



An interpretable deep-learning architecture of capsule networks for identifying cell-type gene expression programs from single-cell RNA-sequencing data

Lifei Wang^{1,2,3}, Rui Nie^{1,2,3}, Zeyang Yu^{1,2,3}, Ruyue Xin⁴, Caihong Zheng^{1,2,3}, Zhang Zhang⁴, Jiang Zhang⁴  and Jun Cai^{1,2,3} 

Single-cell RNA sequencing (scRNA-seq) technologies are used to characterize the heterogeneity of cells in cell types, developmental stages and spatial positions. The rapid accumulation of scRNA-seq data has enabled single-cell-type labelling to transform single-cell transcriptome analysis. Here we propose an interpretable deep-learning architecture using capsule networks (called scCapsNet). A capsule structure (a neuron vector representing a set of properties of a specific object) captures hierarchical relations. By utilizing competitive single-cell-type recognition, the scCapsNet model is able to perform feature selection to identify groups of genes encoding different subcellular types. The RNA expression signatures, which enable subcellular-type recognition, are effectively integrated into the parameter matrices of scCapsNet. This characteristic enables the discovery of gene regulatory modules in which genes interact with each other and are closely related in function, but present distinct expression patterns.

The technology of single-cell RNA sequencing (scRNA-seq) enables the dissection of each heterogeneous cell and its transcriptome profile within a tissue, organ or organism^{1,2}. Many studies, such as [Human Cell Atlas](#), have demonstrated the power of scRNA-seq to enable unprecedented views of cell types or states. The recent rapid accumulation of scRNA-seq data makes it possible to assign each single cell a label that records cell-type identity, developmental stage^{3,4}, spatial position^{5,6} or another biological feature of the cell. To meet the requirement of cell-type labelling, several bioinformatics algorithms for single-cell transcriptome analysis have been developed^{7–11}. For example, Seurat assigns distinct identity tags to cells by clustering scRNA-seq data^{12,13} and Moana constructs cell-type classifiers from heterogeneous scRNA-seq datasets under a hierarchical machine learning framework¹⁴. Approaches such as canonical correlation analysis (CCA) and mutual nearest neighbours (MNN) realize the batch merging of datasets for effective recognition of rare cell states^{12,15}. The existing supervised or unsupervised learning methods embody a robust, quantitative, expression-based definition of cellular identity. However, feature selection to reveal how the classifiers produce each labelling result for a single cell is required. Some preliminary analysis of differential gene expression and co-expression modules has attempted to extract transcriptome features from groups of single cells with the same cell type^{12,13}. Unfortunately, the genes obtained were still inadequate to explain the gene expression programs that could determine the cell-type labelling for each single cell. In particular, why a single cell is labelled to be a certain cell type is not clear because of the lack of interpretability in the ‘black box’ machine learning models¹⁶.

Deep learning neural networks, as state-of-the-art machine learning models, have been used in many successful implementations. These networks typically have the ability to extract insights from biological data as well as image data^{17–21}. However, traditional deep learning network architecture still lacks transparency in the decision-making process^{22–24}. This black box problem has always blurred the interpretability of deep networks, even though the recent advances have introduced a ‘capsule structure’ into the hidden layer. A capsule structure is defined as a neuron vector representing a set of properties of a specific object²⁵. The flexible modular architecture of capsule networks, being composed of capsules in interconnected modules, provides the possibility of opening the black box of deep learning and enabling us to interpret complex biological networks¹⁷.

Here, we have designed an interpretable deep-learning architecture of capsule networks (scCapsNet) and made the decision-making black box transparent by analysing internal weight parameters among capsule structures. We evaluated the values assigned by scCapsNet for single-cell transcriptome analysis using multiple scRNA-seq datasets, such as mouse retinal bipolar cells (mRBC) data and human peripheral blood mononuclear cells (hPBMC) data. We used two-dimensional principal component analysis (PCA) on the internal weight parameters in the feature extraction layer to properly define a set of core genes; this set enables the identification of groups of single cells with same cell type. In addition, the internal weight parameters in the model effectively embed the single-cell expression profile into a low-dimensional vector for each gene. The vector contains both informative gene-expression signatures and properties of cell-type labelling contributed by knowledge of the gene. Thus, the core gene regulatory modules in which genes are

¹China National Center for Bioinformation, Beijing, China. ²Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China. ³University of Chinese Academy of Sciences, Beijing, China. ⁴School of Systems Science, Beijing Normal University, Beijing, China. ✉e-mail: zhangjiang@bnu.edu.cn; juncaij@big.ac.cn

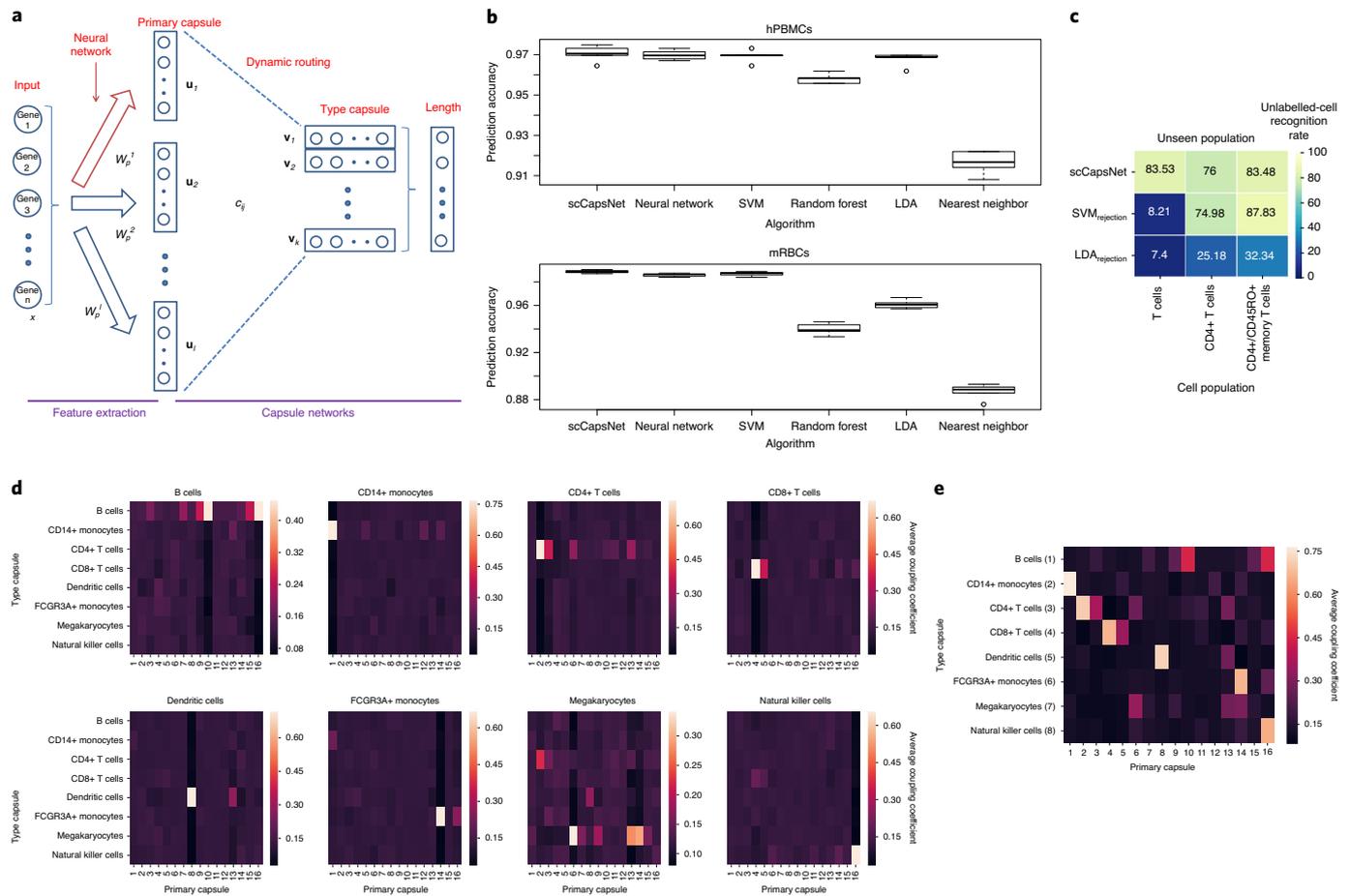


Fig. 1 | The architecture of scCapsNet and its cell-type-recognition characteristics. a, Two-layer architecture of scCapsNet. The first layer consists of l parallel fully connected neural networks for feature extraction from the inputs of single-cell gene expression. The primary capsule of vector u_i is the output of the neural network i . The subsequent layer is a Keras implementation of capsule networks for classification. The parameters of coupling coefficients express the mathematical contributions of the primary capsules to the type capsules for cell-type recognition. The length of each type capsule v_j represents the probability of single cell x belonging to the corresponding cell type. **b**, The performance for single-cell type recognition in the scCapsNet model. With hPBMC and mRBC scRNA-seq data, the cell-type prediction accuracies in five replicates were compared with other algorithms: neural network, support vector machine (SVM), random forest, linear discriminant analysis (LDA) and nearest neighbour. The box-and-whisker plots drawn by boxplot from R show the prediction accuracies in five replicates of each method. **c**, The rejection option evaluation in the unseen population experiment on the scCapsNet, SVM_{rejection} and LDA_{rejection} models. The column labels of the heatmap represent the three sub-tasks in the unseen population experiment: the one with removed T cells, the one with removed CD4+ T cells, and the one with only the CD4+/CD45RO+ memory T cell subpopulation removed from the training set. The heatmap shows the recognition rates of the unlabelled single cells that are not seen during training. **d**, Heatmaps of the matrices of average coupling coefficients for the hPBMC dataset. The heatmaps represent the average coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and natural killer (NK) cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, Overall heatmap of the combined matrix of average coupling coefficients. The combined matrix contains the effective type capsule row in Fig. 1d where its recognition type is in accordance with the type of single cells input.

closely relevant in function but present distinct transcript expression patterns can be recognized in the form of a tiny, tight cluster of embedded vectors.

Results

The architecture of capsule network model and its performance for single-cell type recognition. In our deep learning model scCapsNet to enable single-cell type recognition, we designed the architecture to consist of the feature extraction module and the capsule network module (Fig. 1a). Instead of the convolutional kernels in a traditional capsule network, multiple parallel fully connected neural networks play the part of the feature extractor in the feature extraction module (Fig. 1a). These neural networks convert the input of the single-cell RNA expression profile into ‘primary capsule’

vectors through weight matrices and a rectified linear unit (ReLU) activation function. Then, in the capsule networks, the features are delivered from the primary capsules to the next ‘type capsule’ vectors for cell-type recognition by iterative dynamic routing (Fig. 1a). The coupling coefficient matrices in the hidden layer represent the mathematical contributions of the primary capsules to the type capsules (see Methods for the details).

The scRNA-seq data of hPBMCs and mRBCs from 10x Genomics and Drop-Seq²⁶ platforms were used to evaluate the performance of the scCapsNet model in single-cell-type recognition^{27,28}. The shuffle-split cross-validation and prediction results reveal that the model has a powerful recognition ability of up to 99% and 97% accuracy for the two scRNA-seq datasets, respectively. We further demonstrate the effectiveness of the scCapsNet model by comparing

it with other algorithms, such as the black box neural network, SVM and random forest, using the same scRNA-seq data (Fig. 1b). The evidence supports the conclusion that the scCapsNet method is suitable for single-cell-type labelling with competitive recognition accuracy. Besides our original hPBMC dataset we selected a different hPBMC dataset from the 10x Genomics platform^{11,29}. The models were trained using one dataset and then evaluated using another dataset. The result in Extended Data Fig. 1a shows that scCapsNet is the top-performing classifier for this task. We also tested the classification performance among different models across four human pancreatic cell datasets from different single-cell RNA-seq protocols. These four datasets were those used in Abdelaal et al.^{11,30–33}. The recognition accuracies in Extended Data Fig. 1b correspond to four sub-tasks, in each of which one of the four datasets was used as a test set and the other three datasets were used as training sets. The result suggests that scCapsNet has stable performance of single-cell type recognition in new scRNA-seq datasets.

In a realistic scenario for scRNA-seq data analysis, some single cells have unknown cell-type labels and remain unassigned. The subcellular-type recognition model should incorporate a rejection option evaluation to test whether the classifiers indeed leave these single cells unlabelled. The rejection option evaluation can be implemented using an unseen population experiment and a negative control experiment¹¹. In the unseen population experiment, the classification model is first trained on labelled cell subpopulations from one dataset. A group of differently labelled cells in the same dataset, that are not seen during training, are then used in the test set for rejection. In the negative control experiment, the classifiers are used to predict the single-cell types of one tissue after training on a dataset from a different tissue.

Our scCapsNet model includes a rejection option by setting a threshold of 0.9 to the maximum probability in the output process of cell-type identification. The maximum probability above the threshold determines the cell type that the input single cell belongs to. When the maximum probability value is lower than the threshold, scCapsNet triggers the rejection option and classifies the input single cell as unlabelled. We used the same datasets and sub-tasks as in Abdelaal's benchmark paper in the unseen population experiment and the negative control experiment for the rejection option evaluation on scCapsNet, the SVM_{rejection} and LDA_{rejection} methods. The three sub-tasks in the unseen population experiment were the one with removed T cells, the one with removed CD4+ T cells, and the one with only the CD4+ /CD45RO+ memory T cell subpopulation removed from the training set^{11,27}. The result demonstrates that scCapsNet has a stable and much improved rejection option performance in all three sub-tasks, whereas the SVM_{rejection} and LDA_{rejection} methods fail in at least one sub-task (Fig. 1c). In a realistic scenario for scRNA-seq data analysis, it is important that the algorithm is able to discover a novel cell subgroup in which previously known cell-type labels are unassigned to the single cells. In the unseen population experiment, scCapsNet recognizes almost

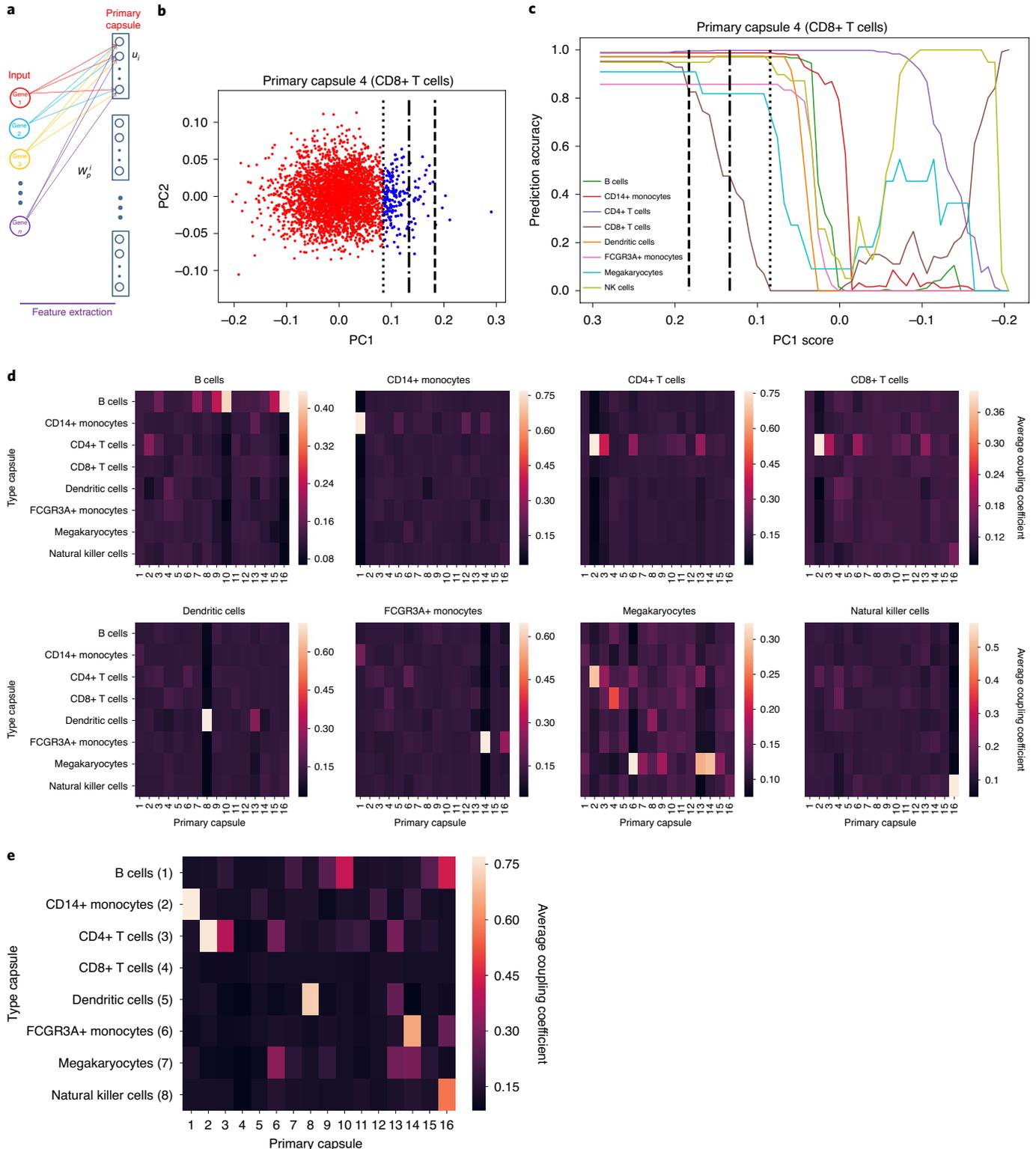
80% of the unlabelled single cells that are not seen during training. This result suggests that if the scRNA-seq dataset has an unknown cell subgroup, scCapsNet could distinguish it clearly from many known subcellular types. The negative control experiment contains two groups of datasets, the group of human cell datasets of PBMC and pancreas tissues, and the group of mouse cell datasets of visual cortex and pancreas tissues^{11,27,30,34}. As desired and expected, scCapsNet rejects close to 100% of the negative control cells (Extended Data Fig. 1c).

Interpretable characteristics in primary and type capsules relative to cell type recognition. The coupling coefficients for each single cell are available in the well trained scCapsNet model using hPBMC scRNA-seq data. They are weight values representing the contribution of the primary capsules to the type capsules. We calculated the matrices of average coupling coefficients for the single cells with the same cell-type label, as shown in the eight heatmaps (Fig. 1d) (see Methods for the detailed calculation of average coupling coefficients). Here we give an example to explain the meaning of the brightness of the element at row 3 (CD4+ T cell row) and column 2 in the first heatmap (Fig. 1d). This brightness value measures the contribution to the CD4+ T cell recognition of the features extracted from the transcriptome of single B cells by the second primary capsule. The meanings of the brightness for other elements in the heatmaps are similar. This first heatmap exhibits an interesting pattern: that the matrix of average coupling coefficients for the B cells has one or very few very-high-value elements in only the B-cell type capsule row. Other heatmaps follow a similar pattern (Fig. 1d). This pattern demonstrates that primary capsules effectively extract informative features from single cells with the same cell-type label, which is responsible for recognition of the cell type. Following this pattern, the effective type capsule row whose recognition cell type is in accordance with the type of input single cells was selected from each heatmap in Fig. 1d. The row was organized into a new matrix, visually represented as an overall heatmap (Fig. 1e). The high-value elements in this matrix are sparse not only in the row but also in the column. The sparsity of the column elements ensures that cell-type recognition for each type capsule is independently determined by one or two different primary capsules. As a typical example, the unique high-value element at row 2 and column 1 specifies that the first primary capsule completely corresponds to the CD14+ monocyte type capsule, which enables recognition of CD14+ monocytes. The above characteristics indicate that the simple connection between primary capsules and type capsules can be decomposed via coupling coefficient analysis. The characteristics of coupling coefficients linking primary capsules with type capsules are similar in the well trained scCapsNet model using mRBC scRNA-seq data (Extended Data Fig. 1d–e). In this way, the primary-type capsule architecture and internal parameter analysis in scCapsNet has made the black box transparent and interpretable in the decision-making process of cell labelling.

Fig. 2 | The identification of the core gene set responsible for recognition of each cell type. **a**, An internal weight matrix of neural networks connecting inputs and each primary capsule in feature extraction module. Each gene can be represented by a real-value low-dimensional column vector of the weight matrix. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the fourth primary capsule. Each dot represents a gene with a rank according to the score of principal components. The three lines represent three sliding values, which cut out some genes along the principal component score. A group of core genes marked in blue are defined by the dotted line. **c**, Curves of cell-type recognition accuracies. The ranked genes are defined by a sliding cutoff value on the principal component score are excluded from the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The degrading recognition accuracies corresponding to the three sliding lines in Fig. 2b are marked. The dotted line defines a group of core genes responsible for CD8+ T cell identification, where the recognition accuracy degrades to close to 0 for CD8+ T cells but only slightly decreases for any other cell type. **d**, The heatmaps of the revised matrices of average coupling coefficients for the hPBMC dataset with the loss of the group of CD8+ T cell core genes in the inputs of the scCapsNet model. The heatmaps represent the revised average coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, Revised overall heatmap of the combined matrix of average coupling coefficients. The combined matrix contains the effective type capsule row in **d** where its recognition type matches the type of single cells input.

Feature extractors in primary capsules determining groups of core genes responsible for subcellular-type identification. As well as coupling coefficients, the internal weight parameters in the networks connecting the inputs with each primary capsule are estimated in the well trained scCapsNet model (Fig. 2a). These parameters form a weight matrix for each primary capsule, which provides a low-dimensional column vector representation for each input gene. Through PCA on these vectors in the weight matrix, the input genes

have a rank according to the scores of principal components for each primary capsule (Fig. 2b and Extended Data Figs. 2b–9b). For example, Fig. 2b shows the plot of two-dimensional PCA on column vectors in the weight matrix for the fourth primary capsule. A cutoff value of the principal component score corresponds to a group of ranked genes. When we exclude this group of genes in the inputs of the scCapsNet model, a new set of cell-type recognition accuracies are obtained. We slide the cutoff values along the principal component



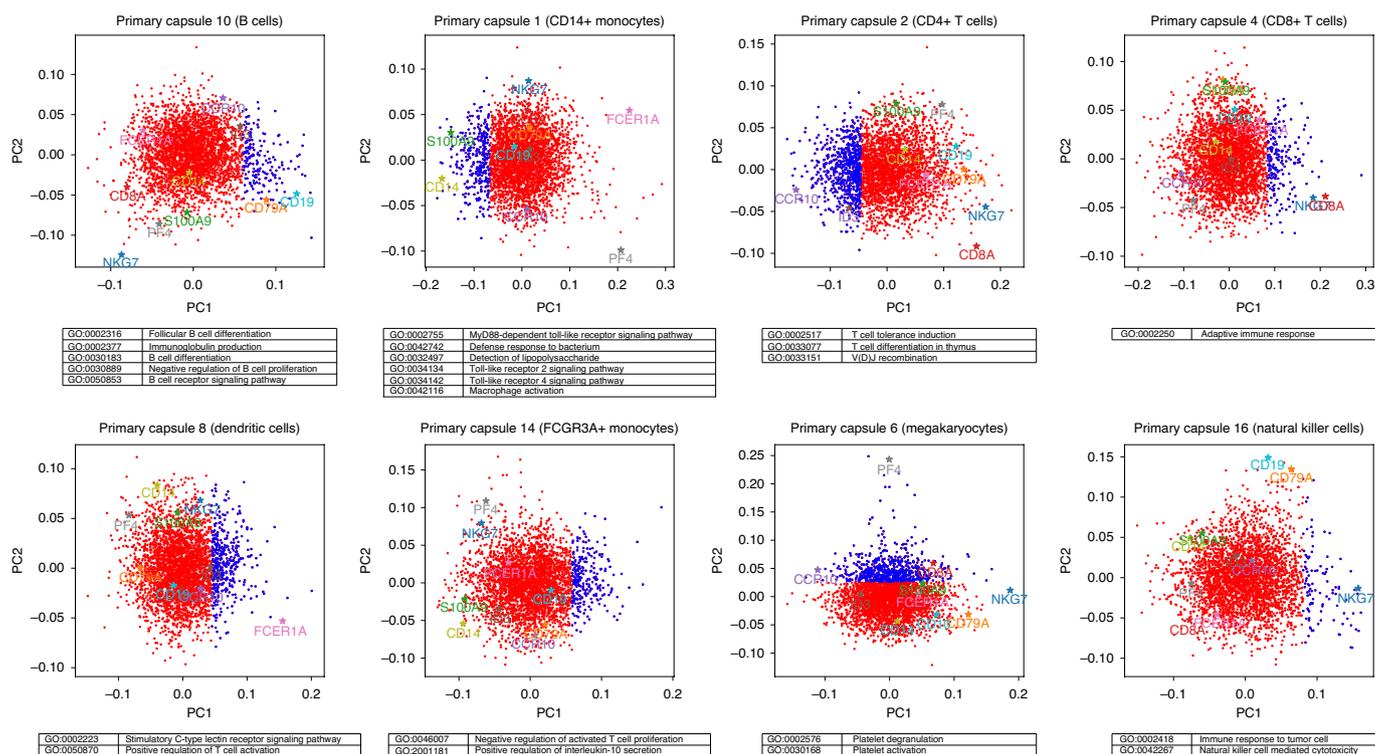


Fig. 3 | The core genes that are essential for the biological functions of different subcellular types. The scatter plots depict the two-dimensional PCA on the column vectors of weight matrices that fully connect inputs and primary capsules 10, 1, 2, 4, 8, 14, 6 and 16. They define the groups of core genes (in blue dots), contributing to the identification of B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. Several well studied cell-type markers are represented as coloured stars with the gene name underneath. Some enriched GO terms corresponding closely to the group of core genes are listed below each scatter plot.

scores in both ascending and descending order. There exists a set of cell-type-recognition-accuracy curves, where the recognition accuracy of one cell type reduces much more sharply than others with the ascending or descending sliding cutoff values. For example, the curves in Fig. 2c reflect the degradation of cell-type-recognition accuracies with the descending sliding cutoff values along the principal component scores in Fig. 2b. The appropriate threshold of the principal component score as a dotted line, where the recognition accuracy of CD8+ T cells is close to 0 but others slightly decrease, corresponds to a group of core genes responsible for CD8+ T cell identification (Fig. 2b,c). The exclusion of this group of core genes in the inputs specially results in the modification of the parameter matrix of average coupling coefficients for the CD8+ T cells (Fig. 1c, Fig. 2d). The new high-value elements in this matrix occur in the CD4+ T cell type capsule row (Fig. 2d). Average coupling coefficients with high values in the CD8+ T cell type capsule row, which were previously responsible for recognition of CD8+ T cells, become much lower (Fig. 1d, Fig. 2d,e). The above interpretation on the modified parameter matrices of average coupling coefficients is consistent with the outputs of the scCapsNet model. First, the loss of the core genes prevents CD8+ T cells from being correctly identification, but rarely affects the recognition of other cell types. Second, CD8 T cells are misclassified as CD4+ T cells.

Similarly, the sixth primary capsule defines another group of core genes responsible for megakaryocyte identification. The recognition accuracy of megakaryocytes decreases rapidly with the sliding cutoff value on the second principal component of its weight matrix (Extended Data Fig. 7). We performed similar procedures for the recognition of the remaining subcellular types. In sum, we performed PCA analysis on the weight matrices for primary capsules 1, 2, 4, 6, 8, 10, 14 and 16 to obtain groups of core genes, which

are responsible for the identification of CD14+ monocytes, CD4+ T cells, CD8+ T cells, megakaryocytes, dendritic cells, B cells, FCGR3A+ monocytes, and NK cells, respectively (Fig. 2, Extended Data Figs. 2–8, and Supplementary Table 1). Additional evidence using mRBC scRNA-seq data also suggests that feature extractors in selected primary capsules in scCapsNet can define groups of core genes responsible for mRBC subcellular-type identification (Extended Data Fig. 9).

Groups of core genes for different subcellular types being essential in biological functions. Typical cell markers and some previously reported cell-type-related genes in hPBMCs are marked with coloured stars in the PCA plots on the weight matrices connecting input genes with primary capsules (Fig. 3). The plots show that the primary capsules of scCapsNet model correctly extract the well known marker genes relevant to different subcellular types, such as *CD19* and *CD79A* for B cells, *CD14* and *S100A9* for CD14+ monocytes, *CCR10* and *ID3* for CD4+ T cells, *CD8A* and *NKG7* for CD8+ T cells, *FCER1A* for dendritic cells, *PF4* for megakaryocytes, and *NKG7* for NK cells²⁷. The cell-type-associated genes, such as the differentially up-expressed genes *CTSL*, *EPS8*, *CKB* and *CIQA* in FCGR3A+ monocytes, are also included in the core gene sets relevant to distinct subcellular types (Extended Data Fig. 10). We further analysed GO enrichment and reactome pathway for the groups of subcellular-type core genes defined by our scCapsNet model. The results demonstrate that each group of core genes is enriched in the special pathways which are closely connected with the corresponding cell type (Fig. 3 and Supplementary Table 2). For example, the GO terms and pathways that are strongly associated with B cell function, such as ‘B cell differentiation’, ‘B cell receptor signalling pathway’, ‘Immunoglobulin production’, and ‘Antibody

maturation relative DNA mismatch repair (MMR), are enriched in the genes responsible for the identification of B cells³⁵. Similarly, the enriched terms of ‘Toll like receptor signaling pathway’, ‘Defense response to bacterium’, ‘Detection of lipopolysaccharide (LPS)’, and ‘Macrophage activation’ in CD14+ monocyte genes describe the traits of CD14+ monocytes. In sum, the groups of core genes responsible for the subcellular-type identification in scCapsNet are essential for the biological functions of different subcellular types.

The alternative embedding representation of each gene in scCapsNet enabling the identification of gene modules. For hPBMC scRNA-seq data, the primary capsules 1, 2, 4, 6, 8, 10, 14 and 16 in scCapsNet effectively extract cell-type features from the input of single-cell RNA expression through the weight matrices. The gene names are used as indexes to concatenate column vectors of these weight matrices. Thus each gene is embedded into an alternative vector representation in a low-dimensional space. We performed PCA on the embedded vectors for all the genes. The two groups of core genes for CD8+ T cells and for CD4+ T cells, as well as the core genes for CD14+ monocyte and FCGR3A+ monocytes, can be separated along the first principal component (PC1) and the second principal component (PC2), respectively (Fig. 4a,b). The two groups of core genes for B cells and NK cells are distinguished along the fourth principal component (PC4) (Fig. 4c). And the two-dimensional PCA visualization between principal component 5 (PC5) and principal component 6 (PC6) describes the differences of core genes between dendritic cells and megakaryocytes (Fig. 4d). The separation in dimension-reduction visualization suggests that the core genes of each subcellular type determined by the scCapsNet model are separable in the embedded vector space. Thus scCapsNet provides a meaningful vector representation for each gene for which the RNA expression signature and its attributes that enable subcellular-type recognition are effectively embedded.

Some gene modules present in the form of tiny compact clusters in the two-dimensional T-SNE (t-distributed stochastic neighbour embedding) visualization of the embedding representation, where the Euclidean distance exhibits the intrinsic similarity between the embedded genes (Fig. 4e). The embedding representations of genes in scCapsNet enable the recognition of some gene regulatory modules in the tiny clusters and the annotation of their connections to cellular phenotypes. The four genes *KIR2DL3*, *KIR2DL1*, *KLRF1* and *GZMB*, defined as the NK cell core genes in scCapsNet, reside closely in the embedding space and form a gene module (Fig. 4e,f). By contrast, they are mostly scattered in a T-SNE visualization of gene expression profiles because of their utterly distinct expression (Fig. 4g,h). Previous literature suggests that *KIR2DL3*, *KIR2DL1*, *KLRF1* and *GZMB* represent simultaneous up-regulation, and initiate the activation of cytotoxic T lymphocytes and/or NK cells in the blood of stroke patients³⁶. Also, *KIR2DL1* and *GZMB* are regulated by the same transcription factor CREB1 as are its target genes and interact with the ubiquitin factor Ube4a^{37,38}. The above evidence suggests that *KIR2DL3*, *KIR2DL1*, *KLRF1* and *GZMB* form a gene module in which they are physically interacting and have similar biological function. The embedding representations of genes in scCapsNet enable the recognition of some gene regulatory modules in which genes are closely relevant in function but present distinct RNA expression patterns.

Discussion

As a black box classification model, deep learning has been applied primarily in biology and medicine. Here we propose scCapsNet, an interpretable deep-learning architecture of capsule networks for exploratory analysis of scRNA sequencing data. In the model architecture, the input of the single-cell transcriptome profile has parallel connections with the primary capsules, which extract and deliver features to the next type capsules for decisions of cell-type labelling. Such an architecture provides the key advantage that scCapsNet has the ability to make the deep-learning black box transparent through the direct interpretation of internal parameters. In this way, scCapsNet can extract core genes and informative gene modules that describe the gene expression programs of distinct subcellular types.

Feature selections, including three strategies of ‘Filter’, ‘Wrapper’ and ‘Embedded’, are effective ways to evaluate how the single cells are labelled with different identities. Independent of the training process, the ‘Filter’ strategy measures the intrinsic properties of the feature subset specific in each subcellular type via univariate statistics. Differential expression analysis, a method of the typical ‘Filter’ strategy, extracts the transcriptome features from groups of single cells with the same cell type. Compared with differential expression analysis, we wondered whether the core genes recognized by scCapsNet can more accurately describe the gene expression programs of different subcellular types. We visualized the separated subcellular-type clusters via the T-SNE calculation on the hPBMC scRNA-seq data with all the genes (Fig. 5a). The T-SNE plot of single-cell transcriptome data without the group of B cell genes defined by scCapsNet is visualized for comparison (Fig. 5b). The separated red B cells move next towards CD4+ T cells, while the position and the shape of other cell clusters remain the same (Fig. 5a,b). Differential expression analysis on the hPBMC scRNA-seq data via the popular Seurat package also extracted an equal amount of differentially expressed B cell genes^{12,13}. The T-SNE plot without these differentially expressed genes shows that it can no longer clearly distinguish among subcellular types because all the single cells move close to each other (Fig. 5c). The result indicates that the groups of genes defined by scCapsNet are essential for recognition of the subcellular-type gene expression program.

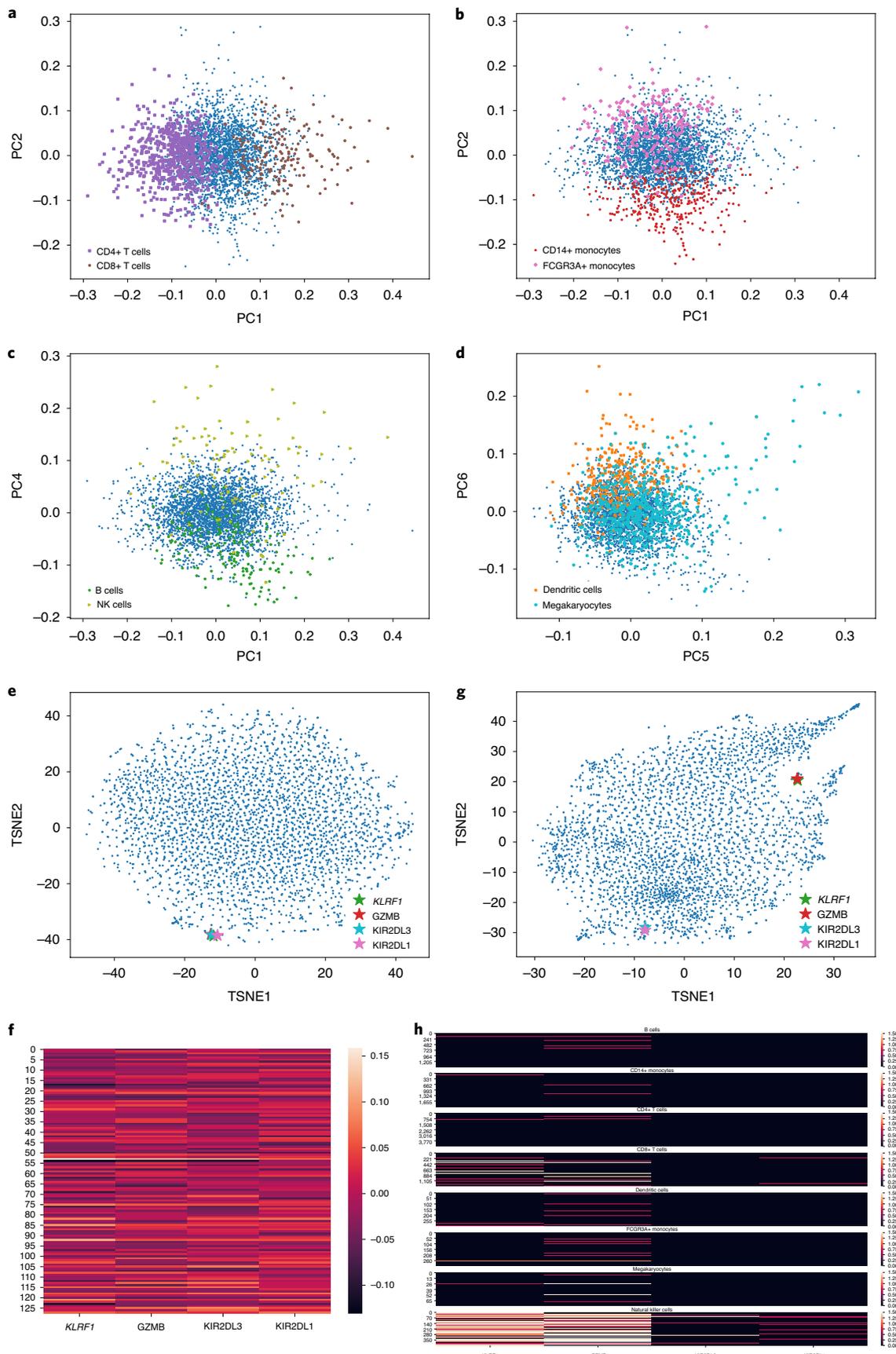
During classification, some existing machine learning algorithms can extract gene features that explain the reasons for cell-type classification. The feature selection in the ‘Embedding’ strategy is contained in the learning process of classifier algorithms such as decision tree or random forest. But this strategy usually optimizes in local space. Moreover, one or two features on each branch of classifier tree determine the classification of several subcellular types. So the subset of gene features responsible for recognition of each subcellular type is indivisible. The feature extractor module is integrated with the classifier in the ‘Wrapper’ strategy. The prediction accuracy of the model is taken as the selection standard of the features. Recursive feature elimination is a typical example of this strategy for feature selection. However, it is computationally more expensive owing to the repeated learning steps and cross-validation required. The strategy for recognition of core genes in scCapsNet is like a combination of ‘Wrapper’ and ‘Embedding’. Unlike the random forest algorithm, feature extractors in primary capsules determine groups of high-quality core genes for each subcellular type. Moreover, the feature extractors in primary capsules effectively

Fig. 4 | An embedding representation of each gene integrating its RNA expression signature and its cell-type-labelling attribute in scCapsNet.

a–d, Visualization of PCA on the embedded gene vectors along different principal components. Each dot represents an embedded vector for a gene and its cell-type-labelling attribute is marked in colour. **e**, The two-dimensional T-SNE (t-distributed stochastic neighbour embedding) visualization for the embedding representation of all the genes. The four genes *KIR2DL3*, *KIR2DL1*, *KLRF1* and *GZMB*, marked with coloured stars, form a tiny compact cluster. **f**, Heatmap of the four embedded vectors of genes *KIR2DL3*, *KIR2DL1*, *KLRF1*, and *GZMB*. **g**, The two-dimensional T-SNE visualization for expression profiles of all the genes. The genes *KIR2DL3* and *KIR2DL1* are separable from the genes *KLRF1* and *GZMB*. **h**, RNA expression heatmap of the genes *KIR2DL3*, *KIR2DL1*, *KLRF1* and *GZMB*.

integrate the RNA expression signature and its cell-type-labelling attribute into the embedding representation of each gene. The embedding representations of genes in scCapsNet enable the

recognition of some gene regulatory modules in the tiny clusters and the annotation of their connection to cellular phenotypes (Fig. 4e–h). Instead of recursively selecting one or a few genes that



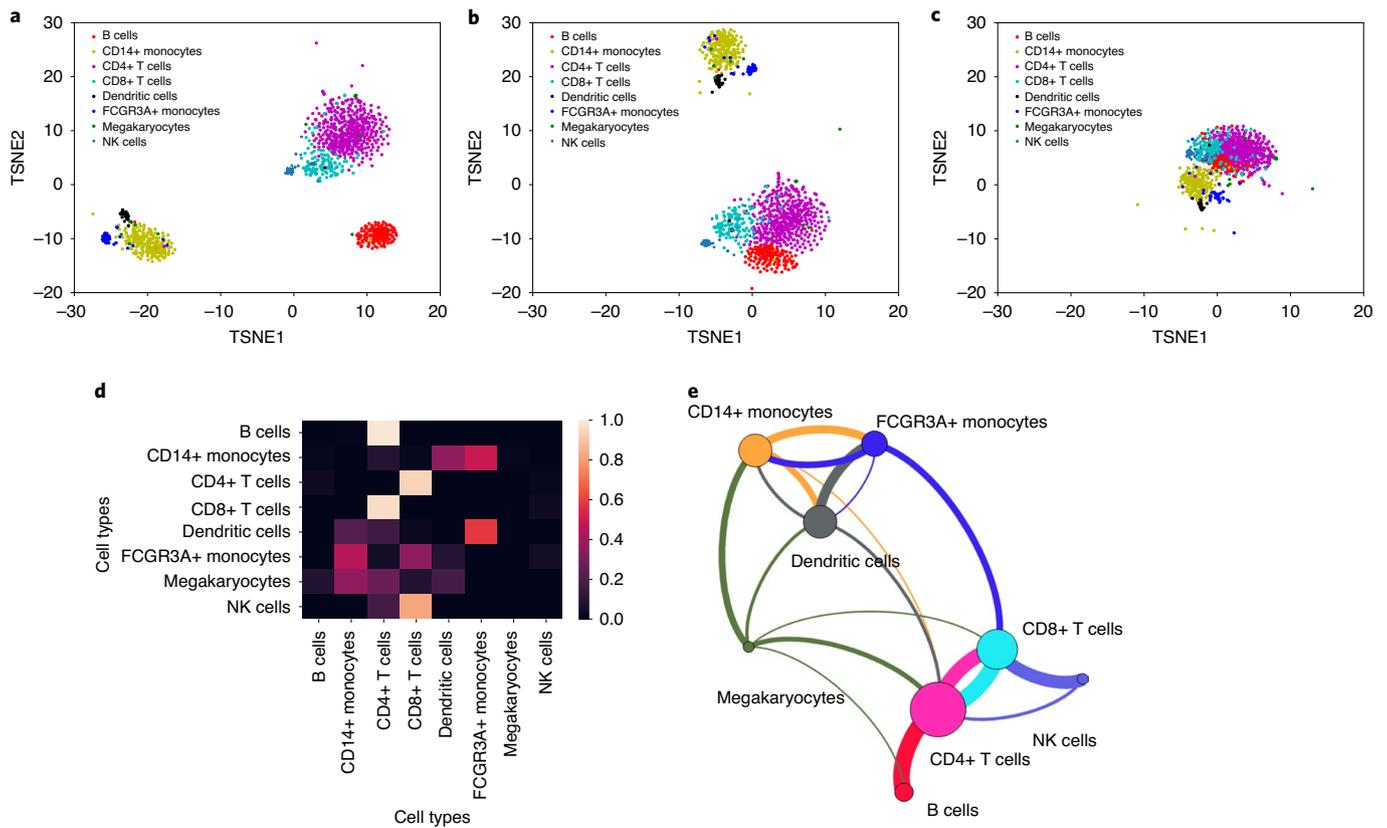


Fig. 5 | Some characteristics of the core genes recognized by scCapsNet. a, The T-SNE plot of the single-cell expression profile of the hPBMC scRNA-seq data. The single cells in different subcellular types are marked with distinct colours. **b**, The T-SNE plot of single-cell transcriptome data without the group of B cell core genes defined by scCapsNet. The view window size is the same as that in **a**. **c**, The T-SNE plot of single-cell transcriptome data excluding the differentially expressed genes in B cells defined by the Seurat package. The view window size is the same as that in **a**. **d**, Heatmap representation for the labelling conversions among different cell types owing to loss of the corresponding core genes. The row represents the true cell type associated with each single cell and the column represents the cell type misidentified by scCapsNet. **e**, The labelling conversions among different cell types due to loss of the core genes, represented as a digraph. The nodes represent different cell types and each edge denotes the occurrence of misidentification between two nodes. The direction of each edge starts from the node in the same colour. The thickness of each edge reflects the misidentification ratio for the cells with the same cell label as the starting node.

slightly reduce recognition accuracy, the cutoff values sliding in the two-dimensional PCA plot in scCapsNet simply select distinct gene sets for the feature elimination experiment. A gene set is defined as the core genes of a subcellular type if the removal of the gene set leads to almost zero recognition accuracy for this subcellular type and unchanged accuracy for other types. The scCapsNet model has high accuracy and low time complexity on feature selection.

Meanwhile, the feature elimination experiment in scCapsNet obtains cell label conversions among different cell types. The removal of the core genes for B cells, dendritic cells, NK cells and megakaryocytes basically results in subcellular-type misidentification from B cells to CD4+ T cells, from dendritic cells to FCGR3A+ monocytes, from NK cells to CD8+ T cells, and from megakaryocytes to CD14+ monocytes, respectively. The two monocyte categories of FCGR3A+ monocytes and CD14+ monocytes, as well as the two T cell types CD8+ and CD4+ are converted into each other, with loss of their respective core genes. The above cell label conversions among different cell types are summarized in Fig. 5d,e in the heatmap and digraph. Interestingly, the cell-type conversions obtained in the feature elimination experiment (Fig. 5e) reflect the relationship of hPBMC lineages to some extent. In the hematopoietic cell lineage, all types of blood cells progress from a hematopoietic stem cell, which can undergo differentiation into a megakaryocyte erythrocyte progenitor, a myeloid progenitor or a lymphoid progenitor.

The megakaryocytes differentiate from the megakaryocyte erythrocyte progenitor. The myeloid progenitor gives rise to the myeloid lineage of monocytes and dendritic cells. And the lymphoid progenitor gives rise to the lymphoid lineage of leukocytes: the NK cells and the T and B lymphocytes. The inter-conversions of cell types occur mainly among cells of the same lineage (such as B cells, T cells and NK cells belonging to the lymphoid lineage) that originate from the same progenitors. Moreover, the cell conversion from NK cells to T cells, rather than to B cells, is consistent with the fact that in the lymphoid lineage CD25 stage pro-T cells, the common T/NK precursor cells, display dual T/NK cell development potential³⁹. This suggests a potential application scenario for scCapsNet, that is, cell lineage construction of single-cell subtypes.

The implementation of capsule networks depends critically upon the availability of large, high-quality datasets. The genes expressed in thousands of single cells can be screened in one scRNA-seq experiment. The data quality is improved with the wide utility of the scRNA-seq technologies. So the capsule network model is particularly suitable for scRNA-seq data analysis. We have designed an interpretable architecture of capsule networks for single-cell-type labelling and subcellular-type gene expression program identification. More custom architecture designs of capsule networks will extend applications making use of supervised machine learning. For instance, the connections between inputs of single-cell

multi-omic sources (for example, transcriptomics, proteomics and metabolomics) and primary capsules would provide the best information possible to feature extraction and multi-omic integration analysis¹⁷.

Methods

The scRNA-seq datasets and data preprocessing. We evaluated the utility of our method for single-cell transcriptome analysis using Drop-Seq single-cell data of mRBCs, two datasets of 10X Genomics single-cell data of hPBMCs, four pancreas datasets (Baron Human by inDrop, about 5,000 cells; Muraro by CEL-Seq2, about 1,500 cells; Segerstople by SMART-Seq, about 1,400 cells, and Xin by SMARTer, about 1,400 cells) and so on^{11,30–33}. We adopt the same scRNA-seq data preprocessing module as in Lopez et al.²⁰. The transcriptome profiles included about 20,000 single mRBCs with on average 13,000 genes from 15 subcellular groups and about 12,000 and 6,000 single hPBMCs with on average 3,300 genes from eight subcellular groups and others (as summarized in Supplementary Table 3). All the data were log-transformed before being used. The ‘computeSumFactors’ method from the R package ‘scran’ was used for data normalization⁴⁰.

The architecture and parameters of the scCapsNet model. In the architecture of our scCapsNet model, we choose *l* parallel fully connected neural networks using a ReLU activation function as the feature extractor.

$$u_i = \text{ReLU}(W_p^i x) \quad i \in [1, 2, \dots, l] \tag{1}$$

where *x* represents input vector with length *g*, *g* is equal to the number of genes and *W_pⁱ* represent weight matrices of neural networks with dimension (*n*, *g*), where the rows are denoted as primary capsules and column vectors (*cv₁ⁱ*, *cv₂ⁱ*... *cv_gⁱ*) are indexed with genes. The output *u_i* of each fully connected neural network *i* (*i* ∈ [1, 2, ... , *l*]) is a vector with length *n*, viewed as the ‘primary capsule’ in the model. The feature extractor module converts the features in RNA expression to activities of output (*u₁*, *u₂*... *u_l*).

The features are subsequently delivered through primary capsule to the capsule in the final layer by ‘dynamic routing’. Each capsule in the final layer, the ‘type capsule’, corresponds to each cell type. They are denoted as vectors *v_j*, where *j* ∈ [1, 2, ... , *k*], *k* is the number of the cell type and *m* is the length of the vectors. The capsule network module is implemented in Keras (<https://github.com/bojone/Capsule>).

Before the ‘dynamic routing’ process, the primary capsules are multiplied by weight matrices *W_{ij}* to produce ‘prediction vectors’ *û_{iji}*.

$$\hat{u}_{iji} = W_{ij} u_i \tag{2}$$

Then the iterative dynamic routing begins. First, the ‘coupling coefficients’ *c_{ij}* are calculated by:

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \tag{3}$$

where *b_{ij}* is an intermediate parameter with initial value of zero, representing the inner product of the prediction vector and type capsule vector.

To compute the *b_{ij}* value for the next-round iteration, the weighted sum *s_j* over all *k* prediction vectors *û_{iji}* is calculated as follows:

$$s_j = \sum_i c_{ij} \hat{u}_{iji} \tag{4}$$

Second, *b_{ij}* is computed by the dot product of *û_{iji}* and *s_j* as the last step of the one-round dynamic routing process:

$$b_{ij} = \hat{u}_{iji} \cdot s_j \tag{5}$$

After several rounds of dynamic routing, the type capsule *v_j* is calculated by a non-linear ‘squashing’ function:

$$v_j = \frac{\|s_j\|^2}{0.5 + \|s_j\|^2 + \|s_j\|} \tag{6}$$

The following pseudocode illustrates the implementation of scCapsNet:

- 1: For all primary capsule *i*: *u_i* = ReLU (*W_pⁱx*)
- 2: For all primary capsule *i* and type capsule *j*: *û_{iji}* = *W_{ij}u_i*
- 3: Procedure ROUTING (*û_{iji}*, *r*)
- 4: For all primary capsule *I* and type capsule *j*: *b_{ij}* ← 0.
- 5: For *r* iterations do
- 6: For all primary capsule *i*: *c_i* ← softmax(*b_i*)
- 7: For all type capsule *j*: *s_j* ← ∑*i* *c_{ij}û_{iji}*
- 8: For all primary capsule *i* and type capsule *j*: *b_{ij}* ← *û_{iji}* · *s_j*
- return *v_j* ← squash(*s_j*)

The implementation of scCapsNet can be found in <https://github.com/wanglf19/scCaps>. We used shuffle-split cross-validation, a data resampling method, to evaluate the accuracy of scCapsNet’s subcellular type recognition. During the process of testing the scCapsNet classifier for recognition of cell-type *C_i*, the samples labelled with cell type *C_i* were the pre-known positives, and the other testing samples were the pre-known negatives. At the *j*th cross-validation, the test set with *N_j* samples was derived from a shuffle split of the original scRNA-seq data. The number of the positive samples correctly classified to be cell type *C_i* was recorded as *N_{ji,C_i}*. The accuracy was calculated by the equation ∑*C_i* ∈ all the types *N_{ji,C_i}* / *N_j*.

Coupling coefficients contributing to type recognition. In the scCapsNet model, the type capsule *v_j* derives from a weighted sum of prediction vectors *û_{iji}*. The weights are the coupling coefficients *c_{ij}* and the magnitude of those coefficients can roughly be regarded as the contribution of the primary capsules *u_i* to the type capsules *v_j*. Each single cell generates its own coupling coefficients. The average coupling coefficients for the single cells with same cell type are calculated by the formula

$$c_{ij}^{\text{type average}} = \frac{\sum_{\text{type}} c_{ij}^{\text{type}}}{\sum_{\text{type}} 1} \tag{7}$$

Function analysis on the core gene sets. A literature search annotated several cell-type-associated genes in the core gene sets related to distinct subcellular types^{241–70}. Gene Ontology (GO) enrichment analysis and reactome pathway analysis were utilized for function annotations on the gene set. The ‘GOenrichment’ function from R package ‘GOSim’ was used for GO enrichment analysis⁷¹, which uses the method from R package ‘topGO’⁷². The reactome web server was used for pathway analysis^{73,74}.

An embedding representation of a gene. The primary capsules 1, 2, 4, 6, 8, 10, 14 and 16 contribute to the recognition of CD14+ monocytes, CD4+ T cells, CD8+ T cells, megakaryocytes, dendritic cells, B cells, FCGR3A+ monocytes and NK cells. Column vectors represent the embedding of the same gene from a different weight matrix are concatenated together to form a long vector, for example [*cv₁¹*, *cv₂²*, *cv₄⁴*, *cv₆⁶*, *cv₈⁸*, *cv₁₀¹⁰*, *cv₁₄¹⁴*, *cv₁₆¹⁶*], as a unique embedding representation of a gene.

Algorithm implementation for comparisons on cell-type-recognition accuracy. A neural network with sigmoid activation function was implemented in Keras. The algorithms SVM, random forest, LDA, nearest-neighbour, SVM_{rejection} and LDA_{rejection} were implemented with the Python package ‘scikit-learn’ (https://github.com/tabdelaal/scRNAseq_Benchmark/).

Data availability

The pre-processed single-cell transcriptome data of mRBCs²⁸ and hPBMCs²⁷ can be downloaded and extracted from Github (RetinaDataset and PurifiedPBMCDataSet, <https://github.com/YosefLab/scV1>)²⁰. Other pre-processed single-cell transcriptome data for the cross-dataset experiment, unseen population experiment and negative control experiment can be downloaded from <https://zenodo.org/record/3357167#.X0kHPPZuJZU>¹¹. All the data used in this Article are summarized in Supplementary Table 3.

Code availability

The implementation of scCapsNet can be found in <https://github.com/wanglf19/scCaps> or <https://zenodo.org/record/4007185#.X0oHPPZuJZU>.

Received: 3 May 2020; Accepted: 28 September 2020;

Published online: 02 November 2020

References

1. Zheng, C. et al. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* **169**, 1342–1356 (2017).
2. Guo, X. et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* **24**, 978–985 (2018).
3. Setty, M. et al. Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat. Biotechnol.* **34**, 637–645 (2016).
4. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
5. Halpern, K. B. et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017).
6. Halpern, K. B. et al. Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. *Nat. Biotechnol.* **36**, 962–970 (2018).
7. Han, X. et al. Mapping the mouse cell atlas by Microwell-Seq. *Cell* **173**, 1307 (2018).

8. de Kanter, J. K., Lijnzaad, P., Candelli, T., Margaritis, T. & Holstege, F. C. P. CHETAH: a selective, hierarchical cell type identification method for single-cell RNA sequencing. *Nucleic Acids Res.* **47**, e95 (2019).
9. Kiselev, V. Y., Yiu, A. & Hemberg, M. scmap: projection of single-cell RNA-seq data across data sets. *Nat. Methods* **15**, 359–362 (2018).
10. Pliner, H. A., Shendure, J. & Trapnell, C. Supervised classification enables rapid annotation of cell atlases. *Nat. Methods* **16**, 983–986 (2019).
11. Abdelaal, T. et al. A comparison of automatic cell identification methods for single-cell RNA sequencing data. *Genome Biol.* **20**, 194 (2019).
12. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
13. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 e1821 (2019).
14. Florian Wagner, P. Y. Moana: a robust and scalable cell type classification framework for single-cell RNA-Seq data. *bioRxiv* <https://doi.org/10.1101/456129> (2018).
15. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018).
16. Almas Jabeen, N. A. & Raza, K. Machine learning-based state-of-the-art methods for the classification of RNA-seq data. *bioRxiv* <https://doi.org/10.1101/120592> (2017).
17. Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C. & Collins, J. J. Next-generation machine learning for biological networks. *Cell* **173**, 1581–1592 (2018).
18. Zhou, J. et al. Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. *Nat. Genet.* **50**, 1171–1179 (2018).
19. Alipanahi, B., Delong, A., Weirauch, M. T. & Frey, B. J. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat. Biotechnol.* **33**, 831–838 (2015).
20. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053 (2018). +
21. Amodio, M. et al. Exploring single-cell data with deep multitasking neural networks. *Nature Methods* **16**, 1139–1145 <https://doi.org/10.1038/s41592-019-0576-7> (2019).
22. Chen, H. H. et al. GSAE: an autoencoder with embedded gene-set nodes for genomics functional characterization. *BMC Syst. Biol.* **12**, 142 (2018).
23. Ding, J., Condon, A. & Shah, S. P. Interpretable dimensionality reduction of single cell transcriptome data with deep generative models. *Nat. Commun.* **9**, 2002 (2018).
24. Lin, C., Jain, S., Kim, H. & Bar-Joseph, Z. Using neural networks for reducing the dimensions of single-cell RNA-seq data. *Nucleic Acids Res.* **45**, e156 (2017).
25. Sabour, S., Frosst, N. & Hinton, G. E. Dynamic routing between capsules. In *Advances in Neural Information Processing Systems 30* (eds Guyon, I. et al.) 3856–3866 (Curran Associates, 2017).
26. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
27. Zheng, G. X. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
28. Shekhar, K. et al. Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323 (2016).
29. Ding, J. et al. Systematic comparative analysis of single cell RNA-sequencing methods. *bioRxiv* <https://doi.org/10.1101/632216> (2019).
30. Baron, M. et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Syst.* **3**, 346–360 (2016).
31. Muraro, M. J. et al. A single-cell transcriptome atlas of the human pancreas. *Cell Syst.* **3**, 385–394 (2016).
32. Segerstolpe, A. et al. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab.* **24**, 593–607 (2016).
33. Xin, Y. et al. RNA sequencing of single human islet cells reveals type 2 diabetes genes. *Cell Metab.* **24**, 608–615 (2016).
34. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
35. Chahwan, R., Edelmann, W., Scharff, M. D. & Roa, S. AIDing antibody diversity by error-prone mismatch repair. *Semin. Immunol.* **24**, 293–300 (2012).
36. Stone, S. F. et al. Changes in differential gene expression during a fatal stroke. *J. Clin. Neurosci.* **23**, 130–134 (2016).
37. Fishilevich, S. et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database J. Biol. Databases Curation* **2017**, bax028 (2017).
38. Stark, C. et al. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535–D539 (2006).
39. Lee, C. K. et al. Cloning thymic precursor cells: demonstration that individual pro-T1 cells have dual T-NK potential and individual pro-T2 cells have dual alphabeta-gammadelta T cell potential. *Cell. Immunol.* **191**, 139–144 (1999).
40. Lun, A. T., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res.* **5**, 2122 (2016).
41. Frankenberger, M. et al. Transcript profiling of CD16-positive monocytes reveals a unique molecular fingerprint. *Eur. J. Immunol.* **42**, 957–974 (2012).
42. Bernal-Quiros, M., Wu, Y. Y., Alarcon-Riquelme, M. E. & Castillejo-Lopez, C. BANK1 and BLK act through phospholipase C gamma 2 in B-cell signaling. *PLoS One* **8**, e59842 (2013).
43. Lapter, S. et al. A role for the B-cell CD74/macrophage migration inhibitory factor pathway in the immunomodulation of systemic lupus erythematosus by a therapeutic tolerogenic peptide. *Immunology* **132**, 87–95 (2011).
44. Huang, X. et al. Downregulation of the B-cell receptor signaling component CD79b in plasma cell myeloma: a possible post transcriptional regulation. *Pathol. Int.* **61**, 122–129 (2011).
45. Stang, S. L. et al. A proapoptotic signaling pathway involving RasGRP, Erk, and Bim in B cells. *Exp. Hematol.* **37**, 122–134 (2009).
46. Shah, R. D. et al. Expression of calgranulin genes S100A8, S100A9 and S100A12 is modulated by n-3 PUFA during inflammation in adipose tissue and mononuclear cells. *PLoS One* **12**, e0169614 (2017).
47. Gren, S. T. et al. A single-cell gene-expression profile reveals inter-cellular heterogeneity within human monocyte subsets. *PLoS One* **10**, e0144351 (2015).
48. Villasenor-Cardoso, M. I., Frausto-Del-Rio, D. A. & Ortega, E. Aminopeptidase N (CD13) is involved in phagocytic processes in human dendritic cells and macrophages. *BioMed Res. Int.* **2013**, 562984 (2013).
49. Munthe-Fog, L. et al. Variation in FCN1 affects biosynthesis of ficolin-1 and is associated with outcome of systemic inflammation. *Genes Immun.* **13**, 515–522 (2012).
50. Li, Y. et al. A possible role of HMGB1 in DNA demethylation in CD4+ T cells from patients with systemic lupus erythematosus. *Clin. Dev. Immunol.* **2013**, 206298 (2013).
51. Chan, D. V. et al. Differential CTLA-4 expression in human CD4+ versus CD8+ T cells is associated with increased NFAT1 and inhibition of CD4+ proliferation. *Genes Immun.* **15**, 25–32 (2014).
52. Alonso, M. A. & Weissman, S. M. cDNA cloning and sequence of MAL, a hydrophobic protein associated with human T-cell differentiation. *Proc. Natl. Acad. Sci.* **84**, 1997–2001 (1987).
53. Cismasiu, V. B. et al. BCL11B participates in the activation of IL2 gene expression in CD4+ T lymphocytes. *Blood* **108**, 2695–2702 (2006).
54. Bade, B. et al. Differential expression of the granzymes A, K and M and perforin in human peripheral blood lymphocytes. *Int. Immunol.* **17**, 1419–1428 (2005).
55. Huang, R. Y. et al. LAG3 and PD1 co-inhibitory molecules collaborate to limit CD8+ T cell signaling and dampen antitumor immunity in a murine ovarian cancer model. *Oncotarget* **6**, 27359–27377 (2015).
56. Stoekle, C. et al. Cathepsin W expressed exclusively in CD8+ T cells and NK cells, is secreted during target cell killing but is not essential for cytotoxicity in human CTLs. *Exp. Hematol.* **37**, 266–275 (2009).
57. Nizzoli, G. et al. Human CD1c+ dendritic cells secrete high levels of IL-12 and potentially prime cytotoxic T-cell responses. *Blood* **122**, 932–942 (2013).
58. Heger, L. et al. CLEC10A is a specific marker for human CD1c(+) dendritic cells and enhances their toll-like receptor 7/8-induced cytokine secretion. *Front. Immunol.* **9**, 744 (2018).
59. Karsunky, H., Merad, M., Cozzio, A., Weissman, I. L. & Manz, M. G. Flt3 ligand regulates dendritic cell development from Flt3+ lymphoid and myeloid-committed progenitors to Flt3+ dendritic cells in vivo. *J. Exp. Med.* **198**, 305–313 (2003).
60. Ohta, M. et al. Immunomodulation of monocyte-derived dendritic cells through ligation of tumor-produced mucins to Siglec-9. *Biochem. Biophys. Res. Commun.* **402**, 663–669 (2010).
61. Chen, Y. J. et al. Eps8 protein facilitates phagocytosis by increasing TLR4-MyD88 protein interaction in lipopolysaccharide-stimulated macrophages. *J. Biol. Chem.* **287**, 18806–18819 (2012).
62. Kitzenberg, D., Colgan, S. P. & Glover, L. E. Creatine kinase in ischemic and inflammatory disorders. *Clin. Transl. Med.* **5**, 31 (2016).
63. Martinez, F. O. The transcriptome of human monocyte subsets begins to emerge. *J. Biol.* **8**, 99 (2009).
64. Zhang, C., Gadue, P., Scott, E., Atchison, M. & Poncz, M. Activation of the megakaryocyte-specific gene platelet basic protein (PBP) by the Ets family factor PU.1. *J. Biol. Chem.* **272**, 26236–26246 (1997).
65. Seo, H. et al. A beta1-tubulin-based megakaryocyte maturation reporter system identifies novel drugs that promote platelet production. *Blood Adv.* **2**, 2262–2272 (2018).
66. Clay, D. et al. CD9 and megakaryocyte differentiation. *Blood* **97**, 1982–1989 (2001).
67. Hickey, M. J., Deaven, L. L. & Roth, G. J. Human platelet glycoprotein IX. Characterization of cDNA and localization of the gene to chromosome 3. *FEBS Lett.* **274**, 189–192 (1990).

68. Kim, T. D. et al. Human microRNA-27a* targets Prfl and GzmB expression to regulate NK-cell cytotoxicity. *Blood* **118**, 5476–5486 (2011).
69. Kuttruff, S. et al. Nkp80 defines and stimulates a reactive subset of CD8 T cells. *Blood* **113**, 358–369 (2009).
70. Sim, M. J. et al. KIR2DL3 and KIR2DL1 show similar impact on licensing of human NK cells. *Eur. J. Immunol.* **46**, 185–191 (2016).
71. Frohlich, H., Speer, N., Poustka, A. & Beissbarth, T. GOSim—an R-package for computation of information theoretic GO similarities between terms and gene products. *BMC Bioinf.* **8**, 166 (2007).
72. Adrian Alexa, J. R. *topGO: Enrichment Analysis for Gene Ontology* R package version 2.34.0 (2018).
73. Fabregat, A. et al. The reactome pathway Knowledgebase. *Nucleic Acids Res.* **44**, D481–D487 (2016).
74. Fabregat, A. et al. Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinf.* **18**, 142 (2017).

Acknowledgements

This work was supported by grants from the National Key R&D Program of China (grant 2018YFC0910402 to J.C.; grant 2018YFC1003102 to C.Z. and grant 2017YFC0908402 to C.Z.); the Strategic Priority Research Program of the Chinese Academy of Sciences (grant E0XD842201 to J.C.); the National Natural Science Foundation of China (grant 32070795

to J.C. and grant 61673070 to J.Z.); and the Open Project of Key Laboratory of Genomic and Precision Medicine, Chinese Academy of Sciences.

Author Contributions

J.C., J.Z. and L.W. envisioned the project. L.W. implemented the model and performed the analysis. L.W. and J.C. wrote the paper. R.N., Z.Y., R.X., C.Z., Z.Z. and J.Z. provided assistance in writing and analysis.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s42256-020-00244-4>.

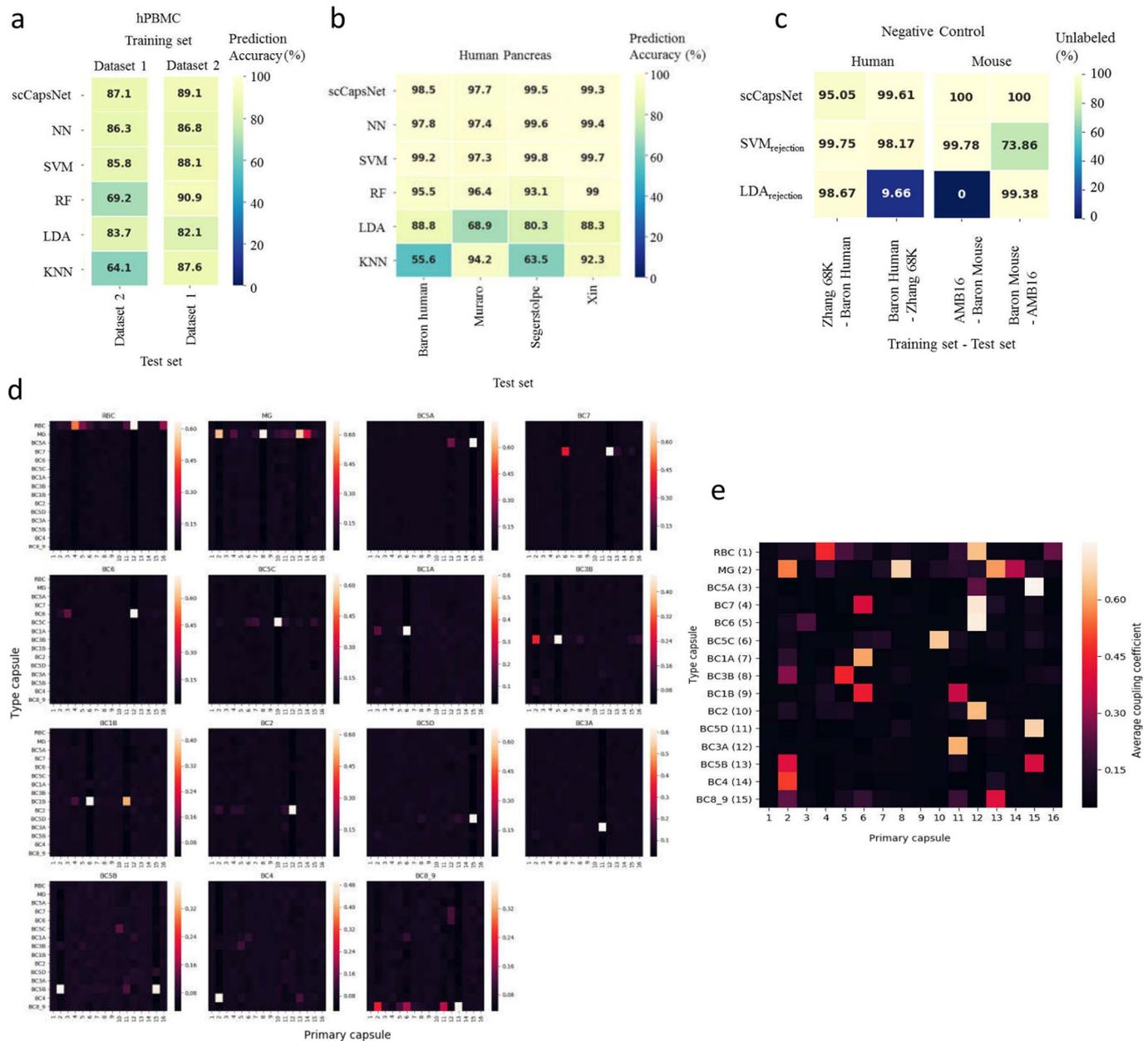
Supplementary information is available for this paper at <https://doi.org/10.1038/s42256-020-00244-4>.

Correspondence and requests for materials should be addressed to J.Z. or J.C.

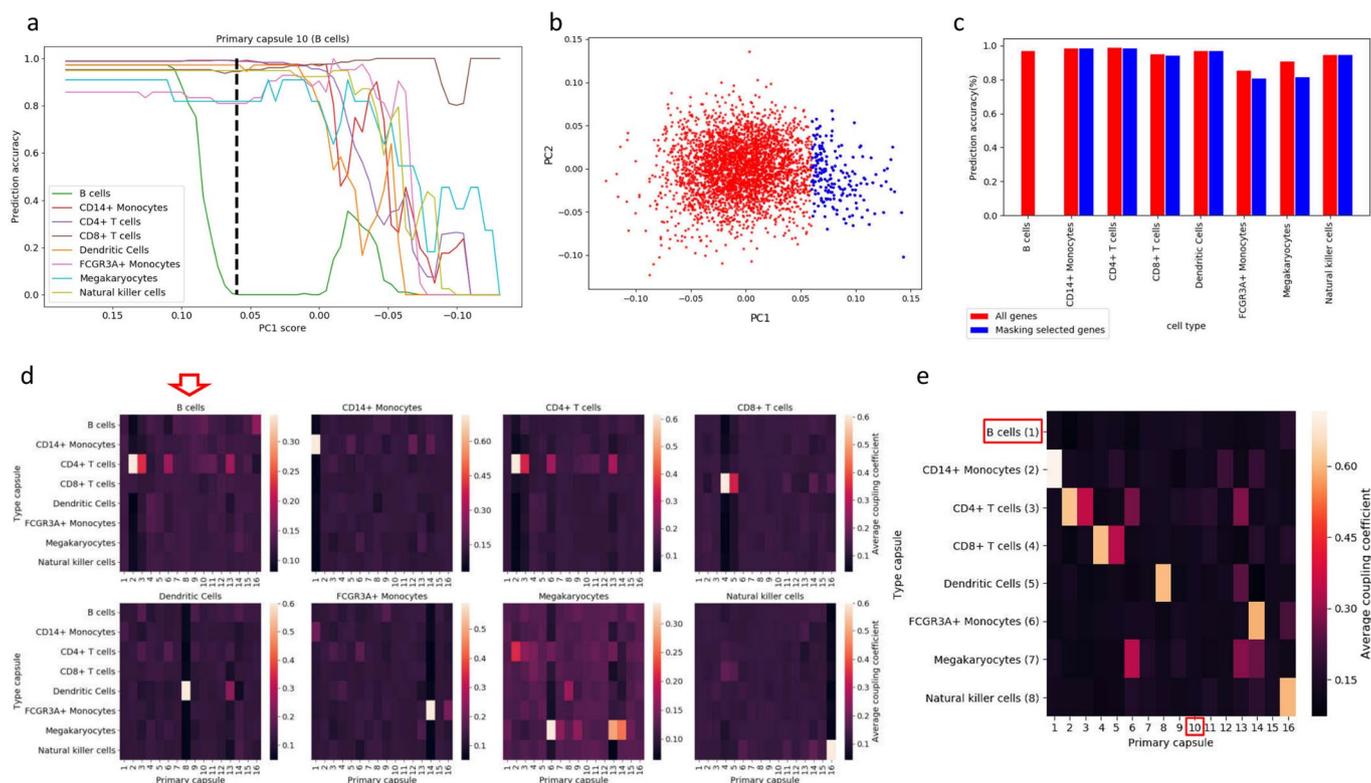
Reprints and permissions information is available at www.nature.com/reprints.

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

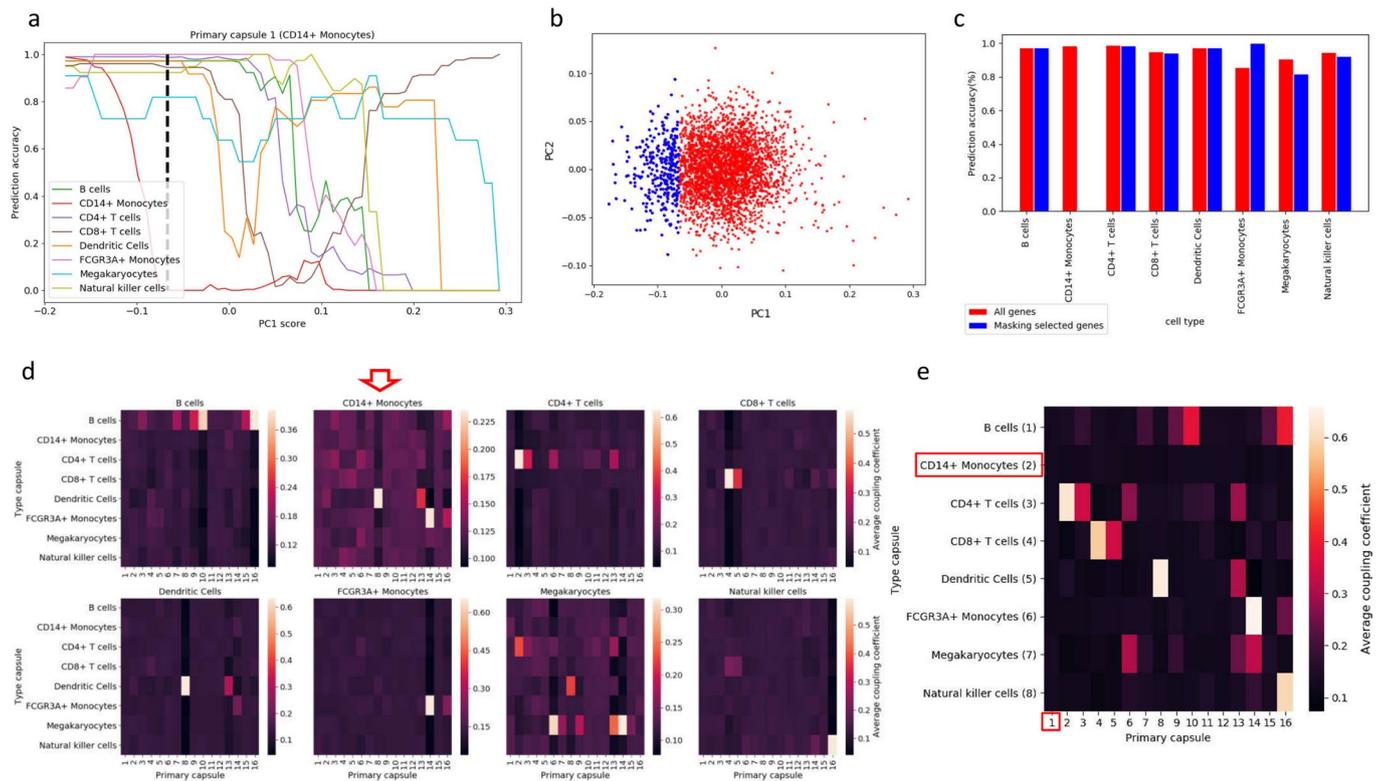
© The Author(s), under exclusive licence to Springer Nature Limited 2020



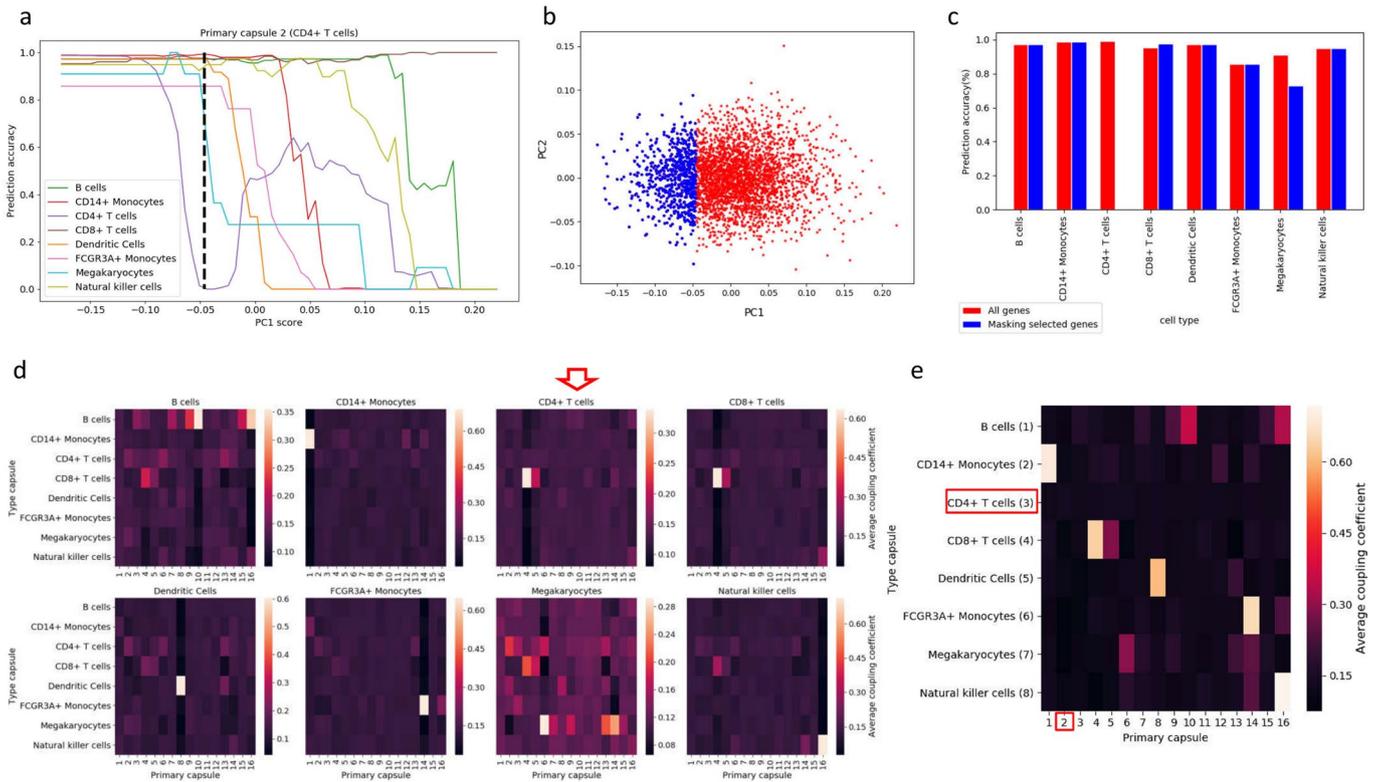
Extended Data Fig. 1 | The performance and its internal parameters of scCapsNet relative with cell type recognition. **a**, The classification performance across two hPBMC datasets from the 10x Genomics platform. We trained scCapsNet and other machine learning methods using one dataset and then evaluated their performances on another dataset. The heatmap shows the prediction accuracy for each classifier. **b**, The classification performance across four human pancreatic datasets from different single-cell RNA-seq protocols. The four datasets are quoted from Abdelal's paper. Each column corresponds to one sub-task in which one of the four datasets was used as a test set and the rest three datasets were used as training. The heatmap shows the prediction accuracy for each classifier. **c**, The rejection option evaluation in the negative control experiment on scCapsNet, SVM_{rejection} and LDA_{rejection} models. There are two groups of datasets, the group of human dataset from PBMC and pancreas tissues, and the group of mouse dataset from visual cortex and pancreas tissues. In each column, the classifiers are used to predict single cell identity of one dataset after training on the paired dataset from another different tissue. The recognition rates of unlabeled single cells as the negative control are shown in the heatmap. The LDA_{rejection} reported error in AMB16-Baron Mouse experiment, so we set the percentage of unlabeled cells to 0. **d**, The heatmaps of the matrices of averaged coupling coefficients for mRBC dataset with cell type listed above. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 1b where its recognition type is in accordance with the type of input single cells.



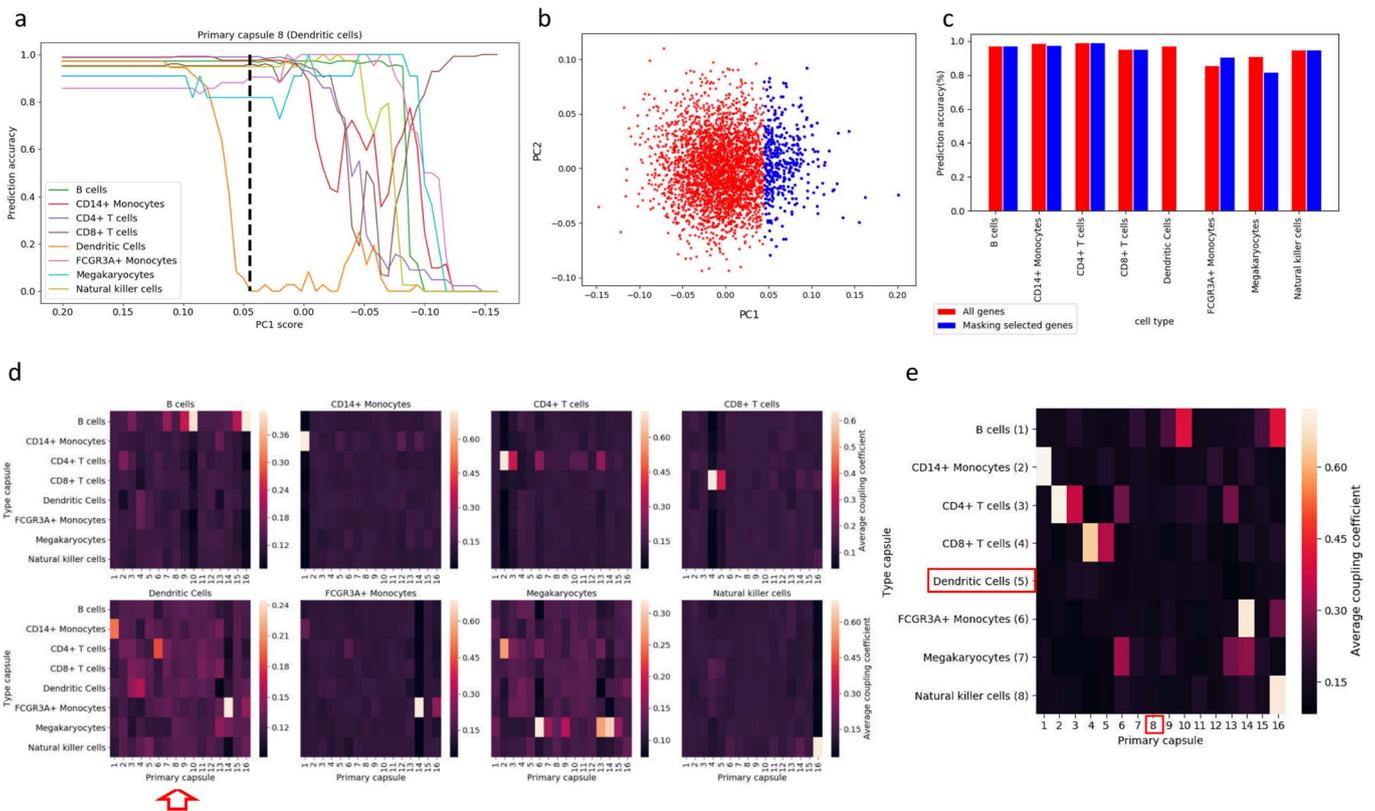
Extended Data Fig. 2 | The identification of the core gene set responsible for recognition of B cells in hPBMC. **a**, The coloured changing curves of cell-type recognition accuracies while the ranking genes defined by a sliding cutoff value on the principal component score were excluded in the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The dotted line defines a group of core genes responsible for B-cell identification, where the recognition accuracy of B cells degrades close to 0 but slightly decreases for any other cell type. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **c**, The comparison of prediction accuracy of each cell type before and after the masking of the B-cell core genes. **d**, The heatmaps of the revised matrices of averaged coupling coefficients for hPBMC dataset with the loss of the group of B-cell core genes in the inputs of the scCapsNet model. The heatmaps in order represent the revised averaged coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 2d where its recognition type is in accordance with the type of input single cells.



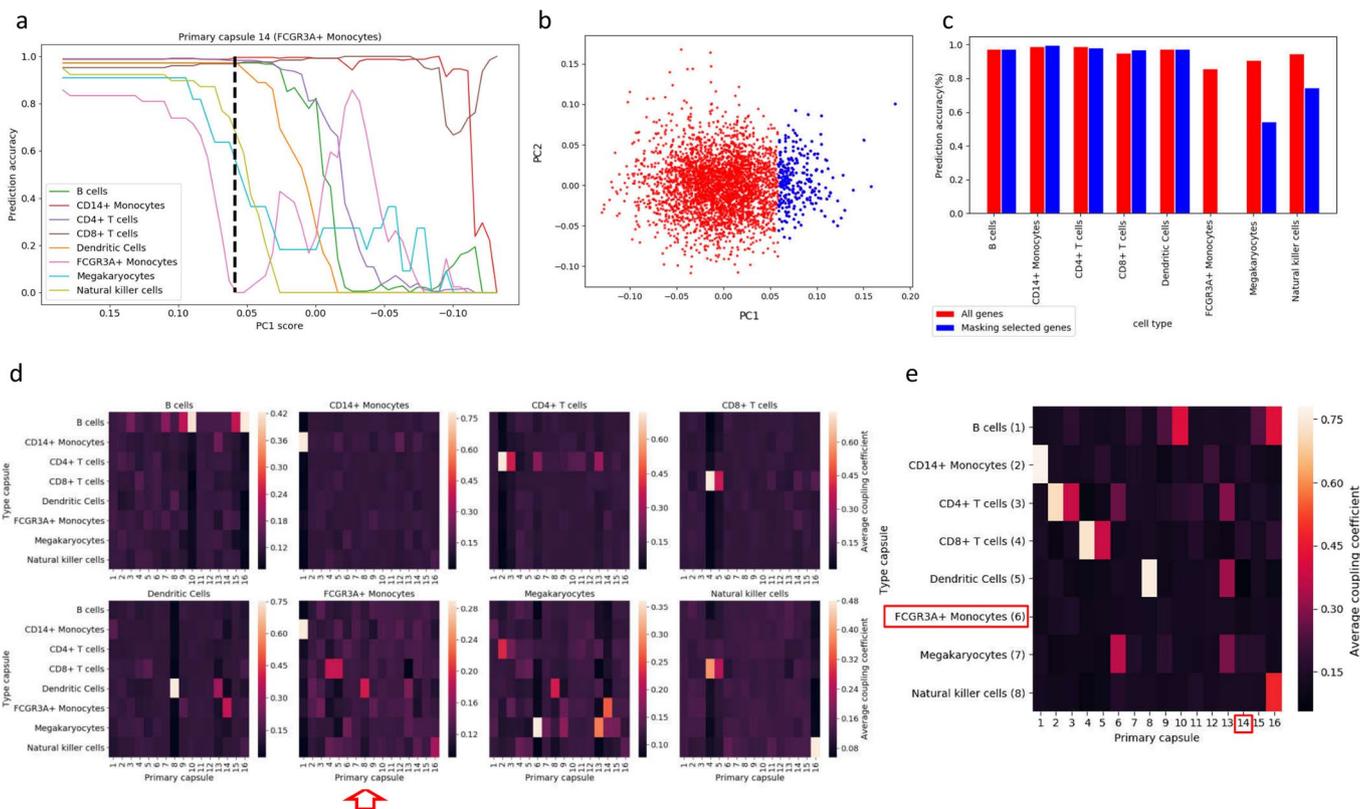
Extended Data Fig. 3 | The identification of the core gene set responsible for recognition of CD14+ monocytes in hPBMC. **a**, The coloured changing curves of cell-type recognition accuracies while the ranking genes defined by a sliding cutoff value on the principal component score were excluded in the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The dotted line defines a group of core genes responsible for CD14+ monocyte identification, where the recognition accuracy of CD14+ monocytes degrades close to 0 but slightly decreases for any other cell type. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **c**, The comparison of prediction accuracy of each cell type before and after the masking of the CD14+ monocyte core genes. **d**, The heatmaps of the revised matrices of averaged coupling coefficients for hPBMC dataset with the loss of the group of CD14+ monocytes core genes in the inputs of the scCapsNet model. The heatmaps in order represent the revised averaged coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 3d where its recognition type is in accordance with the type of input single cells.



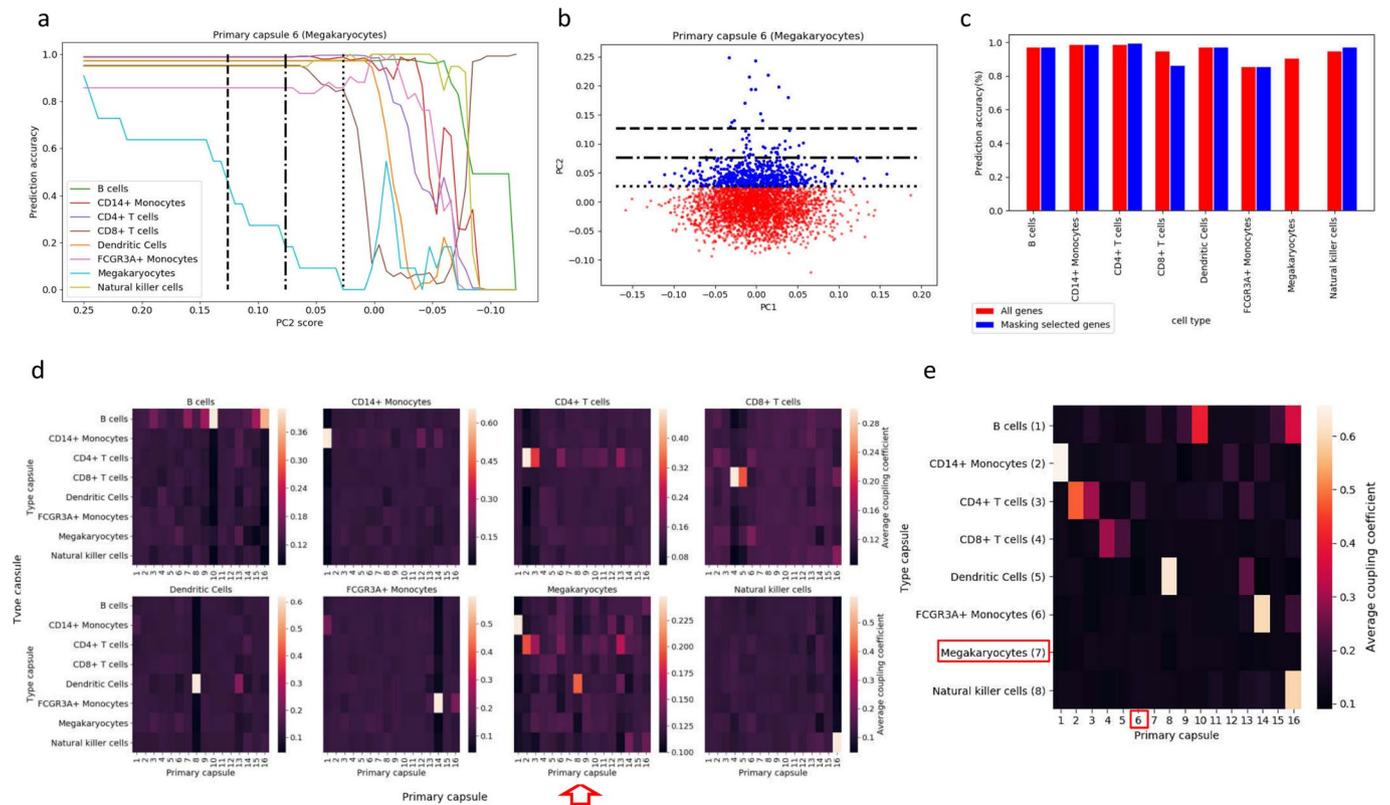
Extended Data Fig. 4 | The identification of the core gene set responsible for recognition of CD4+ T cells in hPBMC. **a**, The coloured changing curves of cell-type recognition accuracies while the ranking genes defined by a sliding cutoff value on the principal component score were excluded in the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The dotted line defines a group of core genes responsible for CD4+ T cell identification, where the recognition accuracy of CD4+ T cells degrades close to 0 but slightly decreases for any other cell type. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **c**, The comparison of prediction accuracy of each cell type before and after the masking of the CD4+ T cell core genes. **d**, The heatmaps of the revised matrices of averaged coupling coefficients for hPBMC dataset with the loss of the group of CD4+ T cell core genes in the inputs of the scCapsNet model. The heatmaps in order represent the revised averaged coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 4d where its recognition type is in accordance with the type of input single cells.



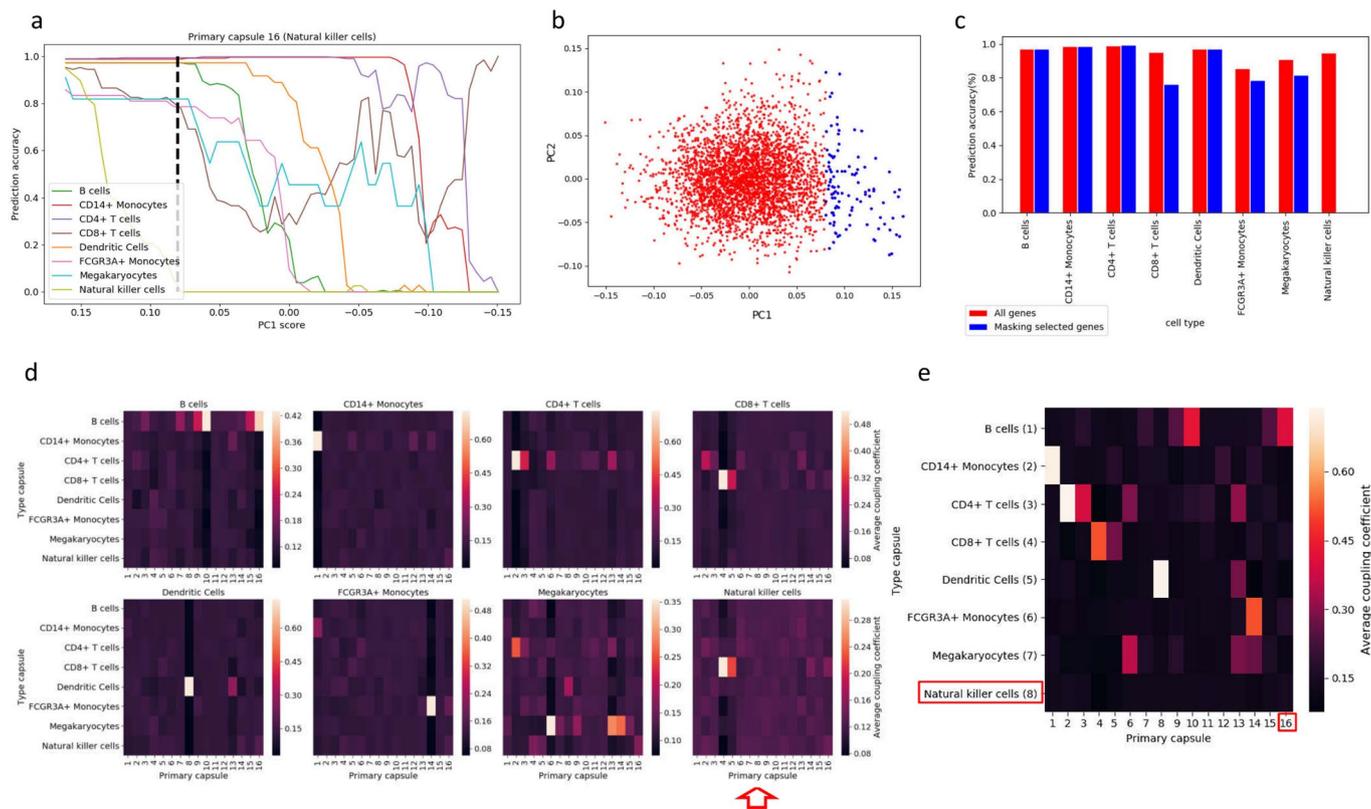
Extended Data Fig. 5 | The identification of the core gene set responsible for recognition of dendritic cells in hPBMC. **a**, The coloured changing curves of cell-type recognition accuracies while the ranking genes defined by a sliding cutoff value on the principal component score were excluded in the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The dotted line defines a group of core genes responsible for dendritic-cell identification, where the recognition accuracy of dendritic cells degrades close to 0 but slightly decreases for any other cell type. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **c**, The comparison of prediction accuracy of each cell type before and after the masking of the dendritic-cell core genes. **d**, The heatmaps of the revised matrices of averaged coupling coefficients for hPBMC dataset with the loss of the group of dendritic-cell core genes in the inputs of the scCapsNet model. The heatmaps in order represent the revised averaged coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 5d where its recognition type is in accordance with the type of input single cells.



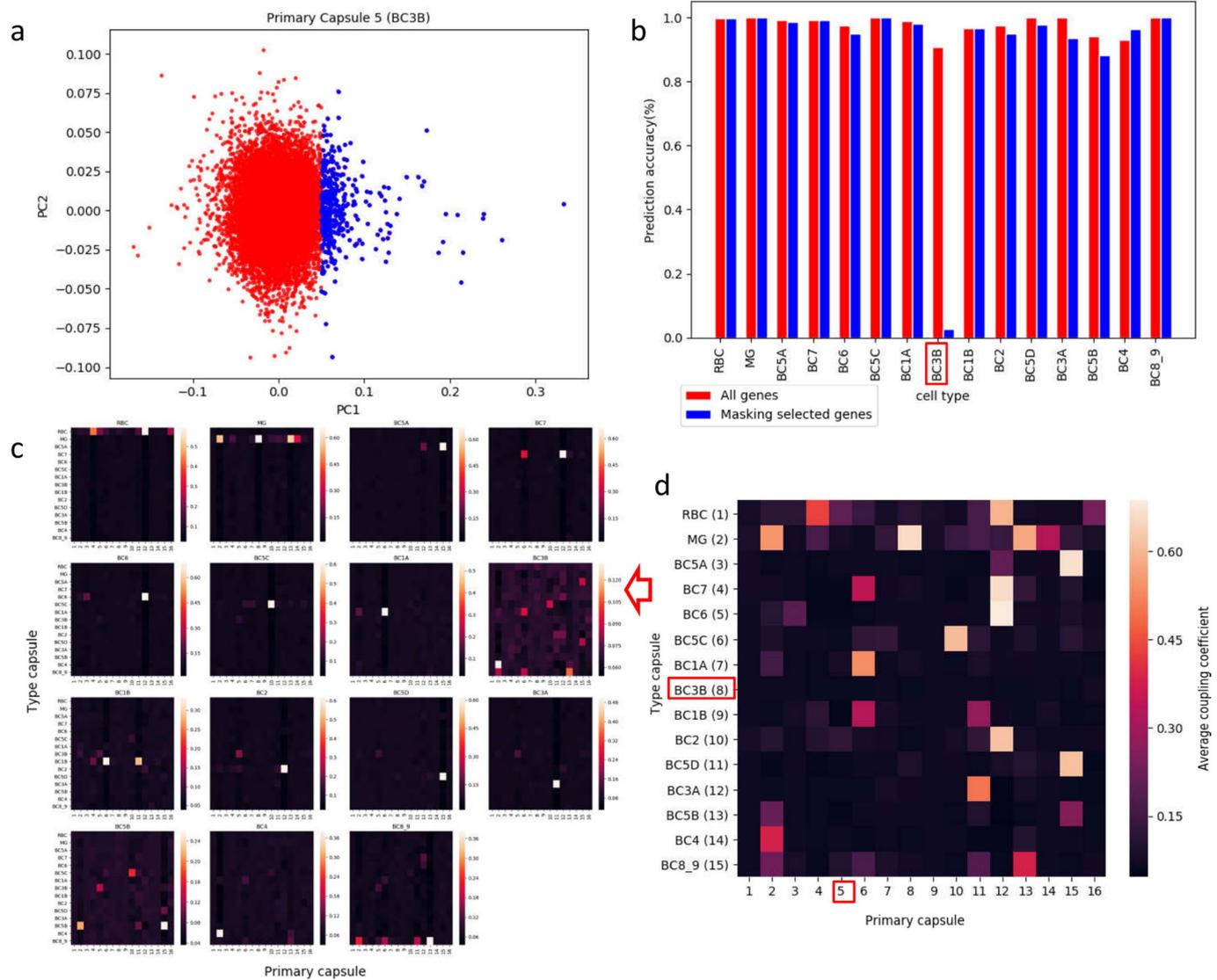
Extended Data Fig. 6 | The identification of the core gene set responsible for recognition of FCGR3A+ monocytes in hPBMC. **a**, The coloured changing curves of cell-type recognition accuracies while the ranking genes defined by a sliding cutoff value on the principal component score were excluded in the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The dotted line defines a group of core genes responsible for FCGR3A+ monocyte identification, where the recognition accuracy of FCGR3A+ monocytes degrades close to 0 but slightly decreases for any other cell type. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **c**, The comparison of prediction accuracy of each cell type before and after the masking of the FCGR3A+ monocyte core genes. **d**, The heatmaps of the revised matrices of averaged coupling coefficients for the hPBMC dataset with the loss of the group of FCGR3A+ monocyte core genes in the inputs of the scCapsNet model. The heatmaps in order represent the revised averaged coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 6d where its recognition type is in accordance with the type of input single cells.



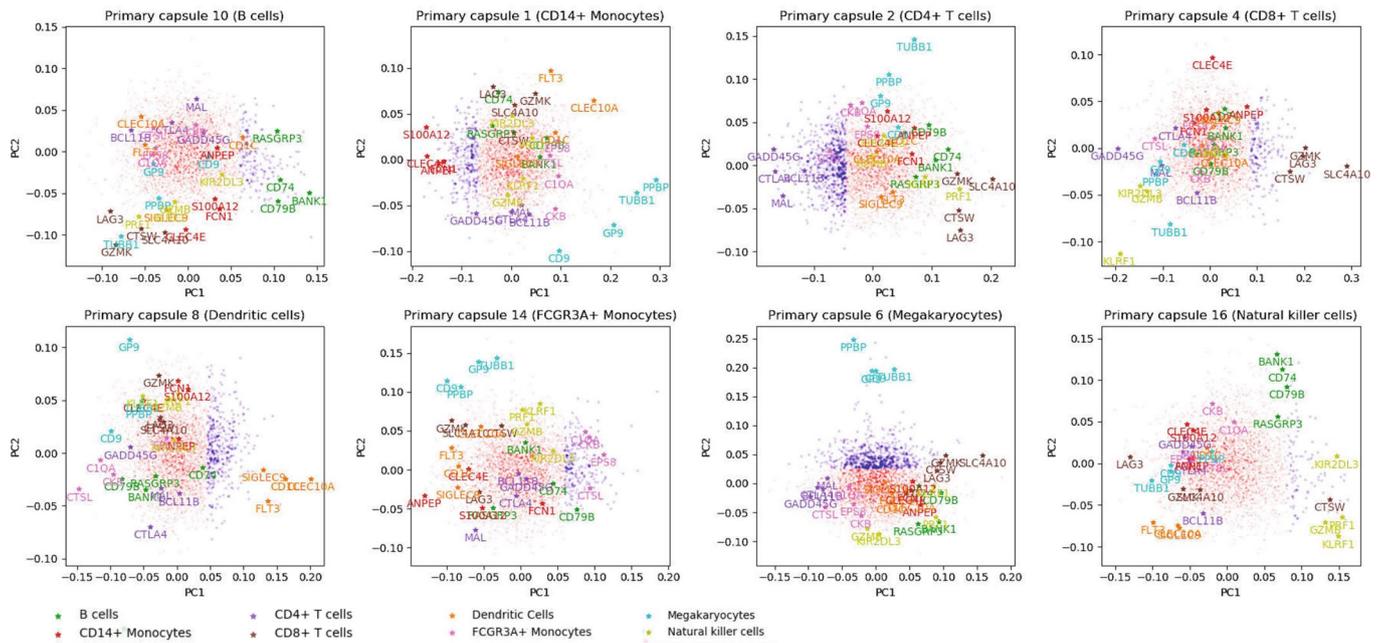
Extended Data Fig. 7 | The identification of the core gene set responsible for recognition of megakaryocytes in hPBMC. a, The coloured changing curves of cell-type recognition accuracies while the ranking genes defined by a sliding cutoff value on the principal component score were excluded in the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The dotted line defines a group of core genes responsible for megakaryocyte identification, where the recognition accuracy of megakaryocytes degrades close to 0 but slightly decreases for any other cell type. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **c**, The comparison of prediction accuracy of each cell type before and after the masking of the megakaryocyte core genes. **d**, The heatmaps of the revised matrices of averaged coupling coefficients for hPBMC dataset with the loss of the group of megakaryocyte core genes in the inputs of the scCapsNet model. The heatmaps in order represent the revised averaged coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 7d where its recognition type is in accordance with the type of input single cells.



Extended Data Fig. 8 | The identification of the core gene set responsible for recognition of NK cells in hPBMC. **a**, The coloured changing curves of cell-type recognition accuracies while the ranking genes defined by a sliding cutoff value on the principal component score were excluded in the inputs of the scCapsNet model. The accuracy curve for each cell type is represented in a distinct colour. The dotted line defines a group of core genes responsible for NK cell identification, where the recognition accuracy of NK cells degrades close to 0 but slightly decreases for any other cell type. **b**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **c**, The comparison of prediction accuracy of each cell type before and after the masking of the NK cell core genes. **d**, The heatmaps of the revised matrices of averaged coupling coefficients for hPBMC dataset with the loss of the group of NK cell core genes in the inputs of the scCapsNet model. The heatmaps in order represent the revised averaged coupling coefficient matrix for the single B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A+ monocytes, megakaryocytes and NK cells. For each heatmap, the row represents type capsules and the column represents primary capsules. **e**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 8d where its recognition type is in accordance with the type of input single cells.



Extended Data Fig. 9 | Identification of the core gene set responsible for recognition of one cell type in hRBC. **a**, The plot depicts the two-dimensional PCA on the weight matrix for the primary capsule five in model trained on mRBC dataset. Each dot represents a gene with a rank according to the score of principal components. A group of core genes marked as blue colour are defined. **b**, The comparison of prediction accuracy of this cell type before and after the masking of the core genes. **c**, The heatmaps of the revised matrices of averaged coupling coefficients for hRBC dataset with the loss of the group of core genes in the inputs of the scCapsNet model. For each heatmap, the row represents type capsules and the column represents primary capsules. **d**, The revision of the overall heatmap of the combining matrix of average coupling coefficient. The combining matrix contains the effective type capsule row in Extended Data Fig. 9c where its recognition type is in accordance with the type of input single cells.



Extended Data Fig. 10 | The well studied cell-type associated genes in the core gene sets relevant to distinct subcellular types. The scatter plots in order depict the two-dimensional PCA on column vectors of weight matrices fully connecting inputs and primary capsules 10, 1, 2, 4, 8, 14, 6, and 16. They defined the groups of core genes (in blue dots), contributing to the identification of B cells, CD14+ monocytes, CD4+ T cells, CD8+ T cells, dendritic cells, FCGR3A + monocytes, megakaryocytes, and NK cells respectively. Several well-studied cell type associated genes are represented as coloured stars with gene name underneath. The colours of the stars represent the cell type of the corresponding gene associated.