Briefings in Bioinformatics, 2026, 27, bbag151
https://doi.org/10.1093/bib/bbag151
Published: 19 April 2026
Problem Solving Protocol

OXFORD

AICellType: a large language model-based platform for accurate cell type annotation

Chuxing Cheng[1,2,3,4,5], Shuo Fang[3,4], Qi Zuo[1,2,3,4,5], JiaHui Sun[3,4], Xiaotong Hu[1,2,4,5], Xiaokun Liu[1,5,6], Meilin Jin[1,2,3,4,5,6,*]

[1]College of Animal Science & Veterinary Medicine, Huazhong Agricultural University, No. 1 Shizishan Street, Wuhan 430070, Hubei, China
[2]National Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, No. 1 Shizishan Street, Wuhan 430070, Hubei, China
[3]Research Institute of Wuhan Keqian Biology Co., Ltd, No. 101 Guanggu 8th Road, Wuhan 430070, Hubei, China
[4]Center for Cross-disciplinary Studies in Animal Epidemic and Public Health, Building A1, R&D Production Center, Health Industry Park, G107 Service Road, Jiangxia District, Wuhan 430200, Hubei, China
[5]Interdisciplinary Frontier Research Center for Animal Diseases and One Health, Huazhong Agricultural University, No. 1 Shizishan Street, Wuhan 430070, Hubei, China
[6]Key Laboratory of Prevention and Control of African Swine Fever and Other Major Swine Diseases, Ministry of Agriculture, No. 1 Shizishan Street, Wuhan 430070, Hubei, China

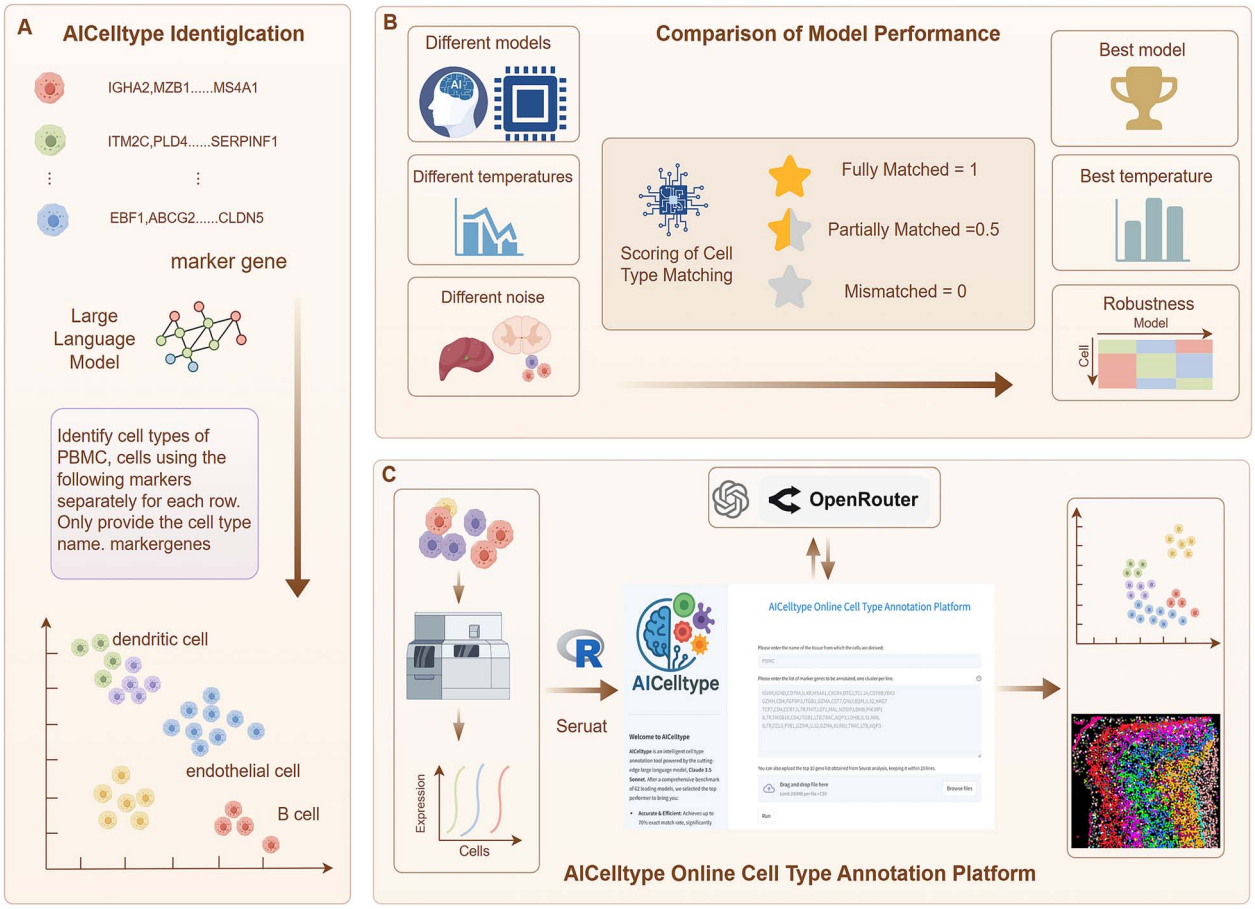
*Corresponding author. College of Animal Science & Veterinary Medicine, Huazhong Agricultural University, No. 1 Shizishan Street, Wuhan 430070, Hubei, China. E-mail: jinmeilin@mail.hzau.edu.cn

Abstract

Accurate cell type annotation is critical for studying cellular heterogeneity in single-cell and spatial transcriptomics. However, existing methods largely rely on static gene markers, limiting adaptability to diverse biological contexts and data types. To overcome this limitation, we systematically benchmarked 79 large language models (LLMs) over 1130 single-cell and spatial transcriptomics datasets using an evaluation framework combining ontology structure and semantic reasoning to quantify model performance in biological relevance and annotation robustness. Claude 3.5 Sonnet achieved the best overall performance, balancing weighted accuracy (76%), robustness, inference speed, and cost-efficiency. Based on these findings, we developed AICellType (https://AICellType.jinlab.online), a free, open-source R package and web platform that integrates seamlessly with Seurat workflows, supports multiple species and tissues, and enables flexible model deployment via OpenRouter or custom APIs. By leveraging LLMs' capacity to interpret marker–cell type associations, AICellType provides a scalable, efficient, and accessible solution for real-world cell annotation in both single-cell and spatial omics research.

Received: September 23, 2025. **Revised:** November 21, 2025. **Accepted:** February 13, 2026

Graphical abstract



Keywords cell type annotation, single-cell RNA-seq, large language model, Claude 3.5 Sonnet

Introduction

Single-cell RNA sequencing has revolutionized our knowledge of cellular heterogeneity [1]. However, accurate cell type annotation remains a major bottleneck in downstream analysis [2, 3]. Manual annotation is labor-intensive, relies heavily on expert knowledge, and lacks reproducibility [4, 5]. In response, various automated methods have been developed, yet they often fall short in terms of scalability, adaptability to diverse biological systems, or cross-platform compatibility.

Recently, large language models (LLMs) have emerged as a promising alternative for biological text and knowledge interpretation, showing potential in cell type annotation tasks [6–9]. Early efforts, such as Cell2Sentence, demonstrated LLM feasibility but were limited by small corpora and narrow task scopes [6]. More advanced LLM-based methods—including those built on GPT-series models—have extended these capabilities, yet practical deployment remains challenging due to high costs, lack of transparency, limited species coverage, and reliance on proprietary commercial application programming interfaces (APIs) [3, 6–8, 10].

Moreover, the fast-paced evolution of LLMs presents a dilemma for researchers: choosing between competing models without objective performance comparisons and navigating integration into common analysis workflows like Seurat. Most existing solutions lack rigorous benchmarking and open-source implementations, further hindering their adoption.

These challenges highlight the urgent need for an open, robust, and cost-effective platform that leverages LLMs for accurate and scalable cell type annotation across species, tissues, and experimental contexts. In this work, we systematically assess the performance of state-of-the-art LLMs for cell annotation and propose a solution that enables seamless integration into mainstream analytical pipelines [10, 11].

Results

To evaluate the accuracy of LLMs in cell type annotation, 79 mainstream LLMs were applied to 1130 single-cell RNA-seq datasets, with expert manual annotations used as the gold standard for comparison. LLM outputs were scored as 1 (full match), 0.5 (partial match), or 0 (mismatch), respectively. For example, when the LLM predicted "Neuron" while the manual reference was "Ganglion cell," a partial match (0.5) was assigned because ganglion cells are a subtype of neurons.

Anthropic's Claude 3.5 Sonnet (version 240620) achieved the highest average score (0.756), significantly outperforming GPT-4 (P = .0137) and SingleR ($P < 2.2 \times 10^{-16}$), while differences relative to o4.mini (P = .87) and Gemini 2.5 (P = .93) were not statistically

significant (Fig. 1a, Supplementary Fig. S2, Supplementary Fig. S3). Claude 3.5 Sonnet also demonstrated favorable cost-effectiveness ($0.41 per 1000 cells, lower than GPT-4's $2.15) and faster annotation speed (median 3.7 s, compared with o4.mini's 94 s) (Fig. 1b and c). Traditional tools such as scType [12] (35.4% weighted accuracy) and CellMarker 2.0 [13] (38.6%) performed substantially worse than leading LLMs (Fig. 1a). Detailed comparative visualization of annotation results across different models on the human immune cell dataset GSE305979 is provided in Supplementary Fig. S3, where AICellType showed high concordance with the ground truth, accurately recovering major immune populations such as B cells, T cells, and myeloid cells, and demonstrating overall superior performance compared with conventional annotation tools.

Notably, in the specific task of cell typing, complex reasoning strategies such as Chain-of-Thought yielded only limited gains in accuracy, and larger models (e.g. Qwen series) did not guarantee improved performance; task-specific tuning and model choice were more critical (Supplementary Fig. S1b).

Based on these advantages, we developed AICellType, a free, open-source cell annotation tool powered by Claude 3.5 Sonnet, accessible via an online platform (https://AICellType.jinlab.online/) and R package. Testing on 54 independent scenarios from diverse sources including Azimuth and HCA, AICellType achieved overall weighted accuracy of 0.76 and score $\geq$0.7 in 74% of test cases (Fig. 1d). AICellType exhibited high robustness to temperature variations, achieving nearly 100% consistency at temperatures 0.1–0.3 (Fig. 1e), and to incomplete or noisy marker gene sets, maintaining $\sim$70% median accuracy even with 80% marker gene removal or 100% noise gene addition (Fig. 1g–j).

We further assessed model stability across varying marker gene counts in the human immune cell single-cell dataset. When only the top five markers were used, annotation performance declined noticeably, whereas increasing the number to 10, 15, or 20 markers produced largely comparable results. However, for highly similar immune subpopulations, more extensive marker sets were occasionally required to achieve optimal discrimination. Accordingly, AICellType incorporates a flexible parameter for marker gene number, allowing users to adjust input size according to dataset complexity and biological context (Supplementary Fig. S4).

To further expand the application of AICellType in other biological contexts, we applied it to two representative datasets: a public 10x Genomics peripheral blood mononuclear cell (PBMC) dataset (Fig. 2a–d) and an in-house generated mouse brain spatial transcriptomics dataset (Fig. 2e–h). In PBMC data, AICellType accurately identified major immune cells with clustering matching Seurat [14]. In mouse brain spatial data, it correctly annotated key neuronal/glial types with spatial patterns aligned to known brain structures, surpassing CellMarker2.0 and scType (Fig. 2f–h).

Notably, a considerable portion of the benchmark datasets employed for AICellType's evaluation predates the knowledge cutoff dates of the tested LLMs. Importantly, both the human immune cell dataset and the porcine spleen dataset used for independent validation were entirely unseen by these models. AICellType not only demonstrated robust performance on these potentially "seen" benchmarks but also achieved excellent annotation accuracy on the unseen human immune cell dataset (Supplementary Fig. S3) and the unpublished porcine spleen dataset (Supplementary Fig. S5) [15]. These findings provide compelling evidence that AICellType's annotation capability reflects its capacity to learn and infer deeper biological associations and patterns for effective generalization, enabling it to effectively address novel biological contexts beyond its training data.

Discussion

In the broader context of existing single-cell analytic approaches, most frameworks emphasize feature extraction and clustering. Compared with previously reported single-cell analysis frameworks, such as scMoMtF, DSINMF, JLONMFSC, and scMCG proposed by Wei Lan *et al.* [16–20], which mainly focus on feature extraction and clustering, AICellType offers an efficient and interpretable cell type annotation capability that complements these upstream analyses.

While efforts to fine-tune commercial LLMs for specialized biological tasks are emerging [8], the selection of a superior base model is paramount for the success of such optimizations. Our study provides a critical framework for this selection, and current findings highlight Claude 3.5 Sonnet as an ideal choice for further fine-tuning in the domain of cell type annotation. Furthermore, for integrating known but unpublished marker gene-cell type associations or dynamically updated proprietary knowledge bases, constructing an auxiliary knowledge system using retrieval-augmented generation (RAG) technology presents a compelling alternative. Compared to model fine-tuning, RAG significantly lower the technical barrier and enables more agile and efficient knowledge integration and application.

Despite the remarkable performance and versatility demonstrated by LLMs in cell type annotation, dependence on commercial LLMs introduces inherent limitations. The underlying architectures and training data of these models are continuously updated by providers, potentially altering inference behaviors and compromising reproducibility. Outputs may vary across versions even for identical inputs, posing challenges for stable biological interpretation. Moreover, reliance on external API access entails risks of service availability, dynamic pricing, and potential latency, all of which can influence the sustainability of academic applications. The opacity of proprietary optimization pipelines and fixed knowledge cutoffs further constrain systematic tracking of their biological knowledge updates. To mitigate these risks, our framework was designed with openness and replaceability in mind: AICellType can seamlessly integrate other open-source or locally deployed LLMs via standardized evaluation interfaces, ensuring cross-model comparability. This strategy not only enhances long-term robustness but also provides a practical foundation for developing domain-specific, community-controlled LLMs tailored for cell biology in the future.

Although AICellType demonstrated high accuracy and robustness across diverse single-cell and spatial omics datasets, the underlying LLM remains intrinsically a "black-box" system [21]. Due to the inaccessibility of the internal parameters and training corpus of commercial LLMs, conventional interpretability techniques—such as SHapley Additive exPlanations (SHAP) analysis or attention-weight mapping—are not directly applicable in this setting. To enhance transparency in biological inference, we implemented a novel interpretability module within AICellType: alongside each predicted cell type, Claude 3.5 Sonnet generates a natural language explanation detailing the reasoning process, including its assessment of the relationships between the provided marker genes and known cell type categories. Researchers can directly inspect these reasoning chains and, if desired, engage in interactive dialog with the model to probe, clarify, or challenge its decision bases. This design provides traceable cognitive pathways,

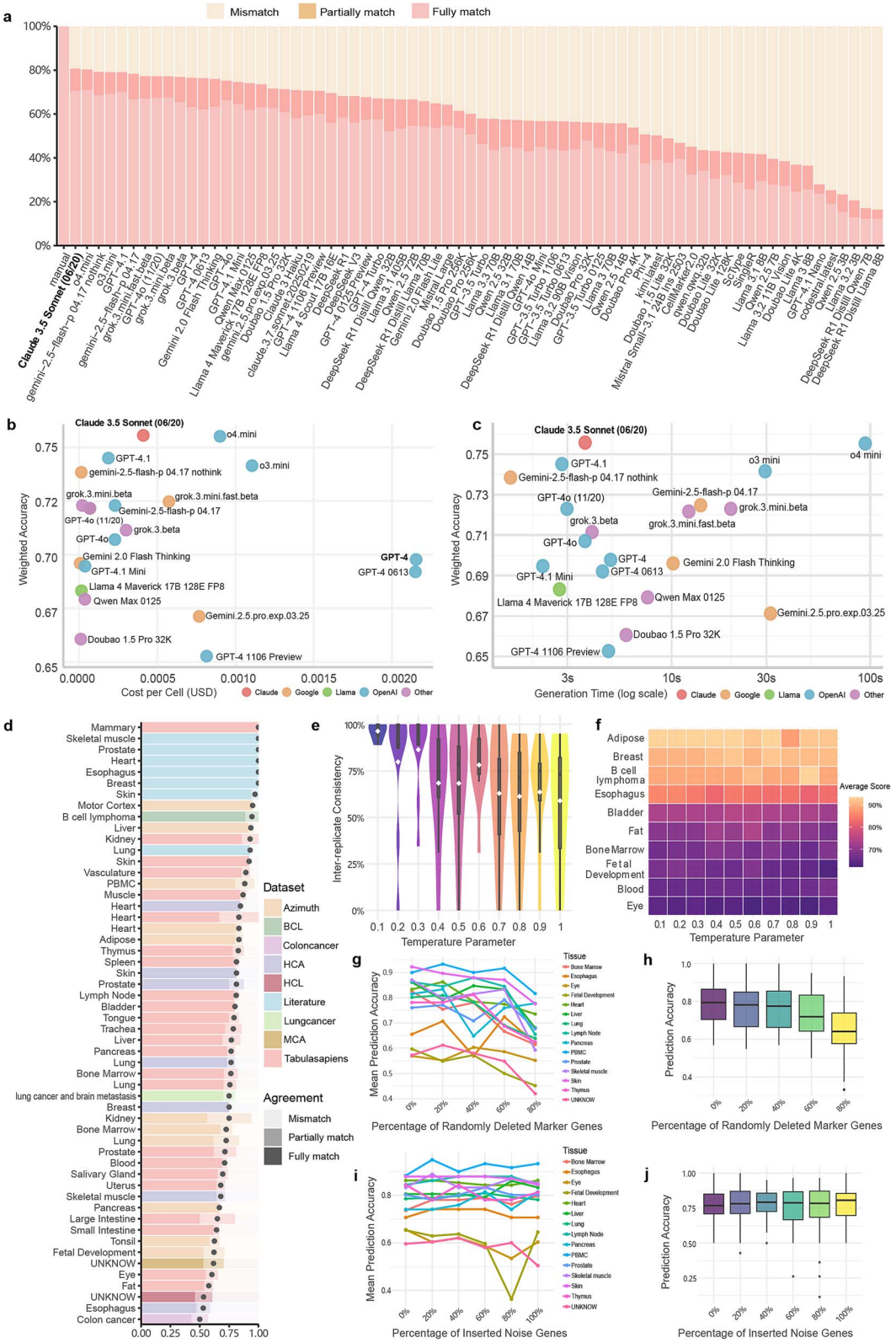


Figure 1 Benchmarking of large language models and AICellType performance in cell type annotation; (a) Cell annotation performance of 79 large language models (LLMs); stacked bars indicate the proportions of fully match, partially match, and mismatch (ordered from bottom to top within each bar), ranked by weighted accuracy; (b, c) weighted accuracy versus (b) annotation cost and (c) generation time for top LLMs; symbols are categorized by model families as defined in the bottom legend; (d) AICellType annotation performance across cell types and datasets; the length of the stacked bars represents the proportions of fully match (darkest shading), partially match (medium shading), and mismatch (lightest shading); dots indicate overall concordance; (e) effect of temperature on AICellType annotation consistency; (f) mean AICellType annotation scores across tissues at varying temperatures; (g–h) robustness to random marker gene deletion; (g) mean accuracy across tissues; (h) overall accuracy distribution; (i, j) robustness to random noisy gene insertion; (i) mean accuracy across tissues; (j) overall accuracy distribution.

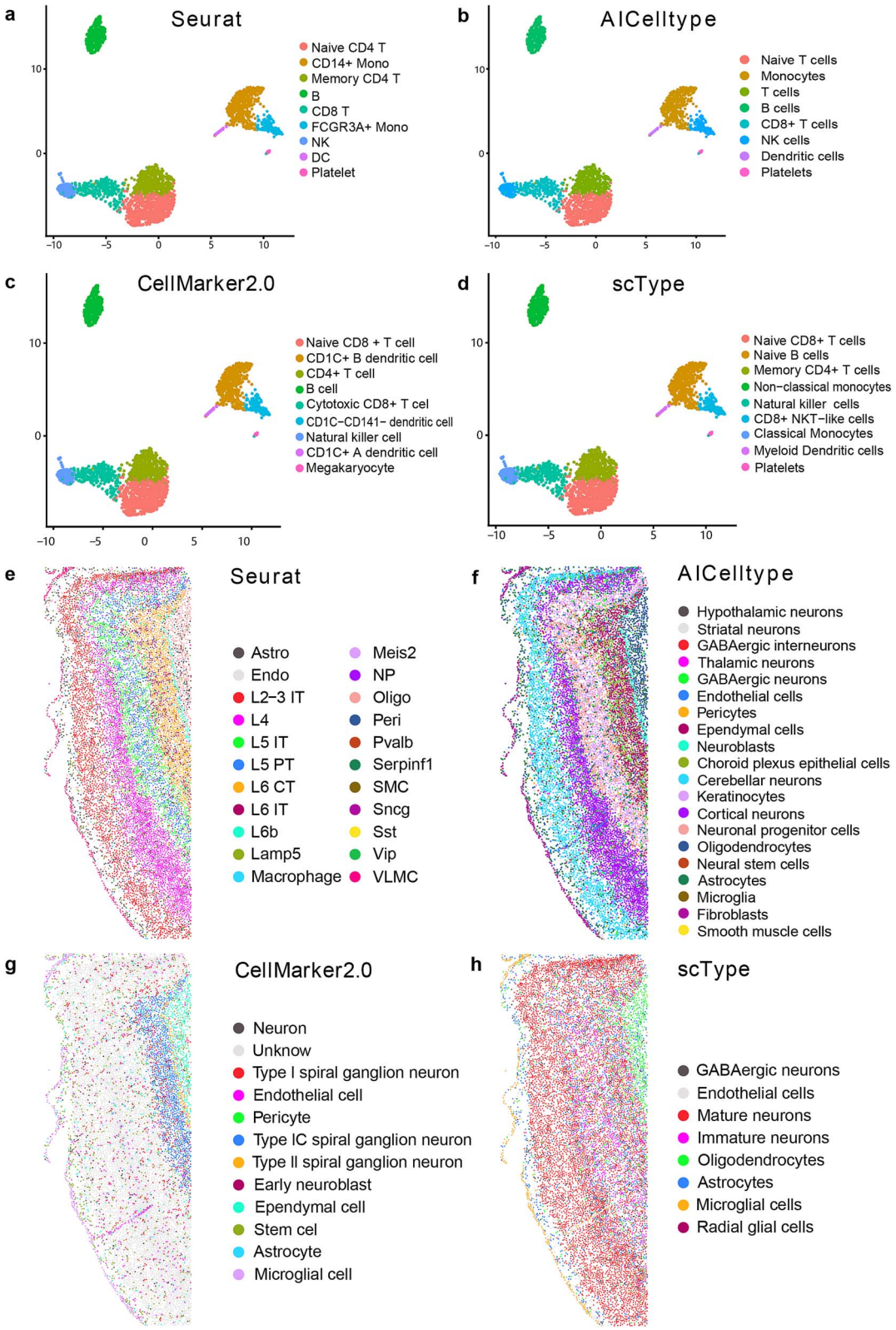


Figure 2 Application of AICellType in single-cell sequencing and spatial transcriptomics; (a–d) cell type annotation in PBMC scRNA-seq data; UMAP visualizations by (a) Seurat reference, (b) AICellType, (c) CellMarker2.0, and (d) scType; (e–h) cell type mapping in mouse brain spatial transcriptomics; annotations by (e) reference, (f) AICellType, (g) CellMarker2.0, and (h) scType.

facilitating both confidence assessment and hypothesis generation for novel marker-cell associations. While this approach cannot substitute for fine-grained attribution analyses based on internal model parameters, it mitigates some limitations of opaque inference and lays a practical foundation for future developments toward interpretable LLMs in biological applications.

Although the performance of AICellType still relies on the quality of input marker genes and may be influenced by the timeliness of embedded biological knowledge, the tool nevertheless provides a powerful framework for accelerating discoveries in cell biology. We anticipate that the application of AICellType will significantly lower the technical barrier for cell annotation, advance the frontiers of single-cell and spatial omics research, and benefit both basic research and clinical translation. By continuously publishing updates on GitHub, we aim to provide the research community with a dynamic evaluation resource, thereby ensuring that AICellType consistently integrates the most advanced models.

Conclusion

In summary, AICellType provides a reliable, efficient, and cost-effective tool for cell type annotation in single-cell and spatial transcriptomics data. LLMs leverage diverse training data to efficiently extract and integrate unstructured biological knowledge, with their knowledge bases expanding and updating at a scale and pace far beyond manual systematic curation. AICellType is a promising tool for cell-type inference.

Materials and methods

Construction of a comprehensive benchmark suite

To systematically evaluate the performance of LLMs in cell type classification, we constructed a comprehensive benchmarking suite. The gold standard for cell types in this study was primarily derived from publicly available and peer-reviewed large-scale cell atlas projects, including HuBMAP Azimuth, GTEx [22], the human cell landscape [23], and Tabula Sapiens [24]. To ensure fair cross-dataset evaluation, all data underwent rigorous standardization: marker gene lists were curated through differential expression analysis and literature review, with annotation consistency ensured by cross-referencing and adaptation of publicly available benchmarks such as GPTCelltype [3].

Gene expression matrices underwent standardized preprocessing, including pseudocount addition, log-transformation, batch effect correction, and high-variance gene selection, to ensure cross-dataset consistency. For spatial transcriptomics and oncological datasets, additional quality control procedures and multi-step data cleaning were implemented.

Cell type annotation protocol

All benchmarking and annotation were performed in the R environment to guarantee reproducibility and methodological transparency. State-of-the-art LLMs supporting OpenAI API interfaces were employed. Providers included OpenRouter (https://openrouter.ai), Aihubmix (https://aihubmix.com), and Deepseek. Because the official OpenAI R package does not yet support customizable base URLs—a limitation in certain experimental scenarios—we utilized the httr2 library for flexible API access, allowing precise endpoint selection and model specification. Throughout our main benchmarks, model inferences were standardized by using the Aihubmix API, ensuring uniform workflows and reproducible results.

Marker gene features for each cell population were derived using a standardized Seurat pipeline. Specifically, the top 10 differentially expressed genes per cluster were selected using FindAllMarkers (Wilcoxon rank-sum test, logFC > 0.25, adj. $P < .01$). This balances informativeness and noise tolerance.

The resulting marker genes were embedded into a standardized prompt template designed for robust JSON output parsing. The prompt was:

> *Identify cell types for '[tissue_name]' tissue based on the following marker genes. Each row represents a distinct cell population: '[marker_gene_list]'. Provide the output strictly as a JSON object with one predicted cell type per row, without additional commentary.*

Here, "tissue_name" refers to the species and tissue origin, and each "marker_gene_list" contains genes characterizing a cell population. All prompts were delivered under consistent system-level instructions to eliminate variability in formatting or verbosity across models. During evaluation, the temperature was fixed at 0.3, and top-k sampling was left unspecified.

Model performance evaluation

For performance evaluation, we implemented a structured scoring scheme that leverages the hierarchical organization of the cell ontology (CL). Both predicted and reference annotations were first normalized through an ontology-mapping pipeline that harmonizes synonyms, aliases, and lineage labels to their corresponding CL terms. Exact matches (1.0 point) were assigned when the normalized terms were identical. Partial matches (0.5 points) were assigned when the predicted label mapped to the direct parent class of the gold-standard term within the CL hierarchy. All other predictions were counted as mismatches (0 points). Although conceptually simple, this scoring system is tightly integrated with our broader methodological framework—including ontology-aware term normalization, model-specific prompt engineering, and LLM-adaptation strategies—which together enable a reproducible and biologically consistent evaluation across heterogeneous datasets.

Large language model cost analysis

To estimate operational costs associated with different LLMs in AICellType, we obtained per-token pricing for input ("prompt") and output ("completion") tokens from official API providers. For each annotation request, we recorded the number of input tokens ($\mathbf{t}_i$) and output tokens ($\mathbf{t_o}$). The cost per request was computed as:

$$\text{Cost} = (\mathbf{t}_i \times \mathbf{P}_i) + (\mathbf{t_o} \times \mathbf{P_o})$$

where $\mathbf{P}_i$ and $\mathbf{P_o}$ correspond to the per-token USD price for input and output, respectively, calculated based on official per-million token rates. Full pricing details and calculation tables are provided in our code repository and Supplementary Table S1.

Robustness and stability analyses

To evaluate the robustness of AICellType against noisy and incomplete marker input, we conducted two series of controlled perturbation

experiments. First, random genes from the human genome (GCA_000001405.29) were inserted into the marker lists at proportions of 20%, 40%, 60%, 80%, and 100%, and model performance rescored using the established framework. Second, we performed marker gene dropout tests by randomly removing 20%, 40%, 60%, or 80% of marker genes from each list. All results were visualized using ggplot2 to facilitate interpretation. Together, these experiments probed different facets of model robustness to data perturbation and incompleteness. We additionally validated model stability under varying marker gene counts (Top 5 – Top 20), confirming consistent annotation performance across feature selection ranges.

Temperature parameter optimization

To determine the optimal model temperature for maximizing stability and consistency, we subjected the LLMs to systematic testing across a range of temperature values (0.1 to 1.0, in 0.1 increments). For each temperature setting, annotation was repeated five times for each input, with all other parameters held constant. Model outputs were scored as previously described; mean and standard deviation of concordance scores were computed per temperature, and inter-replicate label similarity was used as a measure of output stability. All runs were performed in identical computational environments. The results informed default temperature selection in downstream AICellType modules.

Case study

We conducted a systematic benchmark comparison of AICellType against leading cell type annotation tools, including scType and CellMarker 2.0, using two widely adopted datasets: the Seurat PBMC 3k single-cell RNA-seq dataset and the 10x Visium mouse brain spatial transcriptomics dataset. Human single-cell RNA-seq data were obtained from GEO (accession GSE305979, published on 23 October 2025; not seen by the model during training). To reduce computational complexity, 40% of cells were randomly subsampled, followed by standard processing in Seurat (normalization, dimensionality reduction, and clustering). The top 10 marker genes per cluster were identified via Wilcoxon rank-sum tests and used for cell type prediction with AICellType under default parameters. Porcine single-cell datasets comprised raw 10x Genomics spleen samples, sequenced on the Illumina NovaSeq 6000 platform at an average depth of ∼36 500 reads per cell, and processed through FastQC quality assessment, Trimmomatic trimming, and Cell Ranger v7.1.0 for demultiplexing, alignment, and unique molecular identifier (UMI) counting; high-quality cells were retained based on 400–7000 detected genes, ≤50 000 UMIs, and ≤20% mitochondrial RNA, after which the expression matrices were merged and clustered in Seurat and yielded 30 clusters, each annotated using the top 10 Wilcoxon-selected marker genes.

Comparator methods included: SingleR (v2.10), utilizing reference profiles from NovershternHematopoieticData and HumanPrimaryCellAtlasData (HPCA) in the celldex package, assigning cell types via similarity matching; scType (v1.0), applied to the same expression matrices under default settings; scAnnotate (v0.3) [25], trained on HPCA from celldex, intersecting genes between query and reference, then predicting cell types via scAnnotate(); SciBet (v1.0) [26], employing the official "major human cell types" reference model via pro.core() and LoadModel(); Azimuth (v0.5.0) [27], mapping query data to the PBMC reference space using anchor-based matching; CellID (v1.19.0), extracting human blood-specific cell type markers from the PanglaoDB (PanglaoDB_markers_27_March_2020) database [28]; and CellMarker 2.0, serving as an online standardized marker gene repository, providing curated marker sets or Wilcoxon-selected top markers for annotation.

Key Points

- A systematic benchmark of 79 large language models was conducted on 1130 single-cell and spatial transcriptomics datasets.
- Claude 3.5 Sonnet demonstrated superior comprehensive performance, achieving a weighted accuracy of 76% and outperforming both conventional methods and other LLMs.
- AICellType, an open-source R package and web platform, was developed to integrate seamlessly with Seurat workflows.
- The platform enables scalable, efficient, and accessible cell type annotation for diverse single-cell and spatial omics applications.

Author contributions

Chuxing Cheng (conceived and designed the study, implemented the core computational framework, and wrote the manuscript), Shuo Fang (developed the web platform and R package, and participated in the implementation of experimental code), Xiaokun Liu (reviewed the manuscript and provided guidance on data visualization), Qi Zuo (reviewed the manuscript and provided guidance on data visualization), Jiahui Sun (reviewed the manuscript and provided guidance on data visualization), Xiaotong Hu (reviewed the manuscript and provided guidance on data visualization) and Meilin Jin (supervised the study and reviewed the manuscript)

Supplementary material

Supplementary material is available at *Briefings in Bioinformatics* online.

Funding

This work was supported by the Wuhan Science and Technology Plan Project "Creation of Novel African Swine Fever Vaccine" (Project No. 2023020302020573).

Conflict of interest

The authors declare that they have no competing interests.

Data availability

The AICellType package is open source and freely available at https://github.com/mooercccx/AICellType. All scripts and datasets required to reproduce the analyses presented in this study are provided in the GitHub repository. The web-based annotation service can be accessed at https://AICellType.jinlab.online/. Analyses were performed using R version 4.4.3.

References

1. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nat Rev Genet* 2019;**20**:631–56. https://doi.org/10.1038/s41576-019-0150-2

2. Lähnemann D, Köster J, Szczurek E *et al.* Eleven grand challenges in single-cell data science. *Genome Biol* 2020;**21**:31. https://doi.org/10.1186/s13059-020-1926-6
3. Hou W, Ji Z. Assessing GPT-4 for cell type annotation in single-cell RNA-seq analysis. *Nat Methods* 2024;**21**:1462–5. https://doi.org/10.1038/s41592-024-02235-4
4. Zeisel A, Muñoz-Manchado AB, Codeluppi S *et al.* Brain structureCell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 2015;**347**:1138–42. https://doi.org/10.1126/science.aaa1934
5. Chen Y, Zhang S. Automatic cell type annotation using marker genes for single-cell RNA sequencing data. *Biomolecules* 2022;**12**:12. https://doi.org/10.3390/biom12101539
6. Levine D, Lévy S, Rizvi S *et al.* Cell2Sentence: teaching large language models the language of biology. *PMLR* 2024;**235**:27299–325.
7. Tan R, Cui H, Wang B *et al.* Evaluating and interpreting scGPT: a foundation model for single-cell biology in real-world cancer clinical trial data. *Cancer Res* 2024;**84**:A029-A029. https://doi.org/10.1158/1538-7445.PANCREATIC24-A029
8. Chen Y, Zou J. Simple and effective embedding model for single-cell biology built from ChatGPT. *Nature Biomedical Engineering* 2025;**9**:483–93.
9. Wei W, Xia X, Li T *et al.* Shaoxia: a web-based interactive analysis platform for single cell RNA sequencing data. *BMC Genom* 2024;**25**:402. https://doi.org/10.1186/s12864-024-10322-1
10. Lyons A, Brown J, Davenport KM. Single-cell sequencing technology in ruminant livestock: challenges and opportunities. *Curr Issues Mol Biol* 2024;**46**:5291–306. https://doi.org/10.3390/cimb46060316
11. Chen J, Xu H, Tao W *et al.* Transformer for one stop interpretable cell type annotation. *Nat Commun* 2023;**14**:223. https://doi.org/10.1038/s41467-023-35923-4
12. Ianevski A, Giri AK, Aittokallio T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. *Nat Commun* 2022;**13**:1246. https://doi.org/10.1038/s41467-022-28803-w
13. Hu C, Li T, Xu Y *et al.* CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res* 2022;**51**:D870–6. https://doi.org/10.1093/nar/gkac947
14. Stuart T, Butler A, Hoffman P *et al.* Comprehensive integration of single-cell data. *Cell* 2019;**177**:1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031
15. DeMeo B, Nesbitt C, Miller SA *et al.* Active learning framework leveraging transcriptomics identifies modulators of disease phenotypes. *Science* **0**:eadi8577.
16. Lan W, Liu M, Chen J *et al.* JLONMFSC: clustering scRNA-seq data based on joint learning of non-negative matrix factorization and subspace clustering. *Methods* 2024;**222**:1–9. https://doi.org/10.1016/j.ymeth.2023.11.019
17. Lan W, Tang Z, Liu M *et al.* The large language models on biomedical data analysis: a survey. *IEEE J Biomed Health Inform* 2025;**29**:4486–97. https://doi.org/10.1109/JBHI.2025.3530794
18. Wei L, Guo-Hang H, Wei-Hao Z *et al.* scMCG: a method for Analyzing scATAC-seq data based on contrastive learning and generative adversarial network. *J Comput Sci Technol* 2025;**40**:145–62.
19. Lan W, Ling T, Chen Q *et al.* scMoMtF: an interpretable multitask learning framework for single-cell multi-omics data analysis. *PLoS Comput Biol* 2024;**20**:e1012679. https://doi.org/10.1371/journal.pcbi.1012679
20. Lan W, He G, Liu M *et al.* Transformer-based single-cell language model: a survey. *Big Data Mining Anal* 2024;**7**:1169–86. https://doi.org/10.26599/BDMA.2024.9020034
21. Wan H, Yuan M, Fu Y *et al.* Continually adapting pre-trained language model to universal annotation of single-cell RNA-seq data. *Brief Bioinform* 2024;**25**:25. https://doi.org/10.1093/bib/bbae047
22. Eraslan G, Drokhlyansky E, Anand S *et al.* Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. *Science* 2022;**376**:eabl4290. https://doi.org/10.1126/science.abl4290
23. Han X, Zhou Z, Fei L *et al.* Construction of a human cell landscape at single-cell level. *Nature* 2020;**581**:303–9. https://doi.org/10.1038/s41586-020-2157-4
24. The Tabula Sapiens Consortium, Jones RC, Karkanias J *et al.* The tabula sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. *Science* 2022;**376**:eabl4896. https://doi.org/10.1126/science.abl4896
25. Ji X, Tsao D, Bai K *et al.* scAnnotate: an automated cell-type annotation tool for single-cell RNA -sequencing data. *Bioinform Adv* **3**:vbad030.
26. Li C, Liu B, Kang B *et al.* SciBet as a portable and fast single cell type identifier. *Nat Commun* 2020;**11**:1818. https://doi.org/10.1038/s41467-020-15523-2
27. Hao Y, Hao S, Andersen-Nissen E *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021;**184**:3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048
28. Cortal A, Martignetti L, Six E *et al.* Gene signature extraction and cell identity recognition at the single-cell level with cell-ID. *Nat Biotechnol* 2021;**39**:1095–102. https://doi.org/10.1038/s41587-021-00896-6

Briefings in Bioinformatics, 2026, **27**, bbag151
https://doi.org/10.1093/bib/bbag151
Problem Solving Protocol