



家畜单细胞转录组学数据库的研究现状与展望

王 健¹, 李瑞星², 刘翘铭^{3*}, 李远方^{4,5*}

(1. 河南财经政法大学计算机与信息工程学院, 郑州 450000; 2. 河南大学软件学院, 开封 475000;
3. 河南大学人工智能学院, 郑州 450000; 4. 哈尔滨工业大学医学与健康学院, 哈尔滨 150001;
5. 哈尔滨工业大学郑州研究院, 郑州 450000)

摘 要: 本文旨在阐述单细胞组学技术如何推动畜牧学研究, 并系统综述家畜单细胞转录组学数据库的构建流程、应用现状及主要挑战, 以为研究者提供数据参考与分析思路, 促进该技术在家畜研究中的深入应用。通过全面梳理与整合分析现有家畜单细胞转录组数据库资源, 总结其构建方法与应用路径, 并探讨当前面临的技术瓶颈。单细胞转录组技术为解析家畜疾病机制、分子育种及健康管理提供了新的研究视角和丰富的数据资源, 同时也揭示了在数据整合、注释标准化等技术环节仍存在显著挑战。未来应加强高质量数据库构建、开发适用于家畜的分析算法, 并推动多组学整合, 以充分发挥单细胞技术在畜牧业发展中的潜力。

关键词: 家畜; 单细胞转录组学; 数据库

中图分类号: S813

文献标志码: A

文章编号: 0366-6964(2026)01-0001-12

Research Status and Prospects of Livestock Single-cell Transcriptomics Databases

WANG Jian¹, LI Ruixing², LIU Qiaoming^{3*}, LI Yuanfang^{4,5*}

(1. School of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou 450000, China; 2. Software College, Henan University, Kaifeng 475000, China; 3. College of Artificial Intelligence, Henan University, Zhengzhou 450000, China; 4. School of Medicine and Health, Harbin Institute of Technology, Harbin 150001, China; 5. Zhengzhou Research Institute, Harbin Institute of Technology, Zhengzhou 450000, China)

Abstract: This article aims to elucidate how single-cell omics technologies advance livestock research and provides a systematic review of the construction process, current applications, and major challenges of single-cell transcriptomic databases in domestic animals. It seeks to offer data references and analytical frameworks for researchers, thereby promoting deeper application of this technology in livestock studies. This review conducted a comprehensive review and integrated analysis of existing single-cell transcriptomic database resources for livestock, summarizing their construction methodologies and application pathways, while also discussing current technical bottlenecks. The review demonstrates that single-cell transcriptomics offers novel perspectives and abundant data resources for elucidating mechanisms of livestock diseases, molecular breeding, and health management, while also revealing significant challenges in areas such as data integration and

收稿日期: 2025-05-19

基金项目: 河南省科技攻关项目(221100110200); 国家自然科学基金(62402145)

作者简介: 王 健, 博士, 主要从事大数据分析、数据挖掘、群智能算法研究, E-mail: goodjian121@126.com

* 通信作者: 李远方, 主要从事生物信息学、畜禽资源分子评价研究, E-mail: yuanfangli1991@163.com; 刘翘铭, 主要从事生物信息学研究, E-mail: cslqm@henu.edu.cn

annotation standardization. Future efforts should focus on constructing high-quality databases, developing analysis algorithms tailored to livestock species, and promoting multi-omics integration to fully leverage the potential of single-cell technologies in the development of animal husbandry.

Keywords: livestock; single-cell transcriptomics; database

* **Corresponding authors:** LI Yuanfang, E-mail: yuanfangli1991@163.com; LIU Qiaoming, E-mail: cslqm@henu.edu.cn

单细胞转录组学测序技术(scRNA-seq)通过解析复杂组织中单个细胞的转录组信息,揭示了组织中的细胞异质性和细胞间的相互作用,为生命科学研究提供了前所未有的深度和精度^[1]。在家畜研究领域,scRNA-seq的应用日益广泛,尤其在生长发育、繁殖和免疫等关键性状的研究中取得了显著进展^[2]。例如,Cai等^[3]利用几个胚胎阶段(E16、E18、E21和E28)的藏猪和杜洛克×藏猪的体节组织和肌节组织进行了scRNA-seq,通过这些数据识别了肌源性细胞类型和其他细胞类型,构建了猪骨骼肌肌节的分化轨迹,并确定了*EGR1*和*RHOB*是猪胚胎生肌的关键调节因子。Li等^[4]利用大恒肉鸡和藏鸡胚胎阶段(E10、E14和E18)的腿肌组织进行了scRNA-seq,通过这些数据构建了一个鸡胚胎骨骼肌的综合细胞图谱,并揭示了大恒肉鸡和藏鸡之间骨骼肌发育的分子机制差异,这些数据为家畜、家禽骨骼肌发育研究提供了宝贵的信息。Yang等^[5]对成年绵羊睾丸组织进行了scRNA-seq,确定了绵羊生殖细胞的几个特定阶段的标记基因,如*EZH2*和*SOX18*等基因,该研究探索了绵羊精子生成过程中各种细胞类型的转录水平的变化,为绵羊精子生成和精子生成细胞发育提供了新的见解。此外,Zhang等^[6]利用两种具有不同抗病性的肉鸡品种,即北京油鸡和科布肉鸡,通过scRNA-seq表征了两种鸡脾组织的CD45⁺免疫细胞,并比较了北京油鸡和科布肉鸡在细菌感染后的免疫反应,这些发现提供了对不同肉鸡品种的炎症调节机制的见解。由此可见,scRNA-seq的广泛应用不仅推动了畜禽相关的基础生物学研究,还为分子育种和疾病预防的实际应用奠定了坚实基础。

目前,关于家畜单细胞转录组公开的数据已经越来越多,涵盖了猪^[7]、鸡^[8]、牛^[9]、羊^[10]和兔^[11]等重要畜禽业动物。随着单细胞测序数据的快速增长,如何高效存储、整合和挖掘这些海量信息成为新的挑战,而单细胞组学数据库的建立为此提供了有效的解决方案。单细胞组学数据库通过整合已发表

的单细胞测序数据,为科研人员提供了强大的数据共享和分析平台。例如,CellMarker 2.0^[12]、CDSCP^[13]和 PanglaoDB^[14]等数据库,它们不仅包含基因表达谱和细胞类型注释,还涉及许多标记基因等信息。然而,这些数据库主要收集的是人和小鼠物种的单细胞转录组数据,针对畜禽的单细胞转录组学的数据库目前仍相对匮乏。随着畜禽单细胞转录组数据量的快速增长,如何高效管理、整合和利用这些数据,已成为当前亟待解决的问题。因此,本文通过分析关于畜禽单细胞转录组学数据库的研究进展,探讨其应用价值与未来发展方向,旨在为推动畜禽单细胞转录组学数据库的创新提供理论支撑,为家畜和家禽的遗传改良和精准育种实践提供高效的数据资源平台。

1 单细胞转录组数据库的构建

单细胞转录组数据库的构建通常包括3个主要步骤:数据收集,数据分析和数据展示(图1)。

1.1 数据收集

单细胞转录组数据的收集主要通过以下几个公共数据库。

PubMed 数据库^[15]: PubMed (<https://pubmed.ncbi.nlm.nih.gov>)是由美国国家生物技术信息中心(NCBI)维护的免费生物医学文献数据库,该数据库主要收录已发表的文章。近年来,PubMed收录的关于scRNA-seq研究的文章大幅度增长,用户可通过搜索相关领域的关键词和“scRNA-seq”共同检索相关的文章,并通过文章中的提供的数据储存链接获取相应的数据。PubMed作为文献存储平台中心,研究学者通过PubMed能高效获取scRNA-seq从技术细节到应用创新的全链条信息,促进试验设计和数据分析的优化。

GEO(Gene Expression Omnibus)数据库^[16]:由国家生物技术信息中心(NCBI)开发的GEO数据库(<https://www.ncbi.nlm.nih.gov/gds>)主要储存高通量基因表达和其他功能基因组学的数据集。该

数据库主要收录经过处理后的基因表达数据,如矩阵结果等,这些数据可以直接用于下游分析。此外,也有部分FASTQ格式的原始数据储存于该数据库中,这些数据需要先进行序列比对,再进行定量分析。该数据库不仅收录了大量的scRNA-seq公开的数据,还提供了各种数据分析的工具,使用户能够快捷的找到感兴趣的数据集,并进行个性化的数据分析及数据可视化。

GSA (Genome Sequence Archive) 数据库^[17]: GSA 数据库 (<https://ngdc.cncb.ac.cn/gsa/>) 是由中国国家基因组科学数据中心 (National Genomics Data Center) 开发的。GSA 主要储存原始序列数据的数据集,其中大部分为FASTQ、BAM和VCF格式的原始数据,数据库也收录了大量scRNA-seq数据集。用户可以通过指定的BioProject、BioSample、Experiment或Run ID来搜索感兴趣的数据。此外, GSA还允许用户通过输入物种名称、疾病类型、表型特征、组织或细胞类型等进行高级搜索,允许用户浏览所有公开可用的生物项目、生物样本和试验等,为广大研究学者提供一个高效且易于使用的平台。

ArrayExpress 数据库^[18]: 欧洲生物信息学研究所 (EBI) 开发的 ArrayExpress 数据库 (<https://www.ebi.ac.uk/biostudies/arrayexpress>) 是一个重要的公共生物数据库,包含了大量的高通量功能基因组学试验的数据,数据主要以FASTQ格式储存。近年来,数据库中的RNA-seq数据集显著增加,值得关注的是,单细胞层面的试验研究呈现爆发式增长,scRNA-seq正逐步取代传统群体样本的分析模式。针对目前单细胞转录组数据的增加, ArrayExpress还开发了Annotare工具,该工具不仅支持原始数据和预处理数据的提交,更能系统整合并解读试验所需的全套元数据。平台精选数据集经标准化处理后,将整合至姊妹数据库——Expression Atlas分析平台(包含单细胞表达谱的数据库)。这一改进不仅提升数据可解释性,同时构建了严格的质量控制体系。随着ArrayExpress数据库的不断更新,未来将被纳入到一个新的、更通用的储存资源BioStudies数据库,并将保留原有的数据编号和应用程序编程接口,确保科研工作的延续性。

ENA (European Nucleotide Archive) 数据库^[19]: 由欧洲生物信息学研究所 (EMBL-EBI) 开发的

ENA 数据库 (<https://www.ebi.ac.uk/ena/browser/home>) 主要用于高通量数据的管理、储存、出版和传播,目前也存储了大量的scRNA-seq原始数据。在2024年期间,ENA收到了超过12 400项研究、150万个样本、120万个原始读取数据集和50万个基因组集合,ENA已发展成为协调核苷酸数据的全球关键数据平台,并被世界上90%以上的国家使用。总体而言,ENA为生物信息来源的数据提供了一个全面和综合的资源平台。

总的来说,以上这些数据库共同构成了单细胞研究领域的重要数据资源,为研究学者提供了丰富的可重复利用的数据支持,从而高效地推动相关研究领域的发展。

1.2 数据分析

单细胞转录组数据的处理过程主要包括以下3个关键步骤。

数据的预处理:该步骤将测序读数转换为表达矩阵,该矩阵用作所有下游分析的输入。与bulk RNA-seq相比,预处理scRNA-seq数据的主要区别在于必须处理各种DNA条形码,这些条形码有助于将序列读数分配给其原来的细胞,大多数方法都设计用于使用基于唯一分子标识符(UMI)的数据^[20]。常规的scRNA-seq预处理工作流程涉及去多路复用、序列对比、转录本定量、质量控制和数据整合等步骤,目前这些步骤主要使用Cell Ranger工具进行^[21],该工具的分析流程不仅适用于人类或小鼠等模式生物,同样适用于家畜和家禽物种,只需提供相应的物种参考基因组即可实现数据处理。例如,Herrera-Urbe等^[22]利用Cell Ranger软件处理猪外周血免疫细胞的单细胞转录组测序数据,并采用其默认参数生成基因表达矩阵。不过,Cell Ranger是专门为分析10x Genomics平台数据而开发的,不能用于处理其他单细胞技术平台产生的数据。

细胞类型鉴定和差异分析:细胞类型鉴定主要通过无监督聚类(如K-means^[23]、Louvain^[24]算法)将细胞分群。以Seurat工具包中的FindClusters函数为例,其核心即采用Louvain算法进行细胞聚类。Li等^[25]正是通过该函数及其默认参数对猪及其他多种哺乳动物的回肠上皮细胞数据进行聚类分析,成功揭示了细胞类型的保守性与物种特异性。此外,细胞类型注释通常通过结合已知标记基因或使用SingleR等工具调用参考数据集(如人类细胞图

谱 HPCA 或 CellMarker 等)来完成。然而,这些参考数据集主要基于人类或小鼠数据构建,若要准确鉴定家畜和家禽的细胞类型,仍需开发相应物种特异的参考数据集。差异表达分析通常通过统计模型(如 MAST^[26]和 Wilcoxon 检验^[27])筛选不同细胞亚群或试验条件之间的差异表达基因。Seurat 工具包中的 FindAllMarkers 函数默认采用 Wilcoxon 检验,该方式已被广泛应用。例如, Li 等^[4]即利用该函数鉴定出大恒肉鸡与藏鸡腿肌中与骨骼肌发育相关的差异基因。

基因功能富集分析:基因功能分析是在完成单细胞转录组数据的预处理和细胞鉴定及差异分析后的重要环节。基于差异表达基因或特定的基因集,可展开基因功能富集(如 GO、KEGG)分析^[28-29],可用于分析感兴趣的基因集的生物功能。此外,还可进行细胞通讯分析和轨迹推断分析(拟时序分析)等^[30],用于推断细胞的分化路径和发育动态。不过,这些分析通常需要经过前期对数据的筛选和处理,部分复杂的分析不适合在数据库中展示。

综上所述,家畜单细胞转录组数据的处理与分析是一个系统而复杂的过程,需综合运用多种生物信息学工具与统计方法,并充分考虑物种特殊性,才能有效保障数据质量与结果可靠性。

1.3 数据展示

1.3.1 后端架构 数据存储:采用关系型数据库(如 MySQL^[31])管理结构化元数据,确保数据的规范化和高效查询,表结构设计遵循单细胞数据相关元数据标准(如细胞类型、样本来源、基因 ID 等)。对于非结构化数据,如单细胞测序产生的矩阵信息,使用 MongoDB^[32]进行存储。同时,结合云存储服务,将大型文件存储在云端,通过分布式存储和缓存机制,确保数据的高效存储和快速访问。

计算引擎:集成主流单细胞数据分析工具(如 Scanpy、Seurat^[33])的 API 接口,实现无缝调用和高效处理。通过分布式计算框架(如 Apache Spark^[34])优化计算性能,支持大规模数据的并行处理,显著提升数据分析和处理的效率。目前,已有多个分析工具基于 Apache Spark 分布式框架开发,以显著提升大规模单细胞转录组数据的处理能力。例如, scSPARKL^[35]充分利用 Apache Spark 在并行计算、弹性扩展与容错机制方面的优势,实现了对海量 scRNA-seq 数据的高效分析。

数据安全与备份:采用多层安全策略保护数据

隐私和完整性。数据传输采用加密协议(如 TLS),确保数据在传输过程中的安全^[36]。定期对数据进行备份,使用异地备份和快照技术,防止数据丢失和损坏。

可扩展性与维护性:后端架构设计注重可扩展性和维护性。采用微服务架构将不同功能模块解耦,便于独立开发和更新。通过自动化部署工具(如 Ansible)集成到 CI/CD 流程中^[37],实现快速迭代和高效维护,确保数据库的长期稳定运行。

1.3.2 前端展示 交互式查询:常规单细胞转录组学数据库支持按基因、细胞类型、组织器官类型等多维度检索,为用户提供便捷精准的数据查询功能。

可视化模块:目前大部分数据库的数据可视化是利用 R/Shiny 工具进行展示的,截止 2022 年, R/Shiny 已经开发了 470 多个生物网络应用程序^[38]。Shiny 允许用户通过 R 代码创建动态可视化界面,特别适合单细胞数据库数据可视化页面的构建。利用 Shiny 开发的数据可视化界面一般包含以下几个核心模块:a)降维可视化模块;b)细胞类型鉴定模块;c)基因表达分析模块;d)差异分析模块。

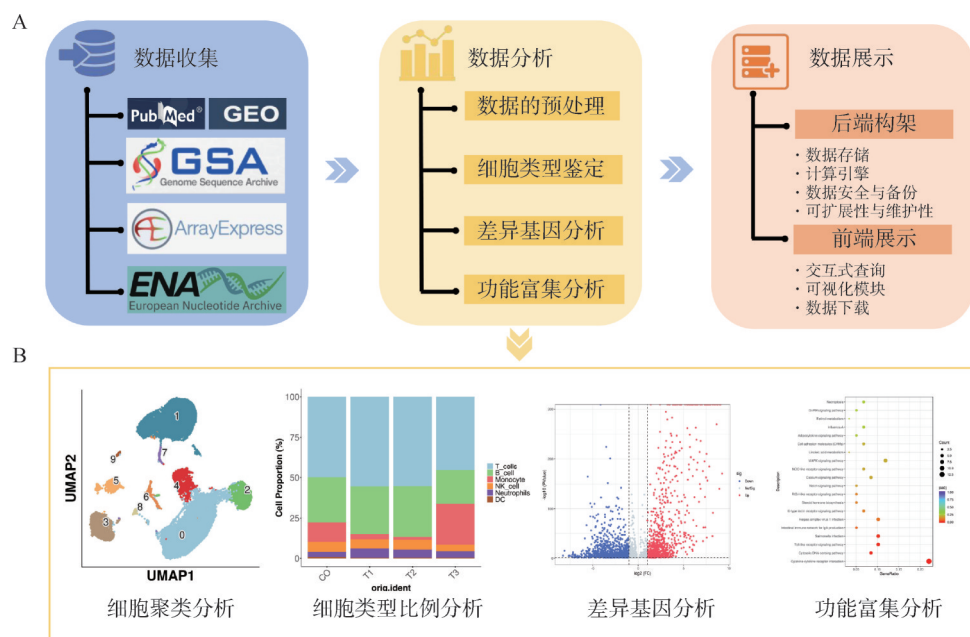
数据下载:大多单细胞转录组数据库提供原始数据、预处理矩阵和分析结果的多种格式导出,满足用户在不同研究阶段对数据的需求。

2 家畜和家禽单细胞转录组数据库构建现状

目前,关于畜禽相关的单细胞数据库数量还是相对匮乏,已有的数据库的物种包括猪、鸡、牛、兔等,大部分数据库的数据是以猪和鸡为主,具体的数据库信息列在表 1 中。

2.1 猪相关单细胞转录组数据库

2021 年, Zhu 等^[33]构建了首个家猪脑组织单细胞数据库平台 PBAtlas,该数据库包括了 3 月龄家猪五个脑区 32 250 个细胞的单核 RNA 测序(snRNA-seq)数据,探究了家猪脑组织的细胞组成和基因表达。该数据库平台分为 4 个功能模块,其中 PBcluster 用于显示猪大脑五个区域内的细胞群; PBtalk 用于展示由配体和受体介导的假定相互作用; PBhunter 用于分析用户自定义提交的细胞和数据库中的细胞类型之间的相关性; PBdisease 用于显示每种细胞类型中神经系统疾病基因的富集水平。该平台为使用家猪进行人类神经系统疾病研究以及动物模型的构建提供了重要的参考信息。



A. 单细胞转录组数据库构建的三大步骤;B. 单细胞转录组数据分析的4个模块

A. The three major steps in constructing a single-cell transcriptomic database; B. The 4 modules of single-cell transcriptomic data analysis

图1 单细胞转录组数据库的构建流程

Fig. 1 The construction process of the single-cell transcriptome database

2.2 鸡相关单细胞转录组数据库

2025年,由中国农业大学、中国农业科学院等多家单位联合开展的ChickenGTEx项目发布了鸡的多组织遗传调控图谱^[39]。该项目整合了7 015个样本的RNA-seq, 127 598个单细胞RNA-seq和2 869个全基因组重测序数据,系统识别了数百万个影响转录及转录后过程的调控变异,揭示了这些变异在分子机制上的独特性及其对复杂性状解析的重要作用。

2.3 牛相关单细胞转录组数据库

由浙江大学构建的牛单细胞图谱数据库CCL (Cattle Cell Landscape)^[40],包含了新生和成年奶牛的13种组织类型的单细胞转录组数据,涵盖瘤胃、肝、肠道等关键组织。该数据库由30个单细胞数据集组成,支持细胞类型注释、比较分析和标记基因查询等功能。可在多个时间和空间尺度上观察细胞类型组成、细胞间通讯的变化,并且提供细胞类型特异性标记基因查询功能,该数据库为奶牛研究学者提供有效的细胞类型注释信息及宝贵的参考资源。此外,Han等^[41]通过生成和分析小牛和成年牛59种组织类型的1 793 854个细胞的scRNA-seq和snRNA-seq数据,为畜牧研究界构建了全面的牛

细胞图谱(CattleCA),揭示了基因表达、转录因子调节和细胞间通信中的组织内和组织间的细胞异质性,并构建了牛泛组织单细胞基因表达信息知识平台CattleCA。

2.4 驴相关单细胞转录组数据库

Tian等^[42]利用scRNA-seq和snRNA-seq数据开发了全球首个驴全身组织器官单细胞图谱DCA (Donkey Cell Atlas),其中包括德州驴的20个主要组织和器官,84种细胞类型以及275 050个高质量的细胞。该数据库允许用户查询特定组织中特定细胞的特定基因表达情况、细胞与细胞之间的通讯情况以及转录因子在不同组织中的表达信息等,该数据库为驴的单细胞转录组分析提供了宝贵的见解。

2.5 多物种单细胞转录组数据库

除了针对某一物种的单细胞转录组数据库外,目前已有多个跨物种单细胞转录组数据库,它们包含了多个家畜相关物种。例如,Li等^[13]构建了一个跨物种的scRNA-seq数据库CDCP。CDCP数据库包括家猪、家鸡、家牛、家兔和绵羊等物种,该数据库共享和集成了641个数据集,2万多个样本,9 800多万万个细胞的单细胞数据集。但是目前CDCP数

数据库中关于家畜的数据集相对较少,大部分研究集中于人和小鼠物种上。此外,该数据库还提供了单细胞分析工具和可视化分析等模块,并允许用户查询特定细胞类型或集群中不同基因的表达情况。总体而言,该数据库功能丰富,为研究人员从已有数据中挖掘有价值信息提供了重要支持。然而,CDCP的标志基因查询功能目前仅覆盖人类和小鼠两个物种,缺乏家畜和家禽动物相关物种的标志基因数据。由于标志基因是单细胞数据精准注释的基础,这一缺陷严重限制了该数据库在家畜和家禽单细胞研究中的应用价值。

目前,收录品种最多的数据库之一 SCEA (Single Cell Expression Atlas)^[43],一共收录了 21 个物种的 scRNA-seq 研究,包括家鸡、家兔和绵羊。该数据库提供了用于 scRNA-seq 分析的代码,集成了许多用于过滤、读数映射、量化表达、聚类以及发现标记基因等现有的分析工具,并提供了用于绘制 t-SNE 图的网页插件。此外,该数据库可以根据特定疾病、组织进行数据集搜索,该数据库十分适合免疫学、生物发育学等相关领域研究学者使用。然而,SCEA 中家畜动物的数据总量仍相对有限,目前仅收录鸡物种数据集 4 个、家兔数据集 2 个和绵羊数据集 1 个。数据覆盖范围的不足显著降低了其在畜禽研究中的参考价值。

SPEED^[44]是一个整合 scRNA-seq 和单细胞全基因组测序(scWGS)数据的平台,SPEED 数据库提供了 127 个物种的单细胞多组学数据,包括多种畜禽相关数据,如家猪、家鸡、家兔、绵羊和家牛等物种,并支持跨物种比较分析。SPEED 数据库还提供了细胞通讯分析(C2C)、基因互作分析(G2G)和跨物种进化分析(S2S)等功能模块,为生物发育与疾病发病的分子机制研究提供了重要的资源。值得注意的是,SPEED 同时整合了家兔的 scRNA-seq 和 scWGS 数据,为该物种的研究提供了宝贵的多组学基础。然而,与上述数据库类似,当前数据库中畜禽相关数据的规模仍显不足,在一定程度上限制了其在畜禽领域的广泛应用。

此外,SCP(Single Cell Portal)^[45]数据库是由 Broad 研究所建立,SCP 数据库的宗旨是加速单细胞研究,这个数据库提供了 18 个物种,831 个数据集,5 754 万多个细胞的分析,包含了家猪、家鸡、家牛和绵羊这几种畜禽动物。SCP 数据库可通过检索相关的研究深入探索数据集所包含的细胞类型。

此外,用户可手动输入基因集查看其在不同细胞类群中的表达情况,并提供详细的细胞聚类情况、基因表达矩阵分析数据和 marker 基因数据信息,且可供用户下载。该数据库比较创新的是,它提供了一个自助服务平台,研究人员可以将自己的数据存入私人 SCP 研究中,并与他们的合作者共享该项研究,与合作者在平台上可一起完善数据分析,且在该数据发表时支持同行评审过程,从而增加研究学者工作的影响力。

总体而言,上述数据库通过整合多种家畜、家禽物种的单细胞数据,为该领域研究提供了重要的数据资源与分析基础。然而,现有平台仍普遍存在家畜、家禽数据类型覆盖有限、标志基因注释不够完善等局限,尚未形成以畜禽为核心、育种应用为导向的综合性数据库。因此,构建一个集成多物种、多组织并具备精准注释功能的畜禽专属单细胞数据库,已成为推动该领域从基础研究向育种实践转化的关键需求。

3 单细胞转录组数据库的应用

单细胞转录组数据库的快速发展为家畜、家禽领域研究提供了前所未有的高分辨率数据资源^[46]。通过整合多物种、多组织的单细胞转录组数据,研究者能够深入解析家畜、家禽器官发育动态、揭示疾病机制,还可进行跨物种进化分析并评估家畜、家禽在医学转化领域的应用价值等^[47]。以下将从四个方面阐述单细胞数据库的重要应用。

3.1 器官发育动态研究

通过整合分析单细胞转录组数据库中的特定组织器官的数据,可以深入了解特定器官在不同发育阶段的细胞分化 and 功能变化,有助于鉴定新的细胞类型,并揭示关键调控机制。目前,研究学者测定的家畜、家禽的单细胞转录组数据可通过与公开的人类或小鼠的数据进行整合分析,从而揭示器官的发育动态变化。例如,Wang 等^[48]利用猪睾丸的 scRNA-seq 数据与公开的人类、小鼠 scRNA-seq 数据进行对比,分析了从幼年期到成年期猪、人类和小鼠的生殖细胞组成。研究发现与小鼠相比,猪的生殖细胞的减数分裂开始的更早。另外,在猪睾丸中发现了一种与人类早期状态精原干细胞(SSCs)高度相似的生精细胞亚型,表明它们可能代表着最原始的生精干细胞。此外,Zhao 等^[49]通过分析公开的猪和人类的睾丸 scRNA-seq 数据,揭示了人与猪

表 1 家畜、家禽相关单细胞转录组数据库及其功能特点

Table 1 Single-cell omics databases and their functions for livestock and poultry

数据库名称 Database name	数据库详情 Database details	物种类型 Species types	功能特点 Functional features	网址 Website	文献 Reference
PBatlas	1个物种,1个样本,5个大脑区域,32 250个细胞	家猪	细胞聚类与基因表达分析、细胞通讯分析、疾病风险基因分析	http://120.79.46.200:82	[33]
CCL	1个物种,30多个数据集,13种组织	奶牛	细胞类型注释、比较分析和标记基因查询	http://cattlecelllandscape.zju.edu.cn	[40]
DCA	1个物种,84种细胞类型,27万多个细胞	驴	基因表达查询、细胞通讯分析、转录因子查询	https://egrb.qau.edu.cn/	[42]
CattleCA	1个物种,59种细胞类型,170万多个细胞	牛	基因表达查询和标记基因查询	https://ngdc.cncb.ac.cn/cattleca/	[41]
ChickenGTEx	1个物种,127 598个细胞	鸡	基因表达查询	https://chicken.farmgtex.org/	[39]
CDCP	641个数据集,2万多个样本,9 800万多个细胞	家猪、家鸡、家牛、家兔和绵羊	允许查询特定细胞类型或集群中不同基因的表达	https://db.cngb.org/cdcp/	[13]
SCEA	21个物种,355个数据集,1 700万多个细胞	家鸡、家兔和绵羊	集成化 scRNA-seq 分析流程、t-SNE 绘图网页部件	https://www.ebi.ac.uk/gxa/sc/home	[43]
SPEED	127个物种,630多万个细胞,85种疾病类型	家猪、家鸡、家兔、绵羊和家牛	细胞通讯分析、基因间互作、跨物种进化分析	http://speedatlas.net	[44]
SCP	831个数据集,5 754万多个细胞	家猪、家鸡、家牛、绵羊	基因表达查询、细胞聚类分析、标记基因查询	https://singlecell.broadinstitute.org/single_cell	[45]

各阶段精子细胞的特异性分子标记,并鉴定到人类与猪的精子形成过程中涉及物种特异的关键基因(猪的 *SNRPD2.1* 和人的 *CATSPERZ* 基因),为精子细胞成熟机制和雄性生殖发育研究提供了新的见解。轨迹分析表明,人与猪在精子发生早期阶段具有较高的保守性,但在精子拉长阶段出现显著分化。通过差异表达分析发现了猪物种中参与圆形精子与拉长精子形成的关键 *PPAR* 信号通路,为解析猪物种生殖细胞成熟机制提供了新依据。上述结果表明,利用单细胞转录组数据库开展跨物种整合分析,不仅能够揭示家畜器官发育中的进化保守性与物种特异性,还可为家畜生殖生物学研究提供新的见解和思路。

3.2 疾病机制研究

单细胞转录组数据库有助于揭示疾病发生发展过程中的细胞异质性及其分子机制,为疾病诊断、治疗和预防提供新的靶点和策略。例如,Qian等^[50]通过整合、共识注释和分析 302 个公开可用的 scRNA-seq 数据集,创建了健康和疾病状态下人类

大脑的单细胞转录组数据集,根据这些数据确定了八种星形细胞亚型,并探索了它们在大脑区域的分布特异性,并进一步确定了不同疾病状态下的细胞之间的相互作用,为中枢神经系统疾病机制研究提供了新的见解。这种分析策略在家畜研究中同样适用,Xu等^[51]利用 3 种猪品种和公开的人类、小鼠肌肉组织的 scRNA-seq 数据集,并对骨骼肌的单细胞转录组特征进行了跨物种比较,确定了不同物种间总共 76 个共享的细胞标记基因和 618 个物种特异性细胞标记基因,为肌肉细胞群的剖析和肌肉发育提供了新的见解。此外,该项研究通过检测 6 种人类肌肉疾病的易感性基因的表达情况,发现猪与人类肌病杜氏肌萎缩(DMD)和脊髓肌肉萎缩(SMA)疾病在单细胞转录表达特征方面具有高度的相似性,揭示猪物种在治疗这类人类肌肉疾病中的潜在应用价值。

3.3 进化生物学分析

此外,基于单细胞转录组数据库的跨物种分析也为进化生物学研究提供了宝贵的数据资源。例

如,Zhu等^[33]通过构建猪大脑scRNA-seq数据,并整合小鼠下丘脑数据集开展跨物种比较分析,系统揭示了小鼠与猪在基因表达模式上的保守性和差异性。研究发现,少突胶质组细胞的经典标志物(如*PDGFRA*、*CSPG4*)在两个物种间表达高度一致,体现了物种间的保守性。然而,在细胞分化调控机制方面却存在显著物种差异,小鼠少突胶质细胞分化过程中关键基因(如*KLF6*、*FOS*和*MAPK3*)呈现高度动态表达模式,其中*KLF6*作为已知的分化加速因子,其表达随分化进程逐渐下调;而在猪中,*KLF6*并未表现出类似的动态变化趋势,表明该基因在猪下丘脑少突胶质细胞成熟过程中可能不具有核心调控作用。Zhang等^[52]通过构建人、老鼠和鸡等6个物种的PBMC的scRNA-seq数据集,并利用公共数据库中的猪、牛和恒河猕猴等6个物种,构建了共12个物种的PBMC图谱。通过scRNA-seq分析和跨物种比较,确定了不同物种之间保守的标记基因。例如,在单核细胞中,*FOSL2*、*FOS*、*TCF7L2*和*SP1*这4个转录因子在12个物种中均呈现激活状态,表明这些因子在单核细胞的功能调控中发挥着核心作用。

3.4 生物医学模型研究

利用单细胞转录组数据库进行跨物种比较分析为评估动物模型的医学适用性提供了重要依据。猪因其与人类在生理和遗传上的高度相似性,已成为生物医学研究的重要模型。例如,Wang等^[53]通过比较猪与公开的人类和小鼠的皮肤组织的单细胞转录组数据,发现猪表皮的细胞亚型组成比小鼠更接近人类表皮,表明猪模型是了解人类皮肤发育的宝贵资源。进一步通过对无毛猪和正常猪的比较研究,发现*OGN*⁺/*UCHL1*⁺细胞在增殖与迁移过程中的异常行为,是导致无毛猪早期毛囊祖细胞形成缺陷的关键原因。Herrera-Urbe等^[22]的研究对猪PBMC进行了scRNA-seq分析,成功鉴定出13种常见免疫细胞类型。通过进一步与人类PBMC的公共scRNA-seq数据集开展跨物种比较,揭示出猪与人类在许多免疫细胞群体间存在高度同源性,再次证实猪PBMC可作为人类生物医学研究的优良模型。上述成果充分表明,基于单细胞转录组数据库的跨物种比较分析,不仅能够为转化医学提供关键数据支持,还有力促进了人类疾病机制研究和临床前模型的开发。

4 总结与展望

家畜相关的单细胞组学数据库建设已取得显著进展,为家畜科学研究提供了重要资源。然而,目前仍面临诸多挑战。首先,现有的数据库数据规模和覆盖范围有限,主要集中在少数畜禽物种以及特定组织中,例如猪和鸡等物种^[44-45],难以全面反映家畜整体的细胞图谱。其次,数据产生过程中因测序平台(如10x Genomics、Drop-seq等^[54])、试验方案与分析方法的不同,导致数据格式不统一、标准化流程缺失,为跨研究数据整合与比较带来显著障碍。例如,单细胞数据本身存在高通量试验引入的技术噪声和批次效应,这些问题在多模态整合分析中会进一步加剧。此外,不同的模态数据通常独立获取,多为非配对数据,且不同模式的测序数据具有不同特征空间,在模态对齐过程中易造成信息损失,从而加剧整合分析的难度^[55]。此外,相比人类和小鼠,家畜的细胞标记基因相关的数据库较少,许多稀有细胞亚群缺乏明确的分子标记。现有的分子标记基因数据库(如CellMarker2.0^[12])主要针对人类和小鼠,家畜和家禽特异的标记基因资源有限,影响细胞类型鉴定的准确性。最后,家鸡等非哺乳动物与哺乳动物(如人和小鼠)的细胞类型进化差异较大^[56],直接使用人类和小鼠的细胞注释参考数据集可能导致误判,关于这类非哺乳动物家禽的细胞类型注释方法及参考数据集仍有待进一步完善,以提高家畜和家禽细胞类型鉴定的准确性和可靠性。因此,针对上述这些问题,期望未来数据库的构建能够有以下新的突破。

构建家畜和家禽专属单细胞组学数据库:开发专门针对家猪、家鸡、家牛等畜禽动物的综合性单细胞数据库,并建立不同畜禽细胞类型标记基因数据库。这将为畜禽研究提供更全面的资源,弥补当前数据规模和覆盖范围的不足,同时为稀有细胞亚群的鉴定提供更准确的分子标记,从而促进畜禽分子育种和疾病研究。

开发多模态数据分析工具:单细胞测量技术的进步使以细胞分辨率获取多模态数据成为可能,包括RNA、DNA以及蛋白等层面。随着数据量的快速增长,开发基于深度学习的多模态整合工具变得尤为迫切,以更精确地解析细胞状态和行为。目前多模态分析主要面临数据整合与跨模态推算的难题。Tang等^[57]提出的无监督学习方法Monae,可同

时实现多模态数据的整合与缺失信息推算,为异构数据融合和互补推断提供了方法论参考。畜禽研究面临高质量配对多模态数据稀缺、成本高昂的瓶颈,严重限制了相关进展。因此,开发适用于畜禽跨模态整合与推算的分析方法,对深入解析畜禽复杂性状和推动分子育种具有重要价值。

开发畜禽单细胞注释工具:目前,单细胞注释方法主要分为无监督聚类与监督学习两类,但现有方法多数基于人类或小鼠数据构建。例如,scDeepSort^[58]以人类和小鼠单细胞转录组图谱为基础数据,针对不同组织分别训练模型;该工具采用加权图神经网络(GNN)作为核心深度学习架构,在免疫细胞亚型注释任务中表现出较高的准确性^[59]。借鉴此类基于深度学习与已知细胞图谱构建监督式注释模型的方法,开发畜禽特异的细胞类型注释工具,有望显著提升细胞类型鉴定的准确性与可靠性,有效减少因物种差异所引起的细胞类型误判,从而为畜禽单细胞数据的应用提供更坚实的技术支撑。

多组学整合数据库构建:构建整合 scRNA-seq、单细胞全基因组(scWGS)、染色质可及性(scATAC-seq)及空间转录组等多模态数据的畜禽多组学数据库,已成为当前研究的重要挑战^[60]。通过多层次数据的融合与联合分析,能够系统解析细胞状态调控机制、基因调控网络及微环境异质性,为重要经济性状的解析、抗逆育种和生殖发育研究提供深入的数据支持与机制见解^[61]。此外,可进一步融入表型组学思路,借鉴 ChickenGTEx^[39]等项目的策略,整合单细胞基因表达与宏观表型数据,建立“基因型-细胞特异性表达-表型”的关联网络,增强对复杂性状遗传机制的多层次解析。

综上所述,畜禽单细胞数据库的开发目前仍处于发展初期,面临着数据整合、注释标准以及计算分析等多方面的挑战。未来,需深度融合多组学技术、人工智能以及跨学科合作,构建高精度的畜禽单细胞组学数据库,为畜禽领域研究提供坚实支撑。同时,建立标准化的数据注释与分析流程,开发更高效的生物信息学工具,将显著提升数据整合与分析的准确性和效率。随着测序技术的持续优化与成本降低,人们有望获取更大规模、更全面的单细胞数据,涵盖更多畜禽物种和组织类型。因此,畜禽单细胞数据库的不断完善以及相关单细胞组学分析技术的持续进步,将使其在器官发育、疾

病机制、遗传资源保护与利用等领域的应用更加深入,为畜禽产业的可持续发展提供有力的数据资源支持。

参考文献(References):

- [1] YAN Y, ZHU S, JIA M, et al. Advances in single-cell transcriptomics in animal research [J]. *J Anim Sci Biotechnol*, 2024, 15(1):102.
- [2] LU Y, LI M, GAO Z, et al. Innovative insights into single-cell technologies and multi-omics integration in livestock and poultry [J]. *Int J Mol Sci*, 2024, 25(23): 12940.
- [3] CAI S, HU B, WANG X, et al. Integrative single-cell RNA-seq and ATAC-seq analysis of myogenic differentiation in pig [J]. *BMC Biol*, 2023, 21(1):19.
- [4] LI J, YANG D, CHEN C, et al. Single-cell RNA transcriptome uncovers distinct developmental trajectories in the embryonic skeletal muscle of Daheng broiler and Tibetan chicken [J]. *BMC Genomics*, 2025, 26(1): 187.
- [5] YANG H, MA J, WAN Z, et al. Characterization of sheep spermatogenesis through single-cell RNA sequencing [J]. *FASEB J*, 2021, 35(2):e21187.
- [6] ZHANG Q, WANG Q, ZHENG J, et al. Single-cell RNA sequencing of the spleen reveals differences in Salmonella typhimurium infection mechanisms between different chicken breeds [J]. *Poult Sci*, 2025, 104(2): 104669.
- [7] FAN B, ZHOU J, ZHAO Y, et al. Small intestine using single-cell RNA sequencing [J]. *J Immunol*, 2023, 210(3):271-282.
- [8] HANEDA Y, MIYAGAWA-TOMITA S, UCHIJIMA Y, et al. Diverse contribution of amniogenic somatopleural cells to cardiovascular development: With special reference to thyroid vasculature [J]. *Dev Dyn*, 2024, 253(1):59-77.
- [9] WANG L, GAO P, LI C, et al. A single-cell atlas of bovine skeletal muscle reveals mechanisms regulating intramuscular adipogenesis and fibrogenesis [J]. *J Cachexia Sarcopeni*, 2023, 14(5):2152-2167.
- [10] HE T, GUO W, YANG G, et al. A single-cell atlas of an early mongolian sheep embryo [J]. *Vet Sci*, 2023, 10(9):543.
- [11] SANTIAGO CP, GIMMEN MY, LU Y, et al.

- Comparative analysis of single-cell and single-nucleus RNA-sequencing in a rabbit model of retinal detachment-related proliferative vitreoretinopathy [J]. *Ophthalmol Sci*, 2023, 3(4):100335.
- [12] HU C, LI T, XU Y, et al. CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data [J]. *Nucleic Acids Res*, 2023, 51(D1):D870-D876.
- [13] LI Y, YANG T, LAI T, et al. CDCP: a visualization and analyzing platform for single-cell datasets [J]. *J Genet Genomics*, 2022, 49(7):689-692.
- [14] FRANZÉN O, GAN LM, BJÖRKEGREN JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data [J]. *Database (Oxford)*, 2019, 2019:baz046.
- [15] WHITE J. PubMed 2.0 [J]. *Med Ref Serv Q*, 2020, 39(4):382-387.
- [16] CLOUGH E, BARRETT T. The Gene Expression Omnibus Database [J]. *Methods Mol Biol*. 2016, 1418:93-110.
- [17] WANG Y, SONG F, ZHU J, et al. GSA: Genome Sequence Archive [J]. *Genom Proteom Bioinf*, 2017, 15(1):14-18.
- [18] ATHAR A, FÜLLGRABE A, GEORGE N, et al. ArrayExpress update-from bulk to single-cell expression data [J]. *Nucleic Acids Res*, 2019, 47(D1):D711-D715.
- [19] BURGIN J, AHAMED A, CUMMINS C, et al. The European nucleotide archive in 2022 [J]. *Nucleic Acids Res*, 2023, 51(D1):D121-D125.
- [20] YOU Y, TIAN L, SU S, et al. Benchmarking UMI-based single-cell RNA-seq preprocessing workflows [J]. *Genome Biol*, 2021, 22(1):339.
- [21] WU Q, HU X, ZHANG X, et al. Single-cell transcriptomics of peripheral blood reveals anti-tumor systemic immunity induced by oncolytic virotherapy [J]. *Theranostics*, 2022, 12(17):7371-7389.
- [22] HERRERA-URIBE J, WIARDA J E, SIVASANKARAN S K, et al. Reference transcriptomes of porcine peripheral immune cells created through bulk and single-cell RNA sequencing [J]. *Front Genet*, 2021, 12:689406.
- [23] TIMMERMAN ME, CEULEMANS E, DE ROOVER K, et al. Subspace K-means clustering [J]. *Behav Res Methods*, 2013, 45(4):1011-23.
- [24] BLONDEL VD, GUILLAUME JL, HENDRICKX JM, et al. Local leaders in random networks [J]. *Phys Rev E Stat Nonlin Soft Matter Phys*, 2008, 77(3 Pt 2):036114.
- [25] LI H, WANG X, WANG Y, et al. Cross-species single-cell transcriptomic analysis reveals divergence of cell composition and functions in mammalian ileum epithelium [J]. *Cell Regen*, 2022, 11(1):19.
- [26] FINAK G, MCDAVID A, YAJIMA M, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data [J]. *Genome Biol*, 2015, 16:278.
- [27] LI S, ZHANG Q, LI J, et al. Comprehensive analysis of autoimmune-related genes in amyotrophic lateral sclerosis from the perspective of 3P medicine [J]. *EPMA J*, 2022, 13(4):699-723.
- [28] The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong [J]. *Nucleic Acids Res*, 2019, 47(D1):D330-D338.
- [29] KANEHISA M, GOTO S. KEGG: kyoto encyclopedia of genes and genomes [J]. *Nucleic Acids Res*, 2000, 28(1):27-30.
- [30] CHENG C, CHEN W, JIN H, et al. A review of single-cell RNA-Seq annotation, integration, and cell-cell communication [J]. *Cells*, 2023, 12(15):1970.
- [31] GOPI K, MAZUMDER D, CRAWFORD J, et al. Developing a MySQL database for the provenance of black tiger prawns (*Penaeus monodon*) [J]. *Foods*, 2023, 12(14):2677.
- [32] D L R, H V A, HONNAVALI PB, et al. The MongoDB injection dataset: a comprehensive collection of MongoDB-NoSQL injection attempts and vulnerabilities [J]. *Data Brief*, 2024, 54:110289.
- [33] ZHU J, CHEN F, LUO L, et al. Single-cell atlas of domestic pig cerebral cortex and hypothalamus [J]. *Sci Bull (Beijing)*, 2021, 66(14):1448-1461.
- [34] ABUÍN JM, LOPES N, FERREIRA L, et al. Big data in metagenomics: Apache Spark vs MPI [J]. *PLoS One*, 2020, 15(10):e0239741.
- [35] ADIL A, BHATTACHARYA N, AADAM, et al. Enabling scalable single-cell transcriptomic analysis through distributed computing with Apache spark [J].

- Sci Rep*, 2025, 15(1):27713.
- [36] TRAN D D, WAI MON T, OGATA K. Transport Layer Security 1.0 handshake protocol formal verification case study: How to use a proof script generator for existing large proof scores [J]. *PeerJ Comput Sci*, 2023, 9:e1284.
- [37] HANWELL M D, DE JONG W A, HARRIS C J. Open chemistry: RESTful web APIs, JSON, NWChem and the modern web application [J]. *J Cheminform*, 2017, 9(1):55.
- [38] JIA L, YAO W, JIANG Y, et al. Development of interactive biological web applications with R/Shiny [J]. *Brief Bioinform*, 2022, 23(1):bbab415.
- [39] GUAN D, BAI Z, ZHU X, et al. Genetic regulation of gene expression across multiple tissues in chickens [J]. *Nat Genet*, 2025, 57(5):1298-1308.
- [40] WU J J, ZHU S, TANG Y F, et al. Age- and microbiota-dependent cell stemness plasticity revealed by cattle cell landscape [J]. *Research (Wash D C)*, 2023, 6:0025.
- [41] HAN B, LI H, ZHENG W, et al. A multi-tissue single-cell expression atlas in cattle [J]. *Nat Genet*, 2025, Epub ahead of print.
- [42] TIAN Y, LIU S, SHI H, et al. Revealing the transcriptional and metabolic characteristics of sebocytes based on the donkey cell transcriptome atlas [J]. *Adv Sci (Weinh)*, 2025, 12(16):e2413819.
- [43] GEORGE N, FEXOVA S, FUENTES A M, et al. Expression Atlas update: insights from sequencing data at both bulk and single cell level [J]. *Nucleic Acids Res*, 2024, 52(D1):D107-D114.
- [44] CHEN Y, ZHANG X, PENG X, et al. SPEED: single-cell pan-species atlas in the light of Ecology and Evolution for Development and Diseases [J]. *Nucleic Acids Res*, 2023, 51(D1):D1150-D1159.
- [45] TARHAN L, BISTLINE J, CHANG J, et al. Single cell portal: an interactive home for single-cell genomics data [J]. *bioRxiv*, 2023, 17:2023. 07. 13. 548886.
- [46] 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(02):421-433.
- LIU Y Y, LI X Y, BA Y N M L, et al. Single-cell transcriptome sequencing technology and its application in animal reproduction [J]. *Acta Veterinaria et Zootechnica Sinica*, 2023, 54(02):421-433. (in Chinese)
- [47] 熊和丽, 沙 茜, 刘韶娜, 等. 单细胞转录组测序技术在动物上的应用研究 [J]. *生物技术通报*, 2022, 38(3):226-233.
- XIONG H L, SHA Q, LIU S N, et al. Application of single-cell transcriptome sequencing in animals [J]. *Biotechnology Bulletin*, 2022, 38(3):226-233. (in Chinese)
- [48] WANG X, WANG Y, WANG Y, et al. Single-cell transcriptomic and cross-species comparison analyses reveal distinct molecular changes of porcine testes during puberty [J]. *Commun Biol*, 2024, 7(1):1478.
- [49] ZHAO Y D, YANG C X, DU Z Q. Integrated single cell transcriptome sequencing analysis reveals species-specific genes and molecular pathways for pig spermiogenesis [J]. *Reprod Domest Anim*, 2023, 58(12):1745-1755.
- [50] QIAN Z, QIN J, LAI Y, et al. Large-scale integration of single-cell RNA-Seq data reveals astrocyte diversity and transcriptomic modules across six central nervous system disorders [J]. *Biomolecules*, 2023, 13(4):692.
- [51] 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 (Weinh)*, 2023, 10(35):e2305080.
- [52] ZHANG S, FANG X, CHANG M, et al. Cross-species single-cell analysis reveals divergence and conservation of peripheral blood mononuclear cells [J]. *BMC Genomics*, 2024, 25(1):1169.
- [53] WANG Y, JIANG Y, NI G, et al. Integrating single-cell and spatial transcriptomics reveals heterogeneity of early pig skin development and a subpopulation with hair placode formation [J]. *Adv Sci (Weinh)*, 2024, 11(20):e2306703.
- [54] VOLDEN R, VOLLMERS C. Single-cell isoform analysis in human immune cells [J]. *Genome Biol*, 2022, 23(1):47.
- [55] CAO Z J, GAO G. Multi-omics single-cell data integration and regulatory inference with graph-linked embedding [J]. *Nat Biotechnol*, 2022, 40(10):1458-1466.
- [56] MURAT F, MBENGUE N, WINGE SB, et al. The molecular evolution of spermatogenesis across

- mammals[J]. *Nature*, 2023, 613(7943): 308-316.
- [57] TANG Z, CHEN G, CHEN S, et al. Modal-nexus auto-encoder for multi-modality cellular data integration and imputation[J]. *Nat Commun*, 2024, 15(1): 9021.
- [58] SHAO X, YANG H, ZHUANG X, et al. scDeepSort: a pre-trained cell-type annotation method for single-cell transcriptomics using deep learning with a weighted graph neural network[J]. *Nucleic Acids Res*, 2021, 49(21): e122.
- [59] FU Q, DONG C, LIU Y, et al. A comparison of scRNA-seq annotation methods based on experimentally labeled immune cell subtype dataset[J]. *Brief Bioinform*, 2024, 25(5): bbae392.
- [60] HU Y, AN Q, SHEU K, et al. Single cell multi-omics technology: methodology and application [J]. *Front Cell Dev Biol*. 2018, 6: 28.
- [61] 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 transcriptome sequencing technology in domesticated animal [J]. *Acta Veterinaria et Zootechnica Sinica*, 2024, 55(8): 3276-3287. (in Chinese)
- (编辑 郭云雁)