

胚胎器官发生图谱构建中单细胞测序的应用

赵洵

中国科学院大学

版本: 0.01

日期: August 10, 2019

摘 要

摘要

关键词: 单细胞测序 器官发育 胚胎发育

1 引言

器官发生是胚胎发育过程中自原肠胚形成直到出生的一段时期 (Gilbert and Barresi, 2017), 原肠胚外胚层、中胚层、内胚层不同位置的细胞逐步发育为不同器官, 这个过程需要各类细胞和组织的紧密协调, 受到多种生长因子通路的调控, 且随着时间不断变化 (Zorn and Wells, 2009)。对器官发生过程中单细胞的基因表达分析有着重要意义, 一方面, 准确区分不同的细胞类型有助于建立完整的细胞间相互调控网络体系 (Dong et al., 2018)。另一方面, 为了解器官在正常和病理条件下发育机制的区别, 解析胚胎中准确细胞表达模式是至关重要的 (Belle et al., 2017)。由于早期胚胎多能干细胞具有分化潜力, 对器官发生过程中各类干细胞的分化过程的分析会有助于细胞治疗的发展 (Lancaster and Knoblich, 2014), 对于特定细胞的表达模式分析, 可以诱导其他细胞表达某些基因使其具有其他类型细胞的功能, 例如仅诱导成纤维细胞表达四个转录因子即可使其转变为干细胞 (Takahashi and Yamanaka, 2006)。

相比于传统研究方法, 单细胞 RNA 测序同时具有高分辨率高通量的优点, 经过几年的发展, 已经有专门针对单细胞 RNA 测序而建立的测序方法、分析流程。在胚胎发育的研究中, 样品缺乏、动态性高, 而单细胞 RNA 测序能够很好的解决这些问题, 目前逐渐出现对胚胎发育过程中的各个时期不同器官的研究, 在细胞标记基因、调控通路、分化路径等方面有了不错的成果。

2 单细胞 RNA 测序

DNA 测序技术在 40 年来快速发展, 从最初的 Sanger 测序到现在第二代和第三代例如 Illumina、PacBio、Nanopore 测序技术, 已经极大的提高了测序通量并且降低了测序价格 (Shendure et al., 2017)。随着表达序列标签 (EST)、DNA 元件百科 (ENCODE) 等计划的启动以及对真核细胞转录组的了解日渐加深 (Birney et al., 2007), 人们意识到已有的 DNA 阵列、标签筛选等方法已经不足以应付真核细胞转录组的复杂性, 于是借助高通量测序技术分析转录组数据的 RNA 测序技术 (RNA-seq) 在 2008 年首次出现 (Cloonan et al., 2008), 在一年之后, 即 2009 年, 一方面为了解决某些实验条件下样品量缺乏的问题, 例如早期胚胎发育的研究中并不能获得足量的实验样品, 另一方面借助 cDNA 扩增技术准确性的突破, 出现了单细胞 RNA 测序 (scRNA-seq) (Tang et al., 2009)。

2.1 测序方法

单细胞 RNA 测序大致分为四步。首先是将样品分为单个细胞或单个细胞核，再经过逆转录并对 cDNA 进行 PCR 扩增，最后使用第二代高通量测序技术对扩增后的 cDNA 进行测序。在单细胞分离阶段，可以使用微液滴、微孔芯片、流式细胞术、微移液管技术等将细胞悬液分为单个细胞 (Hedlund and Deng, 2018)。在逆转录阶段，因为测序只需要 mRNA 数据，所以需要使用寡聚胸腺嘧啶或随机六位核酸来去除 rRNA 的影响（占哺乳动物细胞总 RNA 数的 95%），但是实验发现即使不经过 rRNA 清除步骤也不会带来显著的 rRNA 污染，因为常见实验条件还不足以打开 rRNA 的二级结构 (Fan et al., 2015)。常见的单细胞 RNA 测序方法有多种，一共可以归为两类，一类为对 RNA 全长测序 (full-length)，一类为基于人工添加标签的测序 (tag-based)，在面对不同生物学问题上两种类别各有优劣，全长测序有更高的敏感度和可重复性，标签法测序适用于更廉价的大规模测序 (Ziegenhain et al., 2017)。

在胚胎器官发生的研究中，Drop-seq (Macosko et al., 2015) 是一种非常常用的测序方法，Drop-seq 属于基于人工添加标签的测序方法。在分离细胞的过程中，每个细胞会进入一个油滴，油滴里有结合有引物以及标签的微粒。每个微粒上结合有许多引物，在靠近微粒的一段为逆转录时的引物片段，中间一段为细胞标签 (Cell-barcode)，最外侧为单分子计数标签 (Unique Molecule Index, UMI, Molecule-barcode)。一个微粒上都带有同样的细胞标签，用于区分各个油滴里的单细胞，在最初的文献中默认长度为 12bp，所以最多可以区别 4^{12} 个细胞，而在通常实验的细胞数量级上，则可以保证几乎没有相同重复的标签。而在单个微粒上，各个细胞标签所连的单分子计数标签又是各不相同的，因为在 cDNA 扩增时，并非所有的 cDNA 都可以以相同的倍数扩增，原有的表达量信息经过 cDNA 扩增后就变得没有意义了，所以需要单分子计数标签用于统计在最初细胞分离时，即每种 mRNA 连接的不同的单分子计数标签的数目等价于细胞内该种 mRNA 的数量。

在 Drop-seq 测序时，一端为标签序列，即细胞标签和单分子计数标签，另一端为 mRNA 序列。对于测序数据，Drop-seq 的开发者提供了一个方便易用的 Java 包，首先会提取测序一端的细胞标签和单分子计数标签，作为另一端 mRNA 序列的附加注释。在质控流程中，因为低质量测序的两种标签会影响细胞表达量的统计，而低质量的 mRNA 序列会影响基因组比对的结果，所以需要分别对测序两端质控。之后再对 mRNA 比对到基因组上，则这些比对结果根据细胞标签分为各个细胞，每个细胞每个基因的不同单分子计数标签的数目作为这个细胞这个基因的表达量。

2.2 分析方法

单细胞 RNA 测序基于原有 RNA-seq 技术，可以获取单个细胞的转录组信息，以细胞标签和基因名分别为行或列，以某个细胞某个基因的表达量为值，构成表达矩阵 (Digital Expression Matrix, DEG)。表达矩阵可以用于比较不同细胞的基因表达模式，进行细胞类型的注释，展示不同表达模式的细胞类型在时间和空间上的变化 (Potter, 2018)。现在已经有多种快速且方便的分析软件，能辅助处理测序结果和表达矩阵 (Bacher and Kendzierski, 2016)，例如 scanpy (Wolf et al., 2018)、Seurat (Stuart et al., 2019) 等

数据过滤 在测序以及对表达矩阵构建的过程中，已经有对测序质量控制的步骤，但仍然有一些细胞中只检测到了很少量的基因表达，这样的细胞一般来说需要被去除。另外，也可以根据线粒体 mRNA 所占比例来判断这个细胞是否属于低质量细胞。

降维 未经处理的表达矩阵中的每个细胞可以视为在高维空间中的点，而高维空间中的轴即为不同的基因。为了将这样的复杂数据可视化，就需要将高维空间经过降维变为低维空间。最基本的方法是主成分分析 (Principal Components Analysis, PCA)，算法的目的是使用线性变换，将数据点投影到新的坐标轴上，

使得每个坐标轴互不相关（协方差为 0），并尽量使数据点在坐标轴上投影的方差最大 (Pearson, 1901)。但是在单细胞 RNA 测序中，维度往往过高，主成分分析并不能很好的将数据点分离开来，所以需要非线性变换处理数据点 (Becht et al., 2019)。两种常用的非线性降维方法是 tSNE(Maaten and Hinton, 2008) 和 UMAP(McInnes et al., 2018)，经过比较，UMAP 更能分辨出具有细微差异的细胞类群，并且保持着数据集原有的整体结构，保持细胞类群的连续性，让降维结果更易解读 (Becht et al., 2019)。但是不可否认的是，在 UMAP 出现之前，tSNE 仍然具有很好的降维效果。

聚类 经过降维之后，可以在低维空间上将数据点分别划归到各个类群中。常用的聚类方法有 Louvian(Blondel et al., 2008)、SLM(Traag et al., 2018) 等，目的是将空间上邻近的数据点归为一类。而在表达矩阵中，空间上邻近的点即为表达模式相同的点，例如都高表达某个或某一类基因等。

3 胚胎器官发生转录组分析

早在 2003 年，就已经有发育中胰腺的单细胞分析的研究，对 60 个细胞中的大约 100 个基因做了分析，发现细胞表达的基因和其可能的分化结果有关 (Chiang and Melton, 2003)，但是在这个研究中仅仅使用了常规微阵列荧光杂交来表示细胞的基因表达情况，由于实验技术的限制，细胞数量和基因数量远少于现在的单细胞 RNA 测序。在单细胞 RNA 测序中，可以一次性处理分析大量细胞，并可以覆盖整个基因组的所有已知基因，更容易找到单个细胞类型的表达模式规律 (Hedlund and Deng, 2018)。

3.1 构建细胞类型集

细胞是生命的基本单位，基于细胞的特征对不同细胞的描述、分类显得尤为重要，细胞根据特定基因表达的组合被分为不同的类别。但是在一开始，细胞类型的定义主要根据细胞的位置、形状、细胞组分等，后来利用免疫荧光杂交等技术，但是这些研究都依赖于偶然发现的细胞标志物，导致不同研究队伍使用不同标志物时，难以比较研究结果 (Regev et al., 2017)。所以，在胚胎发育过程中，构建每个器官的细胞类型集可以很大程度上方便对细胞类型的注释，并且也可以统一不同研究者对细胞标志物的选择。

小鼠细胞类型集 (Mouse Cell Atlas, MCA) 是一个可以用于查询在某一组织中，某一基因是否可以作为细胞标志物，若是，则可以获得具体是哪一种细胞类型标志物的网站数据库 (Han et al., 2018)。作者对小鼠胚胎各个器官分别取样，再对各个器官的样品进行单细胞 RNA 测序，一方面对各个组织进行聚类分析，另一方面综合所有表达矩阵，对整个数据集进行聚类分析。作者选择了各个类别有显著差异或具有特异性的表达基因，参照过去证实的各个细胞类型的标志基因，注释各个类别的细胞类型。当整个数据集都被注释完成后，比较不同类型细胞的表达模式，就可以判断一个基因是否在一个类群中显著高表达或者具有特异性，以此建立 MCA 数据库。

3.2 解析细胞分化轨迹

上述提到的细胞类型集的构建仅仅是对于同一时期组织的单细胞 RNA 测序，以不同时期的样品重复单细胞 RNA 测序则可以得到不同时期的细胞类型分布，结合时间变化，可以得到一种细胞是如何分化其他细胞类型的，也能得到不同类细胞数量的变化。例如通过表达基因的改变验证神经元沿着放射状胶质细胞迁移过程中的分化过程 (Zhong et al., 2018)，或心脏发育过程中先完成心室心房分化再完成左右分化的过程 (Cui et al., 2019)，或发现肝脏发育过程中胆管细胞可能是从未成熟肝细胞而不是成肝细胞分化而来 (Su et al.,

2017)。因为胚胎器官发生是一个动态的过程，相比于已经成熟的器官，胚胎中未成熟的器官在短时间内细胞类型有着丰富的变化，单细胞 RNA 测序能准确获得各个时间点的发育情况，综合之后就可以得到动态的器官发生细胞动态分化轨迹。

4 器官发生研究进展

自单细胞 RNA 测序被发明以来，已经有很多器官经过单细胞 RNA 测序分析之后，人们对其组分的细胞类型有了更深的了解，例如肠 (Grün et al., 2015; Haber et al., 2017)、大脑 (Zeisel et al., 2015)、脾脏 (Jaitin et al., 2014)、胰腺 (Baron et al., 2016; Muraro et al., 2016)、肺 (Treutlein et al., 2014)、视网膜 (Macosko et al., 2015)。但是这些研究使用的都是成熟的器官样品，仅有少数研究针对胚胎发育中的器官发生，下面是有关单细胞 RNA 测序在胚胎器官发生研究中的一些最新进展。

4.1 心脏

心脏是胚胎发育过程中第一个形成的功能性器官 (Buckingham et al., 2005) 对人类胚胎心脏发育的单细胞 RNA 测序将样品细胞分为心肌细胞、内皮细胞、心外膜细胞、成纤维细胞、瓣膜细胞以及一些血细胞，并发现了各类细胞的基因表达特点 (Cui et al., 2019)。

一般观点认为，心外膜前体细胞只会发育为心外膜细胞 (Wessels and Pérez-Pomares, 2004)，但是 Cui et al. (2019) 发现，在妊娠第 5 周心脏发育初期时存在高表达心外膜细胞标志基因的心外膜前体细胞，但是这些细胞与成熟的心外膜细胞仍然保留有各自不同的特点，并且这些细胞占据了妊娠第 5 周心脏总细胞数的 20%，远远高于成熟后心外膜细胞的占比，所以作者认为这些心外膜前体细胞不仅可以分化为心外膜细胞，还可以发育为其他细胞。

Cui et al. (2019) 通过分析聚类结果发现了心脏发育中基因表达模式的时间和空间变化，即心肌细胞在 5 周时完成心房心室的区分，此时的心肌细胞只被分为两种不同的类别，之后再形成左右心房心室四个腔体，此时在原有的两种类别内，各自再被分为两种类别，且大类的区别比小类的区别更加明显。

细胞外基质的相关基因在心脏中会影响细胞的增殖、分化、迁移等生命活动 (Doppler et al., 2017)，这些基因的非正常表达会导致心力衰竭等疾病 (Bassat et al., 2017)，原有观点认为在心脏中只有成纤维细胞会表达相关的细胞外基质基因 (Lockhart et al., 2011)，但是 Cui et al. (2019) 发现心肌细胞也会表达细胞外基质基因，这个研究结果可以为对细胞外基质基因和心力衰竭等研究提供有效信息。

4.2 前额叶

前额叶作为认知功能的中心，是高度专门化的大脑区域 (Roth and Dicke, 2012)。即使这些复杂的神经通路往往在胚胎发育末期才形成，但是在发育早期就已经有神经元向合适位置迁移的活动 (O’Rahilly and Mueller, 2008)。Zhong et al. (2018) 使用单细胞 RNA 测序对人胚胎前额叶分析，将细胞分为 35 个小类分属于 6 个大类，即神经前体细胞、兴奋性神经元、中间神经元、星形胶质细胞、少突胶质前体细胞以及小胶质细胞，还进一步定义了每个小类的标志基因，在荧光染色的实验中，这些多样细胞类型的新标志物的发现有助于阐明神经回路发育的更精细的细胞基础。

结合时间分析，Zhong et al. (2018) 发现神经前体细胞在妊娠 9-10 周分化而为胶质细胞和神经元两支不同的谱系，分别进一步分化为星形胶质细胞、少突胶质细胞和神经元。由于小胶质细胞来源于中胚层 (Trapnell et al., 2014)，而中间神经元由神经节隆起迁移而来 (Ma et al., 2013)，所以并不包含在分化路径内。

另外, Zhong et al. (2018) 还发现, 受神经营养因子、轴突引导信号等相关基因的调控, 中间神经元比兴奋性神经元更晚达到发育高峰期, 兴奋性神经元在 16 周达到发育高峰, 而中间神经元在 26 周才达到高峰。由于某些神经障碍、社会认知缺陷等疾病与不同类型神经元的比例有关 (Shi et al., 2005), 所以该研究结果有助于进一步认识不同类型神经元的发育情况, 帮助分析神经障碍等疾病。

4.3 消化道

在妊娠第 4 周, 人类胚胎的原肠可以分为三个部分, 前肠、中肠、后肠, 进而发育为食道、胃、小肠、大肠四个消化道器官 (Kraus and Grapin-Botton, 2012), 消化道发育异常会导致某些新生儿先天性疾病, 例如先天性肠闭锁、坏死性小肠结肠炎等 (Palm et al., 2014)。

在对消化道四个器官的单细胞 RNA 测序分析中, Gao et al. (2018) 分别对每个器官进行了细胞类型分类, 并定义了每种类别的标记基因。在发育过程上, 作者发现食道和胃各自的细胞类型从妊娠 6 周开始出现, 并直到 25 周各种类型所占比例都维持相对不变, 呈现出平稳的发育趋势。不同的是, 在 6-25 周, 小肠的发育过程呈现明显的阶段性, 发育前期以各类前体细胞为主, 后期则以杯状细胞、肠内分泌细胞为主, 并且发育前期的细胞增殖相关基因比后期更加活跃。在大肠的发育过程中, 胚胎时期各细胞类型的表达模式和成熟时期的表达模式有很高的相似性, 通过主成分分析并不能将二者区分开来。

Gao et al. (2018) 还观察到 *HOX* 基因在不同器官中呈现不同的随时间变化的趋势, 符合原有的对 *HOX* 基因表达时空特异性的描述 (Beck, 2002), 另外, Gao et al. (1999) 还观察到了存在于人类小肠中的 Hedgehog 信号通路, 与原有的观点保持一致 (van den Brink et al., 2004)。

4.4 胸腺

胸腺是主要的淋巴器官, 参与建立自身耐受的、功能适应性的免疫系统 (Gruver and Sempowski, 2008), 发育中的胸腺细胞和胸腺上皮细胞以及血细胞的相互接触诱导了胸腺细胞的选择 (Hogquist et al., 2005)。

通过分析小鼠胚胎胸腺发育单细胞 RNA 测序的结果, Kernfeld et al. (2018) 结合全基因组关联性分析后发现, 某些与自身免疫病有关的单核苷酸多态性大多富集在骨髓细胞和非传统淋巴细胞类群中, 而这些类群在胸腺发育早期出现, 所以作者认为自身免疫病的病因可能在胸腺器官发生的早期出现。除此以外, 被观测到的非传统淋巴细胞类群在过去并没有在单细胞和分子水平上被解析 (Brennecke et al., 2015), 所以这个细胞类群的表达模式的发现会有助于以后的研究。

因为过去的研究表明, 在胸腺的发育过程中, 胸腺上皮细胞呈现出多层异质性 (Fischer et al., 2017), 所以在对胸腺上皮的更精细细胞类群划分后, Kernfeld et al. (2018) 发现部分胸腺髓质上皮细胞类群高表达 *Plb1* 基因, *Plb1* 已经被证明和类风湿性关节炎的自身免疫病病因有关 (Okada et al., 2014), 并且高表达 *Plb1* 的类群也同时高表达 *Skint1*, *Ccl19*, *Tnfrsf11b* 基因, 这些基因也已经被证明和 T 细胞选择、T 细胞趋化作用有关 (Hikosaka et al., 2008)。所以当研究类风湿性关节炎时, 这一类胸腺髓质上皮细胞将会是重点研究对象。

4.5 胰腺

胰腺是重要的内分泌器官之一, 过去形态学研究和分子标记研究已经得出胰腺可分为 6 种细胞类型, 其中包括分泌消化酶的外分泌细胞、4 种内分泌细胞、导管细胞 (Edlund, 1998)。由于胚胎发育过程中胰腺组织的细胞异质性以及胰腺内分泌细胞的低含量, 传统的基因表达分析并不能检测到内分泌前体细胞之间的微小表达差异 (Saliba et al., 2014)。

因为在小鼠胚胎发育第 13.5 天, 胰腺上皮层开始了明显的分化并确定各个细胞的谱系 (Pan and Wright, 2011)。所以通过对 13.5 天小鼠胚胎胰腺的单细胞 RNA 测序, Stanescu et al. (2017) 观测到了内分泌细胞的前体细胞, 并发现了它们的特异性基因标记以及和成熟细胞的差别。

通过伪时间发育路径分析, Stanescu et al. (2017) 发现一部分内分泌前体细胞的前体细胞标记基因下调, α 细胞相关基因上调, 说明这部分细胞正在向 α 细胞方向分化。Stanescu et al. (2017) 还发现有一类基因和胰高血糖素的分泌有关, 其中有一个基因 *Scl38a5* 在过去的研究中并没有被提到与 α 细胞的分化有关, 但是对成熟胰腺的转录组分析表明, 成熟的胰腺仍然有 *Scl38a5* 的 mRNA 合成 (Xin et al., 2016), 所以 Stanescu et al. (2017) 认为, 在成熟的胰腺中 *Scl38a5* 的表达产物 SLC38A5 经历了翻译后修饰或并没有锚定在细胞膜上, 所以在原有的实验条件下并不能检测到 *Scl38a5* 的表达。在中枢神经系统中的研究表明, SLC38A5 提哦阿姐了星形胶质细胞在突触周围细胞膜的甘氨酸和谷氨酸的运输 (Hamdani et al., 2012), 所以 Stanescu et al. (2017) 认为 *Scl38a5* 在未成熟的 α 细胞中行使着同样的功能, 并且甘氨酸和谷氨酸都已经被证明可以调节胰高血糖素的分泌 (Gao et al., 1999)。这些新的 α 细胞促成熟因子的发现可以用于优化糖尿病的细胞治疗。

4.6 肝脏

在胚胎发育早期, 肝脏是主要的造血器官, 所以大多数肝脏细胞是血细胞 (Miyajima et al., 2014)。除血细胞以外, 两种主要的肝脏上皮细胞类型是肝细胞和胆管细胞, 它们都由成肝细胞分化而来 (Gordillo et al., 2015)。另外, 由于成熟的肝脏细胞也保留着分化的潜力, 这些肝脏干细胞/前体细胞可以在成熟之后继续分化为肝细胞和胆管细胞 (Miyajima et al., 2014), 所以对胚胎发育过程中成肝细胞的研究可以有助于加深对肝脏干细胞/前体细胞的理解。

过去已经有研究表明, 某些细胞表面的膜蛋白会决定肝脏干细胞/前体细胞的分化命运, 例如可能是 EpCAM、DLK1、CD13 (Okabe et al., 2009; Tanimizu et al., 2003; Kakinuma et al., 2009) 等, 但是真正的决定因子还留有争议。对小鼠胚胎肝脏发育的不同时期的单细胞 RNA 测序表明 (Su et al., 2017), 在胚胎发育的 13.5-16.5 天, 成肝细胞中参与肝细胞功能的基因表达量逐渐增加, 但这是由于向肝细胞分化的细胞数目远多于向胆管细胞分化的细胞数目, 所以平均来看, 会显得成肝细胞只向肝细胞分化。其实, 由于单细胞 RNA 测序数据分析中会遗漏极低占比的细胞种类, Su et al. (2017) 通过寻找已被发现的胆管细胞标记基因 *Epcam* (Miyajima et al., 2014) 在各个细胞中的表达情况, 发现只有少数细胞表达这个基因, 表明这些细胞将来会分化为胆管细胞。同时在这个胆管细胞类群中, 有关 Notch 信号通路的 *Jag1*, *Notch2*, *Hes1* 基因的表达量也比其他类群更高, 说明 Notch 信号通路的激活和胆管细胞的分化有关, 和原有研究结果一致 (Zong et al., 2009)。最后 Su et al. (2017) 发现, 成熟的胆管细胞更接近于 11.5 天时的成肝细胞, 但是具体成肝细胞分化命运决定的时间还需要利用更可靠标记基因做进一步分析。

5 讨论

参考文献

- Bacher, R. and Kendzierski, C. (2016). Design and computational analysis of single-cell RNA-sequencing experiments. *Genome Biology; London*, 17.
- Baron, M., Veres, A., Wolock, S., Faust, A., Gaujoux, R., Vetere, A., Ryu, J., Wagner, B., Shen-Orr, S., Klein, A., Melton, D., and Yanai, I. (2016). A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. *Cell Systems*, 3(4):346–360.e4.

- Bassat, E., Mutlak, Y. E., Genzelinakh, A., Shadrin, I. Y., Baruch Umansky, K., Yifa, O., Kain, D., Rajchman, D., Leach, J., Riabov Bassat, D., Udi, Y., Sarig, R., Sagi, I., Martin, J. F., Bursac, N., Cohen, S., and Tzahor, E. (2017). The extracellular matrix protein agrin promotes heart regeneration in mice. *Nature*, 547(7662):179–184.
- Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W. H., Ng, L. G., Ginhoux, F., and Newell, E. W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, 37(1):38–44.
- Beck, F. (2002). Homeobox genes in gut development. *Gut*, 51(3):450–454. WOS:000177507100032.
- Belle, M., Godefroy, D., Couly, G., Malone, S. A., Collier, F., Giacobini, P., and Chédotal, A. (2017). Tridimensional Visualization and Analysis of Early Human Development. *Cell*, 169(1):161–173.e12.
- Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigó, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E., and Kuehn, M. S. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature; London*, 447(7146):799–816.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *arXiv:0803.0476 [cond-mat, physics:physics]*. arXiv: 0803.0476.
- Brennecke, P., Reyes, A., Pinto, S., Rattay, K., Nguyen, M., Küchler, R., Huber, W., Kyewski, B., and Steinmetz, L. M. (2015). Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. *Nature Immunology*, 16(9):933–941.
- Buckingham, M., Meilhac, S., and Zaffran, S. (2005). Building the mammalian heart from two sources of myocardial cells. *Nature Reviews Genetics*, 6(11):826–835.
- Chiang, M.-K. and Melton, D. A. (2003). Single-Cell Transcript Analysis of Pancreas Development. *Developmental Cell*, 4(3):383–393.
- Cloonan, N., Forrest, A. R. R., Kolle, G., Gardiner, B. B. A., Faulkner, G. J., Brown, M. K., Taylor, D. F., Steptoe, A. L., Wani, S., Bethel, G., Robertson, A. J., Perkins, A. C., Bruce, S. J., Lee, C. C., Ranade, S. S., Peckham, H. E., Manning, J. M., McKernan, K. J., and Grimmond, S. M. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods; New York*, 5(7):613–9.
- Cui, Y., Zheng, Y., Liu, X., Yan, L., Fan, X., Yong, J., Hu, Y., Dong, J., Li, Q., Wu, X., Gao, S., Li, J., Wen, L., Qiao, J., and Tang, F. (2019). Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell Reports*, 26(7):1934–1950.e5.
- Dong, J., Hu, Y., Fan, X., Wu, X., Mao, Y., Hu, B., Guo, H., Wen, L., and Tang, F. (2018). Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. *Genome Biology*, 19(1):31.
- Doppler, S. A., Carvalho, C., Lahm, H., Deutsch, M.-A., Dreßen, M., Puluca, N., Lange, R., and Krane, M. (2017). Cardiac fibroblasts: more than mechanical support. *Journal of Thoracic Disease*, 9(1):S36–S51–S51.
- Edlund, H. (1998). Transcribing pancreas. *Diabetes*, 47(12):1817–1823. WOS:000077199800001.
- Fan, X., Zhang, X., Wu, X., Guo, H., Hu, Y., Tang, F., and Huang, Y. (2015). Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biology*, 16(1):148.
- Fischer, D. S., Fiedler, A. K., Kernfeld, E., Genga, R. M. J., Hasenauer, J., Maehr, R., and Theis, F. J. (2017). Beyond pseudotime: Following T-cell maturation in single-cell RNAseq time series. *bioRxiv*, page 219188.
- Gao, S., Yan, L., Wang, R., Li, J., Yong, J., Zhou, X., Wei, Y., Wu, X., Wang, X., Fan, X., Yan, J., Zhi, X., Gao, Y., Guo, H., Jin, X., Wang, W., Mao, Y., Wang, F., Wen, L., Fu, W., Ge, H., Qiao, J., and Tang, F. (2018). Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. *Nature Cell Biology*, 20(6). WOS:000433237300013.
- Gao, Z. Y., Li, G., Najafi, H., Wolf, B. A., and Matschinsky, F. M. (1999). Glucose regulation of glutaminolysis and its role in insulin secretion. *Diabetes*, 48(8):1535–1542.
- Gilbert, S. F. and Barresi, M. J. F. (2017). Developmental Biology, 11th Edition 2016. *American Journal of Medical Genetics Part A*, 173(5):1430–1430.
- Gordillo, M., Evans, T., and Gouon-Evans, V. (2015). Orchestrating liver development. *Development*, 142(12):2094–2108.
- Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature*, 525(7568):251–255.
- Gruver, A. L. and Sempowski, G. D. (2008). Cytokines, leptin, and stress-induced thymic atrophy. *Journal of Leukocyte Biology*, 84(4):915–923.
- Haber, A. L., Biton, M., Rogel, N., Herbst, R. H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T. M., Howitt, M. R., Katz, Y.,

- Tirosh, I., Beyaz, S., Dionne, D., Zhang, M., Raychowdhury, R., Garrett, W. S., Rozenblatt-Rosen, O., Shi, H. N., Yilmaz, O., Xavier, R. J., and Regev, A. (2017). A single-cell survey of the small intestinal epithelium. *Nature*, 551(7680):333–339.
- Hamdani, E. H., Gudbrandsen, M., Bjørkmo, M., and Chaudhry, F. A. (2012). The system N transporter SN2 doubles as a transmitter precursor furnisher and a potential regulator of NMDA receptors. *Glia*, 60(11):1671–1683.
- Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., Huang, D., Xu, Y., Huang, W., Jiang, M., Jiang, X., Mao, J., Chen, Y., Lu, C., Xie, J., Fang, Q., Wang, Y., Yue, R., Li, T., Huang, H., Orkin, S. H., Yuan, G.-C., Chen, M., and Guo, G. (2018). Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell*, 172(5):1091–1107.e17.
- Hedlund, E. and Deng, Q. (2018). Single-cell RNA sequencing: Technical advancements and biological applications. *Molecular Aspects of Medicine*, 59:36–46.
- Hikosaka, Y., Nitta, T., Ohigashi, I., Yano, K., Ishimaru, N., Hayashi, Y., Matsumoto, M., Matsuo, K., Penninger, J. M., Takayanagi, H., Yokota, Y., Yamada, H., Yoshikai, Y., Inoue, J.-i., Akiyama, T., and Takahama, Y. (2008). The Cytokine RANKL Produced by Positively Selected Thymocytes Fosters Medullary Thymic Epithelial Cells that Express Autoimmune Regulator. *Immunity*, 29(3):438–450.
- Hogquist, K. A., Baldwin, T. A., and Jameson, S. C. (2005). Central tolerance: learning self-control in the thymus. *Nature Reviews Immunology*, 5(10):772–782.
- Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types. *Science*, 343(6172):776–779.
- Kakinuma, S., Ohta, H., Kamiya, A., Yamazaki, Y., Oikawa, T., Okada, K., and Nakauchi, H. (2009). Analyses of cell surface molecules on hepatic stem/progenitor cells in mouse fetal liver. *Journal of Hepatology*, 51(1):127–138.
- Kernfeld, E. M., Genga, R. M. J., Neherin, K., Magaletta, M. E., Xu, P., and Maehr, R. (2018). A Single-Cell Transcriptomic Atlas of Thymus Organogenesis Resolves Cell Types and Developmental Maturation. *Immunity*, 48(6):1258–1270.e6.
- Kraus, M. R. C. and Grapin-Botton, A. (2012). Patterning and shaping the endoderm in vivo and in culture. *Current Opinion in Genetics & Development*, 22(4):347–353. WOS:000308899000007.
- Lancaster, M. A. and Knoblich, J. A. (2014). Organogenesis in a dish: Modeling development and disease using organoid technologies. *Science*, 345(6194):1247125.
- Lockhart, M., Wirrig, E., Phelps, A., and Wessels, A. (2011). Extracellular Matrix and Heart Development. *Birth Defects Research Part a-Clinical and Molecular Teratology*, 91(6):535–550. WOS:000291545100013.
- Ma, T., Wang, C., Wang, L., Zhou, X., Tian, M., Zhang, Q., Zhang, Y., Li, J., Liu, Z., Cai, Y., Liu, F., You, Y., Chen, C., Campbell, K., Song, H., Ma, L., Rubenstein, J. L., and Yang, Z. (2013). Subcortical origins of human and monkey neocortical interneurons. *Nature Neuroscience*, 16(11):1588–1597. WOS:000326205700016.
- Maaten, L. v. d. and Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9(Nov):2579–2605.
- Macosko, E., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A., Kamitaki, N., Martersteck, E., Trombetta, J., Weitz, D., Sanes, J., Shalek, A., Regev, A., and McCarroll, S. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, 161(5):1202–1214.
- McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv:1802.03426 [cs, stat]*. arXiv: 1802.03426.
- Miyajima, A., Tanaka, M., and Itoh, T. (2014). Stem/progenitor cells in liver development, homeostasis, regeneration, and reprogramming. *Cell Stem Cell*, 14(5):561–574.
- Muraro, M., Dharmadhikari, G., Grün, D., Groen, N., Dielen, T., Jansen, E., van Gurp, L., Engelse, M., Carlotti, F., de Koning, E. P., and van Oudenaarden, A. (2016). A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Systems*, 3(4):385–394.e3.
- Okabe, M., Tsukahara, Y., Tanaka, M., Suzuki, K., Saito, S., Kamiya, Y., Tsujimura, T., Nakamura, K., and Miyajima, A. (2009). Potential hepatic stem cells reside in EpCAM+ cells of normal and injured mouse liver. *Development*, 136(11):1951–1960.
- Okada, Y., Diogo, D., Greenberg, J. D., Mouassess, F., Achkar, W. A. L., Fulton, R. S., Denny, J. C., Gupta, N., Mirel, D., Gabriel, S., Li, G., Kremer, J. M., Pappas, D. A., Carroll, R. J., Eyler, A. E., Trynka, G., Stahl, E. A., Cui, J., Saxena, R., Coenen, M. J. H., Guchelaar, H.-J., Huizinga, T. W. J., Dieudé, P., Mariette, X., Barton, A., Canhão, H., Fonseca, J. E., Vries, N. d., Tak, P. P., Moreland, L. W., Jr, S. L. B., Miceli-Richard, C., Choi, H. K., Kamatani, Y., Galan, P.,

- Lathrop, M., Raj, T., Jager, P. L. D., Raychaudhuri, S., Worthington, J., Padyukov, L., Klareskog, L., Siminovitch, K. A., Gregersen, P. K., Mardis, E. R., Arayssi, T., Kazkaz, L. A., and Plenge, R. M. (2014). Integration of Sequence Data from a Consanguineous Family with Genetic Data from an Outbred Population Identifies PLB1 as a Candidate Rheumatoid Arthritis Risk Gene. *PLOS ONE*, 9(2):e87645.
- O’Rahilly, R. and Mueller, F. (2008). Significant features in the early prenatal development of the human brain. *Annals of Anatomy-Anatomischer Anzeiger*, 190(2):105–118. WOS:000257663300003.
- Palm, N. W., de Zoete, M. R., Cullen, T. W., Barry, N. A., Stefanowski, J., Hao, L., Degnan, P. H., Hu, J., Peter, I., Zhang, W., Ruggiero, E., Cho, J. H., Goodman, A. L., and Flavell, R. A. (2014). Immunoglobulin A Coating Identifies Colitogenic Bacteria in Inflammatory Bowel Disease. *Cell*, 158(5):1000–1010. WOS:000340945000007.
- Pan, F. C. and Wright, C. (2011). Pancreas organogenesis: From bud to plexus to gland. *Developmental Dynamics*, 240(3):530–565.
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572.
- Potter, S. S. (2018). Single-cell RNA sequencing for the study of development, physiology and disease. *Nature Reviews Nephrology*, 14(8):479.
- Regev, A., Teichmann, S. A., Lander, E. S., Amt, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., Clevers, H., Deplancke, B., Dunham, I., Eberwine, J., Elis, R., Enard, W., Farmer, A., Fugger, L., Gottgens, B., Hacohen, N., Haniffa, M., Hemberg, M., Kim, S., Klennerman, P., Kriegstein, A., Lein, E. D., Linnarsson, S., Lundberg, E., Lundberg, J., Majumder, P., Marioni, J. C., Merad, M., Mhlanga, M., Nawijn, M., Netea, M., Nolan, G., Pe’er, D., Phillipakis, A., Ponting, C. P., Quake, S., Reik, W., Rozenblatt-Rosen, O., Sanes, J., Satija, R., Schumacher, T. N., Shalek, A., Shapiro, E., Sharma, P., Shin, J. W., Stegle, O., Stratton, M., Stubbington, M. J. T., Theis, F. J., Uhlen, M., Van Oudenaarden, A., Wagner, A., Watt, F., Weissman, J., Wold, B., Xavier, R., and Yosef, N. (2017). The Human Cell Atlas. *Elife*, 6:e27041. WOS:000419858100001.
- Roth, G. and Dicke, U. (2012). Evolution of the brain and intelligence in primates. In Hofman, M. A. and Falk, D., editors, *Evolution of the Primate Brain: From Neuron to Behavior*, volume 195, pages 413–430. Elsevier Science Bv, Amsterdam. WOS:000310995600021.
- Saliba, A.-E., Westermann, A. J., Gorski, S. A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Research*, 42(14):8845–8860.
- Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A., and Waterston, R. H. (2017). DNA sequencing at 40: past, present and future. *Nature; London*, 550(7676):1–9.
- Shi, G. X., Han, J. H., and Andres, D. A. (2005). Rin GTPase couples nerve growth factor signaling to p38 and b-Raf/ERK pathways to promote neuronal differentiation. *Journal of Biological Chemistry*, 280(45):37599–37609. WOS:000233044500036.
- Stanescu, D. E., Yu, R., Won, K.-J., and Stoffers, D. A. (2017). Single cell transcriptomic profiling of mouse pancreatic progenitors. *Physiological Genomics*, 49(2):105–114. WOS:000395792600005.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, 177(7):1888–1902.e21.
- Su, X., Shi, Y., Zou, X., Lu, Z.-N., Xie, G., Yang, J. Y. H., Wu, C.-C., Cui, X.-F., He, K.-Y., Luo, Q., Qu, Y.-L., Wang, N., Wang, L., and Han, Z.-G. (2017). Single-cell RNA-Seq analysis reveals dynamic trajectories during mouse liver development. *BMC Genomics*, 18(1):1–14.
- Takahashi, K. and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4):663–676.
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., and Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, 6(5):377–382.
- Tanimizu, N., Nishikawa, M., Saito, H., Tsujimura, T., and Miyajima, A. (2003). Isolation of hepatoblasts based on the expression of Dlk/Pref-1. *Journal of Cell Science*, 116(Pt 9):1775–1786.
- Traag, V., Waltman, L., and van Eck, N. J. (2018). From Louvain to Leiden: guaranteeing well-connected communities. *arXiv:1810.08473 [physics]*. arXiv: 1810.08473.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S., and Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology*, 32(4):381–U251. WOS:000334340800023.

- Treutlein, B., Brownfield, D. G., Wu, A. R., Neff, N. F., Mantalas, G. L., Espinoza, F. H., Desai, T. J., Krasnow, M. A., and Quake, S. R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature*, 509(7500):371–375.
- van den Brink, G. R., Bleuming, S. A., Hardwick, J. C. H., Schepman, B. L., Offerhaus, G. J., Keller, J. J., Nielsen, C., Gaffield, W., van Deventer, S. J. H., Roberts, D. J., and Peppelenbosch, M. P. (2004). Indian Hedgehog is an antagonist of Wnt signaling in colonic epithelial cell differentiation. *Nature Genetics*, 36(3):277–282. WOS:000189250400017.
- Wessels, A. and Pérez-Pomares, J. M. (2004). The epicardium and epicardially derived cells (EPDCs) as cardiac stem cells. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology*, 276A(1):43–57.
- Wolf, F. A., Angerer, P., and Theis, F. J. (2018). SCANPY : large-scale single-cell gene expression data analysis. *Genome Biology*, 19(1):1–5.
- Xin, Y., Kim, J., Ni, M., Wei, Y., Okamoto, H., Lee, J., Adler, C., Cavino, K., Murphy, A. J., Yancopoulos, G. D., Lin, H. C., and Gromada, J. (2016). Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. *Proceedings of the National Academy of Sciences*, 113(12):3293–3298.
- Zeisel, A., Muñoz-Manchado, A. B., Codeluppi, S., Lönnerberg, P., Manno, G. L., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., Rolny, C., Castelo-Branco, G., Hjerling-Leffler, J., and Linnarsson, S. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*, 347(6226):1138–1142.
- Zhong, S., Zhang, S., Fan, X., Wu, Q., Yan, L., Dong, J., Zhang, H., Li, L., Sun, L., Pan, N., Xu, X., Tang, F., Zhang, J., Qiao, J., and Wang, X. (2018). A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature*, 555(7697):524–528.
- Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., and Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell*, 65(4):631–643.e4.
- Zong, Y., Panikkar, A., Xu, J., Antoniou, A., Raynaud, P., Lemaigre, F., and Stanger, B. Z. (2009). Notch signaling controls liver development by regulating biliary differentiation. *Development (Cambridge, England)*, 136(10):1727–1739.
- Zorn, A. M. and Wells, J. M. (2009). Vertebrate Endoderm Development and Organ Formation. *Annual Review of Cell and Developmental Biology*, 25(1):221–251.