质谱流式技术用于单细胞检测

杨雨 孙传强 龚子珊 杨茹 汪曦*
(天津大学精密仪器与光电子工程学院, 天津 300072)

摘要 质谱流式技术(mass cytometry)是利用质谱原理对单细胞进行多参数检测的流式技术, 能够在单细胞水平实现超过50种标志物的同时测量, 显著增强了对细胞生长进程和复杂细胞系统的评估能力。该文简要介绍了质谱流式技术的基本工作原理, 并从金属元素标记、质量分析器、高维单细胞数据处理等方面展开论述, 阐明设计新型金属元素标签和选择飞行时间质谱的必要性, 归纳分析高维单细胞数据的算法并总结各种算法的优点和局限性。

关键词 质谱流式技术; 金属元素标记; 飞行时间质谱; 高维单细胞数据处理

Application of Mass Cytometry in Single-Cell Detection

YANG Yu, SUN Chuanqiang, GONG Zishan, YANG Ru, WANG Yan*
(*School of Precision Instrument and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China*)

Abstract Mass cytometry is a cytometric technique in which multiple biomarkers of a single cell are detected through mass spectrometry. It can simultaneously measure more than 50 biomarkers at single-cell resolution, effectively augmenting the ability to evaluate cell growth processes and complex cellular systems. This work will briefly introduce the basic principle of mass cytometry and discuss metal labeling, mass analyzer and high-dimensional single-cell data processing. The needs to design new metal tags and select time-of-flight mass spectrometry are clarified, the algorithms which can be used for high-dimensional single-cell data analysis are described. At the same time, the advantages and limitations of these algorithms are discussed.

Keywords mass cytometry; metal labeling; time of flight mass spectrometry; high-dimensional single-cell data processing

流式细胞技术是一种基于荧光标记的单细胞多参数检测技术[1-4], 但荧光基团发射光谱较宽, 相互之间会发生重叠, 因而需要复杂的补偿计算[5]。为了解决流式细胞技术的局限性, 多伦多大学SCOTT TANNER教授团队[6]提出“质谱流式技术”: 将纯化过的稳定金属同位素连接到抗体(取代荧光标记), 利用抗原抗体特异性结合的原理标记细胞标志物, 通过电感耦合等离子体电离细胞, 然后使用飞行时间质谱检测相应金属同位素(舍弃光学检测), 可同时检测超过50种标志物, 极大地增加了可检测通道且无需复杂的补偿计算[7]。在单细胞分辨率下可实现精细免疫细胞分群、细胞周期检测及揭示细胞内信号转导途径等[8-13]。

1 质谱流式技术的基本原理

质谱流式技术的核心是流式细胞技术与电感耦合等离子体飞行时间质谱(inductively coupled plasma-time of flight-mass spectrometry, ICP-TOF-MS)的融合。它既继承了流式细胞仪的单细胞检测技术, 又具有ICP-TOF-MS的高分辨能力, 其工作流程如图1所示。首先, 制备由稳定金属同位素标记的抗体, 然后与细胞表面相应标志物特异性结合, 达到细胞被金属元素标记的目的。被标记的细胞通过雾化器, 形成单细胞液滴流, 经过等离子体炬管时进行去溶剂化和

收稿日期: 2019-06-06 接受日期: 2019-09-03
*通讯作者。Tel: 022-84853550, E-mail: wangyan@tju.edu.cn
Received: June 6, 2019 Accepted: September 3, 2019
*Corresponding author. Tel: +86-22-84853550, E-mail: wangyan@tju.edu.cn
URL: http://www.cjcb.org/arts.asp?id=5189

电离, 接着经过射频四极杆去除常见生物元素(如C、N等), 富集重金属离子, 最后垂直引入飞行时间质量分析器, 按不同质荷比在不同时间被检测器捕获。这样细胞标志物的类型(金属标签元素种类)和表达水平(金属标签信号强度)会变成相应数据, 用于之后的数据分析。

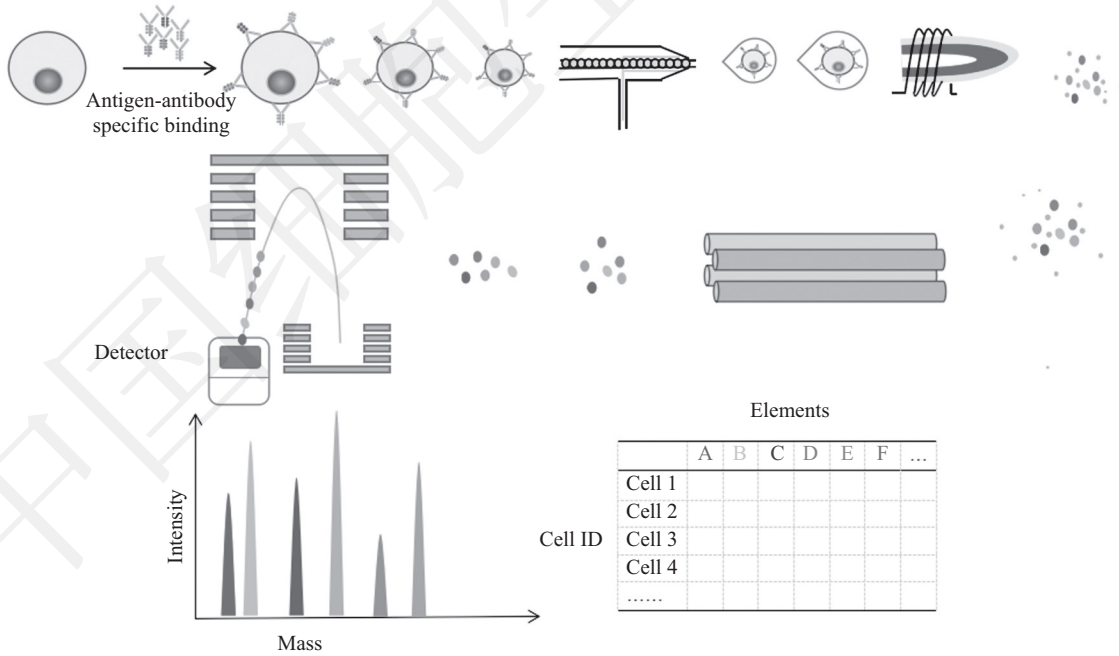
2 基于质谱流式技术的金属元素标记及质量分析器的选择

2001年, 清华大学张新荣课题组[14]通过抗原抗体特异性免疫反应与电感耦合等离子体四极杆质谱(quadropole inductively coupled plasma with mass spectrometry, Q-ICP-MS)联用的方式, 首次建立了基于Eu^{3+}的标记免疫分析方法, 并将其应用于人血清中促甲状腺激素(thyroid stimulating hormone, TSH)的测定。BARANOV等[15]将离心过滤、蛋白质琼脂糖A(protein A-sepharose, PAS)、尺寸排阻凝胶过滤及ELISA等4种不同的免疫分析方法与 Q-ICP-MS 联用, 快速有效地检测低至0.1~0.5 ng/mL的目标蛋白, 线性响应超过3个数量级。QUINN等[16]首次提出了基于ICP-Q-MS的多参数金属元素标记的免疫分析方法, 用Eu^{3+}和1.4 nm Au纳米簇标记Smad 2和Smad 4对应的抗体, 检出限低至0.1~0.5 ng/mL, 线性响应范围为2~100 ng/mL。ORNATSKY等[17]同时使用Sm^{3+}、Eu^{2+}、Tb^{3+}和Au纳米簇4种金属标签对人白血病细胞系的4种目标蛋白进行定量分析。研究表明, 基于Q-ICP-MS的金属元素标记方法在免疫分析方面是可行的, 但在研究过程中发现: (1)所用金属元素标签每个只含有6~9个金属原子, 且未经纯化, 易受同位素干扰; (2)细胞经ICP电离后, 离子云的持续时间为200~300 μs, 而四极杆扫描电压的建立时间为50~200 μs, 难以实现更多细胞标志物检测[6,18]。

2.1 基于镧系金属的标签设计

LOU等[19]研发了一种基于线性丙烯酸聚合物、由马来酰亚胺修饰的金属元素标签, 含有多个纯化过的非放射性稳定金属同位素, 极大地提高了分析灵敏度, 实现了含量相差近500倍的2种细胞表面标志物的同时测量。

制备抗体标签的主要步骤是: (1)通过可逆加



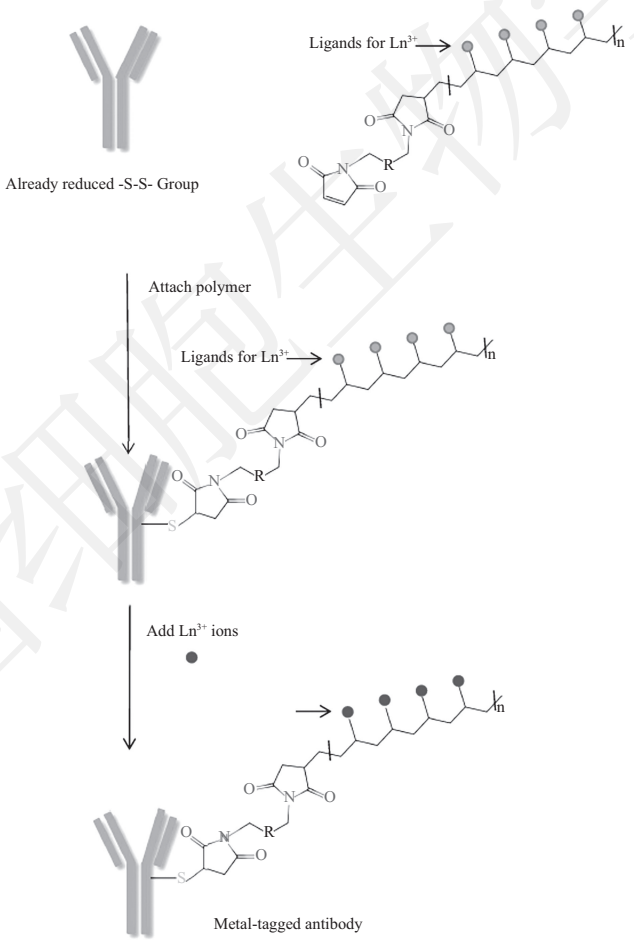
用已被稳定金属同位素标记的抗体对细胞进行染色, 细胞通过雾化器被雾化成单细胞液滴。然后进入电感耦合等离子体被电离, 通过四极杆去除常见的生物元素并富集重金属离子, 最后通过飞行时间质量分析器进行定量。

Cells are stained with epitope-specific antibodies conjugated to transition element isotope reporters, and are nebulized into single-cell droplets as they are introduced into the mass cytometry. They then travel into the ICP (inductively coupled plasma), in which cells are ionized. The ion cloud is filtered by quadrupole to remove common biological elements and enrich the heavy metal reporter ions to be quantified by time-of-flight mass spectrometry.

图1 质谱流式技术的基本原理

Fig.1 Basic principle of mass cytometry

成–断裂链转移(RAFT, reversible addition-fragmentation chain transfer)或阴离子聚合等方法制备线性丙烯酸聚合物, 其链长分散性较低且含有多个活化*N*-羟基琥珀酰亚胺(*N*-hydroxysuccinimide)酯基; (2)官能化的DTPA(diethylenetriaminepentaacetic acid)或DOTA(1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid)与线性丙烯酸聚合物反应, 将数量可重复的金属螯合剂连接至聚合物(例如, DTPA可连接30个金属螯合剂); (3)用硫醇对聚合物进行末端官能化, 然后通过硫醇与双马来酰亚胺反应, 生成既含有金属螯合剂, 又可与还原后的二硫化物(抗体Fc段)连接的聚合物; (4)与抗体连接后, 加入镧系金属氯化物溶液, 使镧系金属同位素螯合到聚合物上。整个流程如图2所示。抗体的Fc段可连接2~4个聚合物, 即每个抗体相当于被60~120个金属原子标记, 极大地提高了分析灵敏度。纯化过的金属标签只含有所选元素的1种同位素, 确保了最低的背景干扰和最小的信号重叠。金属元素选择镧系金属, 因为它们具有可分辨的稳定同位素, 且有相似的化学性质、较低的天然丰度, 能够螯合到相同的标签结构中。ORNATSKY等[20]对上述金属标签的可行性进行进一步验证, 详细讨论了标签元素的选择、抗体的选择、标记的可行性及多参数免疫分析的切实性等, 最后使用7种镧系金属元素标签连接相应的抗体, 对KG-1a和ThP-1细胞系的表面抗原进行分析, 实验结果与常规免疫分析一致。此后, 新标签的设计以多个金属原子标记单个抗体为基本准则, 旨在探索开发基于不同骨架结构的金属元素标签, 并通过质谱流式细胞仪验证标签的



选择性还原目标抗体Fc片段中的-S-S-基团, 产生反应性-SH基团, 能够与未标记的聚合物末端的马来酰亚胺基团反应。用给定的镧系金属离子处理, 每种抗体都用不同的稳定金属同位素标记。

The antibody of interest is subjected to selective reduction of -S-S-groups in the Fc fragment to produce reactive -SH groups, which are reacted with the terminal maleimide groups of the unlabeled polymer tag. The polymer-bearing antibodies are treated with a given lanthanide ion, and each type of antibody is labeled with a different element.

图2 抗体标记流程(根据参考文献[19]修改)

Fig.2 Antibody labeling process (modified from reference [19])

可行性[21-25]。

现阶段, 一般可通过在Fluidigm直接购买已标记抗体或通过Fluidigm MaxPar X8试剂盒标记感兴趣的抗体, 简化抗体标记的流程。标记完成后, 还需使用Nanodrop分光光度计在280 nm、IgG模式下测量标记完成后的抗体浓度C_1, 然后测量体积V_1, 计算回收率, 如公式1所示, 其中M_{Ab}为最初使用的抗体质量。测得C_1后, 需使用抗体稳定缓冲液将抗体稀释至0.4 mg/mL, 便于之后细胞染色。此外, 还需进行标记指数检测(label index), 测得的金属同位素信号强度的中位数M_1与EQ beads中Eu151强度的中位数M_{Eu151}的比值即为相对标记指数, 如公式2所示。回收率和相对标记指数是抗体标记的两个关键指标: 回收率越大, 抗体浓度越高; 相对标记指数越大, 表明螯合的金属越多, 单个抗体的信号强度越高。

完成标记后, 需对抗体最佳使用浓度进行滴定。滴定前, 根据要滴定的抗体选择合适的细胞系及抗体阴阳对照, 通常在滴定时选用的抗体浓度为0.5 μg/mL、1 μg/mL、2 μg/mL、4 μg/mL。以滴定anti-PD1为例, 选用激活T细胞, 阳性对照抗体为CD3, 阴性对照抗体为CD19。在4种不同浓度下, PD-1的表达如图3所示, 抗体浓度为4 μg/mL时, 趋于饱和状态。通常使用4 μg/mL的anti-PD1进行细胞染色。

$$\%\text{Recovery} = \frac{C_1 \times V_1}{M_{Ab}} \quad \text{公式1}$$

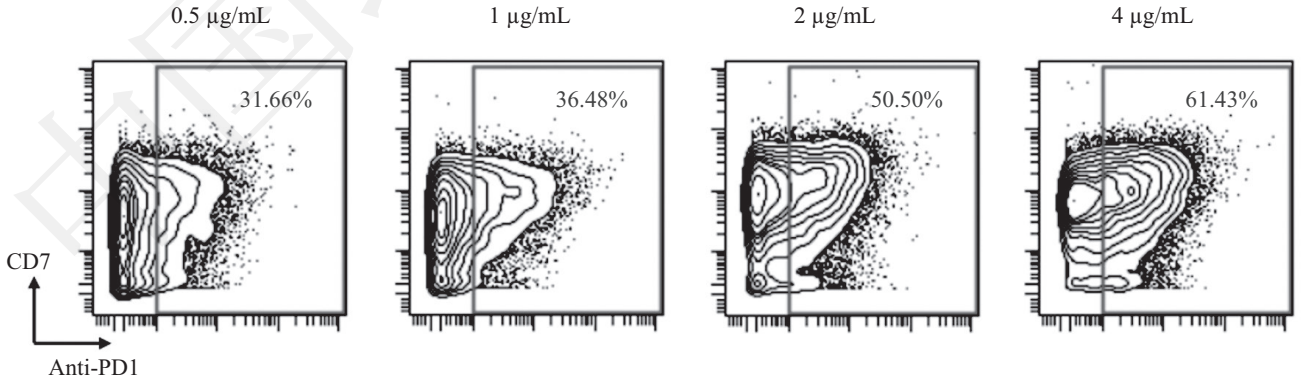
$$\text{Label index} = \frac{M_1}{M_{Eu151}} \quad \text{公式2}$$

2.2 其他用途的金属标记

2.2.1 金属嵌入剂 细胞中DNA的识别和定量在生物分析中不可或缺, 已报道过使用Rh、Ru、Os、Co、Re或Ir等金属嵌入剂作为敏感分子探针或治疗剂[26-28]。金属嵌入剂是一种金属配合物, 能够插入到DNA的核苷酸对之间[29-31]。就免疫测定而言, TANNER等[18]证明, DNA结合的金属嵌入剂与DNA数量成正比。ORNATSKY等[32]开发使用Ir-嵌入剂和Rh-嵌入剂进行DNA检测及细胞计数: 嵌入剂定量地结合在DNA的碱基对之间, 通过质量平衡原理对DNA进行量化, 在细胞计数模式下区分单个细胞。图4左上显示的是金属嵌入剂与DNA的结合[33], 左侧显示插入Rh-嵌入剂, 右侧显示插入Ir-嵌入剂, 右下显示的是细胞悬浮液雾化后观察到的瞬变信号。

Ir-嵌入剂有以下几种优点: (1)Ir具有2个稳定的同位素Ir191和Ir193, 其丰度比约为3:5, 2种同位素的同时嵌入能够更好地测量DNA。(2)Ir具有更高的电离能, 可以反映ICP电离温度的变化, 指示可能的基质效应。因此, 在质谱流式细胞仪单细胞检测中, 多使用Ir191、Ir193嵌入剂识别DNA[34]。如图5所示, 在使用Ir嵌入剂之后, 可以直接门控出单细胞, 便于之后的细胞分群。

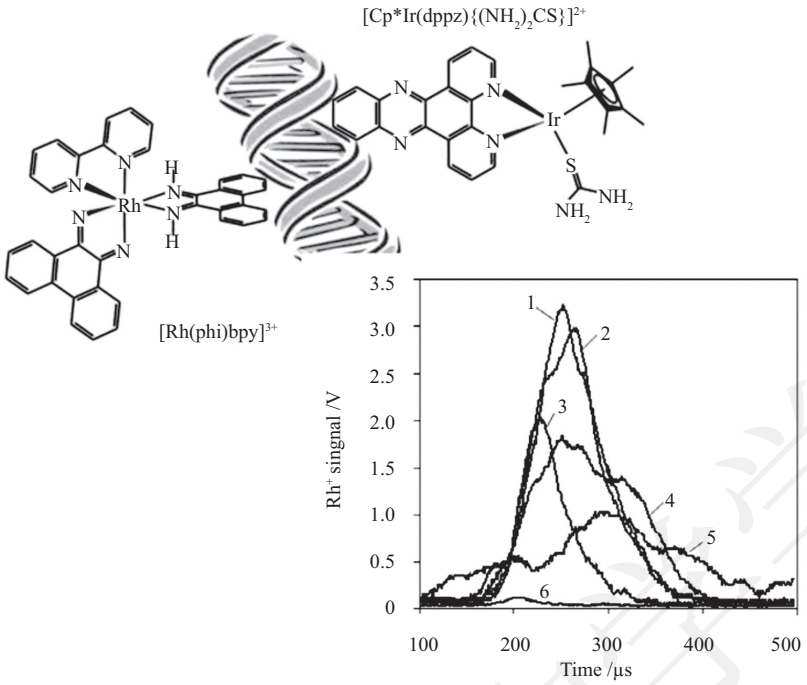
2.2.2 基于不同金属同位素的编码(barcode)分析 barcode分析的主要思路是使用不同的金属同位素标签编码不同的样本, 然后混合在一起同时进行细胞染色及后续的质谱流式分析, 既简化了样本制备时间, 也消除了由于细胞染色和数据收集产生的误



使用4种不同浓度的anti-PD1抗体对激活T细胞染色, 确定抗体饱和时的最佳使用浓度。在接下来的实验中, 应使用4 μg/mL的anti-PD1对细胞染色。
Activated T cells were stained with four different concentrations of anti-PD1 antibodies to confirm the optimal concentration when the antibody was saturated. 4 μg/mL anti-PD1 antibodies were used for cell staining in next experiment.

图3 细胞在不同浓度抗体下的表达

Fig.3 Expression in different concentrations of Antibodies



左上: 含有金属原子的嵌入剂定量地结合在DNA的碱基对之间, 左侧显示含Rh的金属嵌入剂, 右侧显示含Ir的金属嵌入剂。右下: 高斯瞬态。1~3: 显示单细胞通过; 凝聚细胞4和细胞碎片5会导致较长时间的多峰瞬变; 6表示不含金属嵌入剂的细胞电离信号。

Top left: an intercalator containing a metal atom quantitatively binds between the base pairs in the DNA groove: A Rh-containing metallointercalator is shown at the left, and an Ir-containing metallointercalator is shown on the right. Figure below right: whole cells are evidenced by a Gaussian transient 1-3 while cell fragments 5 or agglomerated cells 4 result in multiple-peaked transients of typically longer duration, the transient marked 6 shows no metallointercalation.

图4 金属嵌入剂与DNA的结合(根据参考文献[33]修改)

Fig.4 The combination of metal intercalators and DNA (modified from reference [33])

差。BODENMILLER等[35]开发7种纯化过的镧系金属同位素标签编码不同样本的方法(La139、Pr141、Nd146、Tb159、Ho165、Tm169、Lu175), 使用这7种标签编码96种不同的PBMC样本, 分析PBMC信号动态、细胞间通信及基于PBMC信号响应的小分子药物调节机制。但其使用镧系金属同位素作为标签, 会相应的减少抗体测量通道。MEI等[36]使用Pd104、Pd106、Pd108、Pd110、In113和In115等6种标签标记CD45抗体, 然后用6选3的编码方式对不同的PBMC样本进行编码, 即每个样本被3种不同的CD45抗体标记。这种方法会占用CD45通道, 而CD45是单细胞分析中较为重要的标志物。ZUNDER等[37]开发基于Pd 6种同位素的barcode分析方法, 同样用6选3的方式, 产生20种不同的编码方案, 标记20个不同的样本, 然后混合在一起进行染色及后续的质谱流式分析, 最后再解编码分析数据。相较而言, 这种方法既不会占用镧系金属同位素, 也不会占用较为关键的CD45通道, 因此, 基于Pd的barcode分析已成为质谱流式技术单细胞分析必不可少的步骤[38]。

2.3 质量分析器的选择

QUINN等[16]提出基于Q-ICP-MS的多参数免疫分析方法, 并成功验证其可行性, 但仍然无法在单细胞水平实现超过10种标志物的同时测量。TANNER等[33]首次搭建ICP-TOF-MS平台, 飞行时间质谱的垂直引入脉冲提取频率为55 KHz, 实现对KG1a细胞系20种表面标志物的同时测量。BANDURA等[6]将流式细胞仪的液流系统与质谱检测相结合, 设计首台质谱流式细胞仪, 既拥有单细胞分辨能力, 又能进行多种标志物的同时检测。表1总结了该仪器及后三代商用仪器的主要参数: 质谱流式细胞仪致力于提高灵敏度、拓宽质量范围和线性动态范围。

3 数据处理

质谱流式技术可同时测量超过50种的细胞标志物, 增加了单细胞可测量的信息含量和数据处理的复杂性, 因此, 亟需开发适用于高维质谱流式数据的分析方法。在分析前, 使用基于MATLAB的归一化算法, 消除EQ beads、背景噪声及细胞二聚体等

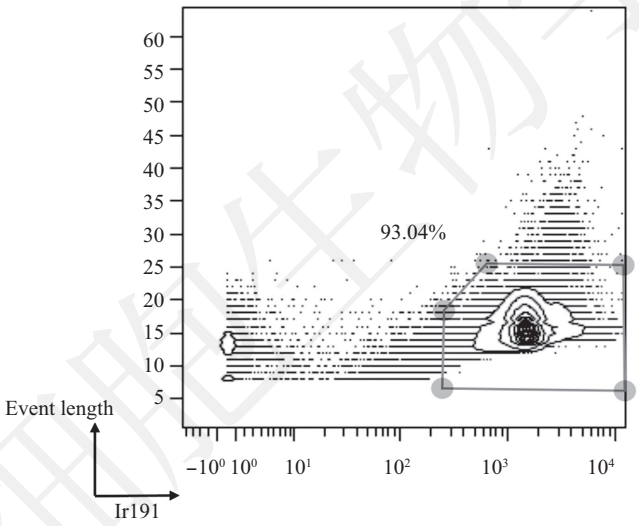
以清洁数据[39]。

3.1 SPADE

SPADE(spanning-tree progression analysis of density-normalized events)是用于处理质谱流式数据的工具, 通过分层聚类的方法将细胞分成不同的类: 表型相似的细胞被限定在相同类, 表型不相似的细胞分配在不同的类中[40]。首先将表型相似的细胞聚成群, 然后按照各群的表型相似度进行聚类分析, 得到树形图, 实现聚类和最小生成树(minimum spanning tree, MST)投影。SPADE树形图上每个节点(node)都是由一群表型相似的细胞构成的, 因而损失了单细胞分辨率, 但能直观展示细胞的亚群构成。BENDALL等[13]使用SPADE聚类分析实现了骨髓细胞等复杂细胞群体的精确免疫分群。SALMON等[41]利用32种细胞标志物进行SPADE分析, 图6A展示的是注射PBS和FLT3L的小鼠黑色素瘤浸润的免疫细胞SPADE图谱, 可以明显看出注射FLT3L的小鼠CD103$^+$树突细胞比例明显增加。SPADE聚类分析的主要优点就是: 通过细胞表型分析能够直接显示样本的细胞亚群组成, 同时可以比较同一样本不同刺激条件下, 细胞亚群的变化情况。

3.2 FlowSOM

FlowSOM提供了与SPADE类似的可视化能力。它用自组织映射(self-organization mapping net, SOM)代替SPADE中的分层聚类, 减小了对密度降采样的依赖, 大大缩短了计算时间[42]。将星图(star charts)引入MST是将FlowSOM与SPADE区别开来的另一个特征。星图允许在MST节点上可视化感兴趣标志



用Ir191和信号长度门控单细胞, 一般情况下单细胞百分比在85%以上。实验中, 单细胞约占93%。

Use Ir191 and Event length to gate singlets and usually the percentage of this population is above 85%. Singlets were about 93% in this work.

图5 Ir嵌入剂用于门控单细胞

Fig.5 Gate singlets with Ir-intercalators

表1 质谱流式细胞仪参数总结

Table 1 Summary of instrument parameters

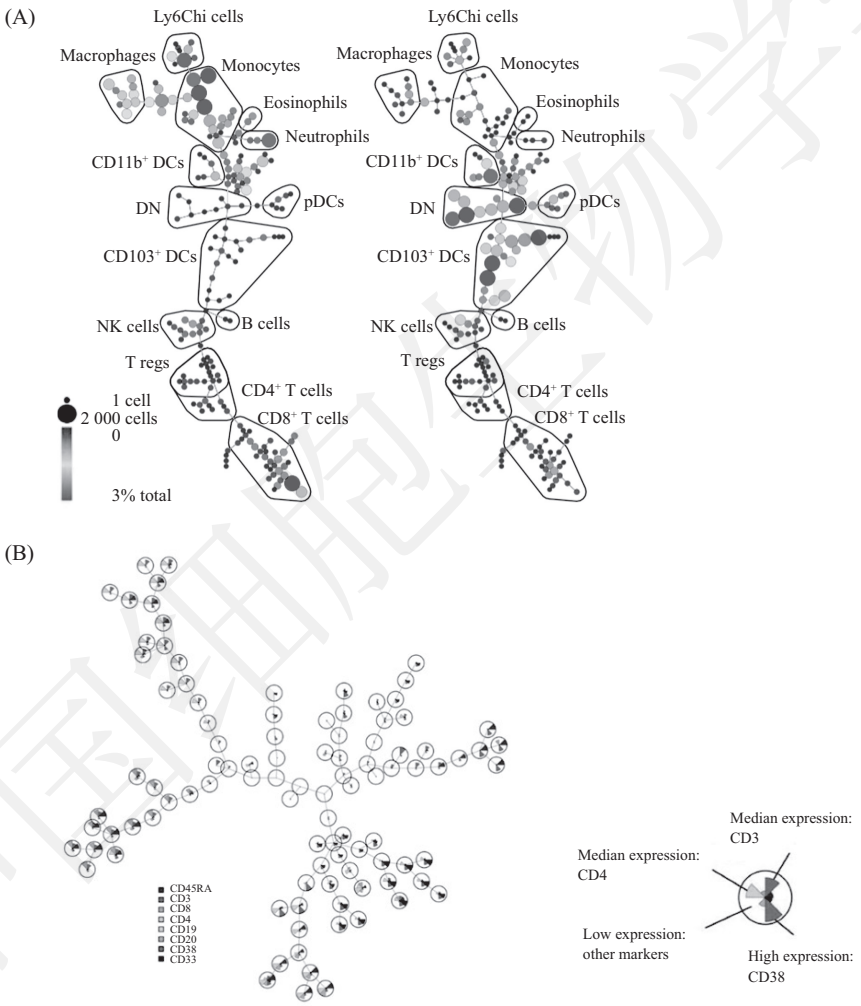
参数 Parameters	首台仪器 First instrument	CyTOF	CyTOF2	Helios
Sensitivity	140 Kcps/ppb for Tb159	200 Kcps/ppb for Tb159	400 Kcps/ppb for Tb159	600 Kcps/ppb for Tb159
Mass range	125-215 amu	103-193 amu	88-207 amu	75-209 amu
Spectrum generetion	76.8 KHz	76.8 KHz	76.8 KHz	76.8 KHz
Resolution	High resolution mode > 900 for Tb159 High sensitivity mode >500 for Tb159	>400 for Tb159	>400 for Tb159	>400 for Tb159
Dynamic range (magnitude)	4.0	4.5	4.5	4.5

物的平均表达水平, 如图6B所示, 与SPADE相比, 单个图中包含更多信息, 避免使用多张图确定不同标记物表达水平的需要, 简化了鉴别细胞亚群的过程。此外, FlowSOM还包含节点的元聚类选项: 元聚类函数可以基于细胞的表型相似性将节点分组在一起, 定义大的细胞群, 如B细胞、$CD4^+$ T细胞和单核细胞等。

SPADE和FlowSOM是强大的可视化工具, 但它们都依赖随机MST, 重复性略低。每进行一次SPADE或FlowSOM分析, 都会产生不同的分支结构和不同的MST。虽然两种算法都具有检测稀有亚群的能力, 但会依赖人为设定的预期亚群数目。

3.3 viSNE

viSNE是一种基于*t*分布随机相邻嵌入算法(*t*-distributed SNE, t-SNE)的高维单细胞数据可视化工具[43-44], 其目的就是将表型相似的细胞分组在一起, 并将表型不同的细胞分开。它是一种非线性降维算法, 能够在单细胞分辨率下, 显示细胞亚群的多样性。AMIR等[45]使用viSNE分析白血病表型的异质性。使用viSNE分析人骨髓样本, 发现健康骨髓样本与白血病样本之间及不同白血病样本之间的细胞亚群组成存在差异。如图7所示, 用viSNE分析不同的

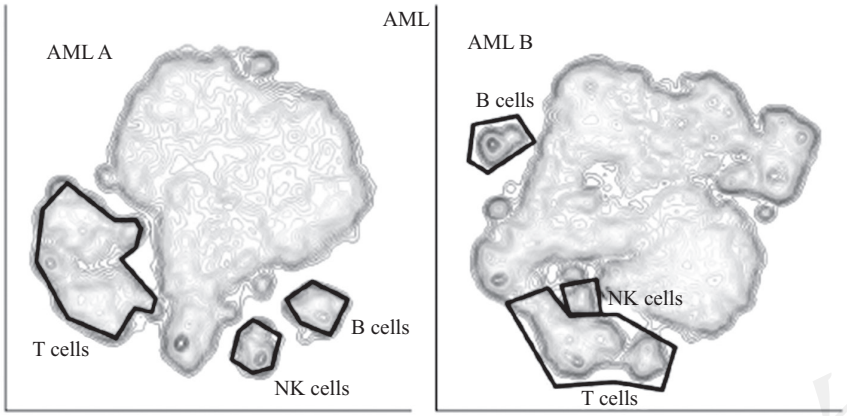


A: 从小鼠经腹腔注射B16肿瘤的第4天起, 连续9天分别注射30 mg PBS或FLT3L, 对第13天的肿瘤浸润免疫细胞进行CyTOF分析, 注射FLT3L小鼠的$CD103^+$树突细胞明显增多。B: 对人骨髓数据分析结果。只在星图中显示选定的标志物。通过比较这些标志物的表达, 可以区分$CD8^+$ T、$CD4^+$ T和B细胞。

A: CyTOF analysis of tumor-infiltrating immune cells on day 13 after B16 tumor challenge in mice treated i.p. for 9 consecutive days starting at day 3 after tumor injection with 30mg FLT3L or PBS. B: result for the human bone marrow dataset. Only a selected number of markers are shown in the star charts for clarity. By comparing these markers, we can identify $CD8^+$ and $CD4^+$ T cells and B cells.

图6 注射PBS或FLT3L后的SPADE视图和人骨髓瘤数据的FlowSOM视图(根据参考文献[41-42]修改)

Fig.6 SPADE visualization after PBS or FLT3L injection and FlowSOM visualization of human bone marrow (modified from references [41-42])



等高线代表细胞密度, 通过检测标志物的表达进行门控, 已标记的细胞亚群(T细胞、B细胞、NK细胞)代表健康免疫亚群。由于不同AML样本之间存在差异, 因此健康亚群分布于viSNE图的不同位置。

The contours represent cell density in each region of the map. Small gated populations represent healthy immune subtypes, as revealed by examination of their marker expression. Because the structure of each tumor dramatically changes between samples, viSNE places the healthy regions in different locations in each map.

图7 用viSNE分析不同AML样本(根据参考文献[45]修改)

Fig.7 Contour plots of the viSNE maps of two different AML (modified from reference [45])

AML样本, 发现T细胞、NK细胞、B细胞亚群存在明显差异。此外, viSNE也可分析白血病复发前后的独特细胞亚群。

3.4 Citrus

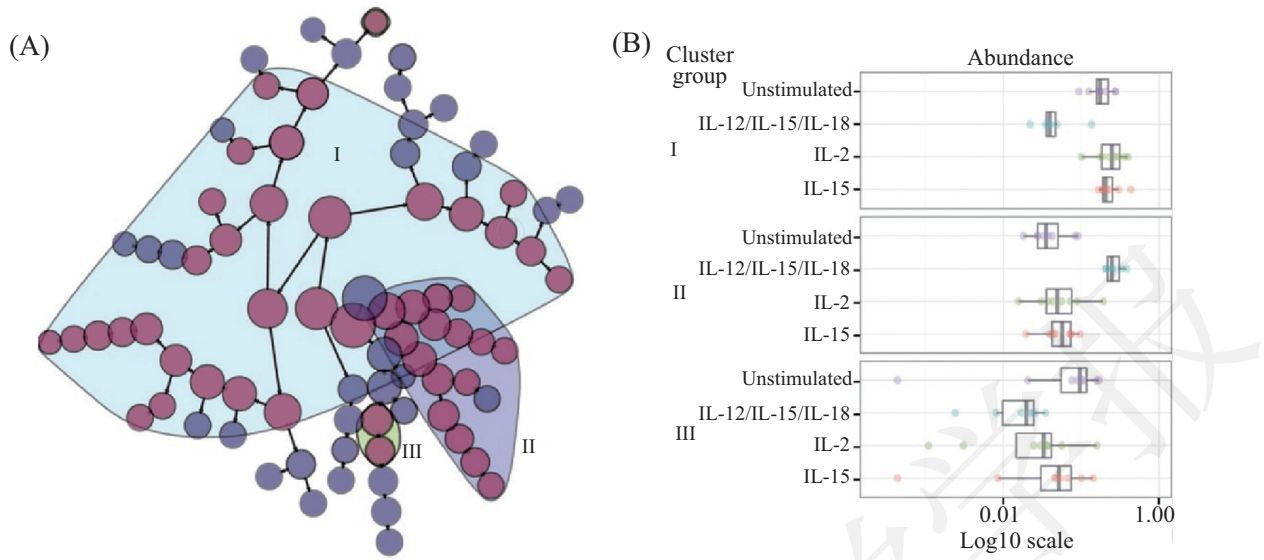
临床样本具有很大的异质性, 因此很难从蛋白的整体表达水平上发现具有统计学意义的差别, 有代表性的差别往往只存在于少数亚群中。viSNE和SPADE可以将样本分成许多亚群, 但是它们很难对细胞亚群中相关的标志物表达水平进行对比及相关性统计。Citrus(cluster identification, characterization, and regression)是一种识别与实验终点是否具有统计学差别的工具, 常见的实验终点包括治疗前或后, 疾病良或恶等。Citrus首先通过分层聚类以无监督的方式识别表型相似的细胞, 并将样本中相关细胞比例和标志物表达水平中值存储到矩阵中, 通过微阵列预测分析、L1惩罚回归等识别相应细胞亚群与实验终点之间的关系[46]。

VENDRAME等[47]使用Citrus分析NK细胞丰度并揭示细胞因子刺激后的差异。实验终点可视为未受刺激的细胞丰度, 然后使用Citrus分析刺激后的3个细胞亚群的丰度与实验终点的统计学差别, 实验结果如图8所示。GAUDILLIERE等[48]使用质谱流式检测髋关节置换手术后的全血样本, 并用Citrus分析患者临床参数(如疼痛、功能障碍等)与表型和功能性免疫反应之间的关系。研究发现, 临床参数与$CD14^+$单核细胞亚群中的STAT3(信号转导和转录激活因子)、CREB(腺苷3',5'-单磷酸反应元件结合蛋白)和NF-κB信号转导响应之间具有相关性。HANSMANN等[49]用质谱流式检测多发性骨髓瘤患者和健康样本外周血细胞标志物的表达, 并使用Citrus分析患者与健康样本之间存在显著差异的生物标志物, 最终发现了2个B细胞相关亚群。在这2个亚群中, CD27在多发性骨髓瘤病人组的表达量要明显高于正常人, 使CD27有希望成为该疾病诊断的标志。

3.5 应用于高维单细胞数据的算法总结

高维单细胞数据的日益复杂化需要不断地研究新的分析策略、开发新的计算工具, 从而收集有用的生物学信息。近几年, 质谱流式技术在高维单细胞分析方面发展迅速, 涌现出了一系列高维数据分析算法。表2中更详细全面地总结了应用于高维单细胞数据的各种算法及其优缺点。

分析高维单细胞数据的目的不仅仅是理解细胞群体各种生物概念之间的关系, 还要产生异质性亚群如何响应各种环境的可测性假设。然而, 用于分析高维单细胞数据的大多数算法都是通过降维、聚类等进行数据可视化或鉴别细胞重要进程来识别全局数据结构(Wanderlust等)。只有非常少的算法用于统计学推断(DREVI)或识别与期望结果具有相关性的细胞特征(如Citrus)。现在使用的算法工具往往是根据生物学知识进行探索和描述, 并不是通过相关生物学假设驱动的。



A: 3种不同的细胞亚群(I、II和III)。B: 在4种不同处理条件下, 细胞丰度显著不同。Citrus识别分层聚类亚群。紫色圆圈代表细胞簇, 蓝色圆圈代表低于阈值的聚类。

A: 3 group of clusters (I, II, III) were noted. B: whose cell abundances in the four treatment conditions are significantly different. Map of hierarchical clusters identified by Citrus. Purple circles represent clusters of cells, blue circles represent clusters below the significance threshold.

图8 Citrus对不同细胞因子刺激后的NK细胞丰度进行分析(根据参考文献[47]修改)

Fig.8 Citrus analysis of NK cell abundance after cytokine treatment (modified from reference [47])

表2 应用于质谱流式数据算法的总结与比较

Table 2 Summary and comparison of current algorithms applied to mass cytometry

处理方法 Process methods	算法 Algorithms	优点 Advantages	缺点 Limitations	参考文献 References
Dimensionality reduction	PCA	Linear dimensionality reduction; a few principal components represent basic information of orginal variables	No consideration of nonlinear relationships between parameters	[50-51]
	viSNE	Single-cell resolution; nonlinear dimensionality reduction	Difficult to compare groups of samples; rare cell subpopulations may be obscured	[45,52]
	ACCENSE	Combine nonlinear dimensionality reduction with density-based cluster; identify new or rare cell subpopulations	Difficult to compare groups of samples	[53]
Cluster	SPADE	Density-based unbiased cluster; visualize subpopulation relationship by tree structure	Loss of single-cell resolution	[13,40-41]
	FlowSOM	SOM; reduce computing time	Loss of single-cell resolution	[42]
	Scaffold map	Unbiased clustering of cell clusters; visualize subpopulation relationship by force-oriented algorithm; integrate biological knowledge to compare new data with references	Not mentioned	[50,54]
Cell state progression	Wanderlust	Define the most possible linear path for cells from beginning to end, and create cell development trajectory over time	The development process must be linear and unbranched, the bifurcated development trajectory cannot be dealt with	[55]
Sample prediction	Citrus	Identify subpopulation associated with experimental endpoints, correlate results with related biological characteristics	Multiple samples are needed for effective comparison	[46-49]
Deduce cell signal network	DREVI	Reveal and quantify the relationship between signaling pathway proteins	Not mentioned	[56-57]

4 总结与展望

质谱流式技术在单细胞检测方面有着独特的优势, 可在单细胞水平下同时检测多种细胞标志物, 增加了揭示各种生物机制的可能性。但目前这种技术仍有一定局限, 如细胞被雾化和电离后不可恢复细胞活性; 质谱流式的通量远远小于流式细胞仪的通量等。然而, 与传统流式细胞仪相比, 它具有金属标签数量多且背景信号低、数据处理方式多样化等优点。可以肯定的是, 随着质谱流式单细胞分析在通量、灵敏度等方面的不断提高, 生物学问题、实验设计和数据分析等连接在一起的总体分析框架, 最终会改变生物医学研究方式并加深我们对细胞生物学的理解。

参考文献 (References)

[1] MANDY F F, BERGERON M, MINKUS T. Principles of flow cytometry [J]. Transfus Sci, 1995, 16(4): 303-14.

[2] WEIR E G, BOROWITZ M J. Flow cytometry in the diagnosis of acute leukemia [J]. Semin Hematol, 2001, 38(2): 124-38.

[3] WOOD B L. Myeloid malignancies: myelodysplastic syndromes, myeloproliferative disorders, and acute myeloid leukemia [J]. Clin Lab Med, 2007, 27(3): 551-75.

[4] CHATTOPADHYAY P K, ROEDERER M. Cytometry: today's technology and tomorrow's horizons [J]. Methods, 2012, 57(3): 251-8.

[5] PERFETTO S P, CHATTOPADHYAY P K, ROEDERER M. Innovation: seventeen-colour flow cytometry: unravelling the immune system [J]. Nat Rev Immunol, 2004, 4(8): 648-55.

[6] BANDURA D R, BARANOV V I, ORNATSKY O I, et al. Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry [J]. Anal Chem, 2009, 81(16): 6813-22.

[7] SPITZER M, NOLAN G. Mass cytometry: single cells, many features [J]. Cell, 2016, 165(4): 780-91.

[8] BEHBEHANI G K, BENDALL S C, CLUTTER M R, et al. Single-cell mass cytometry adapted to measurements of the cell cycle [J]. Cytom Part A, 2012, 81A(7): 552-66.

[9] CORNEAU A, COSMA A, EVEN S, et al. Comprehensive mass cytometry analysis of cell cycle, activation, and coinhibitory receptors expression in cd4 t cells from healthy and HIV-infected individuals [J]. Cytom Part B Clin Cy, 2017, 92(1): 21-32.

[10] BEHBEHANI G K. Cell cycle analysis by mass cytometry In: Lacorazza H. (eds) Cellular Quiescence. Methods in Molecular Biology [M]. New York: Humana Press, 2018, 105-24.

[11] BANDYOPADHYAY S, FISHER D A C, MALKOVA O, et al. Analysis of signaling networks at the single-cell level using mass cytometry [J]. Methods Mol Biol, 2017, 1636: 371-92.

[12] CHOI J, FERNANDEZ R, MAECKER H T, et al. Systems approach to uncover signaling networks in primary immunodeficiency diseases [J]. J Allergy Clin Immun, 2017, 140(3): 881-4.

[13] BENDALL S C, SIMONDS E F, QIU P, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum [J]. Science, 2011, 332(6030): 687-96.

[14] ZHANG C, WU F, ZHANG Y, et al. A novel combination of immunoreaction and ICP-MS as a hyphenated technique for the determination of thyroid-stimulating hormone (TSH) in human serum [J]. J Anal Atom Spectrom, 2001,16(12): 1393-6.

[15] BARANOV V I, QUINN Z, BANDURA D R, et al. A sensitive and quantitative element-tagged immunoassay with ICPMS detection [J]. Anal Chem, 2002, 74(7): 1629-36.

[16] QUINN Z A, BARANOV V I, TANNER S D, et al. Simultaneous determination of proteins using an element-tagged immunoassay coupled with ICP-MS detection [J]. J Anal Atom Spectrom, 2002, 17(8): 892-6.

[17] ORNATSKY O, BARANOV V I, BANDURA D R, et al. Multiple cellular antigen detection by ICP-MS [J]. J Immunol Methods, 2006, 308(1/2): 68-76.

[18] TANNER S D, ORNATSKY O, BANDURA D R, et al. Multiplex bio-assay with inductively coupled plasma mass spectrometry: towards a massively multivariate single-cell technology [J]. Spectrochimica Acta Part B Atomic Spectroscopy, 2007, 62(3): 188-95.

[19] LOU X, ZHANG G, HERRERA I, et al. Polymer-based elemental tags for sensitive bioassays [J]. Angew Chem Int Edit, 2007, 46(32): 6111-4.

[20] ORNATSKY O I, KINACH R, BANDURA D R, et al. Development of analytical methods for multiplex bio-assay with inductively coupled plasma mass spectrometry. [J]. J Anal at Spectrom, 2008, 23(4): 463-9.

[21] THICKETT S C, ABDELRAHMAN A I, ORNATSKY O, et al. Bio-functional, lanthanide-labeled polymer particles by seeded emulsion polymerization and their characterization by novel ICP-MS detection [J]. J Anal at Spectrom, 2010, 25(3): 269-81.

[22] LIN W, MA X, QIAN J, et al. Synthesis and mass cytometric analysis of lanthanide-encoded polyelectrolyte microgels [J]. Langmuir, 2011, 27(11): 7265-75.

[23] MAJONIS D, HERRERA I, ORNATSKY O, et al. Synthesis of a functional metal-chelating polymer and steps toward quantitative mass cytometry bioassays [J]. Anal Chem, 2010, 82(21): 8961-9.

[24] ILLY N, MAJONIS D, HERRERA I, et al. Metal-chelating polymers by anionic ring-opening polymerization and their use in quantitative mass cytometry [J]. Biomacromolecules, 2012, 13(8): 2359-69.

[25] LIN W, HOU Y, LU Y, et al. A high-sensitivity lanthanide nanoparticle reporter for mass cytometry: tests on microgels as a proxy for cells [J]. Langmuir, 2014, 30(11): 3142-53.

[26] CHOW C S, BARTON J K. Transition metal complexes as probes of nucleic acids. [J]. Method Enzymol, 1992, 212(4): 219-42.

[27] CHEN W, TURRO C, FRIEDMAN L A, et al. Resonance raman investigation of Ru(phen) 2 (dppz) $^{2+}$ and related complexes in water and in the presence of DNA [J]. J Phys Chem B, 1997, 101(35): 6995-7000.

[28] DANDLIKER P J, HOLMLIN R E, BARTON J K. Oxidative thymine dimer repair in the DNA helix [J]. Science, 1997, 275(5305): 1465-8.

[29] PYLE A M, LONG E C, BARTON J K. Shape-selective targeting

of DNA by phenanthrenequinone diiminerhodium(III) photocleaving agents [J]. J Am Chem Soc, 1989, 111(12): 4520-2.

[30] PYLE A M, CHIANG M Y, BARTON J K. Synthesis and characterization of physical, electronic, and photochemical aspects of 9,10-phenanthrenequinonediimine complexes of ruthenium(II) and rhodium(III) [J]. Inorg Chem, 1990, 29(22): 4487-95.

[31] BARTON J K. Targeting DNA sites with chiral metal complexes [J]. Pure Appl Chem, 1989, 61(3): 563-4.

[32] ORNATSKY O I, LOU X, NITZ M, et al. Study of cell antigens and intracellular dna by identification of element-containing labels and metallointercalators using inductively coupled plasma mass spectrometry [J]. Anal Chem, 2008, 80(7): 2539-47.

[33] TANNER S D, BANDURA D R, ORNATSKY O, et al. Flow cytometer with mass spectrometer detection for massively multiplexed single-cell biomarker assay [J]. Pure Appl Chem, 2008, 80(12): 2627-41.

[34] KAISER Y, LAKSHMIKANTH T, CHEN Y, et al. Mass cytometry identifies distinct lung $CD4^+$ T Cell patterns in Löfgren's syndrome and non-Löfgren's syndrome sarcoidosis [J]. Front Immuno, 2017, 8: 1130.

[35] BODENMILLER B, ZUNDER E R, FINCK R, et al. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators [J]. Nat Biotechno, 2012, 30(9): 858-7.

[36] MEI H E, LEIPOLD M D, SCHULZ A R, et al. Barcoding of live human peripheral blood mononuclear cells for multiplexed mass cytometry [J]. J Immunol, 2015, 194(4): 2022-231.

[37] ZUNDER E R, FINCK R, BEHBEHANI G K, et al. Palladium-based mass-tag cell barcoding with a doublet-filtering scheme and single cell deconvolution algorithm [J]. Nat Protoc, 2015, 10(2): 316-33.

[38] NASSAR A F, WISNEWSKI A V, RADDASSI K. Automation of sample preparation for mass cytometry barcoding in support of clinical research: protocol optimization [J]. Anal Bioanal Chem, 2017, 409(9): 2363-72.

[39] FINCK R, SIMONDS E F, JAGER A, et al. Normalization of mass cytometry data with bead standards [J]. Cytom Part A, 2013, 83A(5): 483-94.

[40] QIU P, SIMONDS E F, BENDALL S C, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE [J]. Nat Biotechno, 2011, 29(10): 886-91.

[41] SALMON H, IDOYAGA J, RAHMAN A, et al. Expansion and activation of $CD103^+$ dendritic cell progenitors at the tumor site enhances tumor responses to therapeutic PD-L1 and BRAF inhibition [J]. Immunity, 2016, 44(4): 924-38.

[42] VAN GASSEN S, CALLEBAUT B, VAN HELDEN M J, et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data [J]. Cytom Part A, 2015, 87(7): 636-45.

[43] Maaten L, Hinton G E. Visualizing Data Using t-SNE [J]. J Mach Learn Res, 2008, 9(2): 2579-605.

[44] GISBRECHT A, SCHULZ A, Hammer B. Parametric nonlinear dimensionality reduction using kernel t-SNE [J]. Neurocomputing, 2015, 147(1): 71-82.

[45] AMIR E A D, DAVIS K L, TADMOR M D, et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia [J]. Nat Biotechno, 2013, 31(6): 545-52.

[46] BRUGGNER R V, BODENMILLER B, DILL D L, et al. Automated identification of stratifying signatures in cellular subpopulations [J]. Proc Natl Acad Sci USA, 2014, 111(26): 2770-7.

[47] VENDRAME E, FUKUYAMA J, STRAUSS-ALBEE D M, et al. Mass cytometry analytical approaches reveal cytokine-induced changes in natural killer Cells [J]. Cytom Part B Clin Cy, 2017, 92(1): 57-67.

[48] GAUDILLIERE B, FRAGIADAKIS G K, BRUGGNER R V, et al. Clinical recovery from surgery correlates with single-cell immune signatures [J]. Sci Transl Med, 2014, 6(255): 255ra131.

[49] HANSMANN L, BLUM L, HSIN-JU C, et al. Mass cytometry analysis shows that a novel memory phenotype B cell is expanded in multiple myeloma. [J]. Cancer Immunol Res, 2015, 3(6): 650-60.

[50] JACKSON J E. PCA with more than two variables In: A User's Guide to Principal Components. New Jersey: John Wiley & Sons, Inc [M]. 2004: 26-62.

[51] NEWELL E W, SIGAL N, BENDALL S C, et al. Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of $CD8^+$ T cell phenotypes [J]. Immunity, 2012, 36(1): 142-52.

[52] GISBRECHT A, SCHULZ A, HAMMER B. Parametric nonlinear dimensionality reduction using kernel t-SNE [J]. Neurocomputing, 2015, 147(1): 71-82.

[53] SHEKHAR K, BRODIN P, DAVIS M M, et al. Automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE) [J]. Proc Natl Acad Sci USA, 2014, 111(1): 202-7.

[54] SPITZER M H, GHERARDINI P F, FRAGIADAKIS G K, et al. An interactive reference framework for modeling a dynamic immune system [J]. Science, 2015, 349(6244): 1259425.

[55] BENDALL S C, DAVIS K L, AMIR E D, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development [J]. Cell, 2014, 157(3): 714-25.

[56] KRISHNASWAMY S, SPITZER M H, MINGUENEAU M, et al. Conditional density-based analysis of T cell signaling in single-cell data. [J]. Science, 2014, 346(6213): 1250689.

[57] MINGUENEAU M, KRISHNASWAMY S, SPITZEr M H, et al. Single-cell mass cytometry of TCR signaling: amplification of small initial differences results in low ERK activation in NOD mice [J]. Proc Natl Acad Sci USA, 2014, 111(46): 16466-71.