



细胞反卷积算法及其在家养动物研究中的应用进展与挑战

杨彦达¹, 郜鹏举¹, 杨芳¹, 丁小玲³, 高亚辉⁴, 高宁^{1,2}, 何俊^{1,2*}, 张跃博^{1,2*}

(1. 湖南农业大学动物科学技术学院/农业农村部畜禽资源评价与利用重点实验室(猪), 长沙 410128;

2. 岳麓山实验室, 长沙 410128; 3. 安徽农业大学动物科技学院, 合肥 230036;

4. 华南农业大学动物科学学院, 广州 510640)

摘要: 常规RNA测序(bulk RNA sequencing, bulk RNA-seq)在家养动物研究中广泛用于揭示组织水平的转录组差异,但其反映的是群体细胞的平均基因表达,掩盖了组织内的细胞异质性,限制了对复杂性状在细胞层面调控机制的解析。单细胞RNA测序(single cell RNA sequencing, scRNA-seq)具有高精度和高灵敏度,可表征组织的复杂细胞组成并探究细胞异质性背后的遗传调控机制。然而,由于成本高昂、生物系统复杂以及技术操作门槛较高,该技术在动物研究中的规模化应用受限。在此背景下,细胞反卷积(single-cell deconvolution)算法应运而生,根据数据特征差异,该方法基于参考表达谱构建、特征基因筛选或无监督学习等策略,可从bulk RNA-seq中推断细胞组成比例,突破其无法解析细胞异质性的瓶颈;部分算法还能进一步预测细胞类型特异性表达趋势,实现从群体到细胞层面的功能解析。近年来,该方法已逐步应用于猪、牛、鸡等家养动物。本文系统梳理了细胞反卷积算法的基本原理、主流算法和在家养动物中的应用现状,重点分析了参考矩阵构建、数据整合和模型适配等方面的技术瓶颈,并探讨其在多组学融合与细胞功能解析中的拓展潜力,旨在为推动细胞反卷积算法在家养动物研究中的深入应用提供理论框架与方法支持。

关键词: 细胞反卷积;家养动物;转录组;细胞组成

中图分类号: S813.1

文献标志码: A

文章编号: 0366-6964(2026)02-0569-13

Advances and Challenges of Cell Deconvolution Algorithms and Their Applications in Domestic Animal Studies

YANG Yanda¹, GAO Pengju¹, YANG Fang¹, DING Xiaoling³, GAO Yahui⁴, GAO Ning^{1,2},

HE Jun^{1,2*}, ZHANG Yuebo^{1,2*}

(1. Key Laboratory of Livestock and Poultry Resources Evaluation and Utilization (Pig) of Ministry of Agriculture and Rural Affairs, College of Animal Science and Technology, Hunan Agricultural University, Changsha 410128, China; 2. Yuelushan Laboratory, Changsha 410128, China; 3. College of Animal Science and Technology, Anhui Agricultural University, Hefei 230036, China; 4. College of Animal Science, South China Agricultural University, Guangzhou 510640, China)

Abstract: Bulk RNA sequencing (bulk RNA-seq) is extensively used in livestock research to

收稿日期: 2025-07-02

基金项目: 国家重点研发计划项目(2021YFD1301200); 湖南省自然科学基金项目(2024JJ5182); 国家自然科学基金项目(32402730); 湖南省现代农业产业技术体系(HARS-05); 湖南省研究生科研创新项目(CX20240645)

作者简介: 杨彦达, 博士生, 主要从事猪遗传育种研究, E-mail: ydyang@stu.hunau.edu.cn

* 通信作者: 何俊, 主要从事猪遗传育种研究, E-mail: hejun@hunau.edu.cn; 张跃博, 主要从事猪遗传育种研究, E-mail: ybzhangfd@hunau.edu.cn

characterize transcriptomic variation at the tissue level. However, as it captures only the average gene expression across mixed cell populations, it fails to resolve cellular heterogeneity, thereby limiting the dissection of complex trait regulation at the cellular scale. In contrast, single-cell RNA sequencing (scRNA-seq), with its high resolution and sensitivity, enables detailed profiling of cellular composition and uncovers regulatory mechanisms underlying cell-to-cell variability. Despite its advantages, large-scale adoption of scRNA-seq in livestock is hindered by high costs, complex biological systems, and technical barriers. Against this background, single-cell deconvolution algorithms have emerged. Depending on data characteristics, these methods leverage strategies such as reference expression profile construction, marker gene selection, or unsupervised learning to infer cellular composition proportions from bulk RNA-seq data, thereby overcoming its inherent limitation in resolving cellular heterogeneity. Moreover, advanced algorithms can further predict cell-type-specific expression patterns, enabling functional interpretation at single-cell resolution from population-level transcriptomes. In recent years, these methods have been applied in key domestic species such as pigs, cattle, and chickens, supporting studies in traits related to growth, reproduction, and immune function. This review systematically summarizes the principles, mainstream algorithms, and current applications of single-cell deconvolution in livestock, with a particular focus on challenges in reference matrix construction, cross-platform data integration, and model adaptation. Furthermore, we discuss the expanding potential of deconvolution in multi-omics integration and functional cellular analysis, aiming to provide a conceptual and methodological framework for advancing its use in livestock research.

Keywords: single-cell deconvolution; domestic animals; transcriptome; cellular composition

* **Corresponding authors:** ZHANG Yuebo, E-mail: ybzhangfd@hunau.edu.cn; HE Jun, E-mail: hejun@hunau.edu.cn

家养动物是指经人类长期驯化、用于生产、生活或试验研究的经济物种,主要包括猪、牛、羊、鸡等。它们不仅提供优质动物蛋白,还在保障食品安全、生物医学研究和生态系统稳定方面发挥着重要作用。随着动物科学研究从宏观性状逐步迈向细胞与分子层级,如何在组织器官中精准解析不同细胞类型的功能分工,正成为揭示动物生理机制和复杂性状本质规律的关键步骤。为系统解析家养动物中基因表达的遗传调控机制, FarmGTE 项目整合了猪、牛、鸡等多种家养动物的多组织转录组数据与个体水平的基因组变异信息^[1-3],该资源不仅系统鉴定了影响基因表达的遗传变异,还通过跨组织、跨物种分析揭示了调控网络的保守性与特异性,为农业重要性状的分子机制解析和功能基因挖掘提供了高精度的参考体系。然而,该项目仍主要以 bulk RNA-seq 为核心,未曾深度解析细胞类型间的差异与功能异质性^[4-5]。事实上,许多关键的生命活动往往是源自特定细胞亚群的精细表达,如代谢

调控、免疫应答、内分泌调节和疾病进程等^[6]。特别是在复杂组织中,功能稀有或信号活跃的细胞类型,往往对组织状态的变化起到关键作用^[7]。然而,这些细胞亚群所表达的功能信号,往往被 bulk RNA-seq 呈现的“均值效应”所掩盖,进而导致对核心生物过程的误读或遗漏^[8]。因此,从细胞层级出发解析家养动物组织中的细胞构成与功能,不仅有助于识别关键细胞在生命活动中的重要作用,也为揭示遗传调控与疾病发生等提供理论支持,并推动功能基因挖掘与动物模型构建向更高精度发展^[9-12]。

近年来,单细胞 RNA 测序(scRNA-seq)的快速发展为解析细胞异质性提供了强大工具,在家养动物研究中,该技术已成为揭示生长性能、抗病性、肉质等经济性状分子机制的关键手段之一。然而,受限于其高昂成本、技术复杂性及对样本新鲜度的严格要求(图 1),scRNA-seq 尚难以大规模应用于群体遗传分析。为缓解该项难题,细胞反卷积算法应

运而生,该方法可以将 scRNA-seq 数据作为参考,从 bulk RNA-seq 中推断各细胞类型的比例,部分先进算法还可进一步重建细胞类型特异的基因表达谱^[1,13-14]。细胞反卷积技术不仅可以显著降低研究成本、提升分析效率,还能有效挖掘现有 bulk RNA-seq 数据的潜力,拓展应用维度,为家养动物复杂性状的遗传解析与细胞异质性研究提供经济高效的解决方案。该技术已在人类医学与生物学领域广泛应用,并逐步拓展至猪、牛、鸡等家养动物的多种组织^[15-16],成功识别参与泌乳调控、肌肉发育与免疫应答的关键细胞类型及其调控通路^[17-18]。通过细胞反卷积,bulk RNA-seq 数据可被“升维”至细胞水平,展现出在家养动物研究中的巨大潜力。本文将系统综述细胞反卷积的基本原理、主流算法及在家养动物中的应用进展,重点剖析参考矩阵构建、数据整合与模型适配等关键瓶颈,并展望其未来发展方向,旨在为相关研究提供理论参考与方法支持,推动该技术在家养动物领域的深入应用。

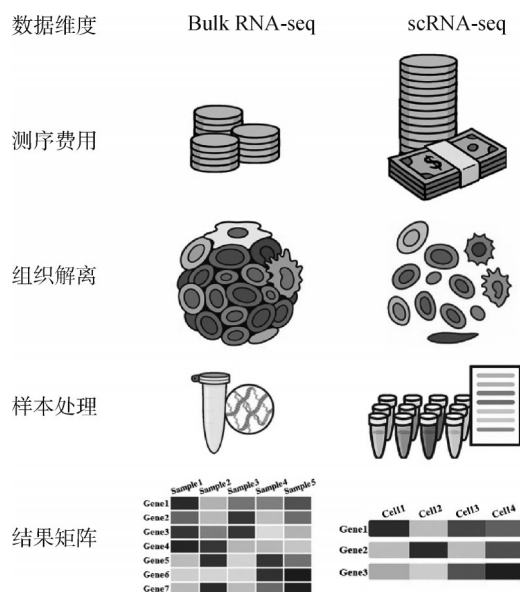


图1 Bulk RNA-seq 与 scRNA-seq 数据获取流程图

Fig.1 Bulk RNA-seq and scRNA-seq data acquisition flow chart

1 细胞反卷积的原理与算法

1.1 细胞反卷积的基本原理

反卷积起源于信号处理领域,最早用于图像还原、音频分离等“信号拆分”任务,其原理在于从混合信号中还原出各组成部分的独立信号^[19-20],后被

引入单细胞转录组学以解析组织内不同细胞类型及其比例^[21-22]。细胞反卷积是一类基于数学模型(如线性回归、非负矩阵分解等)的计算方法,利用 bulk RNA-seq 表达数据与参考表达谱推断细胞组成及类型特异性表达,进而支持生物学深度解析。其核心假设为:组织总表达是各细胞类型表达的线性叠加^[23]。

$$T_i = \sum_{j=1}^k S_{ij} P_j, (S_{ij} \geq 0, P_j \geq 0, \sum_{j=1}^k P_j = 1) \quad (1)$$

$$\hat{S}_{ij} = \frac{T_i P_j}{\sum_{m=1}^k P_m^2} \quad (2)$$

公式(1)中第 i 个基因表示在 bulk RNA-seq 中的总表达量 T_i 是各类细胞在该基因上的表达水平 S_{ij} 与其比例 P_j 的加权求和,而非负性和比例和为 1 的约束条件则保证了生物学意义的可解释性^[24]。部分算法通过反演计算,可进一步估计细胞特异性表达量。具体表现为公式(2)中分母项 $\sum_{m=1}^k P_m^2$ 是采用所有细胞类型比例 P_m 的平方和进行归一化,这一处理在保证估计无偏性的同时,还能避免低丰度细胞表达量被高估,从而实现对不同细胞类型中基因表达的“虚拟预测”。该模型的准确性主要依赖于细胞类型间的表达特异性(如标记基因显著差异)以及数据在物种、组织来源和发育阶段上的匹配程度^[25-26]。通过细胞反卷积推断得到的细胞比例信息,不仅能够精细表征组织结构,还可作为协变量用于性状建模、疾病亚型识别与功能注释等方面,显著提升 bulk RNA-seq 数据的生物学解释力^[27]。

1.2 细胞反卷积的主要算法

细胞反卷积按是否依赖先验细胞类型信息,可分为有参考型(reference-based)、无参考型(reference-free)与半参考型(semi-reference-free) 3 类算法(图 2)^[28]。有参考型算法需使用先验信息(如标记基因、单细胞表达谱等)作为支持,其方法可兼顾参考数据特征、模型优化及生物学原理。根据参考构建需求,可分为基于标记基因、平均表达谱和单细胞数据 3 种类型^[29]。无参考和半参考型算法因无需(或仅需少量)先验信息,其方法设计多依赖于数学分解或统计推断,前者直接从 bulk RNA-seq 推断细胞组成,后者仅需少量标记基因即可估算丰度。

1.2.1 有参考的细胞反卷积算法 基于标记基因的简化模型需要依赖预选细胞类型特异性标记基因构建参考矩阵,通过比较基因在 bulk RNA-seq

的表达水平估算细胞比例^[24,30]。代表性工具包括采用加权最小二乘回归(CLS)算法的DCQ、支持标记基因输入的CIBERSORT^[31]、MIXTURE^[32]和FARDEEP^[33],以及提供多种计算策略的MySort^[34]、quanTIseq^[35]和Bseq-SC^[36]等。部分工具在获得比例后可逆推细胞特异性表达谱,虽然预测精度受标记基因的限制,但基于生物学信息,仍具备解析潜力。该方法适用于细胞类型表达差异显著、标记基因特异性强的场景,如I型(慢肌)与II型(快肌)肌纤维各有一组特异的标记基因,可借此方法快速量化比例差异。因计算高效、操作简便,尤其适合参考数据有限或需大规模群体快速筛查的场景^[37]。

基于平均表达谱的复杂模型主要通过单细胞数据提取各类细胞平均表达谱作为参考,将bulk RNA-seq的表达拟合为参考谱的加权线性组合。其算法演进呈现出多维优化趋势,部分方法聚焦模型泛化与样本适配能力,如CPM^[31]引入样本特异性权重、AutoGeneS^[38]实现参考基因集自动优化、Bisque^[16]和MuSiC^[39]整合多源数据以增强跨样本稳健性;部分致力于提升估计精度与表达重建能力,如DWLS^[40]加权特征基因以识别低丰度群体,DeconPeaker^[41]融合机器学习框架,DeMixT^[26]与BayesPrism^[42]分别基于概率与贝叶斯建模解析异质信号;亦有方法注重计算效率与结果稳定性之间的平衡,如SCDC^[43]与DigitalDLorter^[44]采用约束最小二乘及其变体,在高通量或噪声数据中保持可靠输出。得益于参考矩阵包含完整的基因表达信息,该类算法在复杂组织中可有效捕捉稀有或罕见亚群的微弱信号^[8,45-46],还能解析细胞类型特异的表达谱,为跨组学研究提供了基础。

基于单细胞数据的高分辨率模型是指直接利用原始表达数据构建参考矩阵,保留细胞异质性与动态变化,提升对稀有亚群及连续状态的解析能力。如DESeq2's unmix^[47]和PREDE^[48]是基于最大似然估计框架进行建模,适用于高维表达数据中对细胞比例的稳定估计;EPIC^[49]和dtangle^[50]则通过特征基因表达模式和加权回归策略;ARIC^[51]引入支持向量回归(SVR)以增强对共线性和噪声数据的稳健性;ImmuCellAI^[52]提供内嵌免疫细胞类型参考集合,适用于血液及肿瘤组织的免疫组成估计;TICPE^[53]通过基于评分的估算方法实现多细胞群体的相对量化;EMeth^[54]进一步将细胞反卷积拓展至甲

基化数据,为整合组学提供了可能,DeCompress^[55]通过集成多个细胞反卷积方法结果,提升预测稳健性与可迁移性。该算法虽对计算资源与数据质量要求较高,但凭借其卓越的精确度,在scRNA-seq资源丰富的动物中已逐步成为解析复杂性状调控机制、推动精准育种与疾病建模的核心工具。

1.2.2 无参考的细胞反卷积算法 无参考型算法完全不依赖先验细胞类型信息,而是通过矩阵分解(如NMF、PCA)、贝叶斯推断或深度学习(CNN/RNN)等无监督方法^[56],其假设不同细胞类型在混合信号中仍保留可分离的表达特征,进而对bulk数据进行数据降维、特征提取和信号解构,最终提取潜在的表达模式^[57-59]。代表性工具包括:deconf^[60]利用无监督算法可自动拆解bulk RNA-seq的内在表达特征;ReFACToR^[24]、TOAST^[61]、debCAM^[62]和CellDistinguisher^[63]通过主成分分析和因子提取识别潜在的特征信号^[64];BayesCCE^[65]在无监督的情况下引入贝叶斯建模以增强稳健性;DeconICA^[66]和Linseed^[67]结合线性回归与机器学习实现高效拆分。然而,由于缺乏明确的细胞注释和先验信息,其结果多为数学推断,潜在特征往往难以映射到具体细胞,故无法可靠地重建细胞特异的基因表达谱。该方法适用于高应激、感染或环境极端条件下的组织样本,可提取出最显著的表达模式,有助于揭示应激响应或病理重塑中的主要细胞群体^[68-69]。值得注意的是,尽管无参考和半参考方法不依赖先验参考矩阵,但其本质决定了它们最终会输出一个基因 $\times$ 特征表达模式矩阵(gene $\times$ expression patterns),这与有参考方法中的参考矩阵(gene $\times$ cell types)在形式上具有同构性。这种结构上的相似性为后续整合多方法结果提供了数学基础,可将无参考型识别的模式与已知细胞表达谱结合,提升可读性。

1.2.3 半参考的细胞反卷积算法 半参考型算法介于两者之间,仅需少量标记基因(无需完整表达谱),通过分析Bulk数据中与标记基因协变的表达模式估算细胞丰度。目前常用的半参考软件有:MCP-counter^[70]使用内置的标记基因集(如stromal和immune cell types)对bulk RNA-seq数据进行细胞丰度评分;DSA^[71]和BisqueMarker^[72]则结合主成分分析提取出与标记基因协变的主要特征,用于表征细胞组成变化;Deblender^[73]首先采用无监督聚类将标记基因分组,并基于其表达均值估算各细胞类型的表达特征。半参考型算法部署简便、计算成

本低,特别适用于标记基因已知但缺乏单细胞参考数据的场景,如早期胚胎、稀有组织、非模式或地方品种家养动物。其低资源需求使其成为大规模育种筛选与群体健康监测的理想工具,为数据稀缺研

究提供实用路径。然而,该方法性能高度依赖标记基因的特异性与覆盖完整性,在跨物种迁移或标记集不充分时,易导致细胞类型误判或丰度估计偏差。

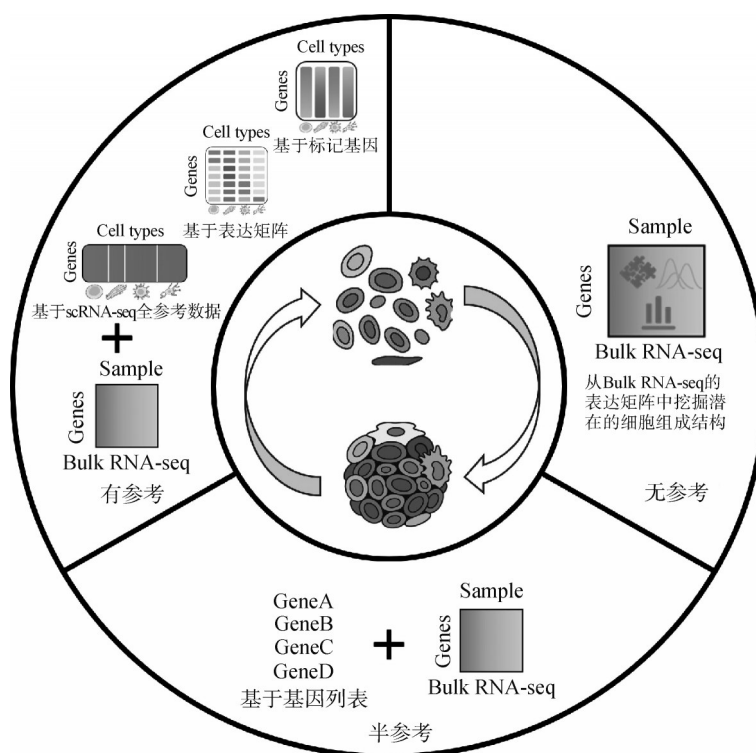


图2 单细胞解卷积方法的计算原理

Fig. 2 The computational methods of single cell deconvolution principle

2 细胞反卷积在家养动物研究中的应用

细胞反卷积的核心价值不在于追求细胞丰度的绝对精确,而在于揭示细胞组成与功能状态的动态变化。相较于成本高昂的 scRNA-seq 和掩盖异质性的 bulk RNA-seq,该方法通过计算建模融合生物学先验,巧妙实现对 bulk RNA-seq 数据的“细胞级解析”。其低成本、高通量特性,尤其适配家养动物研究中长期积累的海量数据资源,赋予其新的生物学维度——不仅为复杂性状机制研究提供高分辨率视角,更架起了从“机制解析”到“育种应用”的转化桥梁,展现出广阔的应用前景。

2.1 解析家养动物生产性状的细胞基础

家养动物的生长速度、胴体组成、繁殖能力与肉品质等核心经济性性状,均受到组织中不同细胞类型的构成比例及其功能状态的影响调控^[74-75]。借助细胞反卷积算法,可基于 bulk RNA-seq 估计关键细

胞类型(如肌源性细胞、脂肪细胞、支持细胞等)的比例变化趋势,进而在细胞水平解析复杂性状的形成机制^[37]。以肉质性状为例,已有研究表明骨骼肌中的慢肌纤维(I型)有助于提升肉的嫩度和风味,而快肌纤维(II型)则更有利于促进生长速度^[76-77]。通过分析肌纤维亚型比例的变化,研究者进一步定位了如 FABP4 和 PPAR γ 等与脂肪酸代谢相关的关键调控因子^[78],为肌肉分子改良与精准育种提供了理论支撑。在分析奶牛乳腺组织时发现,乳腺泡细胞比例与产奶性能呈显著正相关,同时乳脂合成基因(如 PANK3 和 CSN2)在高产个体中显著上调^[79-81]。在生殖系统中,对卵巢与睾丸组织的反卷积分析发现,颗粒细胞、支持细胞等关键功能亚群的比例波动与排卵率、精子活力等繁殖性状密切相关,为提升繁殖效率提供了更精细的视角^[82-86]。通过将细胞类型比例转化为可量化、可遗传的功能表型,细胞反卷积架起了从转录组数据到遗传调控网络的桥

梁,推动畜禽育种从“表型选择”向“机制驱动”的方式转变。

2.2 挖掘家养动物抗性/易感相关细胞群及分子标记

动物对疾病和环境应激的响应高度依赖于免疫系统与组织微环境间的动态互作,而常规转录组却难以区分“基因整体上调”是源于细胞内在激活,还是某类免疫细胞浸润增加所致^[87-88]。细胞反卷积的应用可有效解耦这一混淆,其通过量化外周血或病变组织中免疫细胞亚群的动态重构过程,揭示抗病机制的本质。例如,在猪繁殖与呼吸综合征病毒(PRRSV)感染模型中,CIBERSORT反卷积血液转录组追踪出13种免疫细胞,揭示CD4⁺T与NK细胞早期减少、调节性T细胞恢复期回升的动态轨迹,为疫苗评估提供细胞重构框架^[89-90];在猪肺部感染模型中,巨噬细胞比例升高与病原清除正相关,中性粒细胞扩增则与组织损伤同步,提示其兼具保护与致病双重角色,为平衡型抗炎策略提供依据^[91-92]。此外,在鸡的氧化应激与牛的热应激研究中,库普弗细胞、单核细胞等特定免疫亚群被发现对环境压力高度敏感,其比例变化可作为早期应激响应的生物标志物^[93-94],这些细胞中抗氧化基因(如HMOX1、GPX1)的共表达也进一步揭示协调机制^[95]。在生产实践中,基于“抗病相关细胞群与标志基因”的联合特征识别,可进一步开发细胞类型特异性的分子标记,用于筛选具有优良天然抗病能力和应激耐受性的个体,进而提升畜禽群体的整体健康水平和抗逆性能。

2.3 构建高分辨率家养动物疾病模型细胞图谱

由于在器官结构、生理代谢与免疫反应等方面与人类具有较高的相似性,猪、牛等家养动物已成为重要的人类疾病替代模型^[16,96]。细胞反卷积的引入进一步增强了这些模型在细胞水平的解析能力,尤其适用于探索组织内细胞异质性与疾病进程之间的关联^[97-98]。例如,在针对非酒精性脂肪性肝病(NAFLD)的研究中,研究人员利用猪肝组织的bulk RNA-seq数据结合细胞反卷积分析,发现肝细胞与库普弗细胞比例失衡与纤维化程度密切相关,为病理机制解析与靶点筛选提供了新路径^[99-100]。类似地,在代谢综合征建模中,通过拆分脂肪组织中的免疫细胞构成,研究者揭示了调节性巨噬细胞与脂质代谢网络的协同变动趋势,进一步识别出多个与胰岛素抵抗密切相关的细胞亚群及信号因子,

如CD36⁺巨噬细胞和IL-1 β 活化通路^[101-102]。在异种器官移植与再生医学领域,家养动物模型也成为研究焦点。近年来,结合CRISPR/Cas9等基因编辑手段,有研究团队已构建出具有人源化免疫系统的猪模型^[103-104]。借助细胞反卷积,研究者可实时监测移植组织中免疫细胞群体的动态变化,如调节性T细胞与NK细胞的迁移轨迹,从而辅助评估免疫排斥程度与移植物存活质量^[105-106]。除此之外,在自身免疫类疾病的动物模型研究中(如乳腺炎、病毒性肝炎等),细胞反卷积也已被用于识别关键炎症细胞亚群^[107-108],如活化的中性粒细胞和Th17细胞群体,助力建立更具靶向性的干预策略。

2.4 基于细胞类型特异性遗传关联与功能表型解析的分子育种优化策略

在育种实践中,大量已积累的bulk RNA-seq数据长期停留在“整体表达”层面,未能深入揭示性状形成的细胞机制。细胞反卷积为激活这些历史数据提供了关键突破口——通过整合高质量单细胞参考图谱,无需额外单细胞测序即可从常规转录组中系统推断关键组织的细胞组成与功能状态^[88]。其所获细胞丰度(如I/II型肌纤维、乳腺泡细胞、调节性T细胞等)具有群体可比性与遗传基础,可定义为高分辨率“功能表型”,直接用于遗传定位。这类功能表型不仅可用于定位调控细胞发育与稳态维持的遗传位点(Expression quantitative trait locus, eQTL),还可作为目标性状参与全基因组关联分析(Genome-wide association study, GWAS)或全转录组关联分析(Transcriptome-wide association study, TWAS)。通过批次校正与混合线性模型控制群体结构,细胞比例可稳健融入多组学分析框架,构建“基因—细胞—性状”因果链;多组织/时序数据更可借助多性状GWAS或因子里模型提炼共享细胞维度,提升统计效能。研究表明,基于反卷积构建的细胞类型特异性表达预测模型,能显著提升基因组选择(GS)中功能基因表达的预测精度,推动畜禽育种从“表型关联”迈向“机制驱动”,最终实现从宏观性状到细胞层级的跨尺度精准改良。

3 应用瓶颈与挑战

3.1 生物复杂性对细胞组成精确解析的干扰

家养动物在组织来源、发育阶段和生理状态方面存在高度异质性,即使是同一组织类型,在不同个体间也可能呈现出显著的细胞组成差异。例如,

肌肉在不同日龄、不同营养水平下,肌纤维、脂肪细胞、成纤维细胞的比例变化明显。这种差异使得统一构建参考矩阵或泛用模型变得困难,此外细胞状态的异质性亦是影响细胞反卷积准确性的关键因素^[1,18,109-110]。同一细胞类型在不同环境、免疫应答或应激条件下,可能呈现出截然不同的表达谱。即静息状态与活化状态下的巨噬细胞或T细胞,其功能定位与表达特征呈现显著差异,若在建模时未加区分,易导致细胞类型混淆或估算偏差。而不同品种间的遗传背景和育种目标也会导致细胞层级的调控机制分化,在高产奶牛中,不同品系可能通过激活不同乳腺细胞亚群或分泌通路基因实现乳量提升^[111],尽管表型相似,其背后调控网络却各不相同,进一步削弱了模型的精确解析能力。

3.2 参考资源匮乏限制模型构建与迁移

细胞反卷积的核心在于利用高质量的单细胞参考矩阵来“训练”或“校准”模型,但目前多数家养动物尚无完善的细胞图谱。以山羊、鸭、兔等非模式动物为例,公开可用的单细胞数据资源极其有限^[112-115],研究者往往只能自建小规模数据集或直接借用跨物种表达矩阵,其构建的模型在同一物种内部的拟合性能已显不足,在不同组织类型、发育阶段或跨物种应用时的预测精度更会显著下降。即使在猪、牛等已有初步数据积累的物种,也会因图谱覆盖的组织类型不全、细胞亚群注释粗糙,难以支持针对罕见或功能性特殊亚群的模型训练与验证。除此之外,bulk RNA-seq数据本身常见的测序深度不足、零膨胀偏高、批次效应以及与参考集的技术平台不匹配等因素也进一步削弱了模型对真实细胞比例的重构能力^[116]。当下家养动物的基因组注释尚不全面,基因命名缺乏统一标准,转录本注释亦存在显著空缺,这些导致在特征基因匹配阶段常出现关键标记无法定位或错误对齐的情况,严重削弱了反卷积模型的稳健性和可重复性。此外,现有大多数反卷积软件的输入格式、参数设置与运行流程均针对人类、小鼠等医学模式生物进行了优化,缺乏对家养动物特定数据结构和生物学特性的适配与调整,进一步限制了算法在多物种背景下的应用和性能发挥。

3.3 算法适应性与计算性能瓶颈

目前细胞反卷积方法主要涵盖了线性回归、加权非负最小二乘法、贝叶斯推断、矩阵分解及深度学习等多种方法。尽管各类方法在精度、泛化能力

与计算效率上各有优劣,但其设计初衷大多基于人类或小鼠的临床疾病背景,难以直接迁移至基因组结构复杂、细胞异质性强的家养动物数据中^[117]。以线性模型为例,其假设细胞特征稳定、表达噪声低,适用于组织组成单一的样本,但在涉及多类型细胞共存或稀有亚群信号弱化的家养动物组织中,往往产生拟合偏差^[23,118-119];而贝叶斯类模型尽管具备较高的稳健性,能够整合样本内变异信息,但对表达矩阵质量、先验概率建模与参考结构的依赖性较高,且计算耗时较长,限制了大规模样本处理能力^[120-123]。虽然近年来兴起的深度学习方法在复杂特征提取和非线性拟合方面表现优异,但其对训练数据体量和运算资源要求极高,在计算环境缺失、样本规模有限以及缺乏足量标注的家养动物中难以普及。

4 发展前景与方向

4.1 构建资源化的家养动物单细胞图谱

参考数据缺乏是当前限制细胞反卷积算法在家养动物研究中推广应用的核心问题。未来应重点推动覆盖多物种、多组织、多状态的高质量单细胞图谱建设,特别是在猪、牛、羊、鸡等主要经济动物中,全面采集不同发育阶段、性别、生理和病理状态下的单细胞表达数据,系统绘制主要家养动物的单细胞转录组图谱。目前已有国际合作项目如FarmGTEx和FAANG正在推动从bulk RNA-seq向单细胞分辨率的延伸。同时后续也应进一步完善细胞注释体系、统一基因符号标准、优化表达矩阵格式。这些资源将为构建统一、可共享的家养动物参考矩阵奠定基础,显著提升细胞反卷积模型的准确性、高效性与应用性。

4.2 开发跨物种迁移学习算法模型

面对家养动物的注释稀缺、组织异质性强与表达漂移显著等挑战,亟需新一代算法实现模型的跨物种迁移。通过迁移学习与自适应技术,将在数据丰富物种(如人类、小鼠)上预训练的表达式嵌入模型迁入家养动物中,在算法上引入品系结构、家系信息与组织发育轨迹等辅助变量,增强模型在不同生物背景下的泛化能力。使用图神经网络(GNN)或超图模型将细胞间相互作用、组织拓扑结构与基因共表达网络显式纳入模型架构,增强对稀有亚群与功能状态转换的识别。针对样本量有限的家养动物,可采用少样本(few-shot learning)与元学习

(meta-learning)策略,使模型仅凭少量标注样本即可迅速校准,进一步缓解参考不完备带来的偏倚。最后,应构建模块化、轻量化的跨平台工具,内置自动对齐、自动注释映射与模型压缩功能,以便在普通实验室环境下高效部署。未来的细胞反卷积模型可在多物种、多组织、多平台条件下稳定运行,为家养动物复杂性状与病理机制的细胞级解析提供坚实支撑。

4.3 多组学驱动优化细胞反卷积预测精确度

在生物学视角下,多组学融合为提升家养动物细胞反卷积精度提供了全新思路。在基于转录组、表观组(如单细胞 ATAC-seq 或甲基化)、蛋白质组和空间转录组等多模态数据的基础上,构建“生物学标记-功能网络”的复合签名矩阵,以既保留细胞类型特异的转录本表达,又整合染色质可及性、蛋白修饰和空间分布信息,从而更准确区分高度异质或稀有亚群。此外,该方法通过将细胞类型特异性表达数量性状位点(cell-type-specific eQTL)与基因型 $\times$ 环境互作 QTL ($G \times E$ QTL, Genotype-by-Environment interaction QTL)结果嵌入贝叶斯框架,以基因环境交互效应作为先验信息,加权不同细胞类型在不同遗传背景和应激条件下的实际丰度,为解析表型变异背后的细胞调控机制提供了有力支持。最后,引入伪时序与图神经网络等动态时空建模技术,捕捉细胞构成在发育轨迹或组织微环境中的连续变化。通过“以生物学机制为导向的多组学细胞反卷积”理念,不仅能够精确重建家养动物组织中的细胞比例,还能深入揭示细胞-基因-表型之间的多层次调控关系,为精准育种、健康管理与转化医学模型奠定坚实的数据与机制基础。

5 结 论

细胞反卷积是一种从 bulk RNA-seq 数据中重建组织细胞组成的高效计算方法,有助于从细胞水平解析复杂性状调控机制,提升对生理功能、病理变化及免疫状态的系统认知。尽管该方法在家养动物研究中的应用已初具规模,但其仍面临参考数据缺乏、算法适配性不足与数据异质性高等挑战。未来应加快构建系统化的单细胞参考资源,优化算法以提升跨组织与跨物种的迁移能力,并加强与多组学整合及遗传分析的融合应用。随着数据积累与方法发展的不断推进,细胞反卷积有望成为推动家养动物研究向系统化、精细化发展的关键支撑。

参考文献(References):

- [1] FANG L, CAI W, LIU S, et al. Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle [J]. *Genome Res*, 2020, 30(5):790-801.
- [2] TALAMANCA L, GOBET C, NAEF F. Sex-dimorphic and age-dependent organization of 24-hour gene expression rhythms in humans [J]. *Science*, 2023, 379(6631):478-483.
- [3] CHEN S, LIU S, SHI S, et al. Comparative epigenomics reveals the impact of ruminant-specific regulatory elements on complex traits [J]. *BMC Biol*, 2022, 20(1):273.
- [4] YAO Y, LIU S, XIA C, et al. Comparative transcriptome in large-scale human and cattle populations [J]. *Genome Biol*, 2022, 23(1):176.
- [5] ZENG H, ZHANG W, LIN Q, et al. PigBiobank: a valuable resource for understanding genetic and biological mechanisms of diverse complex traits in pigs [J]. *Nucleic Acids Res*, 2024, 52(D1):D980-D989.
- [6] DING J, LIU R, WEN H, et al. DANCE: a deep learning library and benchmark platform for single-cell analysis [J]. *Genome Biol*, 2024, 25(1):72.
- [7] FUJITA M, GAO Z, ZENG L, et al. Cell subtype-specific effects of genetic variation in the Alzheimer's disease brain [J]. *Nat Genet*, 2024, 56(4):605-614.
- [8] 周素素. 复杂组织细胞组成比例估计的反卷积算法研究 [D]. 抚州:东华理工大学, 2024.
ZHOU S S. Research on deconvolution algorithms for estimating cellular composition proportions in complex tissues [D]. Fuzhou: East China University of Technology, 2024. (in Chinese)
- [9] ZHU J, CHEN F, LUO L, et al. Single-cell atlas of domestic pig cerebral cortex and hypothalamus [J]. *Sci Bull*, 2021, 66(14):1448-1461.
- [10] NATRI H M, DEL AZODI C B, PETER L, et al. Cell-type-specific and disease-associated expression quantitative trait loci in the human lung [J]. *Nat Genet*, 2024.
- [11] ZIMMERMAN K A, BENTLEY M R, LEVER J M, et al. Single-cell RNA sequencing identifies candidate renal resident macrophage gene expression signatures across species [J]. *J Am Soc Nephrol*, 2019, 30(5):767-781.
- [12] ZHANG L, GUO M, LIU Z, et al. Single-cell RNA-seq analysis of testicular somatic cell development in

- pigs[J]. *J Genet Genomics*, 2022, 49(11):1016-1028.
- [13] ZHONG Y, WAN Y W, PANG K, et al. Digital sorting of complex tissues for cell type-specific gene expression profiles [J]. *BMC Bioinformatics*, 2013, 14(1):89.
- [14] 杜文汇. 基于反卷积与自适应图调整的空间转录组数据聚类方法研究[D]. 长春:吉林大学, 2025.
- DU W H. Research on spatial transcriptomics data clustering methods based on deconvolution and adaptive graph adjustment [D]. Changchun: Jilin University, 2025. (in Chinese)
- [15] ZHAO J, CHING W K, WONG C W, et al. BANMF-S: a blockwise accelerated non-negative matrix factorization framework with structural network constraints for single cell imputation[J]. *Briefings in Bioinformatics*, 2024, 25(5):432.
- [16] BARON M, VERES A, WOLOCK S L, et al. A Single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure[J]. *Cell Syst*, 2016, 3(4):346-360.
- [17] TENG J, GAO Y, YIN H, et al. A compendium of genetic regulatory effects across pig tissues [J]. *Nat Genet*, 2024, 56(1):112-123.
- [18] LIU S, GAO Y, CANELA-XANDRI O, et al. A multi-tissue atlas of regulatory variants in cattle [J]. *Nat Genet*, 2022, 54(9):1438-1447.
- [19] WANG K, ABID M A, RASHEED A, et al. DNNGP, a deep neural network-based method for genomic prediction using multi-omics data in plants [J]. *Mol Plant*, 2023, 16(1):279-293.
- [20] MA W, QI X, SUN Y, et al. Computer vision-based measurement techniques for livestock body dimension and weight: a review [J]. *Agriculture*, 2024, 14(2):306.
- [21] O'NEILL N K, STEIN T D, HU J, et al. Bulk brain tissue cell-type deconvolution with bias correction for single-nuclei RNA sequencing data using DeTREM [J]. *BMC Bioinformatics*, 2023, 24(1):349.
- [22] CHENG C, WANG G, ZHU Y, et al. Multiplexed bulk and single-cell RNA-seq hybrid enables cost-efficient disease modeling with chimeric organoids[J]. *Nat Commun*, 2024, 15(1):3946.
- [23] 吕欣荣. 基于转录组测序数据的细胞反卷积算法研究[D]. 济南:山东建筑大学, 2024.
- LYU X R. Research on cellular deconvolution algorithms based on transcriptome sequencing data [D]. Jinan: Shandong Jianzhu University, 2024. (in Chinese)
- [24] CHEN Z, JI C, SHEN Q, et al. Tissue-specific deconvolution of immune cell composition by integrating bulk and single-cell transcriptomes[J]. *Bioinformatics*, 2020, 36(3):819-827.
- [25] DONG M, THENNAVAN A, URRUTIA E, et al. SCDC: bulk gene expression deconvolution by multiple single-cell RNA sequencing references [J]. *Brief Bioinform*, 2021, 22(1):416-427.
- [26] AHN J, YUAN Y, PARMIGIANI G, et al. DeMix: deconvolution for mixed cancer transcriptomes using raw measured data[J]. *Bioinformatics*, 2013, 29(15):1865-1871.
- [27] NEWMAN A M, STEEN C B, LIU C L, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry[J]. *Nat Biotechnol*, 2019, 37(7):773-782.
- [28] BROWN A A, FERNANDEZ-TAJES J J, HONG M G, et al. Genetic analysis of blood molecular phenotypes reveals common properties in the regulatory networks affecting complex traits [J]. *Nat Commun*, 2023, 14(1):5062.
- [29] COBOS F A, PANAH M J N, EPPS J, et al. Effective methods for bulk RNA-seq deconvolution using scRNA-seq transcriptomes [J]. *Genome Biol*, 2023, 24(1):177.
- [30] CHEN C, LEUNG Y Y, IONITA M, et al. Omnibus and robust deconvolution scheme for bulk RNA sequencing data integrating multiple single-cell reference sets and prior biological knowledge [J]. *Bioinformatics*, 2022, 38(19):4530-4536.
- [31] TRAN K A, ADDALA V, JOHNSTON R L, et al. Performance of tumour microenvironment deconvolution methods in breast cancer using single-cell simulated bulk mixtures[J]. *Nat Commun*, 2023, 14(1):5758.
- [32] FERNÁNDEZ E A, MAHMOUD Y D, VEIGAS F, et al. Unveiling the immune infiltrate modulation in cancer and response to immunotherapy by MIXTURE—an enhanced deconvolution method [J]. *Brief Bioinform*, 2021, 22(4):317.
- [33] HAO Y, YAN M, HEATH B R, et al. Fast and robust deconvolution of tumor infiltrating lymphocyte from expression profiles using least trimmed squares [J]. *PLoS Comput Biol*, 2019, 15(5):976.
- [34] CHEN S H, KUO W Y, SU S Y, et al. A gene profiling deconvolution approach to estimating immune cell composition from complex tissues [J]. *BMC*

- Bioinformatics*, 2018, 19(S4):154.
- [35] LOVE M I, HUBER W, ANDERS S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2[J]. *Genome Biol*, 2014, 15(12): 550.
- [36] TRUONG D D, LAMHAMEDI-CHERRADI S E, PORTER R W, et al. Dissociation protocols used for sarcoma tissues bias the transcriptome observed in single-cell and single-nucleus RNA sequencing [J]. *BMC Cancer*, 2023, 23(1):488.
- [37] SOREK G, HAIM Y, CHALIFA-CASPI V, et al. sNucConv: A bulk RNA-seq deconvolution method trained on single-nucleus RNA-seq data to estimate cell-type composition of human adipose tissues [J]. *iScience*, 2024, 27(7):110368.
- [38] ALIEE H, THEIS F J. AutoGeneS: automatic gene selection using multi-objective optimization for RNA-seq deconvolution [J]. *Cell Syst*, 2021, 12(7): 706-715. e4.
- [39] FAN J, LYU Y, ZHANG Q, et al. MuSiC2: cell-type deconvolution for multi-condition bulk RNA-seq data [J]. *Brief Bioinform*, 2022, 23(6):430.
- [40] DU R, CAREY V, WEISS S T. deconvSeq: deconvolution of cell mixture distribution in sequencing data [J]. *Bioinformatics*, 2019, 35(24):5095-5102.
- [41] LI H, SHARMA A, LUO K, et al. DeconPeaker, a deconvolution model to identify cell types based on chromatin accessibility in ATAC-seq data of mixture samples [J]. *Front Genet*, 2020, 11:392.
- [42] JEW B, ALVAREZ M, RAHMANI E, et al. Publisher correction: accurate estimation of cell composition in bulk expression through robust integration of single-cell information [J]. *Nat Commun*, 2020, 11(1):2891.
- [43] CARDOSO-MOREIRA M, HALBERT J, VALLOTON D, et al. Gene expression across mammalian organ development [J]. *Nature*, 2019, 571(7766):505-509.
- [44] ANGERMUELLER C, PÄRNAMAA T, PARTS L, et al. Deep learning for computational biology [J]. *Mol Syst Biol*, 2016, 12(7):878.
- [45] STEEN C B, LIU C L, ALIZADEH A A, et al. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx [M]//KIDDER B L. Stem Cell Transcriptional Networks: Vol. 2117. New York, NY: Springer US, 2020:135-157.
- [46] 聂 斌. 一种细胞类型反卷积集成方法及其应用研究[D]. 合肥:安徽建筑大学, 2025.
- NIE B. An integrated approach for cell-type deconvolution and its applications [D]. Hefei: Anhui Jianzhu University, 2025. (in Chinese)
- [47] 刘源壹, 李昕俞, 巴音那木拉, 等. 单细胞转录组测序技术及其在动物繁殖中的应用进展 [J]. 畜牧兽医学报, 2023, 54(2):421-433.
- LIU Y Y, LI X Y, BAYINNAMULA, et al. Single-cell RNA sequencing technology and its application progress in animal reproduction [J]. *Acta Veterinaria et Zootechnica Sinica*, 2023, 54(2): 421-433. (in Chinese).
- [48] QIN Y, ZHANG W, SUN X, et al. Deconvolution of heterogeneous tumor samples using partial reference signals [J]. *PLoS Comput Biol*, 2020, 16(11):452.
- [49] RACLE J, GFELLER D. EPIC: a tool to estimate the proportions of different cell types from bulk gene expression data [M]//BOEGEL S. Bioinformatics for Cancer Immunotherapy: Vol. 2120. New York, NY: Springer US, 2020:233-248.
- [50] HUNT G J, FREYTAG S, BAHLO M, et al. dtangle: accurate and robust cell type deconvolution [J]. *Bioinformatics*, 2019, 35(12):2093-2099.
- [51] ZHANG W, XU H, QIAO R, et al. ARIC: accurate and robust inference of cell type proportions from bulk gene expression or DNA methylation data [J]. *Briefings in Bioinformatics*, 2022, 23(1):362.
- [52] MIAO Y, ZHANG Q, LEI Q, et al. ImmuCellAI: a unique method for comprehensive T-cell subsets abundance prediction and its application in cancer immunotherapy [J]. *Adv Sci*, 2020, 7(7):1902880.
- [53] XIAO H, ZHANG J, WANG K, et al. A cancer-specific qualitative method for estimating the proportion of tumor-infiltrating immune cells [J]. *Front Immunol*, 2021, 12:672031.
- [54] ZHANG H, CAI R, DAI J, et al. EMeth: an EM algorithm for cell type decomposition based on DNA methylation data [J]. *Sci Rep*, 2021, 11(1):5717.
- [55] BHATTACHARYA A, HAMILTON A M, TROESTER M A, et al. DeCompress: tissue compartment deconvolution of targeted mRNA expression panels using compressed sensing [J]. *Nucleic Acids Res*, 2021, 49(8):48.
- [56] KARIOTIS S, TAN P F, LU H, et al. Omada: robust clustering of transcriptomes through multiple testing [J]. *GigaScience*, 2024, 13:039.
- [57] CHEN G, NING B, SHI T. Single-cell RNA-seq technologies and related computational data analysis [J]. *Front Genet*, 2019, 10:317.

- [58] NGUYEN T, WEI Y, NAKADA Y, et al. Analysis of cardiac single-cell RNA-sequencing data can be improved by the use of artificial-intelligence-based tools[J]. *Sci Rep*, 2023, 13(1):6821.
- [59] 王 芳. 基于深度学习的单细胞异质性分析方法研究[D]. 哈尔滨:哈尔滨工业大学, 2025.
WANG F. Research on single-cell heterogeneity analysis methods based on deep learning[D]. Harbin: Harbin Institute of Technology, 2025. (in Chinese)
- [60] REPSILBER D, KERN S, TELAAR A, et al. Biomarker discovery in heterogeneous tissue samples-taking the in-silico deconvolution approach[J]. *BMC Bioinformatics*, 2010, 11(1):27.
- [61] LI Z, WU H. TOAST: improving reference-free cell composition estimation by cross-cell type differential analysis[J]. *Genome Biol*, 2019, 20(1):190.
- [62] CHEN L, WU C T, WANG N, et al. debCAM: a bioconductor R package for fully unsupervised deconvolution of complex tissues [J]. *Bioinformatics*, 2020, 36(12):3927-3929.
- [63] NEWBERG L A, CHEN X, KODIRA C D, et al. Computational *de novo* discovery of distinguishing genes for biological processes and cell types in complex tissues[J]. *PLoS One*, 2018, 13(3):67.
- [64] MADEN S K, KWON S H, HUUKI-MYERS L A, et al. Challenges and opportunities to computationally deconvolve heterogeneous tissue with varying cell sizes using single-cell RNA-sequencing datasets [J]. *Genome Biol*, 2023, 24(1):288.
- [65] RAHMANI E, SCHWEIGER R, SHENHAV L, et al. BayesCCE: a Bayesian framework for estimating cell-type composition from DNA methylation without the need for methylation reference [J]. *Genome Biol*, 2018, 19(1):141.
- [66] 张肖旭, 李 昊, 冯平捷, 等. 单细胞转录组测序技术在家养动物中的应用[J]. 畜牧兽医学报, 2024, 55(8):3276-3287.
ZHANG X X, LI H, FENG P J, et al. Application of single-cell RNA sequencing in domestic animals [J]. *Acta Veterinaria et Zootechnica Sinica*, 2024, 55(8): 3276-3287. (in Chinese).
- [67] ZAITSEV K, BAMBOUSKOVA M, SWAIN A, et al. Complete deconvolution of cellular mixtures based on linearity of transcriptional signatures[J]. *Nat Commun*, 2019, 10(1):2209.
- [68] AVILA COBOS F, ALQUICIRA-HERNANDEZ J, POWELL J E, et al. Benchmarking of cell type deconvolution pipelines for transcriptomics data [J]. *Nat Commun*, 2020, 11(1):5650.
- [69] JIA S, LYSENKO A, BOROEVICH K A, et al. scDeepInsight: a supervised cell-type identification method for scRNA-seq data with deep learning [R]. *Bioinformatics*, 2023.
- [70] BECHT E, GIRALDO N A, LACROIX L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression[J]. *Genome Biol*, 2016, 17(1):218.
- [71] 魏康康, 马 贵, 李文迪, 等. 单细胞测序技术在绵羊卵巢生长发育过程中的研究进展[J]. 畜牧兽医学报, 2025, 56(7):3080-3087.
WEI K K, MA G, LI W D, et al. Research progress of single-cell sequencing technology in the ovarian growth and development of sheep [J]. *Acta Veterinaria et Zootechnica Sinica*, 2025, 56(7): 3080-3087. (in Chinese)
- [72] JEW B, ALVAREZ M, RAHMANI E, et al. Accurate estimation of cell composition in bulk expression through robust integration of single-cell information [J]. *Nat Commun*, 2020, 11(1):1971.
- [73] DIMITRAKOPOULOU K, WIK E, AKSLEN L A, et al. Deblender: a semi-/unsupervised multi-operational computational method for complete deconvolution of expression data from heterogeneous samples [J]. *BMC Bioinformatics*, 2018, 19(1):408.
- [74] DEKKERS J C M, GIBSON J P, BIJMA P. Design and optimisation of animal breeding programmes [J]. *J Anim Breed Genet*, 2023, 140(3):265-284.
- [75] DU L, LI K, CHANG T, et al. Integrating genomics and transcriptomics to identify candidate genes for subcutaneous fat deposition in beef cattle [J]. *Genomics*, 2022, 114(4):110406.
- [76] CRESPO-PIAZUELO D, ACLOQUE H, GONZÁLEZ-RODRÍGUEZ O, et al. Identification of transcriptional regulatory variants in pig duodenum, liver, and muscle tissues [J]. *GigaScience*, 2022, 12:042.
- [77] XU D, WAN B, QIU K, et al. Single-cell RNA-sequencing provides insight into skeletal muscle evolution during the selection of muscle characteristics [J]. *Adv Sci*, 2023, 10(35):2305080.
- [78] GARIN-SHKOLNIK T, RUDICH A, HOTAMISLIGIL G S, et al. FABP4 attenuates PPAR γ and adipogenesis and is inversely correlated with PPAR γ in adipose tissues [J]. *Diabetes*, 2014, 63

- (3):900-911.
- [79] HUNDE D, TADESSE Y, TADESSE M, et al. Community-based breeding programs can realize sustainable genetic gain and economic benefits in tropical dairy cattle systems [J]. *Front Genet*, 2024, 15:1106709.
- [80] OLASEGE B S, VAN DEN BERG I, HAILE-MARIAM M, et al. Dissecting loci that underpin the genetic correlations between production, fertility, and urea traits in Australian Holstein cattle [J]. *Anim Genet*, 2024;age. 13455.
- [81] BUONAIUTO G, LOPEZ-VILLALOBOS N, COSTA A, et al. Stayability in Simmental cattle as affected by muscularity and body condition score between calvings [J]. *Front Vet Sci*, 2023, 10: 1141286.
- [82] HAO Z, WANG J, LUO Y, et al. MicroRNA-200c affects milk fat synthesis by targeting PANK3 in ovine mammary epithelial cells [J]. *Int J Mol Sci*, 2022, 23 (24):15601.
- [83] LEE J N, WANG Y, XU Y O, et al. Characterisation of gene expression related to milk fat synthesis in the mammary tissue of lactating yaks [J]. *J Dairy Res*, 2017, 84(3):283-288.
- [84] XIA K, LUO P, YU J, et al. Single-cell RNA sequencing reveals transcriptomic landscape and potential targets for human testicular ageing [J]. *Hum Reprod*, 2024:199.
- [85] ZHOU H, YIN Z, NI B, et al. Whole exome sequencing analysis of 167 men with primary infertility [J]. *BMC Med Genomics*, 2024, 17(1):230.
- [86] XU H, ZHANG S, DUAN Q, et al. Comprehensive analyses of 435 goat transcriptomes provides insight into male reproduction [J]. *Int J Biol Macromol*, 2024, 255:127942.
- [87] IKEBUCHI R, FUJIMOTO M, MORIYA T, et al. T cells are the main population in mouse breast milk and express similar profiles of tight junction proteins as those in mammary alveolar epithelial cells [J]. *J Reprod Immunol*, 2020, 140:103137.
- [88] GIORDANO A, PERUGINI J, KRISTENSEN D M, et al. Mammary alveolar epithelial cells convert to brown adipocytes in post-lactating mice [J]. *J Cell Physiol*, 2017, 232(11):2923-2928.
- [89] JANKOVIČOVÁ J, SEČOVÁ P, HOROVSKÁ L, et al. Distribution of tetraspanins in bovine ovarian tissue and fresh/vitrified oocytes [J]. *Histochem Cell Biol*, 2023, 159(2):163-183.
- [90] MUELLER A, LEHNER J, HANCKE K, et al. Fertility preservation: improved neovascularization and follicle viability in cryopreserved bovine ovarian cortex transplants by remaining medulla tissue [J]. *F&S Sci*, 2022, 3(4):340-348.
- [91] ANDERSON T K, CHANG J, ARENDSEE Z W, et al. Swine influenza A viruses and the tangled relationship with humans [J]. *Cold Spring Harb Perspect Med*, 2021, 11(3):737.
- [92] MA W. Swine influenza virus: Current status and challenge [J]. *Virus Res*, 2020, 288:198118.
- [93] XU S, XIAOJING L, XINYUE S, et al. Pig lung fibrosis is active in the subacute CdCl₂ exposure model and exerts cumulative toxicity through the M1/M2 imbalance [J]. *Ecotoxicol Environ Saf*, 2021, 225: 112757.
- [94] HUANG S. The impact of ski servicescape on tourist loyalty: evidence from Chinese market [J]. *Curr Issues Tour*, 2024:1-20.
- [95] MOTA L F M, GIANNUZZI D, PEGOLO S, et al. Genomic prediction of blood biomarkers of metabolic disorders in Holstein cattle using parametric and nonparametric models [J]. *Genet Sel Evol*, 2024, 56(1):31.
- [96] MENG G, TANG W, HUANG E, et al. A comprehensive assessment of cell type-specific differential expression methods in bulk data [J]. *Brief Bioinform*, 2023, 24(1):516.
- [97] FLECK J S, SANCHÍS-CALLEJA F, HE Z, et al. Resolving organoid brain region identities by mapping single-cell genomic data to reference atlases [J]. *Cell Stem Cell*, 2021, 28(6):1148-1159.
- [98] YANG W, YU J, YAO Y, et al. Comparative immune-relevant transcriptome reveals the evolutionary basis of complex traits [J]. *iScience*, 2022, 25(12):105572.
- [99] XIE M, LEE K, LOCKHART J H, et al. TIMEx: tumor-immune microenvironment deconvolution web-portal for bulk transcriptomics using pan-cancer scRNA-seq signatures [J]. *Bioinformatics*, 2021, 37 (20):3681-3683.
- [100] LEI H, GUO X A, TAO Y, et al. Semi-deconvolution of bulk and single-cell RNA-seq data with application to metastatic progression in breast cancer [J]. *Bioinformatics*, 2022, 38:386-394.
- [101] LI C, WU H, GUO L, et al. Single-cell transcriptomics reveals cellular heterogeneity and

- molecular stratification of cervical cancer[J]. *Commun Biol*, 2022, 5(1):1208.
- [102] JIANG J, LI J, HUANG S, et al. CACIMAR: cross-species analysis of cell identities, markers, regulations, and interactions using single-cell RNA sequencing data [J]. *Brief Bioinform*, 2024, 25(4):bbae283.
- [103] PAN Z, YAO Y, YIN H, et al. Pig genome functional annotation enhances the biological interpretation of complex traits and human disease [J]. *Nat Commun*, 2021, 12(1):5848.
- [104] LI J, ZHAO T, GUAN D, et al. Learning functional conservation between human and pig to decipher evolutionary mechanisms underlying gene expression and complex traits [J]. *Cell Genom*, 2023, 3 (10) : 100390.
- [105] ELOSUA-BAYES M, NIETO P, MEREU E, et al. SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes [J]. *Nucleic Acids Res*, 2021, 49 (9) : e50-e50.
- [106] HUANG Z Y, LUO Z Y, CAI Y R, et al. Single cell transcriptomics in human osteoarthritis synovium and in silico deconvoluted bulk RNA sequencing [J]. *Osteoarthritis Cartilage*, 2022, 30(3):475-480.
- [107] XU Z, WANG X, FAN L, et al. Integrative analysis of spatial transcriptome with single-cell transcriptome and single-cell epigenome in mouse lungs after immunization [J]. *iScience*, 2022, 25(9):104900.
- [108] CHARYTONOWICZ D, BRODY R, SEBRA R. Interpretable and context-free deconvolution of multi-scale whole transcriptomic data with UniCell deconvolve [J]. *Nat Commun*, 2023, 14(1):1350.
- [109] ZHAO P, GU L, GAO Y, et al. Young SINEs in pig genomes impact gene regulation, genetic diversity, and complex traits [J]. *Commun Biol*, 2023, 6(1):894.
- [110] 袁心柱, 李玲琴, 杜欢, 等. 基于CIBERSORT反卷积算法研究狼疮肾炎的免疫浸润机制及潜在靶向中药预测 [J]. *中国免疫学杂志*, 2024, 40(10):2121-2129.
- YUAN X Z, LI L Q, DU H, et al. Investigating immune infiltration mechanisms in lupus nephritis and predicting potential targeted traditional Chinese medicines using CIBERSORT deconvolution algorithm [J]. *Chinese Journal of Immunology*, 2024, 40(10):2121-2129. (in Chinese)
- [111] WANG Z, LI X, YANG J, et al. Single-cell RNA sequencing deconvolutes the *in vivo* heterogeneity of human bone marrow-derived mesenchymal stem cells [J]. *Int J Biol Sci*, 2021, 17(15):4192-4206.
- [112] ZHANG Z, SCHAEFER C, JIANG W, et al. A panoramic view of cell population dynamics in mammalian aging [J]. *Science*, 2025, 387(6731):949.
- [113] ÖLING S, STRUCK E, NOREEN-THORSEN M, et al. A human stomach cell type transcriptome atlas [J]. *BMC Biol*, 2024, 22(1):36.
- [114] FANG C, SELEGA A, CAMPBELL K R. Beyond benchmarking and towards predictive models of dataset-specific single-cell RNA-seq pipeline performance [J]. *Genome Biol*, 2024, 25(1):159.
- [115] HAN X, WANG R, ZHOU Y, et al. Mapping the mouse cell atlas by microwell-seq [J]. *Cell*, 2018, 172 (5):1091-1107.
- [116] WANG J, HUANG M, TORRE E, et al. Gene expression distribution deconvolution in single-cell RNA sequencing [J]. *Proc Natl Acad Sci U S A*, 2018, 115(28):E6437-E6446.
- [117] CALDARELLI P, DEININGER L, ZHAO S, et al. AI-based approach to dissect the variability of mouse stem cell-derived embryo models [J]. *Nat Commun*, 2025, 16(1):1772.
- [118] LI Y. Spatial transcriptomic cell-type deconvolution using graph neural networks [J]. *Nat Methods*, 2023, 20(5):678-689
- [119] 潘多. 人类肝脏单细胞RNA测序数据整合及数据库构建 [D]. 南京:东南大学, 2024.
- PAN D. Integration of human liver single-cell RNA sequencing data and database construction [D]. Nanjing: Southeast University, 2024. (in Chinese)
- [120] KO K D, SARTORELLI V. A deep learning adversarial autoencoder with dynamic batching displays high performance in denoising and ordering scRNA-seq data [J]. *iScience*, 2024, 27(3):109027.
- [121] MA Y, ZHOU X. Spatially informed cell-type deconvolution for spatial transcriptomics [J]. *Nat Biotechnol*, 2022, 40(9):1349-1359.
- [122] GROZA T, CAUFIELD H, GRATION D, et al. An evaluation of GPT models for phenotype concept recognition [J]. *BMC Med Inform Decis Mak*, 2024, 24(1):30.
- [123] FELDMANN M J, COVARRUBIAS-PAZARAN G, PIEPHO H P. Complex traits and candidate genes: estimation of genetic variance components across multiple genetic architectures [J]. *G3 (Bethesda)*, 2023:148.