

# CellMarker: a manually curated resource of cell markers in human and mouse

Xinxin Zhang<sup>1,†</sup>, Yujia Lan<sup>1,†</sup>, Jinyuan Xu<sup>1,†</sup>, Fei Quan<sup>1,†</sup>, Erjie Zhao<sup>1,†</sup>, Chunyu Deng<sup>1</sup>, Tao Luo<sup>1</sup>, Liwen Xu<sup>1</sup>, Gaoming Liao<sup>1</sup>, Min Yan<sup>1</sup>, Yanyan Ping<sup>1</sup>, Feng Li<sup>1</sup>, Aiai Shi<sup>1</sup>, Jing Bai<sup>1</sup>, Tingting Zhao<sup>2,\*</sup>, Xia Li<sup>1,\*</sup> and Yun Xiao<sup>1,\*</sup>

<sup>1</sup>College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, Heilongjiang 150081, China and <sup>2</sup>Department of Neurology, The First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang 150001, China

Received August 14, 2018; Revised September 20, 2018; Editorial Decision September 23, 2018; Accepted September 25, 2018

## ABSTRACT

One of the most fundamental questions in biology is what types of cells form different tissues and organs in a functionally coordinated fashion. Larger-scale single-cell sequencing and biology experiment studies are now rapidly opening up new ways to track this question by revealing substantial cell markers for distinguishing different cell types in tissues. Here, we developed the CellMarker database (<http://biocc.hrbmu.edu.cn/CellMarker/> or <http://bio-bigdata.hrbmu.edu.cn/CellMarker/>), aiming to provide a comprehensive and accurate resource of cell markers for various cell types in tissues of human and mouse. By manually curating over 100 000 published papers, 4124 entries including the cell marker information, tissue type, cell type, cancer information and source, were recorded. At last, 13 605 cell markers of 467 cell types in 158 human tissues/sub-tissues and 9148 cell makers of 389 cell types in 81 mouse tissues/sub-tissues were collected and deposited in CellMarker. CellMarker provides a user-friendly interface for browsing, searching and downloading markers of diverse cell types of different tissues. Furthermore, a summarized marker prevalence in each cell type is graphically and intuitively presented through a vivid statistical graph. We believe that CellMarker is a comprehensive and valuable resource for cell researches in precisely identifying and characterizing cells, especially at the single-cell level.

## INTRODUCTION

Development of single-cell RNA sequencing technologies has opened up a hot field that holds tremendous potential to study cell phenotype and cell behaviour at single cell resolution (1,2). A large number of single-cell sequencing studies focusing on resolving tumour heterogeneity, uncovering cell lineage relationships and identifying functional states of individual cells were performed (3–7). One of the most attractive applications of single-cell RNA sequencing technique is to decode complex cellular heterogeneity and create comprehensive reference maps of all cell types in different tissues/organs (8,9). The cell landscapes for various organisms, such as human cell atlas, are gradually being disclosed, accompanied by generation of myriad cell markers for distinct cell types (10,11). For example, Bendall *et al.* identified 13 core surface markers to accurately characterize hematopoietic cell in bone marrow by ‘spanning-tree progression analysis’ for high-dimensional single-cell RNA sequencing data (12). Jaitin *et al.* used unsupervised classification to accurately identify 580 cell markers for five immune cells (B cell, Natural killer cell, Macrophage, Monocyte and Plasmacytoid dendritic cell) by parallel single-cell RNA sequencing (13).

Indeed, many cell markers have been identified through cell biology experiments in recent decades (14,15). These known cell markers were extensively used to identify or isolate cell types of interest by biological techniques (such as fluorescence-activated cell sorting) (16,17). For example, Hermann *et al.* isolated pancreatic stem cells using the cell surface marker CD133 by standard flow cytometry (18). Moreover, the known cell marks are extremely valuable to single-cell sequencing studies, in which they were widely used to label cell types for individual cells, enabling capture

\*To whom correspondence should be addressed. Tel: +86 451 86615922; Fax: +86 451 86615922; Email: xiaoyun@ems.hrbmu.edu.cn  
Correspondence may also be addressed to Xia Li. Email: lixia@hrbmu.edu.cn  
Correspondence may also be addressed to Tingting Zhao. Email: ztt.1984@163.com

†The authors wish it to be known that, in their opinion, the first five authors should be regarded as Joint First Authors.

of the initial viewing of cell compositions. However, there is no database available for cell markers identified by single-cell RNA sequencing and experimental studies. The large amount of cell markers from single-cell sequencing and experimental researches are scattered over thousands of publications. This would make it difficult for researchers to collect cell markers for cells of interest and reliably apply markers to their studies for comprehensive understanding of cell compositions.

To fill this gap, we present CellMarker, a manually curated resource that provides a comprehensive overview of cell markers from single-cell sequencing researches, experimental researches and other resources in both human and mouse. The current version of CellMarker includes more than 13 605 manually curated cell markers in human, involving 467 cell types and 158 tissues/sub-tissues, and 9148 markers in 389 cell types from 81 tissues/sub-tissues in mouse. This database provides a powerful platform that allows for quick retrieval of cell markers of particular cell types in any tissue of interest. Furthermore, a summarized marker prevalence in each cell type is presented, which produces an intuitive view of cell markers used in the cell type. We hope that, in the future, the database will serve as a valuable resource for helping researchers to precisely recognize cell types of interest and further analyze specific biological functions for individual cells in the single-cell sequencing studies.

## DATA COLLECTION AND CONTENT

### Data collection

In view of the important contribution of single-cell RNA sequencing for the identification of cell markers, we searched PubMed using a list of keywords, such as ‘single cell sequencing’, ‘single cell RNA sequencing’, ‘single cell RNA-seq’ and ‘scRNA seq’, and obtained more than 1000 single-cell sequencing related publications. Then, we carefully read all of the full texts and checked their relevant supplement tables, and then extracted the list of cell markers identified in single-cell sequencing studies if the cell types were confidently confirmed.

Furthermore, to obtain cell markers that have been confirmed by flow cytometry, immunostaining, or experimental studies, we searched PubMed by keywords such as ‘cell marker(s)’, ‘cellular marker(s)’, ‘surface marker(s)’, ‘cell specific marker(s)’, ‘cell signature(s)’, ‘cellular signature(s)’, ‘surface signature(s)’ and ‘cell specific signature(s)’. A total of 30 597 publications were obtained. We manually surveyed these publications with their abstracts downloaded through NCBI E-utilities and selected those experimental studies associated with cell markers that were identified by biological experiments or were used to isolate, identify and classify cells. Then, we further checked the full texts of the selected publications ‘by hand’ to confirm the cell markers with reliable evidence and extracted the pertinent information including cell marker name, cell type, tissue type, and species. Among these cell marker-related publications, 2140 review publications that summarized cell markers of corresponding cell types in specific tissues within tables were separately conducted. The relevant cell marker lists were manually extracted.

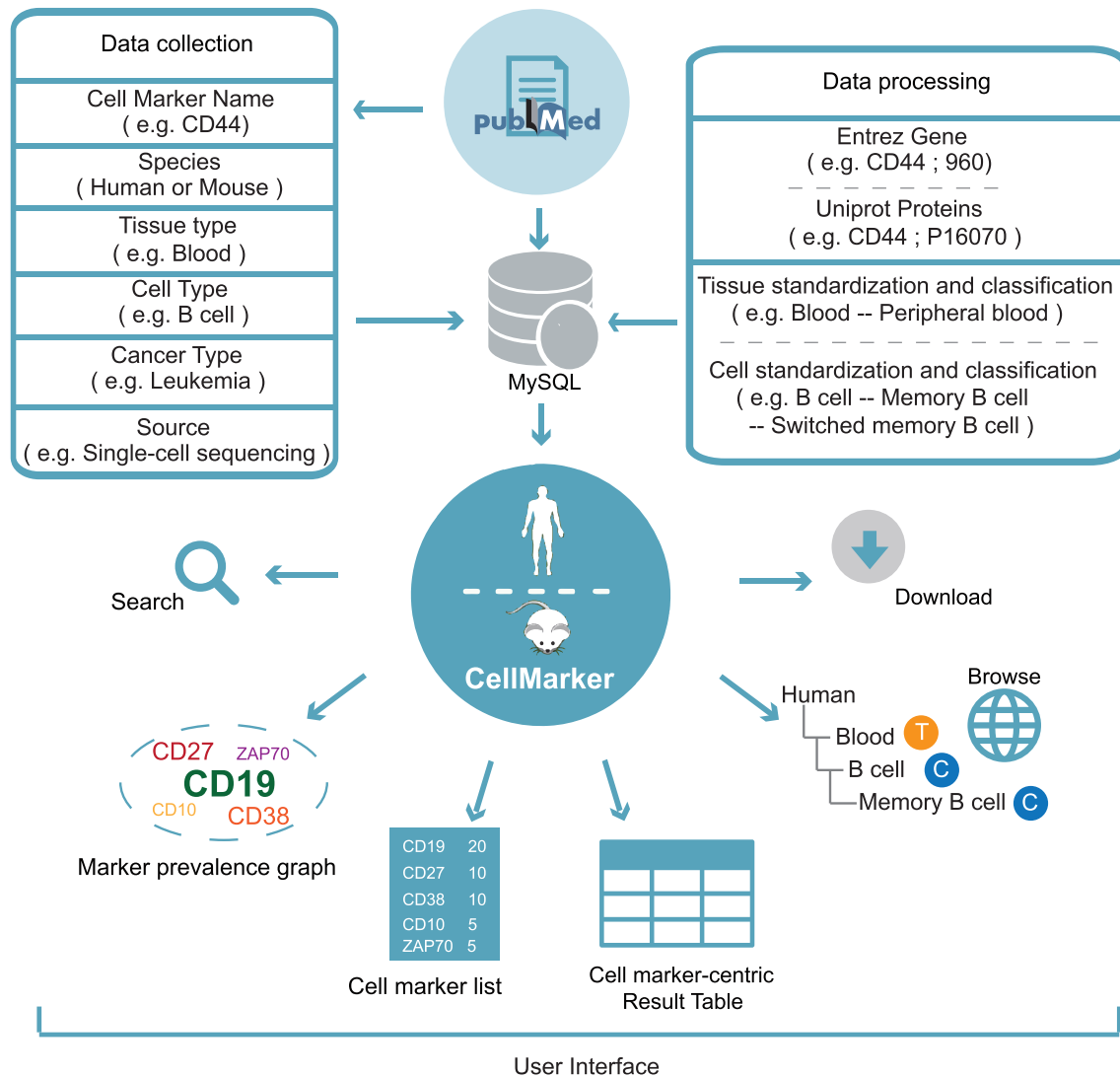
In addition, we applied another search strategy to further complement cell markers that may be lost in the previous search strategy. For each cell type in each type of tissues that were collected and organized in the previous search, we used keywords, including ‘marker(s)’ and the names of tissue type and cell type as well as species information (‘Human’ OR ‘Mouse’ OR ‘mice’) to search for associated cell markers. Like the previous search, we manually extracted the relevant information of cell markers from abstracts of 70 266 experimental publications. The records of cell markers obtained in the two search strategies were then merged.

To reliably collect the information that is presented in distinct and unexportable formats, such as tables encoded into PDF documents and figures in the main text as well as supplementary materials, we carefully scrutinized the corresponding figures and tables in the full text and the supplementary data. There are thirty biological researchers participating in reviewing the full texts of all the papers, with each paper being scrutinized by two or more researchers. Also, we collected cell markers by handbooks or instruction of website from 8 companies (BIO-Rad, Labome, BD biosciences, R&D Systems, BioLegend, abcam, Miltenyi Biotec and Thermo Fisher Scientific).

### Processing and annotation

We carefully curated essential information for each cell marker entry, including species (human or mouse), tissue type, cell type, cell marker name, source (including single-cell sequencing, experiment, review and company) and detailed publication information (i.e. title, PubMed ID, journal and publication year). We also recorded whether cell markers were derived from tumor cells. For example, for cell markers of cancer stem cells in breast cancer, the cancer information was assigned to the corresponding entry. In addition, to avoid duplication of entries, we merged entries with identical cell marker(s), cell type, tissue type, species and cancer information. Through combining these cell marker entries, we built a repertoire of single cell marks for each cell type in each tissue type and recorded the number of supports.

To ensure the unification of tissues and cells curated from the publications, we normalized the names of tissues into the standard tissue list from UniProt (19) and cell names into an integrated reference list based on the Human Cell Atlas (11), CELLPEDIA (20), the list of distinct cell types from Wikipedia and relevant review papers. In detail, based on the tissue list from UniProt, tissue synonyms and tissues with ill-formed names were manually standardized to tissue identifiers. Moreover, tissues were classified into several categories according to anatomical location. If there was no clear clarification of the tissue source, the tissue type would be recorded as ‘Undefined’. This situation for cell types is further complicated by the prevalence of ambiguous names in the literature, which may result from the lack of naming conventions. To address the issue, we carefully checked them in the paper context and uniformed them according to the integrated reference of formal names of cell types. In addition, we filtered entries with cell types that are unlikely to be found in the corresponding tissues. The cell types were further grouped into a hierarchical structure according to their



**Figure 1.** Overview of data collection, annotation and database interface.

localization, morphology and functionality. For example, T helper1 (Th1) cell and T helper2 (Th2) cell were both categorized as T helper cell. T helper cell, Memory T cell, CD4+ T cell and other congeneric cells were further classified as T cell.

The cell markers in the database mainly involve two types: gene and protein. To avoid the inclusion of duplicate vocabulary items, such as different spellings or synonyms, we unified all cell marker names based on the official gene symbols and IDs provided by Entrez Gene (<http://www.ncbi.nlm.nih.gov/gene>) and further converted them into protein names and IDs by UniProt database (19). If the cell marker is a protein complex, we listed all protein subunits and marked them in brackets. In addition, we excluded some types of cell markers, such as monoclonal antibody, fluorescent protein and phytohemagglutinin.

Finally, the cell markers and all the related information were loaded into MySQL. Figure 1 schematically illustrates the general workflow and features of the resource.

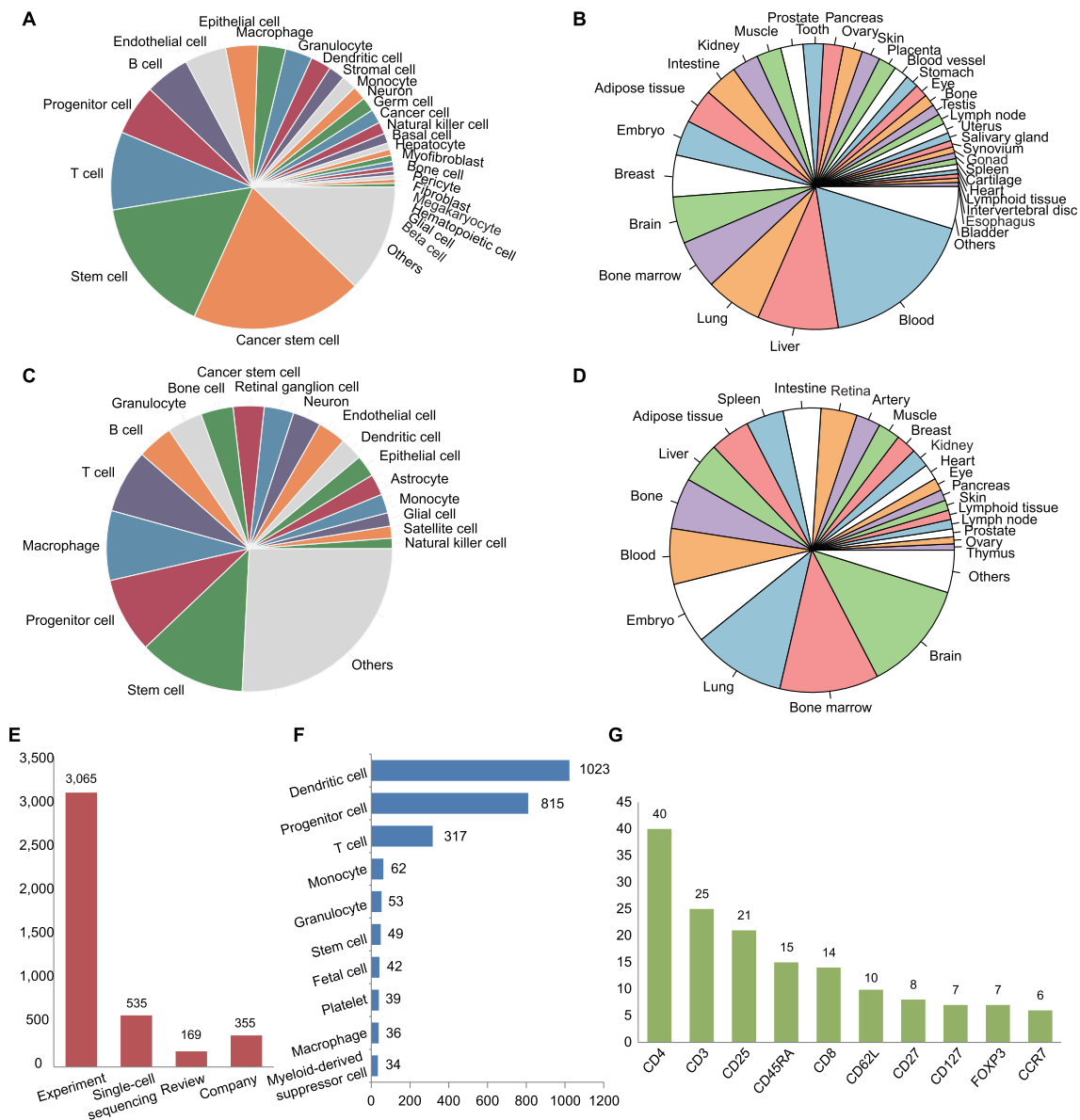
## DATA ACCESS

The CellMarker website is available online at <http://biocc.hrbmu.edu.cn/CellMarker> or <http://bio-bigdata.hrbmu.edu.cn/CellMarker/> and requires no registration. It provides a user-friendly interface to the scientific community to browse, search and download. We have tested it in Mozilla Firefox, Google Chrome and Apple Safari browsers. All cell markers in human and mouse are available as tab-delimited files on the 'Download' page. The CellMarker resource will be updated four times per year, depending on the number of newly published literature, especially those with single-cell RNA sequencing experiments.

## RESULTS

### CellMarker statistics

CellMarker contains a map of a total of 13 605 and 9148 cell markers in human and mouse, respectively, involving



**Figure 2.** Statistics of cell markers in CellMarker. (A) Distribution of cell markers for different cell types in human. (B) Distribution of cell markers for different tissue types in human. (C) Distribution of cell markers for different cell types in mouse. (D) Distribution of cell markers for different tissue types in mouse. (E) The numbers of cell marker entries from four different sources. (F) The top 10 widely studied cell types for cell marker research in blood. (G) The top 10 frequently used cell markers in T cell.

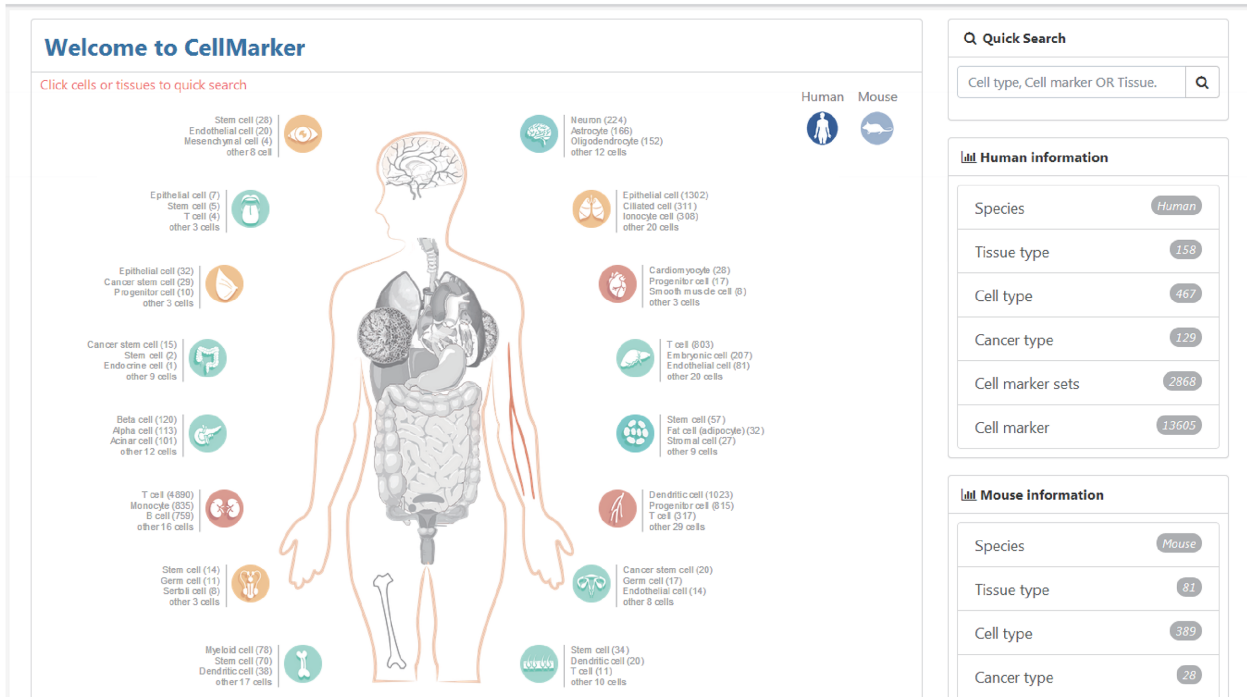
467 cell types from 158 tissues in human (Figure 2A and B) and 389 cell types from 81 tissues in mouse (Figure 2C and D). These markers are derived from 4124 cell marker entries (2868 in human and 1256 in mouse) collected from more than 2300 studies, with 3065 cell marker entries collected from experimental literature, 535 identified using single-cell sequencing technology, 169 summarized in reviews and 355 derived from companies (Figure 2E). Notably, ~19.6% of the cell marker entries are derived from cancer studies, involving 131 cancer types, and the most cancer cell type is cancer stem cell. On average, around 6.5 and 7.1 cell types with available cell marks are included in each tissue of human and mouse, with blood having the most cell types in human and lung having the most in mouse. Taking Blood

in human as an example, the number of cell markers in cell types ranges from 1 to 1023. The top 5 cell types with the most cell markers are dendritic cell, progenitor cell, T cell, monocyte and granulocyte (Figure 2F). Among them, T cell is the cell with the most amount of entries (with 317 cell markers from 98 entries) and the top 5 most frequently used cell markers of T cell are CD4, CD3, CD25, CD45RA and CD8 (Figure 2G).

### Web interface

CellMarker's public user interface allows intuitive browsing and searching of any data in the database. In the 'Home' page, users can quickly explore cell markers by clicking on hyperlinks embedded in the web images - 'Anatomical loca-





**Welcome to CellMarker**

Click cells or tissues to quick search

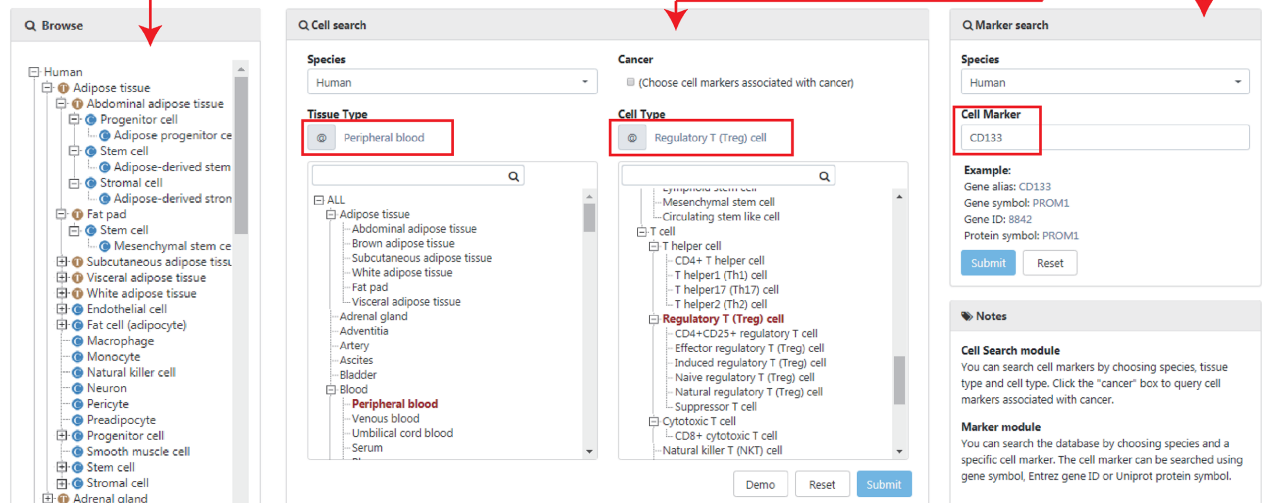
Human Mouse

**Human information**

Species	Human
Tissue type	158
Cell type	467
Cancer type	129
Cell marker sets	2868
Cell marker	13605

**Mouse information**

Species	Mouse
Tissue type	81
Cell type	389
Cancer type	28



**Q Browse**

- Human
  - Adipose tissue
    - Abdominal adipose tissue
      - Progenitor cell
      - Adipose progenitor ce
      - Stem cell
      - Adipose-derived stem
      - Adipose-derived stron
    - Fat pad
      - Stem cell
      - Mesenchymal stem ce
    - Subcutaneous adipose tiss
    - Visceral adipose tissue
    - White adipose tissue
    - Endothelial cell
    - Fat cell (adipocyte)
    - Macrophage
    - Monocyte
    - Natural killer cell
    - Neuron
    - Pericyte
    - Preadipocyte
    - Progenitor cell
    - Smooth muscle cell
    - Stem cell
    - Stromal cell
    - Adrenal gland

**Q Cell search**

Species: Human

Tissue Type: **Peripheral blood**

Cancer: (Choose cell markers associated with cancer)

Cell Type: **Regulatory T (Treg) cell**

**Q Marker search**

Species: Human

Cell Marker: **CD133**

Example:  
Gene alias: CD133  
Gene symbol: PROM1  
Gene ID: 3842  
Protein symbol: PROM1

Submit Reset

Notes

**Cell search module**  
You can search cell markers by choosing species, tissue type and cell type. Click the "cancer" box to query cell markers associated with cancer.

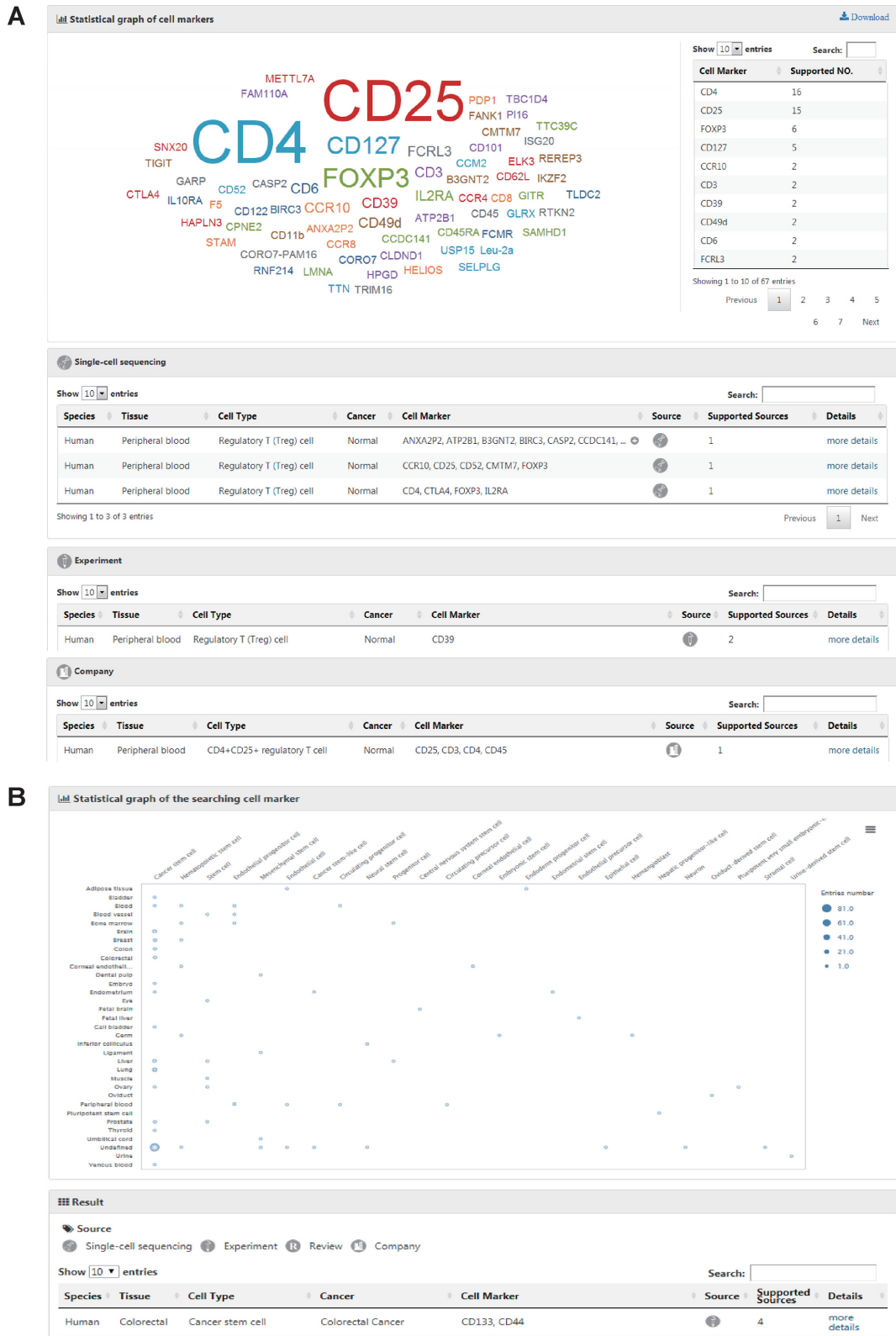
**Marker module**  
You can search the database by choosing species and a specific cell marker. The cell marker can be searched using gene symbol, Entrez gene ID or Uniprot protein symbol.

**Figure 3.** A schematic workflow of CellMarker. (A) The web images in the home page allow to quick search for cell markers of cells in different tissues. (B) The 'Browse' and 'Search' pages allow the users to browse and search cell markers.

tion of human or mouse cells' (Figure 3A). After clicking on the icons of different tissues, users can further select relevant cell types in a pop-up window to retrieve the corresponding cell markers. In addition, the home page provides a quick search utility which can be used to query the database for cell marker, cell or tissue type. In the 'Browse' page, to easily access cell markers of diverse cell types in different tissues, hierarchical structures of cells and tissues for human and mouse are shown in the right tree-based panel. Users can

browse cell markers by clicking cell types in different tissues of human (or mouse), and the complete list of matched cell marker entries can be returned.

In the 'Cell search' page, users can search for cell markers of cells of interest by selecting species from the pull-down menu, tissue type from the hierarchical tree of tissues and cell type from the hierarchical tree of cell types (Figure 3B). In particular, cell markers associated with cancers in the corresponding tissue will be returned if choosing the



**Figure 4.** Search result pages of CellMarker. (A) Results of a search for cell markers of Regulatory T (Treg) cell in blood. The cell markers of this cell type are presented in an intuitive statistical graph of cell marker prevalence. An integrative cell marker list for this cell type is provided. The detailed tables of cell marker entries derived from different sources are obtained in the below. (B) The results of the search for cell marker CD133. An interactive bubble chart and a table of comprehensive information of the cell marker are shown.

‘Cancer’ box. After clicking on the ‘Submit’ button, a result page is displayed. The result page contains an intuitive statistical graph of cell marker prevalence, which is quantified by the number of supportive evidence (i.e., publications) and an integrative cell marker list for the queried cell type (Figure 4A). The graph of cell marker prevalence allows users to easily understand the frequently used cell markers. Also, the result page displays four detailed tables of cell marker entries derived from different sources including single-cell RNA sequencing, experiment, review and company. By clicking on the hyperlink ‘more details’ for an individual entry in these tables, users can obtain more detailed information including gene symbols, gene IDs, protein names, protein IDs, publication information (i.e. title, PubMed ID, journal and publication year) and the cross references to external databases. In addition, CellMarker provides several features for browsing the query results. Users can filter entries by typing terms in the ‘Search’ box at the top right of the table, sort columns in an ascending (or a descending) fashion by clicking the arrows in the column headers and select table size per page by using the ‘Show entries’ pull-down menu on the top left side above the table.

Query on the ‘Marker search’ section, users can search any gene of interest by entering gene symbol, gene ID or protein name to query in which cell types of which tissues the specific gene can act as cell marker (Figure 3B). After clicking the ‘Submit’ button, the search engine will return an interactive bubble chart and a table showing comprehensive information of cell marker entries (Figure 4B). The interactive bubble chart exhibits how often the gene of interest is used as a cell marker in different cells of different tissues.

In addition, users can download all data and submit new data via the ‘Download’ and ‘Submit’ page, respectively. CellMarker also provides a detailed tutorial for the usage of the database on the ‘Help’ page.

### Applications of CellMarker

The CellMarker database provides users with a function to extract cell markers of interest in specific cell types of different tissues. For example, for regulatory T (Treg) cell in human blood, users can investigate the relevant cell markers. In the ‘Cell search’ section, by selecting the species ‘Human’, tissue type ‘Blood’ and cell type ‘Regulatory T (Treg) cell’, users will obtain an intuitive summarized statistical graph of cell marker prevalence and a matched table showing all the cell markers (Figure 4A). Furthermore, the detailed tables show all cell markers of regulatory T (Treg) cell in blood from 3 single-cell RNA sequencing studies, 19 experimental researches and five companies. Among these markers, CD4 and CD25 are the most widely used cell markers for regulatory T (Treg) cell in human blood, with the supportive evidence from two single-cell sequencing studies, 10 experiments, and five companies (Figure 4A). Based on the query results, users can select the top several widely used cell markers, such as CD4, CD25, FOXP3, and CD127, to support their researches.

### SUMMARY AND FUTURE PERSPECTIVES

Cell markers are widely used to isolate, identify and classify cells. Recently, single-cell RNA sequencing studies have suc-

cessfully revealed extensive cellular heterogeneity and accurately characterized cell markers that are unique to different cell types. Therefore, a comprehensive and reliable resource of cell markers that are identified by single-cell RNA sequencing and experimental studies is urgently needed. In this study, we developed CellMarker, a manually curated resource of cell markers collected from single-cell RNA sequencing researches, experimental researches and other resources in human and mouse. Construction of a map of cell markers of distinct cell types from different tissues is the key benefit of our database. CellMarker builds a uniform annotation for names of cell markers with a variety of formats in the original publications and organizes cell and tissue types into hierarchical structures. Importantly, CellMarker also provides the prevalence of cell markers in each cell type through a vivid statistical graph, which gives an intuitionistic exhibition of cell markers used in the cell type.

In summary, CellMaker will be an informative and valuable resource for researchers to identify cell types of interest and further analyze specific biological functions for individual cells in the single-cell RNA sequencing studies. Further extensions will be conducted in the following aspects. First, more and more cell markers of diverse cells (such as some new/rare cell types) will be identified by single-cell RNA sequencing technology. Therefore, we will continue to track the single-cell sequencing studies and update the database by frequent additions of new cell markers. Second, cell markers of other species (such as zebrafish, *Drosophila* and *Caenorhabditis elegans*) will be added to in a future version of CellMarker to provide more comprehensive cell marker information for users. Third, markers of cell states will be added to the CellMarker, since the single-cell RNA sequencing technology allows researchers to systematically and precisely characterize cell states in various cell populations. Fourth, other types of cell markers, such as non-coding RNAs and epigenetic modification of functional elements (e.g. enhancers) (21), will be also collected to the database.

### FUNDING

National High Technology Research and Development Program of China (863 Program) [2014AA021102, in part]; National Program on Key Basic Research Project (973 Program) [2014CB910504]; National Natural Science Foundation of China [61473106, 61573122, 31601076]; China Postdoctoral Science Foundation [2016M600260]; Wu lien-teh Youth Science Fund Project of Harbin Medical University [WLD-QN1407]; Special Funds for the Construction of Higher Education in Heilongjiang Province [UNPYSCT-2016049]; Heilongjiang Postdoctoral Foundation [LBHZ16098]. Funding for open access charge: National High Technology Research and Development Program of China (863 Program) [2014AA021102, in part]; National Program on Key Basic Research Project (973 Program) [2014CB910504]; National Natural Science Foundation of China [61473106, 61573122, 31601076]; China Postdoctoral Science Foundation [2016M600260]; Wu lien-teh Youth Science Fund Project of Harbin Medical University [WLD-QN1407]; Special Funds for the Construction of Higher Education in Heilongjiang Province [UNPYSCT-

2016049]; Heilongjiang Postdoctoral Foundation [LBH-Z16098].

*Conflict of interest statement.* None declared.

## REFERENCES

1. Yuan,G.C., Cai,L., Elowitz,M., Enver,T., Fan,G., Guo,G., Irizarry,R., Kharchenko,P., Kim,J., Orkin,S. *et al.* (2017) Challenges and emerging directions in single-cell analysis. *Genome Biol.*, **18**, 84.
2. Poirion,O.B., Zhu,X., Ching,T. and Garmire,L. (2016) Single-Cell transcriptomics bioinformatics and computational challenges. *Front. Genet.*, **7**, 163.
3. Han,X., Chen,H., Huang,D., Chen,H., Fei,L., Cheng,C., Huang,H., Yuan,G.C. and Guo,G. (2018) Mapping human pluripotent stem cell differentiation pathways using high throughput single-cell RNA-sequencing. *Genome Biol.*, **19**, 47.
4. Cochain,C., Vafadarnejad,E., Arampatzi,P., Pelisek,J., Winkels,H., Ley,K., Wolf,D., Saliba,A.E. and Zernecke,A. (2018) Single-Cell RNA-Seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ. Res.*, **122**, 1661–1674.
5. Zheng,H., Pomyen,Y., Hernandez,M.O., Li,C., Livak,F., Tang,W., Dang,H., Greten,T.F., Davis,J.L., Zhao,Y. *et al.* (2018) Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma. *Hepatology*, **68**, 127–140.
6. Puram,S.V., Tirosh,I., Parikh,A.S., Patel,A.P., Yizhak,K., Gillespie,S., Rodman,C., Luo,C.L., Mroz,E.A., Emerick,K.S. *et al.* (2017) Single-Cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell*, **171**, 1611–1624.
7. Guo,J., Grow,E.J., Yi,C., Mlcochova,H., Maher,G.J., Lindskog,C., Murphy,P.J., Wike,C.L., Carrell,D.T., Goriely,A. *et al.* (2017) Chromatin and Single-Cell RNA-Seq profiling reveal dynamic signaling and metabolic transitions during human spermatogonial stem cell development. *Cell Stem Cell*, **21**, 533–546.
8. Buettner,F., Natarajan,K.N., Casale,F.P., Proserpio,V., Scialdone,A., Theis,F.J., Teichmann,S.A., Marioni,J.C. and Stegle,O. (2015) Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.*, **33**, 155–160.
9. Buenrostro,J.D., Corces,M.R., Lareau,C.A., Wu,B., Schep,A.N., Aryee,M.J., Majeti,R., Chang,H.Y. and Greenleaf,W.J. (2018) Integrated Single-Cell analysis maps the continuous regulatory landscape of human hematopoietic differentiation. *Cell*, **173**, 1535–1548.
10. Han,X., Wang,R., Zhou,Y., Fei,L., Sun,H., Lai,S., Saadatpour,A., Zhou,Z., Chen,H., Ye,F. *et al.* (2018) Mapping the mouse cell atlas by Microwell-Seq. *Cell*, **172**, 1091–1107.
11. Regev,A., Teichmann,S.A., Lander,E.S., Amit,I., Benoist,C., Birney,E., Bodenmiller,B., Campbell,P., Carninci,P., Clatworthy,M. *et al.* (2017) The human cell atlas. *Elife*, **6**, e27041.
12. Bendall,S.C., Simonds,E.F., Qiu,P., Amir el,A.D., Krutzik,P.O., Finck,R., Bruggner,R.V., Melamed,R., Trejo,A., Ornatsky,O.I. *et al.* (2011) Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science*, **332**, 687–696.
13. Jaitin,D.A., Kenigsberg,E., Keren-Shaul,H., Elefant,N., Paul,F., Zaretzky,I., Mildner,A., Cohen,N., Jung,S., Tanay,A. *et al.* (2014) Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*, **343**, 776–779.
14. Ceder,J.A., Jansson,L., Helczynski,L. and Abrahamsson,P.A. (2008) Delta-like 1 (Dlk-1), a novel marker of prostate basal and candidate epithelial stem cells, is downregulated by notch signalling in intermediate/transit amplifying cells of the human prostate. *Eur. Urol.*, **54**, 1344–1353.
15. Ma,S., Chan,K.W., Hu,L., Lee,T.K., Wo,J.Y., Ng,I.O., Zheng,B.J. and Guan,X.Y. (2007) Identification and characterization of tumorigenic liver cancer stem/progenitor cells. *Gastroenterology*, **132**, 2542–2556.
16. Dekaney,C.M., Rodriguez,J.M., Graul,M.C. and Henning,S.J. (2005) Isolation and characterization of a putative intestinal stem cell fraction from mouse jejunum. *Gastroenterology*, **129**, 1567–1580.
17. Roy,N.S., Wang,S., Jiang,L., Kang,J., Benraiss,A., Harrison-Restelli,C., Fraser,R.A., Couldwell,W.T., Kawaguchi,A., Okano,H. *et al.* (2000) In vitro neurogenesis by progenitor cells isolated from the adult human hippocampus. *Nat. Med.*, **6**, 271–277.
18. Hermann,P.C., Huber,S.L., Herrler,T., Aicher,A., Ellwart,J.W., Guba,M., Bruns,C.J. and Heeschen,C. (2007) Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell*, **1**, 313–323.
19. Apweiler,R., Bairoch,A., Wu,C.H., Barker,W.C., Boeckmann,B., Ferro,S., Gasteiger,E., Huang,H., Lopez,R., Magrane,M. *et al.* (2004) UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.*, **32**, D115–D119.
20. Hatano,A., Chiba,H., Moesa,H.A., Taniguchi,T., Nagaie,S., Yamanegi,K., Takai-Igarashi,T., Tanaka,H. and Fujibuchi,W. (2011) CELLPEDIA: a repository for human cell information for cell studies and differentiation analyses. *Database*, **2011**, bar046.
21. Zhang,G., Shi,J., Zhu,S., Lan,Y., Xu,L., Yuan,H., Liao,G., Liu,X., Zhang,Y., Xiao,Y. *et al.* (2018) DiseaseEnhancer: a resource of human disease-associated enhancer catalog. *Nucleic Acids Res.*, **46**, D78–D84.