

doi: 10.13241/j.cnki.pmb.2018.01.034

· 专论与综述 ·

单细胞转录组测序数据分析算法在干细胞研究中的应用*

杜璐 王少军 彭广华[△]

(中国人民解放军总医院眼科 北京 100853)

摘要:细胞的转录组决定其生理状态,每个细胞的转录组都是唯一的。借助单细胞转录组测序可分析单个干细胞的转录组特征,通过进一步的运算方法可以根据转录组特征对细胞进行细胞状态测定以及系谱分化特征的重建,在干细胞及组织发育研究中发挥了强大的作用,推动了其快速发展,加速了对干细胞分化及组织发育的相关过程及调控路径的认识。尤其是在干细胞领域的应用,得益于新算法的发展,单细胞转录组测序分析可用于阐述干细胞的起源、异质性,尤其是对干细胞的分化过程进行连续观察。本文主要对应用于干细胞分化相关研究的单细胞转录组测序数据新的算法及其应用进行了综述。

关键词:干细胞;分化;转录调控网络;单细胞转录;组测序;算法

中图分类号:Q-33;R331.2 **文献标识码:**A **文章编号:**1673-6273(2018)01-155-05

Application of Algorithms in the Analysis of the Stem Cell
ScRNA-Seq data*DU Lu, WANG Shao-jun, PENG Guang-hua[△]

(Department of Ophthalmology, Chinese PLA General Hospital, Beijing, 100853, China)

ABSTRACT: The transcriptome of a cell can be interpreted as a fingerprint, revealing its identity. Recent advances in single-cell RNA-sequencing (scRNA-seq) provide the feasibility of measuring the cellular heterogeneity and dynamic changes of individual cells during stem cell differentiation. ScRNA-seq will be useful for a better understanding of differentiation dynamics in a variety of stem cells. Several algorithms have been developed to quantitatively measure differentiation states of individual cells and reconstruct their lineages from scRNA-seq data. In this article, we mainly review the typical algorithms used in stem cell research.

Key words: Stem cell; Differentiation; Transcriptional regulatory networks (TRNs); Single cell RNA sequencing; Group sequencing; Algorithm

Chinese Library Classification(CLC): Q-33; R331.2 **Document code:** A

Article ID: 1673-6273(2018)01-155-05

前言

皮肤、小肠以及血液等组织终生都在持续更新,组织内干细胞有序更新及分化发挥重要作用。鉴定干细胞并阐明其分化相关的机制对于揭示组织稳态、损伤修复及疾病发生的机制具有重要理论价值^[1-3]。干细胞是存在于组织内数量稀少的一类细胞群体,特异性标记物的缺乏导致了纯化困难,阻碍了揭示干细胞的生物学特征^[4,5]。以往对于干细胞状态的鉴定主要依赖于已知的分化相关基因表达情况及细胞形态来确定^[6,7]。干细胞分化过程是一个非常复杂、受多基因网络协调控制的生物学过程。研究表明干细胞分化过程中大多数基因呈现出细胞间高度异质性,即使是关键的分化调控因子也不例外^[8-10]。以往实验方法不能解析干细胞分化过程异质性的内在机制^[11,12]。令人兴奋的是干细胞的异质性可被单细胞转录组测序 (single cell RNA sequencing, ScRNA-Seq) 技术解析^[13]。随着 ScRNA-Seq 技术的

发展,可同时对成千上万个细胞进行转录组测序,用于不同类型干细胞,比如胚胎干细胞、诱导多能干细胞、造血干细胞及肺泡祖细胞等的异质性分析^[10,12,14-16]。干细胞启动分化后经历一系列连续细胞状态的改变^[9],分析分化过程中细胞间状态差别,可揭示细胞间的调控基因的表达差异^[17]。

应用 ScRNA-Seq 技术结合正确的计算方法,对每一个细胞的转录组数据进行深度挖掘分析。在不使用传统方法纯化出特定类型细胞的基础上就可以分析某类细胞的生物学特征,确定干细胞所处的状态、分化时谱系选择以及分化过程的关键调控基因及信号网络^[18]。阐明单个细胞的分化及干细胞状态有利于预测调控细胞在器官形成、修复及疾病状态下的动态基因调控网络(transcriptional regulation networks, TRNs)。目前,计算方法的改进极大地促进了 ScRNA-Seq 在干细胞分化过程中的应用。因此,本文主要综述了近年应用于干细胞 ScRNA-Seq 数据的具有代表性的计算算法。

* 基金项目:国家重点基础研究发展计划“973 计划”资助项目(2013CB967001);国家自然科学基金项目(81501090)

作者简介:杜璐(1984-),博士后,主要研究方向:视网膜感光细胞发育, E-mail: 13594018494@163.com

△ 通讯作者:彭广华(1958-),博士生导师,教授,主要研究方向:视网膜感光细胞发育, E-mail: peng63088@163.com, 电话:010-66931194

(收稿日期:2017-07-30 接受日期:2017-08-25)

1 Monocle 根据转录组相似度进行干细胞分化排序分析

使用 Monocle 分析 ScRNA-Seq 数据时,将选取的不同分化阶段的细胞重新根据转录组的相似度进行排列,根据细胞的分化程度排列干细胞的分化顺序,而非按照取样时间排列,能够显示出分化过程中细微的基因表达动态变化。Monocle 的优势是不用经过传统的基于免疫技术纯化细胞,可避免改变细胞外的环境,从而重现干细胞在体内与微环境的相互作用。

应用 Monocle 分析经典的成肌分化模型,精准地分辨出单个细胞的转录组在时间轴上的动态变化特征,可以把骨骼肌干细胞分化过程中产生的各类细胞根据分化进程进行排序,揭示调控因子的动态变化。深度挖掘数据后发现关键调控因子表达水平呈现开关样改变,以及基因表达随时间的序贯波形改变,并发现新的调控细胞分化的分子。分化过程中很多抑制性调控因子通过与原癌基因共用调控原件发挥调控下游基因的功能也被发现。骨骼肌干细胞在分化过程中全转录组在逐渐变化,伴随细胞分化进程,细胞命运可能分叉为两个甚至更多种系谱。Monocle 根据干细胞分化的程度排列细胞,重建分化过程的轨迹,发现系谱关系,可用于绘制干细胞分化的基因调控网络^[19]。

Wanderlust 算法、SCUBA 算法和 Waterfall 算法均与 Monocle 具有类似的运算原理^[6,11,20]。结合 ScRNA-Seq 数据与 Wanderlust 算法精确推测了人 B 淋巴细胞的生成过程,重建了从造血干细胞到幼稚 B 淋巴细胞的分化全过程,并且绘制出基因调控网络的调控关键点^[6]。运用 SCUBA 算法来提取 ScRNA-Seq 数据中的细胞系谱关系,重建小鼠早期胚胎发育过程中的细胞分化及命运决定的轨迹,预测出关键调控因子在决定细胞谱系选择中的作用^[20]。还有学者使用 Waterfall 算法发现调控成年动物神经元新生的分子机制^[11]。

值得注意的是,上述基于维度减少的假设时间轴的算法尚存在一些缺陷:首先,这类算法不能精确定量细胞的状态,根据转录组的相似度而非细胞的真实状态,在假设的分化序列时间轴上对细胞进行定位、分类排序,也不足以估测细胞早期分化阶段的系谱特征。第二,解释简化的分化路径的生物学意义存在困难。第三,在简化的空间内细胞的排列会被噪音及与分化不相关的基因表达干扰,影响数据分析的准确性。第四,除 Monocle 算法,其他几种算法还要依赖于某些额外的信息,比如细胞分化的时间信息、细胞特征或者标记基因的表达来决定一个动态过程的起点与终点^[19]。

2 依赖于亚群的 TRNs 算法

干细胞由多个亚细胞群组成,学者认为每种亚群的细胞状态由不同的内在 TRNs 决定,不同的 TRNs 决定了干细胞分化时其命运倾向于某一方向^[12]。两个研究小组通过重建单个 TRNs,预测出系谱选择相关的特异性决定因子^[21,22]。因此,根据不同的细胞亚群构建出不同的 TRNs 可能更加适合分析干细胞分化过程中的转录调控机制。首先,重建细胞亚群特异性的

TRNs。然后,建立模型(每个母体干细胞亚群处于一个特定的 TRN 调控下的稳定状态),根据这个模型预测系谱特异的决定因子。这种母体干细胞稳定状态的维持取决于朝相反方向系谱分化的一对基因的互相牵制。进一步,根据统计方法来确定分化时相反的系谱决定因子的比例发生的改变。学者运用该算法分析了三种细胞分化系统的 ScRNA-Seq 数据,成功预测并捕捉到新的系谱决定因子。这三组数据分别是内细胞群细胞分化为原始内胚层或者外胚层^[9],造血细胞的分化过程^[23],肺泡祖细胞分化过程^[16]。在第一个例子中,预测出 Gata6 为原始外胚层, Klf2 为外胚层的特异性系谱决定因子,这个结果与以往的实验报道一致^[24-26]。而且,造血系统很多已知的系谱决定因子都被成功预测。该算法适用于鉴定分化过程中系谱决定的特异性调控因子以及相关的 TRNs,解析干细胞分化过程的系谱选择机制以及阐明再生医学相关的分子调控机制^[27]。令人感兴趣的是该方法有助于设计出更高效的干细胞分化策略及步骤,可促进再生医学的发展^[13]。

3 StemID 根据转录组特征建立单个细胞标签

细胞的转录组可以作为识别细胞的标记,基于转录组的差异分析可以区分出细胞的类型。通过转录组建立干细胞的标签,进一步可以推测组织的干细胞特征^[28,29]。目前的算法假设在细胞分化进程中转录组呈现出连续动态变化,然后重新建立出系谱树形拓扑结构^[6,19,30]。有学者发明了一种新的算法 StemID,通过更多的初始分类步骤引入 RaceID 的升级版视图^[31]。把 k-均值替换为 k-中位数可以提升数据的更多集群能力。为了推测出分化轨迹,根据细胞的位置,把每个细胞与其起源的亚群以及其他亚群假设为特殊的关联。这样就能充分反映出细胞从一个亚群的状态到另一种细胞亚群移动了多少距离。当一个已知明显的关联程度降低时,提示一个细胞可能向另外一种命运震荡。

运用该算法可重建出干细胞系谱分化的分化发育树。通过分析分化发育树的拓扑结构以及转录组组成,StemID 可鉴定出多种混杂细胞类型中的干细胞群体。StemID 分数评分与细胞的多能性呈正相关。通过运用该算法,学者成功鉴定出两种重要的干细胞:小肠的 Lgr5⁺ 细胞和骨髓内的造血干细胞。小肠的 Lgr5⁺ 细胞和骨髓内的造血干细胞的 StemID 分数评分在相应的组织细胞中最高,这也证明了该算法的准确性。进一步在肺上皮和人放射状胶质细胞中也验证了该算法的准确性^[16,32]。StemID 算法能加速发现新的干细胞,发现确切的标记物^[1]。值得注意的是,运用该算法学者还发现新的干细胞分化路径, Lgr5⁺ 细胞直接分化为潘氏细胞,而不通过 Dll1⁺ 祖细胞的中间过渡态^[33]。StemID 推论出小肠 Lgr5⁺ 细胞和骨髓来源的造血干细胞的系谱分化发育树的精确程度比以往其他方法的精确度更高^[19,30]。推测的骨髓来源的造血干细胞系谱分化发育树提示,与经典的祖细胞二分法分化模式不同的是,细胞的命运在 HSC 阶段就已经出现了分化命运的偏好性^[34,35]。StemID 算法对于更好地了解器官内组织更新及干细胞的动态过程及内在机制有很重要的价值。

StemID 这个算法的缺陷在于缺少中间过渡态祖细胞以及可能重现出不相关的细胞类型。由于缺少中间过渡态祖细胞, StemID 推论出与更高潜能的细胞群假阳性关联。比如,骨髓中的 B 淋巴细胞直接与 HSCs 相关联。不相关细胞类型的混杂可以导致假阳性关联的推导。

4 SLICE 根据转录组对细胞状态定量分析、推测细胞分化轨迹

SLICE (Single Cell Lineage Inference Using Cell Expression Similarity and Entropy) 是根据 ScRNA-Seq 数据来定量检测单个细胞的分化状态以及重建系谱分化特征。该算法中熵数是一个测定细胞异质性的指标: 熵数低的细胞表现出窄谱的、特征明确的基因表达方式, 而熵数高的细胞具有广泛的、多样的基因表达方式。因此, 熵数与干细胞的分化状态呈现负相关: 熵数高的细胞具有更高的分化潜能, 而熵数低的细胞分化潜能低。SLICE 主要包括两个主要功能: 根据计算得出的细胞熵数来定量测定细胞的分化状态, 进一步重建细胞分化轨迹来推测细胞分化的谱系^[36]。通过运用三组独立的且已明确分化谱系及发育时间节点的单细胞转录组测序数据验证了该算法的精确度, 均能成功检测出单个细胞的分化状态, 发现干细胞熵数在分化过程中逐渐降低, 并且根据细胞熵数准确重建出细胞分化发育树^[16,19,37]。例如, 以往对于小鼠肺间质细胞分化发育过程认识尚不清楚, 运用 SLICE 处理胚鼠来源的肺组织 ScRNA-Seq 数据, 发现了新的肺间质细胞系谱关系。鉴定出肺囊纤维化的形成时肺泡祖细胞分化生成的两个分支的 5 种类型细胞, 并确定了细胞分化状态的相对顺序^[38,39]。

SLICE 基于 ScRNA-Seq 数据基础上预测细胞分化状态及重建细胞分化发育系谱结构具有广泛的应用范围及高度的预测准确性, 可以在不需要额外信息的前提下确定细胞的分化状态以及分化时的系谱路径选择, 远优于以往的算法。学者比较了 SLICE 与 Monocle 在处理单细胞转录组测序数据中的优劣^[36]。当数据来源于同一均质群体时, 这两种算法的表现相当。当处理无时间信息的断层数据时, SLICE 能够无偏倚地鉴定出分化的终点以及路径的转换方向; 当处理异质性细胞群体的数据时, SLICE 能正确地发现系谱分支以及根据每个分支重建出细胞状态的转换路径。在疾病、损伤及其他特定情况下, 不同的细胞可能对刺激的反应呈现出不同的动态变化, 而且关于疾病病程相关的标记物也不清楚。因此, SLICE 可用来分析调控细胞命运及功能的信号途径, 疾病相关的分子调控机制, 为开发出新疗法提供帮助。SLICE 可能推广到除了发育学研究以外的领域, 比如肿瘤、创伤以及其他疾病的研究。

然而, 同时我们也认识到 GO terms 的不完善, 有自身的局限性。在以后的研究中还需要测试联合使用其他分类算法, 比如 K- 均值来验证这些方法提升单细胞的基因功能分类^[40]。

5 Div-Seq 跟踪分析中枢神经系统新生神经元

在成年哺乳动物中枢神经系统分离出新生神经元非常困难, 单细胞化过程可能会损伤神经元胞体完整性, 很难观测到

神经生成的动态过程。学者通过结合 sNuc-Seq 与 EdU 标记分裂的神经元设计出 Div-Seq。与遗传标记方法不同的是, EdU 可以随时标记增殖的细胞, 可以在时间上精准地标记干细胞及其子代细胞^[11,41]。使用 Div-Seq 来分析细胞分裂后 14 天内多时间点的神经生成区域 --DG 区^[42], 包含了多种新生神经元类型及神经生成阶段, 跨越了神经干细胞增殖到新生神经元的各个阶段^[42]。BiSNE 算法连续分析 DG 新生神经元的分化轨迹的神经元系谱的细胞核。细胞核在分化轨迹的排列与 EdU 标记时间相一致, 提示这个轨迹本质上是神经元成熟的过程^[11,43,44]。sNuc-Seq 对于细胞核的分析能够忠实反应组织的 RNA 水平。sNuc-Seq 数据表明特征性的细胞核的分群对应于已知的细胞类型及在海马内的解剖学定位。分析结果与纤维切割、原位杂交及基于群体的 RNA 测序结果一致^[45]。这个算法能够敏感地鉴定出海马相关细胞类型以及追踪成年海马神经发生区域的新生神经元的动态转录组变化。随着技术的进一步发展, 其敏感性、通量及能鉴定出的细胞类型均会显著增加。Div-Seq 开辟了分析中枢神经系统新生神经元多样性及动态过程的道路, 该方法还可以用于追踪神经元的成熟过程^[46]。

6 SCOUP 优于 Monocle, 尤其适用于分析分化的最早期事件

学者开发了一个新的基于 Ornstein-Uhlenbeck (OU) process 的算法 SCOUP, 认为一个变量朝向一个引力点做 Brownian 运动。该模型的干细胞分化过程, 吸引力被认为是细胞分化后基因表达模式的稳定状态。因此, OU 过程是适合描述干细胞分化过程的基因动态表达变化。由于 OU 过程只能假设一种吸引力, 不能分析多谱系的分化过程, 学者扩展了经典的 OU 过程, 开发了混合的 OU 过程来代表每一种细胞的命运及系谱关系。

SCOUP 能够直接描述分化过程中基因表达动态变化过程, 用于分析细胞分化全过程, 包括个别细胞的命运决定。与以往的基于维度减法相关的算法不同的是, SCOUP 直接描述分化过程中基因表达的动态变化, 包括细胞的分化程度及细胞命运决定, 并精确地估算出细胞的系谱, 比 Monocle 的精确度更高, 尤其是具有分析细胞分化早期命运的选择性优势。同时, SCOUP 还能用于一系列的下游分析, 阐明基因间的调控关系。运用这个方法分析 ScRNA-Seq 数据, 检测出分化调控的关键调控因子, 而且把他分类到相关的网络。对数据进行标准化之后, 然后计算标准数据间的关系, 而这种方法检测的协方差, 其他模型中不能检测出。使用该方法发现了 ScRNA-Seq 数据中一些调控分化的关键调控因子, 而使用常规的方法未能检出这些分子。基于 SCOUP 开发的原理基础考虑, 这个算法不仅适用于分析 ScRNA-Seq 数据的假设时间推测以及细胞谱系的推算, 还适用于基因调控网络的推测。SCOUP 能反映出连续时间随机的动态变化, 适用于分析时间序列上的一系列数据组, 尤其是干细胞分化进程的研究^[47]。

7 dpath 基于 meta-gene 熵数确定细胞状态, 预测分化

轨迹,发现中途退出的事件

学者发展了一个观点 meta- 基因熵,把 self-organizing map (SOM) and random walk with restart (RWR)(SOM 和 RWR)结合起来发明了一个新的命名为 dpath 的算法。运用该算法把干细胞与分化的子代细胞分离出来,并且无偏倚重建细胞谱系的层次关系。与现在一些算法进行比较,发现该算法可以还原解析并定量细胞的状态,特别是会考虑到中途退出的事件。Etv2 在小鼠胚胎发育过程中的表达窗口期非常短,大概是从 E7.0 到 E8.5 之间^[48,49]。通过分析 Etv2-EYFP 转基因小鼠的胚胎细胞 ScRNA-Seq 数据,发现了特异的主导造血-内皮系谱分化的分子机制。但是研究表明 Etv2 在间质层分化过程中发挥关键作用,尤其是血液、内皮及心脏的谱系选择中发挥关键作用。使用 dpath 算法破译了 Etv2 阳性细胞的三种主要谱系,鉴定出内皮及血液系谱的关键基因及祖细胞的信号调控路径。目前常规应用细胞表面标记物来分选细胞,但是把干细胞分化的祖细胞与子代细胞分离存在困难。运用这个算法分析每一个细胞的转录组,可提供足够的信息发现新的标记物,来提升细胞亚群的纯度并便于功能分析。Dpath 必将易化并解析控制干细胞及祖细胞群体的生物特征的内在机制。这个算法必将能发现新的调控因子及新的调控路径,推进干细胞的研究^[50]。

8 小结与展望

单细胞转录组测序分析技术可定量分析单个细胞所处的状态,尤其是可分析干细胞的分化状态。因此,该技术出现后被广泛应用于研究组织的发育分化及干细胞的相关研究,对于我们进一步认识涉及到发育及干细胞自我更新、分化的相关基因调控机制提供很好的平台。目前随着技术平台的改进,极大地增强了测序通量与精准度,但是由此也产生了海量数据,因此,亟待开发出有效的分析算法来深度挖掘测序产生的数据库代表的生物学意义,进而阐释相关的生物进程中的具体调控机制。虽然,目前有几种算法可分析单细胞转录组数据,但是仍然有改善并开发新算法的空间,来更精准地分析细胞分化状态及分析细胞谱系的产生过程。随着算法的改进,必将加速我们全面认识机体组织发育以及干细胞分化的过程,揭示其分子调控网络机制,为研究疾病的发生提供理论依据,并有助于开发出新的治疗方案。

参考文献(References)

- [1] Grun D, Muraro MJ, Boisset JC, et al. De novo prediction of stem cell identity using single-cell transcriptome data[J]. Cell Stem Cell, 2016, 19(2): 266-277
- [2] Barker N. Adult intestinal stem cells: Critical drivers of epithelial homeostasis and regeneration[J]. Nat Rev Mol Cell Biol, 2014, 15(1): 19-33
- [3] Busch K, Klapproth K, Barile M, et al. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo [J]. Nature, 2015, 518(7540): 542-546
- [4] Eaves CJ. Hematopoietic stem cells: Concepts, definitions, and the new reality[J]. Blood, 2015, 125(17): 2605-2613
- [5] Buczacki SJ, Zecchini HI, Nicholson AM, et al. Intestinal label-retaining cells are secretory precursors expressing lgr5 [J]. Nature, 2013, 495(7439): 65-69
- [6] Bendall SC, Davis KL, Amir el AD, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human b cell development[J]. Cell, 2014, 157(3): 714-725
- [7] Chalut KJ, Ekpenyong AE, Clegg WL, et al. Quantifying cellular differentiation by physical phenotype using digital holographic microscopy [J]. Integrative biology: quantitative biosciences from nano to macro, 2012, 4(3): 280-284
- [8] Shalek AK, Satija R, Adiconis X, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells [J]. Nature, 2013, 498(7453): 236-240
- [9] Guo G, Huss M, Tong GQ, et al. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst[J]. Dev Cell, 2010, 18(4): 675-685
- [10] Tang F, Barbacioru C, Bao S, et al. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell rna-seq analysis [J]. Cell Stem Cell, 2010, 6(5): 468-478
- [11] Shin J, Berg DA, Zhu Y, et al. Single-cell rna-seq with waterfall reveals molecular cascades underlying adult neurogenesis[J]. Cell Stem Cell, 2015, 17(3): 360-372
- [12] Moignard V, Macaulay IC, Swiers G, et al. Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis[J]. Nat Cell Biol, 2013, 15(4): 363-372
- [13] Okawa S, del Sol A. A computational strategy for predicting lineage specifiers in stem cell subpopulations[J]. Stem Cell Res, 2015, 15(2): 427-434
- [14] Guo G, Luc S, Marco E, et al. Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire[J]. Cell Stem Cell, 2013, 13(4): 492-505
- [15] Buganim Y, Faddah DA, Cheng AW, et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase[J]. Cell, 2012, 150(6): 1209-1222
- [16] Treutlein B, Brownfield DG, Wu AR, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell rna-seq[J]. Nature, 2014, 509(7500): 371-375
- [17] Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science [J]. Nat Rev Genet, 2013, 14(9): 618-630
- [18] Wu AR, Wang J, Streets AM, et al. Single-cell transcriptional analysis [J]. Annual review of analytical chemistry, 2017, 10(1): 439-462
- [19] Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells[J]. Nat Biotechnol, 2014, 32(4): 381-386
- [20] Marco E, Karp RL, Guo G, et al. Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape [J]. Proc Natl Acad Sci U S A, 2014, 111(52): E 5643-5650
- [21] Xu H, Ang YS, Sevilla A, et al. Construction and validation of a regulatory network for pluripotency and self-renewal of mouse embryonic

- stem cells[J]. PLoS computational biology, 2014, 10(8): e1003777
- [22] Moignard V, Woodhouse S, Haghverdi L, et al. Decoding the regulatory network of early blood development from single-cell gene expression measurements[J]. Nat Biotechnol, 2015, 33(3): 269-276
- [23] Guo G, Luc S, Marco E, et al. Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire[J]. Cell Stem Cell, 2013, 13(4): 492-505
- [24] Fujikura J, Yamato E, Yonemura S, et al. Differentiation of embryonic stem cells is induced by gata factors [J]. Genes Dev, 2002, 16(7): 784-789
- [25] Yeo JC, Jiang J, Tan ZY, et al. Klf2 is an essential factor that sustains ground state pluripotency[J]. Cell Stem Cell, 2014, 14(6): 864-872
- [26] Gillich A, Bao S, Grabole N, et al. Epiblast stem cell-based system reveals reprogramming synergy of germline factors [J]. Cell Stem Cell, 2012, 10(4): 425-439
- [27] Li VC, Kirschner MW. Molecular ties between the cell cycle and differentiation in embryonic stem cells [J]. Proc Natl Acad Sci U S A, 2014, 111(26): 9503-9508
- [28] Brennecke P, Anders S, Kim JK, et al. Accounting for technical noise in single-cell rna-seq experiments [J]. Nature methods, 2013, 10(11): 1093-1095
- [29] Grun D, Kester L, van Oudenaarden A. Validation of noise models for single-cell transcriptomics [J]. Nature methods, 2014, 11(6): 637-640
- [30] Haghverdi L, Buettner F, Theis FJ. Diffusion maps for high-dimensional single-cell analysis of differentiation data [J]. Bioinformatics, 2015, 31(18): 2989-2998
- [31] Grun D, Lyubimova A, Kester L, et al. Single-cell messenger rna sequencing reveals rare intestinal cell types [J]. Nature, 2015, 525(7568): 251-255
- [32] Pollen AA, Nowakowski TJ, Chen J, et al. Molecular identity of human outer radial glia during cortical development [J]. Cell, 2015, 163(1): 55-67
- [33] Farin HF, Karthaus WR, Kujala P, et al. Paneth cell extrusion and release of antimicrobial products is directly controlled by immune cell-derived ifn-gamma [J]. The Journal of experimental medicine, 2014, 211(7): 1393-1405
- [34] Giebel B, Punzel M. Lineage development of hematopoietic stem and progenitor cells[J]. Biological chemistry, 2008, 389(7): 813-824
- [35] Paul F, Arkin Y, Giladi A, et al. Transcriptional heterogeneity and lineage commitment in myeloid progenitors [J]. Cell, 2015, 163(7): 1663-1677
- [36] MacArthur BD, Lemischka IR. Statistical mechanics of pluripotency [J]. Cell, 2013, 154(3): 484-489
- [37] Yan L, Yang M, Guo H, et al. Single-cell rna-seq profiling of human preimplantation embryos and embryonic stem cells [J]. Nature structural & molecular biology, 2013, 20(9): 1131-1139
- [38] Du Y, Guo M, Whitsett JA, et al. 'Lunggens': A web-based tool for mapping single-cell gene expression in the developing lung [J]. Thorax, 2015, 70(11): 1092-1094
- [39] Guo M, Wang H, Potter SS, et al. Sincera: A pipeline for single-cell rna-seq profiling analysis [J]. PLoS computational biology, 2015, 11(11): e1004575
- [40] Guo M, Bao EL, Wagner M, et al. Slice: Determining cell differentiation and lineage based on single cell entropy [J]. Nucleic Acids Res, 2017, 45(7): e54
- [41] Llorens-Bobadilla E, Zhao S, Baser A, et al. Single-cell transcriptomics reveals a population of dormant neural stem cells that become activated upon brain injury[J]. Cell Stem Cell, 2015, 17(3): 329-340
- [42] Ming GL, Song H. Adult neurogenesis in the mammalian brain: Significant answers and significant questions [J]. Neuron, 2011, 70(4): 687-702
- [43] Tasic B, Menon V, Nguyen TN, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics[J]. Nat Neurosci, 2016, 19(2): 335-346
- [44] Schouten M, Buijink MR, Lucassen PJ, et al. New neurons in aging brains: Molecular control by small non-coding rnas [J]. Frontiers in neuroscience, 2012, 6: 25
- [45] Lein ES, Hawrylycz MJ, Ao N, et al. Genome-wide atlas of gene expression in the adult mouse brain [J]. Nature, 2007, 445(7124): 168-176
- [46] Habib N, Li Y, Heidenreich M, et al. Div-seq: Single-nucleus rna-seq reveals dynamics of rare adult newborn neurons [J]. Science, 2016, 353(6302): 925-928
- [47] Matsumoto H, Kiryu H. Scoup: A probabilistic model based on the ornstein-uhlenbeck process to analyze single-cell expression data during differentiation[J]. BMC bioinformatics, 2016, 17(1): 232-248
- [48] Rasmussen TL, Kweon J, Diekmann MA, et al. Er71 directs mesodermal fate decisions during embryogenesis [J]. Development, 2011, 138(21): 4801-4812
- [49] Koyano-Nakagawa N, Kweon J, Iacovino M, et al. Etv2 is expressed in the yolk sac hematopoietic and endothelial progenitors and regulates lmo2 gene expression[J]. Stem Cells, 2012, 30(8): 1611-1623
- [50] Gong W, Rasmussen TL, Singh BN, et al. Dpath software reveals hierarchical haemato-endothelial lineages of etv2 progenitors based on single-cell transcriptome analysis [J]. Nat Communication, 2017, 8: 14362-14374