503.
19. Ravichandran S, Hartmann A, del Sol A. SigHotSpotter: scRNA-seq-based computational tool to control cell subpopulation phenotypes for cellular rejuvenation strategies. *Bioinformatics* 2019;**36**:1963–5.
20. Ashburner M, Ball CA, Blake JA, *et al*. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nat Genet* 2000;**25**:25–9.
21. The Gene Ontology Consortium. The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res* 2019;**47**:D330–8.
22. Yu G, Wang L-G, Yan G-R, *et al*. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* 2015;**31**:608–9.
23. Zemmour D, Zilionis R, Kiner E, *et al*. Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat Immunol* 2018;**19**:291–301.
24. Pisco AO, McGeever A, Schaum N, *et al*. A single cell transcriptomic atlas characterizes aging tissues in the mouse. *bioRxiv* 2020;661728.
25. Kovacic JC, Dimmeler S, Harvey RP, *et al*. Endothelial to mesenchymal transition in cardiovascular disease. *J Am Coll Cardiol* 2019;**73**:190–209.
26. Franscini N, Bachli EB, Blau N, *et al*. Gene expression profiling of inflamed human endothelial cells and influence of activated protein C. *Circulation* 2004;**110**:2903–9.
27. Dai X, Okon I, Liu Z, *et al*. A novel role for myeloid cell-specific neuropilin 1 in mitigating sepsis. *FASEB J* 2017;**31**:2881–92.
28. Rőszer T. Understanding the mysterious M2 macrophage through activation markers and effector mechanisms. *Mediators Inflamm* 2015;**2015**:1–16.
29. Schlaepfer E, Rochat M-A, Duo L, *et al*. Triggering TLR2, −3, −4, −5, and −8 reinforces the restrictive nature of M1- and M2-polarized macrophages to HIV. *J Virol* 2014;**88**:9769–81.
30. Zhao Z, Hou X, Yin X, *et al*. TNF induction of NF-κB RelB enhances RANKL-induced Osteoclastogenesis by promoting inflammatory macrophage differentiation but also limits it through suppression of NFATc1 expression. *PLoS One* 2015;**10**:e0135728.
31. Wu J, Wang Y, Jiang Z. TNFSF9 is a prognostic biomarker and correlated with immune infiltrates in pancreatic cancer. *J Gastrointest Cancer* 2020.
32. Huang Y, Du KL, Guo PY, *et al*. IL-16 regulates macrophage polarization as a target gene of mir-145-3p. *Mol Immunol* 2019;**107**:1–9.
33. Xia F, Deng C, Jiang Y, *et al*. IL4 (interleukin 4) induces autophagy in B cells leading to exacerbated asthma. *Autophagy* 2018;**14**:450–64.
34. Chen M, Hong MJ, Sun H, *et al*. Essential role for autophagy in the maintenance of immunological memory against influenza infection. *Nat Med* 2014;**20**:503–10.
35. Yan M, Brady JR, Chan B, *et al*. Identification of a novel receptor for B lymphocyte stimulator that is mutated in a mouse strain with severe B cell deficiency. *Curr Biol* 2001;**11**:1547–52.
36. Schneider P, MacKay F, Steiner V, *et al*. BAFF, a novel ligand of the tumor necrosis factor family, stimulates B cell growth. *J Exp Med* 1999;**189**:1747–56.
37. Schubert D, Klein M-C, Hassdenteufel S, *et al*. Plasma cell deficiency in human subjects with heterozygous mutations in Sec61 translocon alpha 1 subunit (SEC61A1). *J Allergy Clin Immunol* 2018;**141**:1427–38.
38. Holmes ML, Huntington ND, Thong RP, *et al*. Peripheral natural killer cell maturation depends on the transcription factor Aiolos. *EMBO J* 2014;**33**:2721–34.
39. Lunemann S, Langeneckert AE, Martrus G, *et al*. Human liver-derived CXCR6 + NK cells are predominantly educated through NKG2A and show reduced cytokine production. *J Leukoc Biol* 2019;**105**:1331–40.
40. Abel AM, Yang C, Thakar MS, *et al*. Natural killer cells: development, maturation, and clinical utilization. *Front Immunology* 2018;**9**:1869.
41. Janssens S, Beyaert R. Role of toll-like receptors in pathogen recognition. *Clin Microbiol Rev* 2003;**16**:637–46.
42. Parameswaran N, Patial S. Tumor necrosis factor-α Signaling in macrophages. *Crit Rev Eukaryot Gene Expr* 2010;**20**: 87–103.
43. Mauer J, Chaurasia B, Goldau J, *et al*. Signaling by IL-6 promotes alternative activation of macrophages to limit endotoxemia and obesity-associated resistance to insulin. *Nat Immunol* 2014;**15**:423–30.
44. Mingari MC, Gerosa F, Carra G, *et al*. Human interleukin-2 promotes proliferation of activated B cells via surface receptors similar to those of activated T cells. *Nature* 1984;**312**:641–3.
45. Le Gallou S, Caron G, Delaloy C, *et al*. IL-2 requirement for human plasma cell generation: coupling differentiation and proliferation by enhancing MAPK–ERK Signaling. *J Immunol* 2012;**189**:161–73.
46. Mcheik S, Van Eeckhout N, De Poorter C, *et al*. Coexpression of CCR7 and CXCR4 during B cell development controls CXCR4

responsiveness and bone marrow homing. *Front Immunol* 2019;**10**:2970.
47. Vogel MJ, Xie L, Guan H, *et al*. FOXO1 repression contributes to block of plasma cell differentiation in classical Hodgkin lymphoma. *Blood* 2014;**124**:3118–29.
48. Vazquez MI, Catalan-Dibene J, Zlotnik A. B cells responses and cytokine production are regulated by their immune microenvironment. *Cytokine* 2015;**74**:318–26.
49. Dunston CR, Griffiths HR. The effect of ageing on macrophage toll-like receptor-mediated responses in the fight against pathogens. *Clin Exp Immunol* 2010;**161**:407–16.
50. Salameh A, Galvagni F, Anselmi F, *et al*. Growth factor stimulation induces cell survival by c-Jun·ATF2-dependent activation of Bcl-X L. *J Biol Chem* 2010;**285**:23096–104.
51. Sturtzel C, Testori J, Schweighofer B, *et al*. The transcription factor MEF2C negatively controls Angiogenic sprouting of endothelial cells depending on oxygen. *PLoS One* 2014;**9**:e101521.
52. Goodwin JE, Feng Y, Velazquez H, *et al*. Endothelial glucocorticoid receptor is required for protection against sepsis. *Proc Natl Acad Sci* 2013;**110**:306–11.
53. Nielsen SS, Christensen GL, Holst JJ, *et al*. Regulation of pancreatic α-cell function and proliferation by bone morphogenetic protein 4 (BMP4) in vitro. *Endocrinology* 2016;**157**:3809–20.
54. Ramachandran P, Dobie R, Wilson-Kanamori JR, *et al*. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* 2019;**575**:512–8.
55. Stanford KI, Bishop JR, Foley EM, *et al*. Syndecan-1 is the primary heparan sulfate proteoglycan mediating hepatic clearance of triglyceride-rich lipoproteins in mice. *J Clin Invest* 2009;**119**:3236–45.
56. Koch KS, Leffert HL. Ectopic expression of CD74 in Ikkβ-deleted mouse hepatocytes. *Acta Histochem* 2011;**113**:428–35.
57. Otte C, Otte J-M, Strodthoff D, *et al*. Expression of leptin and leptin receptor during the development of liver fibrosis and cirrhosis. *Exp Clin Endocrinol Diabetes* 2004;**112**:10–7.
58. Afford SC, Randhawa S, Eliopoulos AG, *et al*. CD40 activation induces apoptosis in cultured human hepatocytes via induction of cell surface Fas ligand expression and amplifies Fas-mediated hepatocyte death during allograft rejection. *J Exp Med* 1999;**189**:441–6.
59. Fernandez-Rojo MA, Ramm GA. Caveolin-1 function in liver physiology and disease. *Trends Mol Med* 2016;**22**:889–904.
60. Schrier SB, Hill AS, Plana D, *et al*. Synergistic communication between CD4+ T cells and monocytes impacts the cytokine environment. *Sci Rep* 2016;**6**:34942.
61. Shalek AK, Satija R, Shuga J, *et al*. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 2014;**510**:363–9.
62. Xue Q, Lu Y, Eisele MR, *et al*. Analysis of single-cell cytokine secretion reveals a role for paracrine signaling in coordinating macrophage responses to TLR4 stimulation. *Sci Signal* 2015;**8**:ra59–9.
63. Vishwakarma A, Rouwkema J, Jones PA, *et al*. The need to study, mimic, and target stem cell niches. *Biol Eng Stem Cell Niches* 2017;3–13.