Learning Class-Consistent Metacell Prototypes via Quantization for Few-Shot Single-Cell Annotation

1st Yiwei Li
College of Computer Science
Sichuan University
Chengdu, China
woshili18wei@163.com

2nd Yongqing Zhang
School of Computer Science
Chengdu University of Information Technology
Chengdu, China
zhangyq@cuit.edu.cn

3th Zixuan Wang
College of Electronics and Information Engineering
Sichuan University
Chengdu, China
175422573@qq.com

4th Junlong Cheng
College of Computer Science
Sichuan University
Chengdu, China
cjl951015@163.com

5th Tianhao Li*
School of Computer Science
Chengdu University of Information Technology
Chengdu, China
1791027235@qq.com

6th Min Zhu*
College of Computer Science
Sichuan University
Chengdu, China
zhumin@scu.edu.cn

***Abstract*—Single-cell annotation plays a critical role in developmental biology, disease research, and precision medicine. However, most existing approaches rely on large-scale labeled datasets for supervised training and exhibit limited generalization in few-shot settings due to the scarcity of rare cell samples, sequencing-induced technical noise, pronounced intra-class representation bias, and weak inter-class separability. To address these challenges, we propose COMET, a few-shot single-cell annotation method based on class-consistent single-cell quantization. By introducing class-consistency constraints into the quantization process, COMET maintains mutually decoupled and structured latent code domains for each cell type, discretizes continuous single-cell representations in latent space, and constructs high-purity metacell prototypes through representation purification. This process effectively compresses redundant intra-class variation while amplifying key discriminative features, endowing metacells with stronger transferability and prior representational capacity under few-shot learning scenarios. Extensive experiments on eight real-world single-cell transcriptomic datasets demonstrate that COMET significantly outperforms non-quantized few-shot annotation baselines under limited supervision, particularly in rare cell recognition tasks. Moreover, the purified metacell representations exhibit superior structural consistency in clustering evaluations compared with raw single-cell representations, validating the robustness and reliability of COMET in complex single-cell analysis settings. Code is available in https://github.com/lilyway18/COMET.**

***Index Terms*—Few-shot cell annotation, Class-consistent learning, Prototype quantization, Metacell representation, Label-efficient annotation**

I. Introduction

Single-cell annotation aims to assign explicit cell-type labels to individual cells based on gene expression profiles, providing critical insights into cellular heterogeneity, functional states, and biological processes. Accurate annotation enables the identification of cell types in complex tissues and facilitates advances in developmental biology, disease research, and precision medicine [1], [2].

Most existing annotation methods rely on large-scale labeled reference datasets and perform prediction through supervised learning or reference mapping strategies [3], [4]. However, in practical scenarios, newly generated single-cell data often contain only a very limited number of labeled samples and are affected by technical noise, batch effects, and the high-dimensional sparsity of gene expression data [5]–[8]. Under such few-shot conditions, learned representations tend to be unstable and sensitive to data perturbations, resulting in significant fluctuations in class structures and degraded annotation performance. Therefore, developing robust learning strategies that can generalize effectively under extremely limited supervision is essential for reliable single-cell annotation.

To alleviate the impact of noise and instability, metacell techniques aggregate multiple similar single cells into higher-level representational units, thereby reducing stochastic variation while preserving key biological structures. Existing methods typically construct metacells based on expression similarity using clustering or graph-based partitioning approaches [9]–[11]. While effective in noise reduction, these approaches ignore class information during aggregation, which may introduce heterogeneous cell mixtures within metacells. Under few-shot settings, such impurity can propagate to downstream tasks and weaken the discriminative power of metacell representations [12], [13].

To address these challenges, we propose COMET, a novel few-shot single-cell annotation framework based on class-consistent metacell quantization. The core idea is to incorporate class consistency into the quantization process, enabling the construction of more structured and semantically coherent metacell representations. Specifically, COMET consists of two

This work was supported by the National Natural Science Foundation of China under Grant Nos. 62272067, and by the Sichuan Science and Technology Program under Grant No. 2026NSFSC0419;
* Corresponding author: Tianhao Li and Min Zhu.

key components. (1) A class-consistent quantization module, which introduces class-specific constraints during latent space learning to improve the structural consistency and purity of metacells. (2) A metacell prototype learning module, which aggregates quantized representations within each class to form a compact set of metacell prototypes, thereby reducing noise and enhancing representation stability. By shifting the learning granularity from single cells to metacells, COMET effectively mitigates intra-class variability and improves generalization under few-shot conditions. In addition, the framework is designed to be plug-and-play and can be seamlessly integrated into existing annotation models.

The main contributions of this work are summarized as follows:

- We propose a class-consistent quantization mechanism that explicitly incorporates class information into metacell construction, improving representation purity and structural consistency.
- We develop a metacell-level few-shot annotation framework that suppresses noise and intra-class variability, leading to more stable and generalizable representations.
- Extensive experiments demonstrate that the proposed method can be readily integrated into existing annotation pipelines and consistently improves performance in both rare-cell recognition and few-shot learning scenarios.

II. Method

A. Problem Formulation

Given a single-cell dataset $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^{N}$, where only a small subset is labeled, we split the data into a labeled support set $\mathcal{S}$ and an unlabeled query set $\mathcal{Q}$. Each cell $\mathbf{x}_i \in \mathbb{R}^D$ is mapped into a latent representation $\mathbf{z}_i \in \mathbb{R}^d$ via an encoder $f_\theta(\cdot)$.

To improve representation quality under limited supervision, COMET performs class-consistent quantization in the latent space. Specifically, for each class c, we maintain a set of metacell prototypes $\mathcal{M}_c$, and assign each cell embedding to its nearest prototype, yielding compact metacell-level representations.

The model is trained in two stages. First, we learn the encoder and metacell prototypes by minimizing a representation loss $\mathcal{L}_{\text{meta}}$. Then, we perform few-shot classification on metacell prototypes using a classifier trained with $\mathcal{L}_{\text{cls}}$, and propagate the predicted labels back to individual cells.

B. Overview of COMET

We propose COMET, a few-shot single-cell annotation framework based on class-consistent metacell quantization. It consists of two modules: (1) a class-consistent quantization module and (2) a metacell prototype learning module.

The quantization module maps cells into a latent space and enforces class-specific codebooks, ensuring that representations are discretized within class-consistent domains. The prototype learning module then aggregates these discrete representations into a set of metacell prototypes.

This design shifts the learning granularity from individual cells to metacells, leading to more stable representations under limited supervision. The method assumes a closed-set setting with known cell types, while open-set scenarios are left for future work.

C. Class-Consistent Quantization

Conventional metacell construction methods typically aggregate cells solely based on expression similarity, which may introduce cross-class mixing and degrade metacell purity. To explicitly enforce category-level consistency during discretization, COMET introduces a class-consistent quantization mechanism that constrains representation learning within class-specific latent domains.

Given a normalized gene expression vector x_i, a shared encoder maps it into a continuous latent representation:

$$z_i = f_{\text{enc}}(x_i), \qquad z_i \in \mathbb{R}^d, \tag{1}$$

where the encoder parameters are shared across all cells to capture global transcriptional patterns.

To incorporate category information into the quantization process, COMET maintains a collection of class-specific latent codebooks:

$$\mathcal{E} = \{\mathcal{E}_c\}_{c\in\mathcal{Y}}, \qquad \mathcal{E}_c = \{e_{c,1}, e_{c,2}, \ldots, e_{c,K_c}\}, \tag{2}$$

where each codebook $\mathcal{E}_c$ is exclusively associated with cell type c and mutually decoupled from others.

For a labeled cell with class label $y_i = c$, the discrete latent representation is obtained via the proposed class-consistent quantization function:

$$\hat{z}_i = \mathcal{Q}_{\text{cc}}(z_i, c) = \arg\min_{e\in\mathcal{E}_c} \|z_i - e\|_2 , \tag{3}$$

which restricts quantization within the corresponding class-specific codebook.

The class-consistent quantization is optimized using a VQ-style objective:

$$\mathcal{L}_{\text{ccq}} = \|\text{sg}(z_i) - \hat{z}_i\|_2^2 + \beta \|z_i - \text{sg}(\hat{z}_i)\|_2^2 , \tag{4}$$

where $\text{sg}(\cdot)$ denotes the stop-gradient operator and β controls the commitment strength.

This enforces structured and class-consistent discrete representations.

D. Metacell Prototype Learning

To enable gradient propagation through the discrete quantization bottleneck, we adopt the straight-through estimator, which is defined as

$$\tilde{z}_i = z_i + \text{sg}(\hat{z}_i - z_i), \tag{5}$$

where $\text{sg}(\cdot)$ denotes the stop-gradient operator. This formulation ensures that the quantized discrete representation is used in the forward pass, while gradients are propagated through the continuous latent variable z_i during backpropagation.

A decoder reconstructs the input as $\hat{x}_i = f_{\text{dec}}(\tilde{z}_i)$, optimized by: $\hat{x}_i = f_{\text{dec}}(\tilde{z}_i)$, and the reconstruction objective is defined as

$$\mathcal{L}_{\text{rec}} = \|x_i - \hat{x}_i\|_2^2 . \tag{6}$$

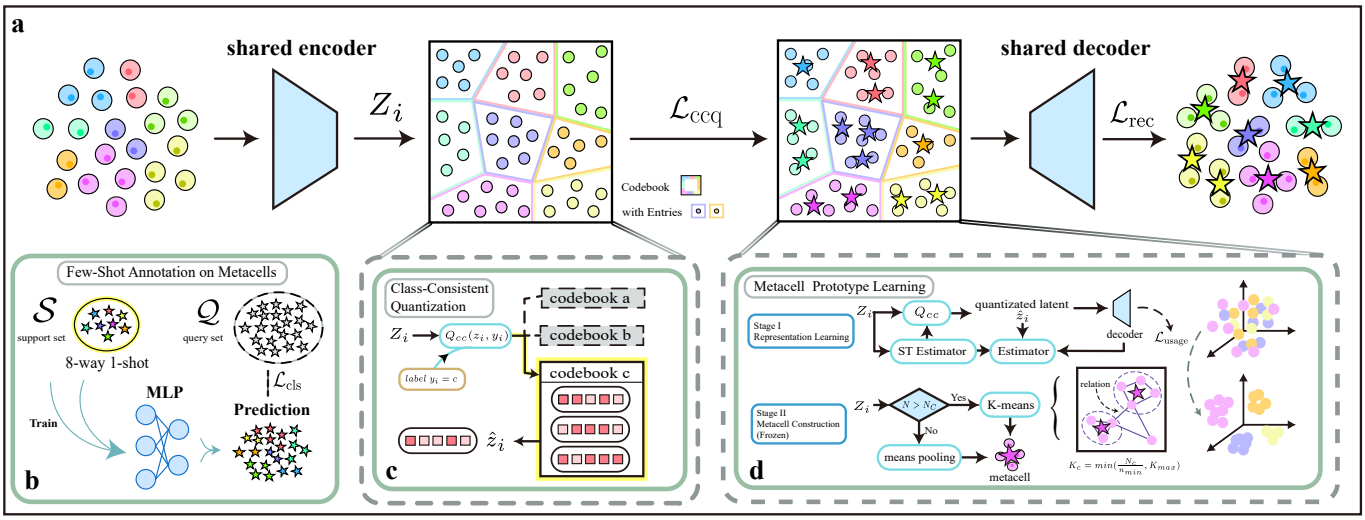

Fig. 1. (a) Pipeline from class-consistent quantization to metacell construction, illustrating how class-consistent constraints guide latent representation learning and subsequent metacell formation. (b) Few-shot annotation on metacells, where support and query sets are constructed at the metacell level for prediction. (c) Expanded view of class-consistent constraint, detailing how labeled information enforces consistency during the quantization and representation learning process. (d) Expanded view of metacell construction (metacell-level prototype learning), showing how original cells are aggregated to learn representative metacell prototypes.

To prevent codebook collapse, we introduce a usage regularization:

$$\mathcal{L}_{\text{usage}} = \sum_{c\in\mathcal{Y}} \sum_{k} p_{c,k} \log p_{c,k}, \tag{7}$$

where $p_{c,k}$ denotes the empirical usage frequency of the k-th code vector $e_{c,k}$ within class c during training. This regularization term encourages balanced utilization of code vectors inside each class, thereby promoting diversity and stability of the learned latent representations.

Overall, the training objective of COMET in the representation learning stage is formulated as

$$\min_{\theta} \lambda_1 \mathcal{L}_{\text{rec}} + \lambda_2 \mathcal{L}_{\text{ccq}} + \lambda_3 \mathcal{L}_{\text{usage}}, \tag{8}$$

where $\mathcal{L}_{\text{ccq}}$ denotes the class-consistent VQ-style quantization loss, and λ_1, λ_2, and λ_3 control the relative contributions of the reconstruction loss, the quantization loss, and the code usage regularization, respectively.

After the VQ-VAE training converges, the encoder and the class-specific codebooks are fixed and used to generate the final discrete latent representations for all single cells:

$$\hat{z}_i = \mathcal{Q}_{cc}(f_{\text{enc}}(x_i), y_i)\,, \tag{9}$$

where $Q_{cc}(\cdot)$ denotes the class-conditional quantization operator that maps a continuous latent embedding to its nearest code vector within the codebook corresponding to class y_i.

Based on the obtained discrete latent representations, COMET constructs metacell prototypes in the quantized latent space.

For each cell type c, let $\{\hat{z}_i^{(c)}\}_{i=1}^{N_c}$ denote the set of quantized latent representations belonging to this class. When the number of cells in class c is small, a single metacell prototype is constructed by simple averaging:

$$m_c = \frac{1}{N_c} \sum_{i=1}^{N_c} \hat{z}_i^{(c)}. \tag{10}$$

When N_c is large, in order to further capture intra-class heterogeneity, K-means clustering is performed within the class-specific discrete latent space to refine metacell construction. The number of metacells for class c is determined adaptively as

$$K_c = \min\left(\left\lfloor \frac{N_c}{n_{\min}} \right\rfloor, K_{\max}\right), \tag{11}$$

where $n_{\min}$ denotes the minimum number of cells contained in each metacell, and $K_{\max}$ represents the maximum allowable number of metacells for a single class. Each cluster centroid is treated as a metacell prototype.

These metacell prototypes provide stable and compact representations for downstream few-shot annotation.

E. Few-Shot Annotation on Metacells

After completing class-consistent quantization and metacell prototype construction, COMET performs few-shot annotation at the metacell level. Query cells are first mapped into the latent space and quantized to obtain their corresponding metacell representations. Based on these metacell embeddings, an N-way K-shot classification task is constructed. For each cell type, a subset of labeled metacell prototypes is randomly selected as the support set, while the remaining metacells form the query set. A lightweight supervised classification head $g_\phi(\cdot)$, implemented as a simple multilayer perceptron (MLP), is trained on the support set, which outputs class probabilities via a softmax layer. The classifier is optimized using the standard cross-entropy loss:

$$\mathcal{L}_{\text{cls}} = -\frac{1}{|\mathcal{M}_{\mathcal{S}}|} \sum_{(\mathbf{m}_i, y_i)\in\mathcal{M}_{\mathcal{S}}} \log\big(p_\phi(y_i \mid \mathbf{m}_i)\big), \tag{12}$$

where $\mathcal{M}_{\mathcal{S}}$ denotes the set of labeled metacell prototypes in the support set and $p_\phi(y_i \mid \mathbf{m}_i)$ is the predicted probability of class y_i given by the classifier.

TABLE I

COMPARISON OF FEW-SHOT CELL TYPE ANNOTATION METHODS ACROSS MULTIPLE SCRNA-SEQ DATASETS. "QUANT." INDICATES WHETHER A METHOD INCLUDES SINGLE-CELL QUANTIFICATION INTO METACELL. N-WAY-K-SHOT DENOTES A SETTING WHERE N IS THE NUMBER OF CELL TYPES AND K IS THE NUMBER OF LABELED SAMPLES PER TYPE. ACC↑ AND F1↑ REPRESENT ACCURACY AND F1-SCORE, RESPECTIVELY, WHERE ↑ INDICATES THAT HIGHER VALUES ARE BETTER. ALL METRICS ARE REPORTED IN PERCENTAGE (%).

Methods	Quant.	N-Way-1-Shot		N-Way-3-Shot		N-Way-5-Shot		Quant.	N-Way-1-Shot		N-Way-3-Shot		N-Way-5-Shot	
		ACC↑	F1↑	ACC↑	F1↑	ACC↑	F1↑		ACC↑	F1↑	ACC↑	F1↑	ACC↑	F1↑
		Zeisel [14], *N=39*							***He*** [15], *N=14*					
scPred [16]	✓	48.92	56.91	70.13	72.63	78.01	75.32	✓	**71.26**	69.75	75.31	**72.67**	**86.15**	74.52
	✗	65.27	58.89	86.26	79.75	89.05	83.57	✗	**53.83**	**64.08**	58.22	**72.68**	66.52	77.36
scTransSort [17]	✓	40.35	37.78	**72.69**	69.01	79.74	75.61	✓	56.25	52.35	72.78	63.47	83.18	78.30
	✗	63.23	33.19	80.49	40.19	80.12	41.10	✗	45.29	48.80	50.69	53.96	60.87	62.81
CIForm [8]	✓	35.07	34.30	55.83	52.75	53.25	46.52	✓	60.35	53.60	59.96	55.63	78.38	69.68
	✗	36.40	27.21	**87.79**	49.75	32.32	28.87	✗	48.24	53.55	55.68	61.45	65.03	57.04
scKAN [18]	✓	38.03	45.42	67.69	68.92	62.06	58.23	✓	56.11	54.97	68.97	70.49	79.95	69.57
	✗	41.84	40.25	82.47	76.66	72.42	69.87	✗	48.52	57.44	57.61	70.37	**70.06**	**79.50**
Our	✓	**55.70**	**60.51**	71.73	**73.46**	**84.96**	**81.75**	✓	69.30	**70.90**	**75.44**	70.38	83.52	**84.62**
	✗	**68.38**	**59.07**	86.06	**80.38**	**89.93**	**84.77**	✗	48.08	58.30	**60.08**	70.78	67.65	77.81
		Madisson [19], *N=25*							***Zheng68k*** [20], *N=11*					
scPred [16]	✓	46.97	**51.22**	72.58	65.94	**79.40**	**77.54**	✓	34.07	33.53	47.23	**45.92**	52.56	48.78
	✗	**61.78**	**53.02**	67.34	60.97	**74.64**	68.04	✗	29.65	31.75	43.90	**44.79**	**44.01**	45.03
scTransSort [17]	✓	23.30	21.69	72.62	68.37	77.68	73.42	✓	30.08	25.17	**47.78**	17.29	46.84	43.39
	✗	48.06	37.80	66.53	48.46	71.75	54.43	✗	36.68	32.78	**51.44**	44.74	43.35	41.05
CIForm [8]	✓	38.79	33.32	67.55	63.18	75.64	65.99	✓	32.42	29.56	40.12	37.61	52.08	44.59
	✗	52.90	43.26	54.85	41.62	71.26	61.09	✗	**36.40**	**32.91**	34.44	35.78	37.06	38.63
scKAN [18]	✓	36.29	39.09	66.23	56.58	74.60	69.41	✓	34.26	32.18	44.70	43.45	49.46	47.53
	✗	59.26	47.10	67.53	60.49	71.12	64.29	✗	27.02	30.08	28.45	25.82	40.75	43.88
Our	✓	**47.47**	48.49	**72.93**	**69.28**	78.47	75.26	✓	**39.04**	**37.88**	46.19	43.36	**53.84**	**51.11**
	✗	57.96	49.48	**72.20**	**64.37**	74.32	**68.59**	✗	30.59	30.29	42.70	42.22	**44.01**	**45.80**
		Stewart [21], *N=43*							***Vento*** [22], *N=32*					
scPred [16]	✓	52.52	52.45	73.55	70.92	78.81	73.95	✓	59.60	63.51	84.47	**83.19**	87.21	**89.02**
	✗	41.11	**42.00**	62.20	61.64	68.71	67.43	✗	64.45	**54.86**	76.31	**72.78**	80.41	**77.27**
scTransSort [17]	✓	45.33	39.26	**77.73**	64.33	84.02	**78.64**	✓	31.35	25.95	**84.85**	49.54	**89.36**	85.99
	✗	42.57	19.78	58.84	32.79	64.18	37.94	✗	61.41	15.78	72.14	17.32	75.26	19.20
CIForm [8]	✓	48.95	41.79	75.23	68.50	82.97	77.38	✓	51.54	32.97	77.22	73.64	84.04	55.12
	✗	**43.97**	27.50	51.78	31.45	**75.17**	56.70	✗	**65.23**	27.81	76.22	43.52	**81.28**	21.64
scKAN [18]	✓	41.90	43.64	68.36	65.29	81.58	77.54	✓	54.68	61.27	80.41	80.05	80.32	80.61
	✗	40.16	41.65	**63.11**	60.11	67.22	64.99	✗	62.39	55.44	**77.18**	72.41	79.61	76.31
Our	✓	**58.02**	**56.03**	74.80	**71.70**	**84.24**	77.99	✓	**61.12**	**64.86**	82.79	82.64	85.06	84.52
	✗	35.68	39.56	59.28	**61.95**	71.17	**70.09**	✗	60.45	55.31	76.69	72.54	80.16	77.02

III. EXPERIMENTS

A. Experiment Settings

1) Benchmark: We evaluate the proposed method on eight widely used single-cell transcriptomic datasets, including Zeisel [14], He [15], Zheng68k [20], Madisson [19], Stewart [21], and Vento [22]. These datasets cover diverse biological conditions, large-scale settings, and imbalanced cell-type distributions.

Additionally, two multimodal datasets, BMMC [23] and ISSAAC-seq [24], are used to assess cross-modal performance, providing paired RNA and chromatin accessibility data.

2) Evaluation System: Few-shot annotation is evaluated under the standard N-way K-shot protocol with 1-shot, 3-shot, and 5-shot settings. Experiments include performance comparison, ablation studies, rare cell identification, and analysis of metacell representation quality.

We compare COMET with four representative methods: scPred, scTransSort, CIForm, and scKAN. All methods are evaluated on both original single-cell representations and quantized metacell representations. Performance is measured using Accuracy and Macro-F1.

3) Implementation Details: Metacell representations are learned using a class-consistent VQ-VAE on standardized gene expression data. The dataset is randomly split into two subsets (50%/50%), where one subset is used to learn class-consistent codebooks, while few-shot supervision is only applied at the downstream metacell-level classifier.

The model is trained using Adam with a learning rate of 0.003 for up to 200 epochs, early stopping with a patience of 25 epochs, and a batch size of 512. All experiments are conducted on an NVIDIA RTX 4090 GPU.

B. Quantitative Comparison

Table I summarizes the few-shot performance across multiple datasets. COMET achieves the best or second-best results in most 1-shot, 3-shot, and 5-shot settings, consistently outperforming single-cell baselines without quantization.

On the Zheng68k dataset, metacell quantization yields stable improvements, with most methods gaining over 3% in both

TABLE II
ABLATION RESULTS OF COMET ON BMMC AND ISSAAC-SEQ. RESULTS ARE REPORTED UNDER N-WAY 1-SHOT AND 3-SHOT SETTINGS, WHERE N DENOTES THE NUMBER OF CELL TYPES IN EACH BENCHMARK. "QUANT." INDICATES VECTOR QUANTIZATION, "C-C" INDICATES CLASS-CONSISTENT PROTOTYPE LEARNING, AND "RP" DENOTES THE PROPORTION OF REFERENCE DATA USED FOR PROTOTYPE LEARNING. ACC AND MACRO-F1 ARE REPORTED IN PERCENTAGE (%).

Baseline	Quant.	C-C	RP	BMMC [23]		ISSAAC-seq [24]	
				ACC	F1	ACC	F1
N-Way-1-Shot				***N=22***		***N=22***	
VAE Latent (w/o Quant.)	✗	✗	–	45.75	36.02	24.21	21.91
VQ (w/Quant., w/o C-C)	✓	✗	50	38.73	27.34	33.94	30.34
COMET (w/RP-10)	✓	✓	10	48.13	46.84	33.30	26.19
COMET (w/RP-20)	✓	✓	20	49.66	48.68	34.05	35.19
COMET (w/RP-30)	✓	✓	30	52.31	52.65	38.85	**36.89**
COMET (w/RP-40)	✓	✓	40	56.07	51.21	41.38	36.29
COMET (w/RP-50)	✓	✓	50	**61.72**	**58.02**	**42.82**	36.36
N-Way-3-Shot				***N=22***		***N=22***	
VAE Latent (w/o Quant.)	✗	✗	–	59.03	57.21	48.84	46.05
VQ (w/Quant., w/o C-C)	✓	✗	50	49.12	43.69	33.54	26.37
COMET (w/RP-10)	✓	✓	10	63.39	63.43	42.52	40.34
COMET (w/RP-20)	✓	✓	20	64.64	66.48	56.23	48.98
COMET (w/RP-30)	✓	✓	30	64.71	65.71	56.26	42.38
COMET (w/RP-40)	✓	✓	40	66.96	65.93	60.01	49.37
COMET (w/RP-50)	✓	✓	50	**70.18**	**71.77**	**64.47**	**54.25**

Accuracy and Macro-F1, and scKAN improving by more than 10% under 3-shot.

In some cases, scPred matches or slightly surpasses COMET, likely due to its SVD-based feature representation. Notably, COMET employs a simple MLP classifier, indicating that the performance gains primarily stem from the proposed quantization module rather than classifier complexity.

Overall, these results demonstrate that COMET provides a general, plug-and-play solution that consistently improves few-shot annotation across diverse models and datasets.

C. Ablation Study

Table II reports the ablation results on vector quantization (Quant.), class-consistent prototype learning (C-C), and reference proportion (RP).

Removing quantization leads to significant performance degradation (e.g., ACC drops to 24.21% on ISSAAC-seq), indicating the instability of continuous representations under few-shot settings. Introducing quantization improves performance, while the full COMET model further achieves the best results, demonstrating the effectiveness of combining quantization with class-consistent prototype learning.

The reference proportion (RP) also impacts performance. Increasing RP improves accuracy in low-shot settings, while results become stable when RP $\geq$ 20, indicating robustness to the scale of reference data.

Overall, these results confirm that class-consistent metacell prototypes effectively reduce noise and improve representation quality in few-shot annotation.

D. Rare Cell Identification

Our method shows clear advantages in identifying rare cell types (<2%). As shown in Table III, under 1-shot conditions,

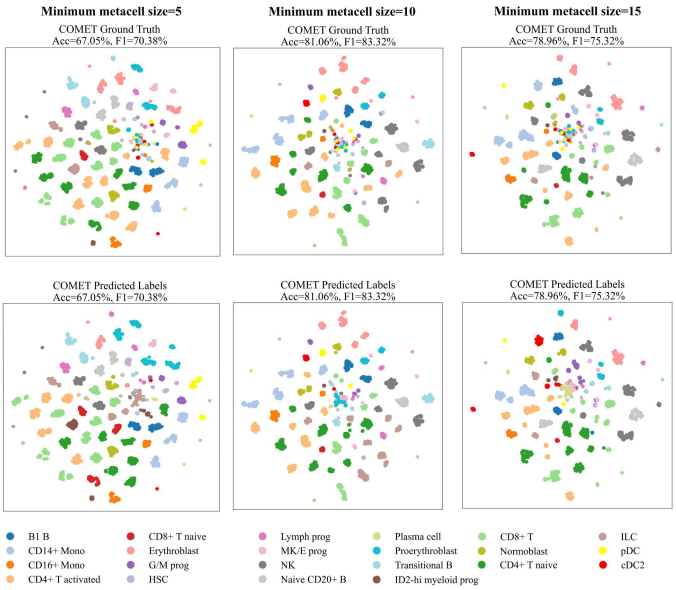

Fig. 2. Visualizations using t-SNE on the BMMC dataset under different quantization scales, where rows denote minimum metacell sizes of 5, 10, and 15 cells, and columns show original metacell representations (top) and predicted results (bottom). Intermediate scales yield more compact and well-separated clusters with higher prediction accuracy, whereas overly coarse or fine scales reduce visual clarity and degrade performance.

COMET achieves the highest accuracy in 5/9 rare types in BMMC and 1/2 in ISSAAC-seq.

For extremely rare populations such as ID2-HI MYELOID PROG (0.37%) and MK/E PROG (1.34%), COMET outperforms the second-best method by 77.78% and 56.91%, respectively. These results demonstrate that our metacell-based, class-consistent prototype learning effectively aggregates sparse signals and constructs stable, high-purity expression prototypes, enabling robust and accurate identification under few-shot conditions.

E. Quantization Scale Analysis

Finally, we investigate the impact of different quantization scale settings on downstream few-shot annotation performance in BMMC dataset. By varying the number of metacells per class, we construct three quantization configurations with different granularities and conduct few-shot annotation and visualization analyses in the corresponding metacell spaces. As shown in Fig.2, an overly coarse quantization scale leads to excessive aggregation of distinct cell subtypes, resulting in the loss of discriminative information, whereas an overly fine quantization scale produces an excessive number of metacells, thereby weakening the noise suppression effect. In contrast, an intermediate quantization scale achieves the best trade-off in terms of visualization structure as well as ACC and F1 scores, indicating that an appropriate quantization granularity is crucial for fully exploiting the advantages of metacells in few-shot learning.

IV. CONCLUSION

COMET achieves robust few-shot single-cell annotation via class-consistent metacell quantization, reducing single-cell

TABLE III

EVALUATION OF RARE CELL TYPE CLASSIFICATION UNDER FEW-SHOT SETTINGS. RARE CELL TYPES ACCOUNT FOR LESS THAN 2% OF ALL CELLS. RESULTS ARE REPORTED UNDER 1-SHOT AND 3-SHOT SETTINGS AT METACELL AND SINGLE-CELL SCALES. OUR METHOD USES METACELL-QUANTIZED FEW-SHOT ANNOTATION, WHILE BASELINES OPERATE AT THE SINGLE-CELL LEVEL. "ALL" DENOTES THE AVERAGE OVER RARE CELL TYPES. BEST AND SECOND-BEST RESULTS IN ALL ARE SHOWN IN BOLD AND UNDERLINE, RESPECTIVELY. ACC (%) IS USED FOR EVALUATION.

Benchmark	Rare Cell Types	Rate (≤2.0%)	*N-Way-1-Shot*						*N-Way-3-Shot*					
			COMET	CIForm	scTransSort	scKAN	scPred	w/o Quant.	COMET	CIForm	scTransSort	scKAN	scPred	w/o Quant.
BMMC	CD8+ T NAIVE	0.68	**100.00**	98.49	74.37	93.97	93.97	87.94	**100.00**	96.98	90.45	**100.00**	**100.00**	98.99
	CDC2	1.17	71.43	77.33	61.92	**83.43**	55.52	4.65	**82.35**	78.49	82.27	79.94	50.29	64.53
	G/M PROG	1.25	72.34	60.87	41.03	**78.80**	40.49	50.82	71.05	**79.35**	55.43	71.74	73.91	67.39
	HSC	0.97	**54.29**	44.95	38.68	19.86	31.71	52.26	**82.14**	47.39	59.58	45.64	48.78	41.81
	ID2-HI MYELOID PROG	0.37	**88.89**	11.11	9.26	8.33	14.81	42.59	**88.89**	**88.89**	62.96	51.85	61.11	58.33
	LYMPH PROG	1.81	45.28	27.56	35.57	42.64	55.12	**56.05**	64.15	68.72	52.70	13.59	29.98	**88.27**
	MK/E PROG	1.34	**63.27**	6.36	6.11	0.76	1.27	2.29	**63.16**	43.26	20.36	17.05	27.48	49.87
	PDC	1.42	81.40	71.09	**87.68**	52.84	82.46	56.87	81.40	78.67	81.04	86.49	**86.97**	75.12
	PLASMA CELL	1.02	**92.11**	13.71	57.86	41.81	33.44	32.78	90.32	82.61	68.90	**91.30**	86.62	47.49
	ALL	10.03	**74.33**	45.72	45.83	46.94	45.42	42.92	**80.38**	73.82	63.74	61.96	62.79	65.76
ISSAAC-seq	A18	1.59	57.14	31.93	27.11	59.64	**66.27**	45.78	78.79	57.23	36.14	**80.01**	15.66	74.70
	A19	1.59	**44.44**	26.22	42.68	29.27	12.20	9.76	**53.33**	31.71	28.05	21.95	21.34	33.33
	A20	1.30	**82.35**	55.56	70.37	11.11	40.74	47.41	66.67	77.04	80.00	77.04	**89.63**	84.44
	A21	0.87	81.82	**98.91**	94.57	96.74	95.65	97.83	94.44	**100.00**	94.57	84.78	94.57	77.78
	ALL	5.35	**66.44**	53.16	58.68	49.19	53.72	50.20	**73.31**	66.50	59.69	65.95	55.30	67.56

noise and improving rare cell recognition. Further analysis suggests that the method has the potential to extend to more extreme weakly supervised settings, although this remains to be systematically validated. We emphasize that COMET is designed as an augmentation to existing metacell construction methods rather than a standalone alternative, with experiments focusing on isolating the effect of high-purity metacell representations on few-shot annotation. The current study operates under a closed-set few-shot assumption and does not explicitly address open-set recognition or unknown cell-type discovery.

Overall, COMET provides an effective approach for few-shot single-cell annotation and demonstrates stable performance under weakly supervised and label-scarce conditions.

REFERENCES

[1] M. D. Luecken and F. J. Theis, "Current best practices in single-cell rna-seq analysis: a tutorial," *Molecular Systems Biology*, vol. 15, no. 6, p. e8746, 2019.

[2] A. Regev, S. A. Teichmann, E. S. Lander *et al.*, "The human cell atlas," *eLife*, vol. 6, p. e27041, 2017.

[3] V. Y. Kiselev, A. Yiu, and M. Hemberg, "scmap: projection of single-cell rna-seq data across data sets," *Nature Methods*, vol. 15, no. 5, pp. 359–362, 2018.

[4] H. A. Pliner, J. Shendure, and C. Trapnell, "Supervised classification enables rapid annotation of cell atlases," *Nature Methods*, vol. 16, no. 10, pp. 983–986, 2019.

[5] C. Jia, Y. Hu, D. Kelly, J. Kim, M. Li, and N. R. Zhang, "Accounting for technical noise in differential expression analysis of single-cell rna sequencing data," *Nucleic Acids Research*, vol. 45, no. 19, pp. 10 978–10 988, 2017.

[6] L. Haghverdi, A. T. L. Lun, M. D. Morgan *et al.*, "Batch effects in single-cell rna-sequencing data are corrected by matching mutual nearest neighbors," *Nature Biotechnology*, vol. 36, no. 5, pp. 421–427, 2018.

[7] R. Lopez, J. Regier, M. B. Cole *et al.*, "Deep generative modeling for single-cell transcriptomics," *Nature Methods*, vol. 15, no. 12, pp. 1053–1058, 2018.

[8] J. Xu, A. Zhang, F. Liu, L. Chen, and X. Zhang, "Ciform as a transformer-based model for cell-type annotation of large-scale single-cell rna-seq data," *Briefings in Bioinformatics*, vol. 24, no. 4, p. bbad195, 2023.

[9] Y. Li, H. Li, Y. Lin *et al.*, "Metaq: fast, scalable and accurate metacell inference via single-cell quantization," *Nature Communications*, vol. 16, no. 1, p. 1205, 2025.

[10] F. Zeng and J. Han, "Building a single cell transcriptome-based coordinate system for cell id with sure," *bioRxiv*, p. 2024.11.13.623403, 2024, preprint.

[11] M. Bilous, L. Tran, C. Cianciaruso *et al.*, "Metacells untangle large and complex single-cell transcriptome networks," *BMC Bioinformatics*, vol. 23, no. 1, p. 336, 2022.

[12] S. Persad, Z.-N. Choo, C. Dien *et al.*, "Seacells infers transcriptional and epigenomic cellular states from single-cell genomics data," *Nature Biotechnology*, vol. 41, no. 12, pp. 1746–1757, 2023.

[13] P. Liu and J. J. Li, "mcrigor: a statistical method to enhance the rigor of metacell partitioning in single-cell data analysis," *Nature Communications*, vol. 16, no. 1, p. 8602, 2025.

[14] A. Zeisel, H. Hochgerner, P. Lönnerberg, A. Johnsson, F. Memic, J. Van Der Zwan *et al.*, "Molecular architecture of the mouse nervous system," *Cell*, vol. 174, no. 4, pp. 999–1014, 2018.

[15] J. He, J. Yan, J. Wang, L. Zhao, Q. Xin, Y. Zeng *et al.*, "Dissecting human embryonic skeletal stem cell ontogeny by single-cell transcriptomic and functional analyses," *Cell Research*, vol. 31, no. 7, pp. 742–757, 2021.

[16] J. Alquicira-Hernandez, A. Sathe, H. P. Ji, Q. Nguyen, and J. E. Powell, "scpred: accurate supervised method for cell-type classification from single-cell rna-seq data," *Genome Biology*, vol. 20, no. 1, p. 264, 2019.

[17] L. Jiao, G. Wang, H. Dai, X. Li, S. Wang, and T. Song, "sctranssort: Transformers for intelligent annotation of cell types by gene embeddings," *Biomolecules*, vol. 13, no. 4, p. 611, 2023.

[18] H. He, Z. Tang, G. Chen, Z. Shen, Y. Yang, C. Yang *et al.*, "sckan: interpretable single-cell analysis for cell-type-specific gene discovery and drug repurposing via kolmogorov-arnold networks," *Genome Biology*, vol. 26, no. 1, p. 300, 2025.

[19] E. Madissoon, A. Wilbrey-Clark, R. J. Miragaia, K. Saeb-Parsy, K. T. Mahbubani, N. Georgakopoulos *et al.*, "scrna-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation," *Genome Biology*, vol. 21, no. 1, p. 1, 2019.

[20] G. X. Zheng, J. M. Terry, P. Belgrader, P. Ryvkin, Z. W. Bent, R. Wilson *et al.*, "Massively parallel digital transcriptional profiling of single cells," *Nature Communications*, vol. 8, no. 1, p. 14049, 2017.

[21] B. J. Stewart, J. R. Ferdinand, M. D. Young, T. J. Mitchell, K. W. Loudon, A. M. Riding *et al.*, "Spatiotemporal immune zonation of the human kidney," *Science*, vol. 365, no. 6460, pp. 1461–1466, 2019.

[22] R. Vento-Tormo, M. Efremova, R. A. Botting, M. Y. Turco, M. Vento-Tormo, K. B. Meyer *et al.*, "Single-cell reconstruction of the early maternal–fetal interface in humans," *Nature*, vol. 563, no. 7731, pp. 347–353, 2018.

[23] D. B. Burkhardt, M. D. Lücken, C. Lance, R. Cannoodt, A. O. Pisco, S. Krishnaswamy *et al.*, "Single-cell multiomics bone marrow mononuclear cells dataset for neurips 2021 multimodal data integration challenge," Dataset, GEO accession GSE194122, 2022.

[24] W. Xu, Y. Hu, Y. Zhang, P. M. Schnepp, L. M. Lo, Q. Zhang *et al.*, "Issaac-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells," *Nature Methods*, vol. 19, pp. 1243–1249, 2022.