

单细胞与空间转录组学 Reference Mapping 方法回顾与比较 (2016–2026)

执行摘要

单细胞与空间转录组学的 **reference mapping** (将“query”数据投影到既有“reference/atlas”并继承其结构与注释) 在 2016–2026 的方法学演进呈现出清晰的主线：从早期以线性对齐与近邻匹配为核心的“整合—再注释”流程，逐步走向 **可复用的压缩参考 (reference compression)**、**生成式潜空间 (probabilistic latent space)**、**空间任务特化 (deconvolution vs. single-cell localization vs. gene imputation)**，并在 2023–2026 明显强化了 **开放集/层级标签、置信度与不确定性、以及可共享的预训练模型仓库** 等工程化与统计严谨性要素。¹

从任务角度看，reference mapping 至少包含三类输入输出范式：

- 1) **scRNA/scMultiome → scRNA atlas**：核心是低维表示对齐 (批次校正/整合) + 标签转移 (label transfer)。代表性的复用式映射体系包括基于 anchor 的映射 (如 Seurat 的 FindTransferAnchors/TransferData 思路)、基于参考压缩的秒级映射 (如 Symphony)，以及基于条件 VAE 的“architecture surgery/冻结—增量适配” (如 scArches、scPoli)。²
- 2) **scRNA → ST (空间)**：可再细分为 (a) **spot/capture location 的细胞类型比例推断 (deconvolution)**，以及 (b) **单细胞定位到空间坐标 (localization/mapping)**。这两类任务的建模假设与评价指标不同，导致“最佳方法”常随任务而变，且需要分别 benchmark。³
- 3) **ST gene panel/低覆盖 → ST 全转录组增强 (gene imputation/enhancement)**：典型如 SpaGE (domain adaptation + kNN 预测未测基因)、以及基于生成模型的多种增强方法。⁴

在基准评测方面，单细胞整合领域形成了较为公认的 scIB 评测框架 (多指标同时衡量 batch correction 与生物信号保真)；空间任务则出现了面向“sc ↔ ST 集成/解卷积”的系统 benchmark (如对 transcript distribution prediction 与 cell type deconvolution 分别评测)，并进一步发展出面向空间解卷积的实践指南型大评测。⁵

研究范围与术语

本文将 **reference mapping** 定义为：在存在参考图谱 (reference atlas) 的前提下，将新的 query 数据 (单细胞或空间) 映射到参考的表示空间 (latent/embedding/graph) 并最大化继承参考的结构信息 (聚类、轨迹、邻域) 与语义信息 (细胞类型、层级标签、基因程序/状态)，同时尽可能抵抗技术与批次差异。⁶

为避免将“整合 (integration)”与“映射 (mapping)”混为一谈，建议按 **是否需要重训练/重整合 reference** 区分：

- **re-integration 型**：将 reference 与 query 合并后重新整合 (常见于早期流程)。⁷
- **projection/mapping 型**：reference 预先建好 (可能被压缩或以模型参数形式存在)，query 仅做投影与增量适配，从而实现更快、更稳定、且更易复现的映射。⁸

下图给出方法族与任务间的结构关系 (同一方法可能覆盖多个步骤，但通常有明确主战场)。⁹

flowchart LR

A[Reference atlas / model] --> B[Query data]

subgraph sc_to_sc[scRNA / scMultiome → sc atlas]

B --> C1[Preprocess & feature harmonization]

C1 --> C2[Latent alignment / batch correction]

C2 --> C3[Label transfer + uncertainty]

C3 --> C4[Project to reference UMAP/graph & downstream]

end

subgraph sc_to_st[scRNA → Spatial transcriptomics]

B --> D1[Shared genes / signatures]

D1 --> D2a[Deconvolution: proportions per spot]

D1 --> D2b[Localization: cell-to-location mapping]

D1 --> D2c[Gene enhancement: predict unmeasured genes]

end

%% 대표 방법

C2 --> M1[CCA/RPCA anchors]

C2 --> M2[MNN / neighbor matching]

C2 --> M3[Harmony / embedding correction]

C2 --> M4[Probabilistic VAE latent: scVI/scANVI]

C2 --> M5[Reference compression: Symphony]

C2 --> M6[Architecture surgery: scArches/scPoli]

D2a --> S1[Bayesian/NB deconvolution: cell2location, stereoscope, CARD, RCTD]

D2b --> S2[Optimization/OT/metric learning: Tangram, CytoSPACE, SpaOTsc, CellTrek]

D2c --> S3[Domain adaptation/regression: SpaGE, gimVI]

方法学演进与优化方向

可复用 reference 与快速 query 投影

核心改进方向：将“reference 的构建”与“query 的映射”彻底解耦，使 reference 成为可复用资产（对象、参数、压缩表示或可下载模型），从而提升复现性、计算效率与跨研究共享能力。 ⁸

- **代表思路与技术细节**

- **Anchor/近邻为中心的参考映射：**以线性降维（PCA/CCA 或其变体）为基础，通过互为近邻（mutual nearest neighbors）寻找跨数据集“锚点（anchors）”，再基于锚点进行 label transfer 与投影。Seurat 系列在文档与实现层面对 reference/query 的映射接口做了体系化封装（如 FindTransferAnchors 的多种 reduction 模式：pcaproject/rpca/cca 等）。 ¹⁰
- **Reference compression（参考压缩）：**Symphony 在“先整合、再压缩”的策略下，将大规模 reference 变为便携形式，使 query 映射可在秒级完成，同时保持在稳定低维参考嵌入中的定位，从而支持可重复的注释转移与下游推断。 ¹¹
- **Model-as-reference（模型即 reference）：**scArches 将 reference 表示为（条件）生成模型参数，并通过“architecture surgery”（冻结大部分权重、仅对新条件/批次相关部分做增量学习）完成 query 适配，避免从头训练与重新整合。 ¹²

- **可发布、可下载、可即用的 reference 模型库**：scvi-tools 生态在 2022 后逐步将“概率模型 + 规范接口”工程化，scvi-hub 进一步强调“预训练模型仓库化”，以降低在超大 reference 上反复训练的存储与算力负担。¹³

• 优缺点与适用性

- 优点：可复现与可共享（尤其是压缩 reference 或模型权重形式）、适合百万级以上数据、支持统一 reference 作为“坐标系”。¹⁴
- 缺点：对 **基因集合一致性/特征对齐** 的依赖更强；例如 scArches 的典型流程要求 query 与 reference 具有相同基因维度（缺失基因需以零填充等工程处理），这在跨平台/跨物种/空间 panel 场景中会成为瓶颈。¹⁵
- 典型适用：大 reference（器官图谱、人群队列图谱）上对新样本进行快速对照、疾病样本投影到健康 reference、跨实验重复注释。¹⁶

从线性对齐到概率生成潜空间

核心改进方向：用显式概率模型（如负二项/零膨胀负二项）描述计数数据生成过程，在潜空间实现批次因素的可控回归，并在标签转移时输出概率与不确定性。¹⁷

• 代表方法与关键算法演进

- scVI 以深度生成模型（VAE/CVAE）对单细胞计数数据进行概率建模，输出可用于整合与下游分析的低维表示。¹⁸
- scANVI 在 scVI 基础上引入半监督标注机制，使“整合 + 标签预测”在同一概率框架内联合优化。¹⁹
- scPoli 进一步强调 **样本级嵌入（condition embeddings）** 与 **prototype-based label transfer**，并在大规模人群图谱上展示可扩展性与多尺度分析（细胞与样本双视角）。²⁰

• 优缺点与适用性

- 优点：更贴近原始计数统计性质；可自然给出概率输出（便于不确定性与开放集识别）；对复杂批次结构更灵活。²¹
- 缺点：训练与调参成本 higher；对超大规模数据需要更强工程与硬件支持；生成模型的潜空间可解释性与校准性并非自动保证。²²

空间任务特化：deconvolution、单细胞定位与基因增强三分化

空间场景的 reference mapping 不是单一问题。至少存在三条主线，彼此在输入假设与输出目标上差异显著，因此需要分开选择方法与评价指标。²³

- **Deconvolution（spot 细胞类型比例）**：以 spot/capture location 为单位估计细胞类型构成。代表路线包括概率模型与空间平滑（例如 cell2location、CARD）、以及针对平台噪声建模与鲁棒分解（如 stereoscope、spatialDWLS、RTC D 等）。²⁴
- **Localization/mapping（细胞→空间坐标）**：直接输出细胞到空间位置的映射（硬/软分配），或给出细胞在空间中最可能位置。代表路线包括深度对齐（Tangram）、线性规划/最小代价匹配（CytoSPACE），以及最优传输（SpaOTsc、novoSpaRc）与度量学习/共嵌入（CellTrek）。²⁵
- **Gene enhancement（空间基因补全）**：面向“空间技术基因测量不足（panel/覆盖有限）”的问题，利用 scRNA reference 预测未测基因。SpaGE 以 domain adaptation（PRECISE）提取共享子空间，再以 kNN 预测未测基因；gimVI 则以联合生成模型处理 unpaired 的 scRNA 与空间数据用于缺失基因补全。²⁶

标签转移方法综述与对比

标签转移 (label transfer) 可理解为：给定参考数据中的标签体系（离散细胞类型、层级标签、连续状态或多标签），将其转移到 query 细胞/spot 上，并尽可能输出可靠的置信度或拒识 (unassigned)。²⁷

常用性能指标 在文献中较为一致：Accuracy、Macro-F1/Weighted-F1（处理类不平衡）、未标注比例 (unassigned rate)、以及运行时间与可扩展性；在存在 ontology/层级体系时，还会加入“父/子/兄弟节点匹配”等层级一致性指标。²⁸

基于最近邻与投影的标签转移

这一类方法以“在对齐空间中寻找最相近参考样本”为核心，常见做法包括：选取特征基因 → 降维 (PCA/CCA 等) → kNN/互为近邻 → 投票或加权概率输出。²⁹

代表方法：- **scmap (2018)**：将 query 投影到参考细胞或参考簇 (cluster) 层面进行匹配，核心是跨实验的投影与近邻判定，强调在实验/分析差异下进行可比映射。³⁰

- **Seurat anchor-based transfer (2019 起)**：利用 anchors（在 CCA/RPCA 等空间中）建立 reference-query 对应关系，并通过 TransferData 等接口输出标签与分数。¹⁰

- **Symphony (2021)**：将 Harmony/embedding correction 后的 reference 进行压缩，query 通过映射函数快速定位到 reference latent space，并进行注释转移，强调秒级映射与稳定嵌入。¹¹

优缺点：实现简单、速度快、对大 reference 友好；但对“参考覆盖不足/新细胞类型出现/强 domain shift”敏感，且近邻法的置信度往往需要额外校准或经验阈值。³¹

基于图与图神经网络的标签转移

图方法通常把细胞视作节点、相似性视作边，通过图传播或 GNN 学习更稳健的表示与分类器；其优势在于能利用邻域结构与半监督信号。³²

代表性方法（按“图用于何处”区分）：- **构图 + 图传播/层级推断**：例如层级/树结构的 classifier 能在不同粒度标签间传播与拒识（如 CHETAH 的层级分类与 unassigned 输出）。³³

- **GNN 直接做分类/映射**：如 scDeepSort 与 scGCN 将表达数据转换为图结构并用图卷积进行细胞类型预测（典型属于 GNN 注释器体系）。³⁴

- **空间 reference mapping 的图模型**：SageNet 用 GNN 在空间 reference 上学习概率映射，将解离单细胞映射回空间组织来源（更接近“空间定位”但输出可被用于标签转移与空间域判定）。³⁵

局限：对邻接图构造与超参敏感；跨平台/跨批次时图结构本身可能被 domain shift 扭曲，需要与整合/对齐模块配合。³⁶

对抗学习与域自适应

这类方法把 reference→query 的差异视为 domain adaptation 问题，通过对抗训练与半监督学习减少域间可分性，从而提升跨实验标签转移。³⁷

代表方法：- **scNym (2021)**：半监督 + 对抗式神经网络，用带标签的 source 与无标签的 target/query 联合训练，实现跨数据集的注释迁移。³⁸

优势：在强 batch/domain shift 条件下可利用 target 的结构信息；局限：训练更复杂，且不确定性/校准往往需要单独评估。³⁹

概率模型与显式不确定性输出

概率模型将标签预测视为后验推断，天然支持概率与置信区间；在开放集与不确定性方面更易与统计准则结合。⁴⁰

代表方法： - **scANVI (半监督 VAE)**：在生成式潜空间中联合学习表示与分类器，适用于“部分标签 reference + 大量未标注 query”。¹⁹
- **CellAssign (2019)**：基于 marker gene 先验的概率模型，对预定义细胞类型进行分配，适合 marker 清晰且希望避免依赖大 reference 的场景。⁴¹
- **不确定性量化专门研究**：已有工作指出 atlas-level label transfer 的校准与鲁棒性并非充分解决，并系统比较不同模型类的不确定性可用性（属于“需要进一步工程化落地”的关键方向）。⁴²

集成与共识投票

集成方法通过组合多个注释器缓解单模型偏差，并将“低共识”作为可操作的不确定性信号。⁴³

代表方法： - **scClassify (2020)**：以 ensemble learning + cell type hierarchy 进行多尺度分类，并强调样本量估计与层级输出。⁴⁴
- **popV (2024)**：以多注释器共识为主，结合层级 ontology 路径传播形成“共识分数”，并报告共识分数与预测准确度高度相关，便于定位低可信细胞群。⁴³

方法对比表（标签转移）

下表面向“reference mapping 语境下的 label transfer”，优先列出有公开实现与常用 benchmark 证据链的方法。对于“性能基准”列：若存在统一 benchmark 则引用对应 benchmark；若无统一 benchmark，则标注为“无统一 benchmark”，并在后文给出建议。⁴⁵

方法	方法类别	输入数据类型	输出标签类型	性能基准（指标/数据集）	代码与实现	发表年份
scmap	最近邻/投影	scRNA (reference+query)	离散细胞类型 (cell/cluster) + 相似度	大型细胞注释 benchmark 覆盖 (Accuracy、unassigned%、time) ⁴⁶	官方实现 ⁴⁷	2018 ⁴⁸
SingleR	参考相关性/迭代精炼	scRNA (reference:纯细胞类型/伪bulk; query:单细胞)	离散类型 + 分数	多方法对比与实际案例广泛；亦被纳入综合 benchmark ⁴⁹	官方实现 ⁵⁰	2019 ⁵¹
Seurat label transfer (anchors)	近邻 + anchors	scRNA (或多模态) reference + query	离散标签 + prediction score	常用于图谱映射；作为对比基线出现在多个研究（含共识评测） ⁵²	官方文档/实现 ⁵³	2019+ ⁵⁴

方法	方法类别	输入数据类型	输出标签类型	性能基准（指标/数据集）	代码与实现	发表年份
Symphony	参考压缩 + 最近邻定位	scRNA reference（预处理/压缩）+ scRNA query	离散标签（可扩展到连续表型）	秒级映射与可重复定位；论文给出多场景示例 ⁵⁵	官方实现 ⁵⁶	2021 ⁵⁵
scNym	对抗学习/半监督域自适应	有标签 source + 无标签 target/query	离散标签 + 概率	原文强调跨实验迁移；常与多注释器对比 ⁵⁷	官方实现 ⁵⁸	2021 ⁵⁹
CHETAH	层级模型/拒识	scRNA reference + query	层级标签（含 intermediate/unassigned）	有独立论文与实现；亦被后续评测文献引用 ⁶⁰	官方实现 ⁶¹	2019 ⁶²
scClassify	集成学习+层级	scRNA reference（单/多 reference）+ query	多尺度层级标签	论文强调层级分类与样本量估计；作为层级注释代表方法 ⁶³	官方实现 ⁶⁴	2020 ⁶³
OnClass	ontology-aware 分类	有标签训练集 + query	ontology 节点（可含 unseen）	以 Cell Ontology 处理 unseen 类型；被共识方法用于层级传播 ⁶⁵	官方实现 ⁶⁶	2021 ⁶⁷
popV	多模型共识/集成	多注释器输出 +（可选）ontology	离散+层级一致性评分	提供 ontology-aware accuracy，并展示共识分数与准确度关联 ⁴³	论文提供可复现模块说明 ⁴³	2024 ⁴³
scANVI / scvi-tools	概率生成模型（半监督）	scRNA（可含 batch covariates）	离散标签概率 + latent	在整合 benchmark 与共识评测中常作为强基线 ²²	官方生态实现 ⁶⁸	2020–2021 ⁶⁹
scArches	迁移学习/architecture surgery	预训练 reference 模型 + query	离散标签（依赖 backbone）+ 映射	强调 reference mapping 与模型更新；并指出基因集合一致性要求 ¹²	官方实现 ⁷⁰	2022 ¹⁵
scPoli	生成式 + prototype	scRNA（含样本/批次嵌入）	离散标签（prototype）+ 不确定性	论文以 weighted/macro F1 与整合指标做对比 ⁷¹	实现集成于 scArches ⁷²	2023 ⁷¹

UMAP/潜空间回归与对齐方法比较

本节将“UMAP/latent regression”聚焦为：**用于 reference mapping 的嵌入对齐与批次校正**（以及由此支持的 label transfer、空间映射与下游分析）。在实践中，UMAP 更多承担“可视化与局部邻域保持的二维投影”，而 mapping 的统计保证主要由上游 latent alignment/graph alignment 提供；因此方法选择应优先围绕“对齐策略与损失函数”而非 UMAP 本身。⁷³

线性回归与互相关对齐

- **CCA / RPCA (anchor-based 系列)**：CCA 通过寻找最大相关的线性投影来对齐数据集，RPCA 通过互投影的方式更保守、速度更快地寻找 anchors，在大数据集与差异较大场景被广泛使用。⁷⁴
- **适用场景**：细胞类型相对重叠、且希望保持细胞类型结构（避免过度整合）；对“reference mapping + label transfer”的工程链路成熟。⁷⁵
- **限制**：线性相关对跨平台非线性偏差的表达能力有限；当 query 出现 reference 不含的新状态/新类型时容易发生“硬投影”误导，需要结合拒识与不确定性策略。⁷⁶

近邻匹配与局部校正

- **MNN / fastMNN**：通过寻找跨批次互为近邻的细胞对，构造局部校正向量来消除 batch effect，是许多后续 anchor/近邻策略的重要思想基础。⁷⁷
- **Scanorama**：以高效近邻搜索识别共享细胞类型并进行数据集“拼接式”整合，强调在异质数据集上的可扩展性与效率。⁷⁸
- **BBKNN**：通过“batch-balanced”的近邻图构建实现快速对齐，强调在超大规模数据中的速度优势。⁷⁹

嵌入校正式 (embedding correction)

- **Harmony**：在低维空间中迭代聚类并估计/移除 batch 对聚类中心的影响，实现快速嵌入级校正；在规模化整合中常作为强基线，并被用于后续 reference compression（如 Symphony pipeline 中包含 Harmony 步骤）。⁸⁰

概率生成潜空间 (VAE 系列) 与映射扩展

- **scVI/scANVI**：用生成式潜空间同时承担“对齐表示 + (可选) 半监督分类”，在数据整合 benchmark 中作为重要方法类出现。⁸¹
- **scArches**：把“reference mapping”显式建模为对预训练变量的增量适配 (surgery)，在保持 reference 稳定的同时适配 query 条件。¹²
- **scPoli**：在 CVAE 框架上引入连续 condition embedding 与 prototype loss，并用 integration score 与 F1 等指标对比，体现“映射 + 注释 + 多尺度”一体化。⁷¹

空间任务中与 latent/对齐相关的代表模型

- **Tangram (2021)**：在共享基因上学习 scRNA 与空间数据的对齐，输出细胞到空间位置的映射（软分配），并支持将细胞注释投影回空间；其论文明确面向多类空间技术并支持利用图像信息。⁸²
- **SpaGE (2020)**：以 domain adaptation 学得共享空间 (PRECISE) 后用 kNN 预测未测基因，强调稳健、可解释的 gene enhancement。⁸³
- **gimVI (2019/后续实现)**：联合建模 unpaired scRNA 与空间数据，实现缺失基因补全（更偏“生成式增强/翻译”）。⁸⁴

是否存在 benchmark 对比下游任务表现

- **单细胞整合/对齐 benchmark**: scIB 对多方法与预处理组合做系统 benchmark, 强调需要同时衡量 batch correction 与生物变异保真 (多指标综合), 并指出“没有单一方法在所有任务上都占优”, 因此应按下游任务选择方法。⁸⁵
- **sc ↔ ST 集成 benchmark**: 有研究将评测明确拆成两个任务: (1) transcript distribution prediction (更偏 gene pattern/空间表达预测), (2) cell type deconvolution (更偏 spot 组成推断), 并在同一框架比较不同方法族, 强调“任务决定指标、指标决定方法选择”。⁸⁶
- **空间解卷积 benchmark 与指南**: 大型 benchmark 给出实践导向的指南式结论, 强调不同数据分辨率、参考构建策略与空间相关性假设会显著改变方法表现。⁸⁷

结论性归纳 (跨 benchmark 的稳定共识):

- 若目标是 **高质量 label transfer + reference 坐标系复用**, 应优先选“projection/mapping 型”方法族 (Seurat mapping、Symphony、scArches/scANVI/scPoli), 并用与任务一致的指标评估 (宏平均 F1、拒识率、校准误差等)。⁸⁸
- 若目标是 **空间 spot 比例**, 应优先选 deconvolution 特化模型 (cell2location、CARD、stereoscope、spatialDWLS 等), 并在空间相关/技术噪声假设上做匹配。⁸⁹
- 若目标是 **单细胞定位到空间**, 应优先选 localization/assignment/OT/metric learning 路线 (Tangram、CytoSPACE、SpaOTsc、CellTrek 等), 并用定位误差、空间域一致性等与“定位”一致的指标评估。⁹⁰
- 若目标是 **空间基因补全**, 应选 SpaGE/gimVI 等专注 gene enhancement 的方法, 并以 held-out gene 的相关系数、空间自相关结构保持等指标进行评估。⁹¹

已发表的优秀架构与可复用 workflow

本节按“架构 (architecture) = 输入输出 + 训练/推理流程 + 可复用组件”组织, 选取 2016–2026 中被广泛复用或在方法论上具有代表性的体系。

Anchor-based 映射架构 (Seurat mapping family)

- **核心组件**: 特征选择与标准化 → 线性降维 (PCA/CCA/RPCA) → anchor/近邻匹配 → 标签转移 (分数/概率) → 投影到参考嵌入并进行下游分析。¹⁰
- **输入/输出**: 输入为 reference+query (表达矩阵与相交基因); 输出为 query 的标签、置信分数与在参考嵌入中的坐标。⁹²
- **优势**: 工具链成熟、与常用分析对象 (整合、聚类、可视化) 无缝; 适用于多数常规 scRNA reference mapping。⁹³
- **限制**: 对强 domain shift 与新类型出现仍需配合拒识/不确定性; 对特征选择与 anchor 质量敏感。⁷⁶

Reference compression 架构 (Symphony)

- **核心组件**: reference 构建 (常含嵌入校正) → 压缩 reference → query 映射 (定位到稳定低维嵌入) → 注释转移与下游推断。¹¹
- **优势**: 极强的推理效率 (秒级映射), 适配大 reference; 嵌入稳定利于跨研究复现。⁵⁵
- **限制**: 压缩 reference 的质量依赖上游整合质量; 对 reference 覆盖不全的 cell state 仍可能误投影。⁹⁴

Architecture surgery 架构 (scArches)

- **核心组件**: 以 CVAE/深度生成模型为 backbone, 把 reference 表示为模型参数; query 通过冻结权重并只学习新条件相关部分完成增量映射。¹²

- **输入/输出**：输入为预训练模型（reference）+ query；输出为 query latent 表示、可转移标签与（可选）缺失模态/空间信息推断。 ¹²
- **已知限制**：强调基因集合一致性需求（缺失基因需工程处理），在跨平台/跨基因 panel 场景需额外对齐策略。 ¹⁵

Prototype + 多尺度条件嵌入架构（scPoli）

- **核心组件**：conditional generative model + 连续 condition embedding + prototype loss；在 reference mapping 时冻结主干并为 query 增加新条件嵌入，标签通过“最近 prototype”转移并输出不确定性。 ²⁰
- **性能证据**：论文以 integration score 与 F1 等指标对比并展示在大规模 PBMC/肺图谱上的应用。 ⁷¹

层级 reference 与开放集更新架构（treeArches）

- **核心组件**：整合（常借助 scVI/scArches）+ 细胞类型术语匹配与层级构建（如结合层级学习工具）+ atlas 扩展与 unseen 类型识别。 ⁹⁵
- **定位**：解决“不同数据集标签粒度不一致/术语不一致”的现实痛点，强调可更新的 cell-type hierarchy。 ⁹⁶

Bayesian 空间 deconvolution 架构（cell2location / CARD / stereoscope）

- **cell2location**：层级贝叶斯模型，显式建模技术变异并在空间位置间借力（统计强度共享），目标是解析细粒度细胞类型并构建组织细胞地图。 ⁹⁷
- **CARD**：将空间相关性纳入细胞类型比例估计（conditional autoregressive 相关结构），面向“空间平滑 + 组成推断”的组合需求。 ⁹⁸
- **stereoscope**：提出概率模型用单细胞 reference 解卷积空间 spot 的细胞类型比例，是早期以概率推断实现 sc↔ST 结合的重要代表。 ⁹⁹

这些方法的共性优势是：输出可解释（比例/后验）且更适合与空间先验结合；共性限制是：模型假设（计数分布、参考 signature 可靠性、空间相关性形式）会显著影响结果，且不同空间技术下需要针对性校准。 ¹⁰⁰

单细胞定位到空间的对齐/优化架构（Tangram / CytoSPACE / SpaOTsc / CellTrek）

- **Tangram**：以深度学习对齐 scRNA 与空间数据，输出细胞到位置的映射，并支持将细胞注释与全转录组信息投影回空间。 ¹⁰¹
- **CytoSPACE**：将细胞-spot/位置分配形式化为优化问题（最小化相关性代价的匹配/分配），强调在多平台与不同组织上的鲁棒与速度。 ¹⁰²
- **SpaOTsc / novoSpaRc**：基于最优传输框架进行空间重建或空间关系推断，适合“有部分空间测量/ marker 约束”的映射与空间结构恢复。 ¹⁰³
- **CellTrek**：通过共嵌入与度量学习将单细胞映射回空间坐标，强调“区别于纯 deconvolution”的单细胞级空间定位分析范式。 ¹⁰⁴

未解决挑战与未来研究方向

文献尚无统一结论或缺乏统一 benchmark 的问题

- **Atlas-level label transfer 的置信度校准与开放集识别**：已有系统研究指出现有模型在校准、鲁棒性与可操作不确定性上仍有不足，且不同模型类表现差异明显；整体仍属于“未充分研究/缺乏统一 benchmark”的方向，需要更标准化的评测协议。 ¹⁰⁵

- “空间定位 (cell localization)” 的统一评价指标：不同论文常用模拟数据、ISH 验证、空间域一致性等各自定义指标，跨方法可比性有限；目前更接近“任务分裂、指标分散”的状态。¹⁰⁶
- 跨平台/跨基因集合差异的 reference mapping：例如生成式映射体系对基因集合一致性要求较高，空间 panel 数据与 scRNA 全转录组之间存在结构性不匹配，仍需要更通用的特征对齐与缺失机制。¹⁰⁷

可用数据集与建议 benchmark 设计（精简版）

建议将 benchmark 切成三条轨道，并对每条轨道固定输入输出与指标，以减少“混评”造成的结论不稳：

1) sc → sc reference mapping:

- 数据：跨批次/跨研究 scRNA 常用集合（免疫、胰腺、肺等）+ 人群大图谱切分的 reference/query 设定。¹⁰⁸
- 指标：macro-F1、weighted-F1、unassigned rate、校准误差（ECE/Brier）、以及 scIB 一类的生物保真与 batch 混合指标。¹⁰⁹

2) sc → ST deconvolution:

- 数据：同组织配对的 scRNA+ST（不同技术平台）、以及模拟/合成 spot（可控制组成与噪声）。¹¹⁰
- 指标：细胞类型比例的相关系数、RMSE/MAE、稀有类型召回、空间平滑一致性（是否与已知组织结构一致）。¹¹¹

3) sc → ST localization / mapping:

- 数据：有高分辨率空间真值或强约束的场景（例如高通量 FISH 作为 reference），并明确“位置是单细胞级还是区域级”。¹¹²
- 指标：位置误差（若有真值）、空间域分类准确度、以及对下游空间邻域/细胞互作推断的影响。¹¹³

未来方向

- 更可操作的不确定性与拒识机制：不仅输出概率，还应保证可校准、可解释，并能在新类型出现时触发“需要人工复核”的信号（共识分数/ontology 距离等是可用方向）。¹⁰⁵
- 更强的“基因集合不一致”处理：面向空间 panel 与跨平台缺失特征，需要在模型结构上显式建模缺失机制或进行更稳健的共享子空间学习。¹¹⁴
- 模型共享与隐私友好 reference：通过发布预训练模型而非原始数据来实现 reference 复用（如 scvi-hub 的思路）会越来越重要。¹¹⁵
- 更大规模与更通用的表征学习：2023–2025 的单细胞 foundation model（如 scGPT、scFoundation、Geneformer）显示了“统一表征”的潜力，但其在零样本泛化、可解释性与严谨 benchmark 上仍有不确定性，需要对 reference mapping 的实际增益给出更严格证据链。¹¹⁶

结论与建议

在 2016–2026 的方法谱系下，reference mapping 的方法选择应以“任务—输出—指标”三者一致为前提，而不是追求单一“最强算法”。¹¹⁷

- 若目标是把新 scRNA query 放进既有 atlas 坐标系，并继承其标签体系：优先考虑 Seurat mapping 或 Symphony（追求速度与稳定嵌入），以及 scANVI/scArches/scPoli（追求概率建模、半监督与不确定性能力），并用 macro-F1 + 校准指标评估。¹¹⁸
- 若目标是空间 spot 的细胞组成：优先选择 cell2location/CARD/stereoscope/spatialDWLS 等 deconvolution 体系，并在空间相关性假设、参考 signature 构建方式上做敏感性分析。¹¹⁹
- 若目标是单细胞定位到空间：优先选择 Tangram/CytoSPACE/SpaOTsc/CellTrek/SageNet 等定位或映射体系，并避免用“比例推断指标”替代“定位质量指标”。¹²⁰

- 若目标是 **空间基因补全**：SpaGE/gimVI 等更契合，并应以 held-out 基因预测与空间结构保持为核心指标，而不是仅报告可视化改善。 ⁹¹

¹ <https://www.nature.com/articles/nbt.4096>

<https://www.nature.com/articles/nbt.4096>

² ⁵³ ⁸⁸ ⁹² ¹¹⁸ <https://cran.r-project.org/web/packages/Seurat/Seurat.pdf>

<https://cran.r-project.org/web/packages/Seurat/Seurat.pdf>

³ ²³ ⁸⁶ ¹⁰⁶ ¹¹⁰ A comprehensive benchmarking with practical guidelines for cellular deconvolution of spatial transcriptomics | Nature Communications

<https://www.nature.com/articles/s41467-023-37168-7>

⁴ ²⁶ ⁸³ ⁹¹ <https://pmc.ncbi.nlm.nih.gov/articles/PMC7544237/>

<https://pmc.ncbi.nlm.nih.gov/articles/PMC7544237/>

⁵ ²² ³⁶ ⁸⁵ ¹⁰⁸ ¹¹⁷ <https://www.nature.com/articles/s41592-021-01336-8>

<https://www.nature.com/articles/s41592-021-01336-8>

⁶ ⁸ ¹¹ ¹⁴ ⁵⁵ ⁹⁴ <https://www.nature.com/articles/s41467-021-25957-x>

<https://www.nature.com/articles/s41467-021-25957-x>

⁷ ⁹ ¹⁰ ²⁷ ⁵⁴ <https://www.cell.com/cell/fulltext/S0092-8674%2819%2930559-8>

<https://www.cell.com/cell/fulltext/S0092-8674%2819%2930559-8>

¹² ¹⁵ ¹⁰⁷ ¹¹⁴ <https://www.nature.com/articles/s41587-021-01001-7>

<https://www.nature.com/articles/s41587-021-01001-7>

¹³ ²¹ <https://yoseflab.github.io/software/scvi-tools/>

<https://yoseflab.github.io/software/scvi-tools/>

¹⁶ ²⁰ ⁷¹ <https://www.nature.com/articles/s41592-023-02035-2>

<https://www.nature.com/articles/s41592-023-02035-2>

¹⁷ ¹⁸ ⁸¹ <https://www.nature.com/articles/s41592-018-0229-2>

<https://www.nature.com/articles/s41592-018-0229-2>

¹⁹ ⁴⁰ ⁶⁹ <https://pmc.ncbi.nlm.nih.gov/articles/PMC7829634/>

<https://pmc.ncbi.nlm.nih.gov/articles/PMC7829634/>

²⁴ ⁸⁹ ⁹⁷ ¹¹⁹ <https://www.nature.com/articles/s41587-021-01139-4>

<https://www.nature.com/articles/s41587-021-01139-4>

²⁵ ⁸² ⁹⁰ ¹⁰¹ ¹¹² ¹²⁰ <https://www.nature.com/articles/s41592-021-01264-7>

<https://www.nature.com/articles/s41592-021-01264-7>

²⁸ ³¹ ⁴⁵ ⁴⁶ ⁴⁹ ⁷⁶ ¹⁰⁹ <https://link.springer.com/article/10.1186/s13059-019-1795-z>

<https://link.springer.com/article/10.1186/s13059-019-1795-z>

²⁹ ³⁰ ⁴⁸ <https://www.nature.com/articles/nmeth.4644>

<https://www.nature.com/articles/nmeth.4644>

³² ³⁵ <https://github.com/MarioniLab/sagenet>

<https://github.com/MarioniLab/sagenet>

³³ ⁶⁰ ⁶² <https://pmc.ncbi.nlm.nih.gov/articles/PMC6895264/>

<https://pmc.ncbi.nlm.nih.gov/articles/PMC6895264/>

34 <https://www.semanticscholar.org/paper/A-joint-model-of-unpaired-data-from-scRNA-seq-and-Lopez-Nazaret/9991c277fd6a0e2363631846e9a7ea40abc9b7dc>
<https://www.semanticscholar.org/paper/A-joint-model-of-unpaired-data-from-scRNA-seq-and-Lopez-Nazaret/9991c277fd6a0e2363631846e9a7ea40abc9b7dc>

37 38 39 57 59 <https://pmc.ncbi.nlm.nih.gov/articles/PMC8494222/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC8494222/>

41 <https://pmc.ncbi.nlm.nih.gov/articles/PMC7485597/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC7485597/>

42 105 <https://arxiv.org/abs/2211.03793>
<https://arxiv.org/abs/2211.03793>

43 52 <https://www.nature.com/articles/s41588-024-01993-3>
<https://www.nature.com/articles/s41588-024-01993-3>

44 63 <https://pmc.ncbi.nlm.nih.gov/articles/PMC7306901/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC7306901/>

47 <https://github.com/hemberg-lab/scmap>
<https://github.com/hemberg-lab/scmap>

50 <https://github.com/dviraran/SingleR>
<https://github.com/dviraran/SingleR>

51 <https://pmc.ncbi.nlm.nih.gov/articles/PMC6340744/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC6340744/>

56 <https://github.com/immunogenomics/symphony>
<https://github.com/immunogenomics/symphony>

58 <https://github.com/calico/scnym>
<https://github.com/calico/scnym>

61 <https://github.com/jdekanter/CHETAH>
<https://github.com/jdekanter/CHETAH>

64 <https://github.com/SydneyBioX/scClassify>
<https://github.com/SydneyBioX/scClassify>

65 67 <https://www.nature.com/articles/s41467-021-25725-x>
<https://www.nature.com/articles/s41467-021-25725-x>

66 <https://github.com/wangshenguiuc/OnClass>
<https://github.com/wangshenguiuc/OnClass>

68 <https://github.com/scverse/scvi-tools>
<https://github.com/scverse/scvi-tools>

70 <https://github.com/theislab/scarches>
<https://github.com/theislab/scarches>

72 https://github.com/theislab/scarches/blob/master/scarches/models/scpoli/scpoli_model.py
https://github.com/theislab/scarches/blob/master/scarches/models/scpoli/scpoli_model.py

73 75 93 https://satijalab.org/seurat/articles/get_started.html
https://satijalab.org/seurat/articles/get_started.html

- 74 https://satijalab.org/seurat/articles/integration_rpca.html
https://satijalab.org/seurat/articles/integration_rpca.html
- 77 <https://abdn.elsevierpure.com/en/publications/batch-effects-in-single-cell-rna-sequencing-data-are-corrected-by/>
<https://abdn.elsevierpure.com/en/publications/batch-effects-in-single-cell-rna-sequencing-data-are-corrected-by/>
- 78 <https://www.nature.com/articles/s41587-019-0113-3>
<https://www.nature.com/articles/s41587-019-0113-3>
- 79 <https://academic.oup.com/bioinformatics/article/36/3/964/5545955>
<https://academic.oup.com/bioinformatics/article/36/3/964/5545955>
- 80 <https://portals.broadinstitute.org/harmony/>
<https://portals.broadinstitute.org/harmony/>
- 84 <https://docs.scvi-tools.org/en/1.3.3/references.html>
<https://docs.scvi-tools.org/en/1.3.3/references.html>
- 87 111 A comprehensive benchmarking with practical guidelines ...
https://www.nature.com/articles/s41467-023-37168-7?utm_source=chatgpt.com
- 95 96 <https://academic.oup.com/nargab/article/5/3/lqad070/7231336>
<https://academic.oup.com/nargab/article/5/3/lqad070/7231336>
- 98 <https://www.nature.com/articles/s41587-022-01273-7>
<https://www.nature.com/articles/s41587-022-01273-7>
- 99 <https://www.nature.com/articles/s42003-020-01247-y>
<https://www.nature.com/articles/s42003-020-01247-y>
- 100 <https://pmc.ncbi.nlm.nih.gov/articles/PMC9294426/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC9294426/>
- 102 <https://www.nature.com/articles/s41587-023-01697-9>
<https://www.nature.com/articles/s41587-023-01697-9>
- 103 113 <https://www.nature.com/articles/s41467-020-15968-5>
<https://www.nature.com/articles/s41467-020-15968-5>
- 104 <https://pmc.ncbi.nlm.nih.gov/articles/PMC9673606/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC9673606/>
- 115 <https://www.nature.com/articles/s41592-025-02799-9>
<https://www.nature.com/articles/s41592-025-02799-9>
- 116 <https://www.nature.com/articles/s41592-024-02201-0>
<https://www.nature.com/articles/s41592-024-02201-0>