博奥晶典
CapitalBio Technology

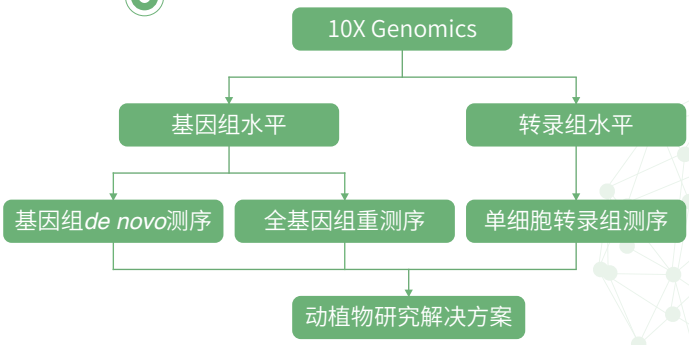
10X Genomics
动植物研究解决方案

北京博奥晶典生物技术有限公司

10X Genomics 新技术
全面解决动植物研究难题！

10X Genomics技术一经推出便被评为2015年生物技术十大创新之首！该平台解决了目前Illumina平台读长短、常规单细胞平台通量低的难题：1)通过特异识别序列barcode的加入，将分散的短reads序列进行连接，产生全新的数据形式Linked-Reads，这使获得测序序列的长片段信息成为轻而易举的事情；2)该平台一次性最多可以完成80,000个细胞的分离和单细胞转录组测序，能充分解析组织样品的细胞异质性。

研究策略



优势总结

基因组研究优势
长片段DNA（>50Kb）建库，以超低成本获得大结构变异和高质量的物种参考基因组

单细胞研究优势
一次单细胞转录组建库测序最多可获得10,000个细胞的单细胞转录组数据

项目经验优势
作为国内首家引进10X Genomics平台的服务商，共构建基因组和单细胞文库600+个

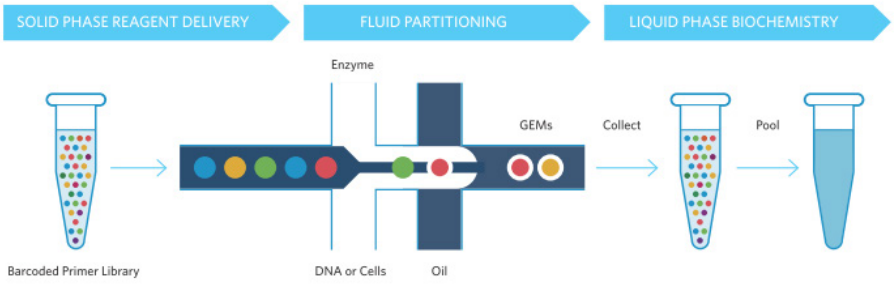
文章成果优势
支持客户发表国内首篇文章PNAS（IF=9.423）和国内最高分值文章Cell Stem Cell（IF=23.290）

关于 10X Genomics 平台建库系统

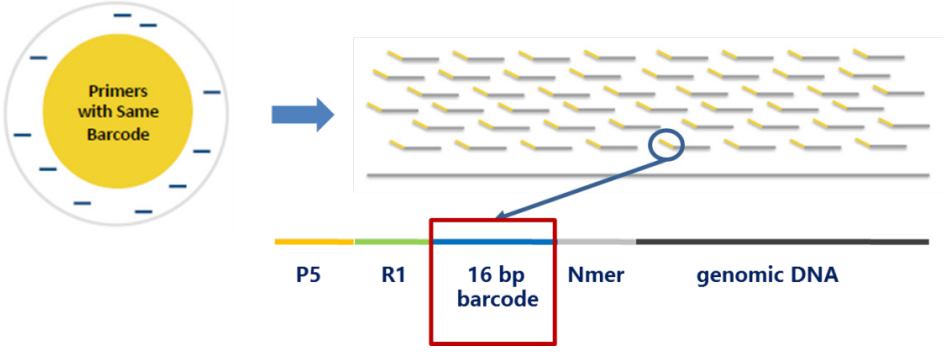
基本技术原理

通过微流体芯片将长片段 DNA（>50Kb）或单个细胞分配到单个油包水液滴中（液滴总数量大于 400 万个）独立进行反应。每个液滴使用唯一 barcode 标签序列进行标记，清楚区分测序 reads 来源的大片段 DNA。

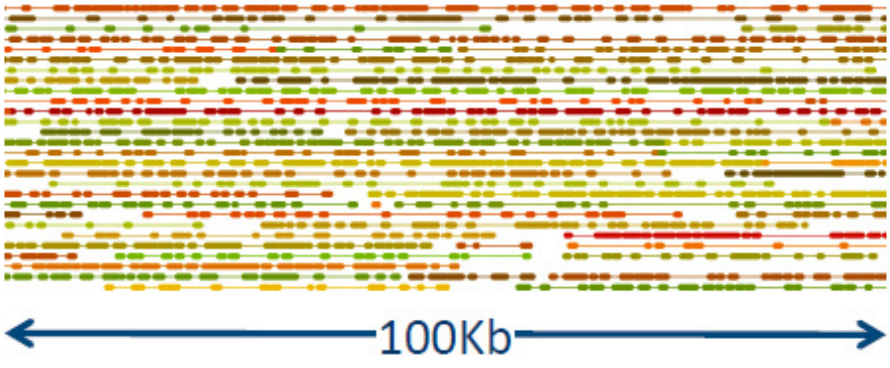
- 微流体芯片分配大片段 DNA 或单个细胞，形成油包水液滴



- Linked-Reads 形成原理



每个液滴具有唯一 barcode 序列，来源于相同 DNA 片段的 reads 具有相同的 barcode。



通过将来自同一长片段模板具有相同 barcode 的 reads 进行连接，输出超长 Linked-Reads。

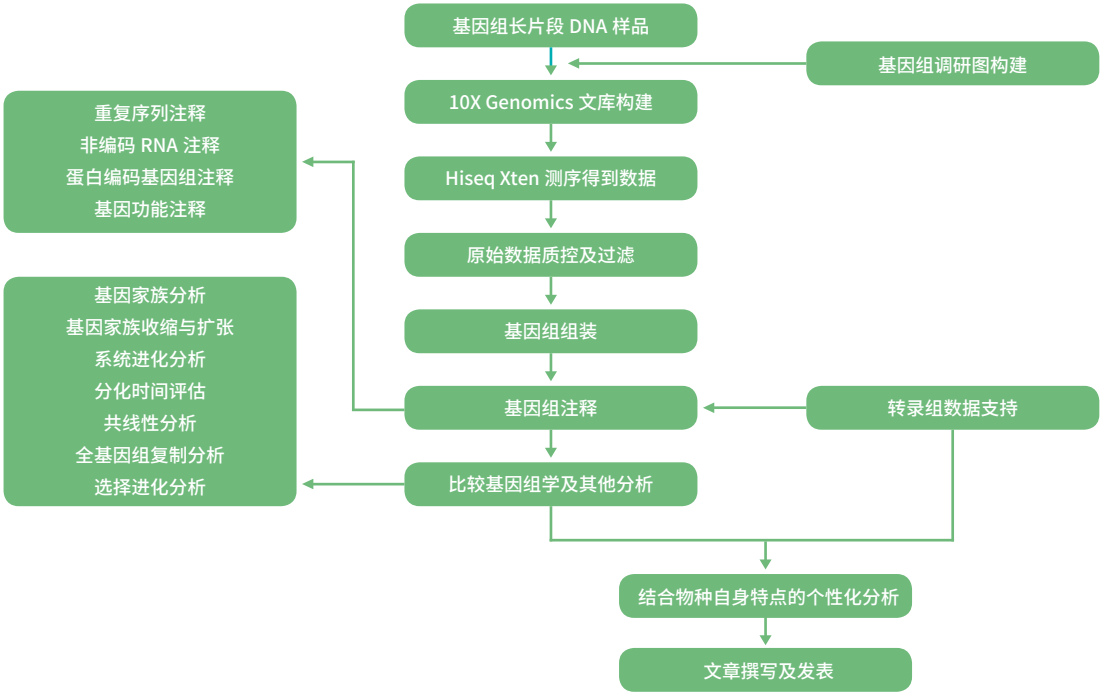
1 解决方案一：10X Genomics 基因组 *de novo* 测序

利用 10X Genomics 的 barcode 序列信息，并利用 Supernova ™ Assembler 软件，从而实现对动植物基因组的从头组装。该组装策略突破传统二代测序组装方法，以其超长 Linked-Reads 的优势，可实现低成本、低 DNA 输入，短周期内获得该物种的基因组。对获得的基因组进行注释、比较基因组分析，可实现大片段结构变异、分子系统进化、功能基因挖掘的目的，从而为特定物种的生物学问题提供相应解决方案，也为该物种的后基因组学研究奠定基础。

方案优势

- 大跨度 Linked-Reads 为 Contig 准确组装成 Scaffold 提供充分支持
- 仅需构建一个 10X Genomics 文库，成本超低
- 组装算法的优化大幅缩短了组装周期
- 作为国内第一家引进 10X Genomics 平台的服务商，博奥晶典具有丰富的项目经验
- 对已有基因组序列，能够有效延长 Scaffold 长度

方案信息分析内容



方案案例

10X Genomics 官方对部分具有代表性物种基因组进行了组装，该技术对哺乳动物、鸟类、爬行动物、昆虫和植物均能具有很好的组装效果（表 1.1）。另外，博奥晶典已完成大量动植物的 10X Genomics 基因组组装，在动植物基因组组装方面积累了丰富的项目经验（表 1.2）。

表 1.1 10X Genomics 官方完成部分物种基因组组装结果

Genome	Size (Gb)	DNA size (Kb)	Contig N50 (Kb)	Scaffold N50 (Mb)
人（HGP）	3.27	139	162	45.6
人（NA12878）	3.39	95	165.5	39.9
人（NA19238）	3.28	119	159	49.4
狗	24	81	100.4	36.2
蜂鸟	1.1	66	175	31.8
科莫多巨蜥	1.8	85	95	10.2
斑马鱼	1.68	89	20.5	4.0
果蝇	0.14	68	166.5	20.5
飞蛾	0.33	22	107.8	6.7
葡萄	0.6	74	55.7	2.3
辣椒	3.21	45	167.2	13.6

表 1.2 博奥晶典 10X Genomics 基因组组装案例（部分）

物种	Size（Gb）	Contig N50（Kb）	Scaffold N 50（Mb）	组装软件
水稻	0.38	56	1.45	Supernova 1.2
某果树	0.78	45	1.15	Supernova 1.2
某鱼类	0.43	42	5.38	Supernova 2.0
某鸟类	1	72.6	9.1	Supernova 1.2
某哺乳动物 1	2.5	82.6	12.4	Supernova 1.2
某哺乳动物 2	2.3	108	40.8	Supernova 2.0

方案应用思路

目前，一个物种的研究仅依赖一个参考基因组能解决的生物学问题越来越有限，在一个研究中构建多个个体基因组图谱越来越受到研究者和高分文章的青睐，也逐渐成为研究的新方向。

利用 10X Genomics 技术构建物种多个基因组，在完成基因组注释和比较基因组分析基础上，充分汇总各个物种 / 品种 / 个体的遗传信息，充分挖掘其中的特有基因和共有基因信息，并挖掘其变异信息（尤其是大的结构变异），可用来探究不同物种 / 品种 / 个体间的特有性状形成机制，同时也可对不同物种间的进化问题进行深入解读。

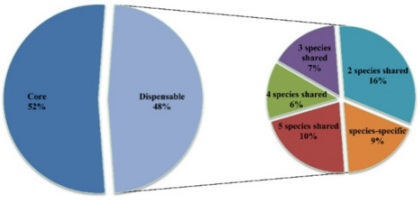


图 1.1 不同柑橘中基因组中特异基因和共有基因占比 [1]

2017 年发表在 Nature Genetics 上的 4 个柑橘泛基因组，对柑橘基因组的基因组资源进行了充分挖掘，结果表明 4 种柑橘中共有基因仅占 58%，近一半（48%）的基因资源为不同柑橘所特有。

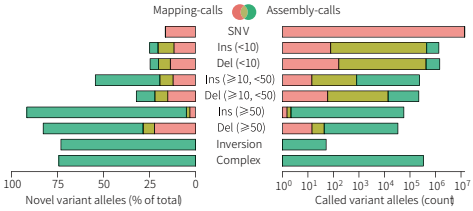


图 1.2 基于 mapping 和基于 *de novo* 组装的变异检测效果对比 [2]

2017 年发表在 Nature 上的丹麦人泛基因组研究中，研究者发现，随着变异信息复杂度的增加，基于 mapping 的检测效果急剧下降，对于倒位和复杂变异检测率几乎为 0，而基于 *de novo* 组装检测效果却很好。

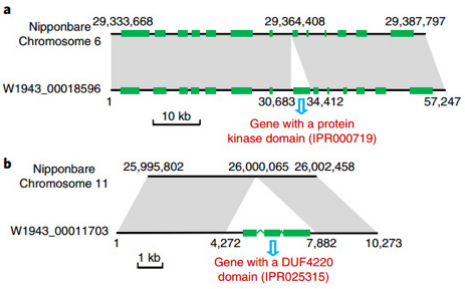


图 1.3 基于序列组装的水稻 SV 检测 [3]

2018 年发表在 Nature Genetics 上的 66 个水稻泛基因组，基于组装后的基因组序列充分挖掘了水稻属的变异位点和基因资源，弥补了日本晴参考基因组中丢掉的信息。该研究还解析了野生稻和栽培稻的驯化机制，鉴定到许多先前没有发现的与重要性状相关的 QTL 位点和基因。

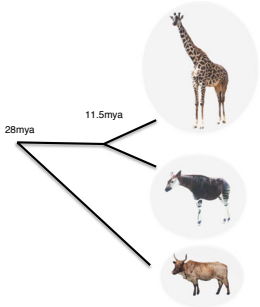


图 1.4 长颈鹿和霍加狒系统进化及分化时间分析 [4]

2016 年发表在 Nature Communications 上的长颈鹿基因组研究中，通过构建长颈鹿和其近缘种霍加狒基因组，发现两种动物发育的差别与一组控制身体和四肢发育的基因表达相关，揭示了长颈鹿独特体型可能与其生理循环系统的演化有关系，此外也明晰了二者的进化地位和分化时间。

方案发表文献

- **7 个人样品基因组图谱构建** [5]

2017 年 4 月发表在 Genome Research 上的研究，利用 10X Genomics 组装了 7 个人的基因组，每个样品仅需 1.25ng DNA，构建 1 个文库，测序深度 56X。其结果 Contig N50 > 106Kb，Scaffold N50 > 15Mb，并对基因组进行了分型，Phase block N50 : 2.7 ~ 9.3Mb。

表 1.3 7 个不同人基因组 10X Genomics 组装结果

Input							Continuity			
Id[a]	Sample[b]	Ethnicity[c]	Sex[d]	Data description[e]	X[f]	F[g]	N50 contig (kb)[h]	N50 phase block (Mb)[i]	N50 scaffold (Mb)[j]	Gappiness[k]
A	NA19238	Yoruban	F	One 10× library	56	115	114.6	8.0	18.7	2.1
B	NA19240	Yoruban	F	One 10× library	56	125	118.8	9.3	16.4	2.3
C	HG00733	Puerto Rican	F	One 10× library	56	106	123.6	3.4	17.8	2.0
D	HG00512	Chinese	M	One 10× library	56	102	113.2	2.7	15.4	2.2
E	NA24385	Ashkenazi	M	One 10× library	56	120	106.4	4.2	15.1	2.6
F	HGP	European	M	One 10× library	56	139	120.2	4.5	18.6	2.5
G	NA12878	European	F	One 10× library	56	92	118.5	2.8	16.4	2.0

僧海豹基因组图谱构建 [6]

来自约翰霍普金斯大学医学院的研究团队利用 10X Genomics 构建了僧海豹的基因组，同时也用常规二代构建其基因组。10X Genomics 组装版本 Scaffold N50 高达 22.23Mb，加入 BioNano 提升后，Scaffold N50 更是提高到 29.63Mb，为常规二代版本的 215 倍，几近染色体水平（表 1.4）。对三个版本得到的基因组所有序列长度进行统计比较发现，10X Genomics 组装版本序列长度远高于常规二代版本（图 1.5）。

表 1.4 僧海豹基因组组装结果统计

Assembly	BNG_10X_hybrid Supernova 1.0	BNG_10X_hybrid Supernova 1.1	10X Supernova v1.1	DISCOVAR
Total scaffolds	216	170	7,932	437,230
Largest scaffold	70.55 Mbp	84.77 Mbp	84.06 Mbp	1.10 Mbp
N50	20.87 Mbp	29.65 Mbp	22.23 Mbp	137,851
N50 Fold Improvement	105	**215**	161	1
L50	49	26	36	5100
# N's per 100 kbp	254.65	2,183.45	2,195.29	96.06
Est. total assembly length	2,318 Mbp	2,360 Mbp	2,400 Mbp	2,462 Mbp

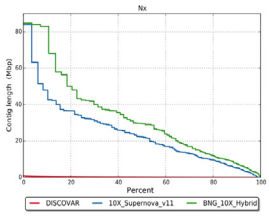


图 1.5 僧海豹基因组三种组装策略序列长度统计

辣椒基因组图谱构建 [7]

加州大学戴维斯分校的研究团队利用 10X Genomics 重新构建了辣椒基因组，在此之前已经有 2 篇辣椒基因组图谱构建文章分别发表在 Nature Genetics 和 PNAS 上，涉及三个版本的辣椒基因组，其中最大的 Contig N50 为 55Kb，最大的 Scaffold N50 为 2.47Mb，利用 10X Genomics 构建的辣椒基因组 Contig 为先前版本的两倍多，图 1.6 中看出，新版本的 Contig 长度明显增长，25Kb 以上的 Contig 占基因组的比例远高于其他三个发表版本，该版本基因组成为迄今最为连续的辣椒基因组。此外本研究构建基因组后找到了辣味相关基因 *pun1*，发现 *pun1* 基因 2.57Kb 的缺失，导致了辣椒辣味的丧失（图 1.7）。

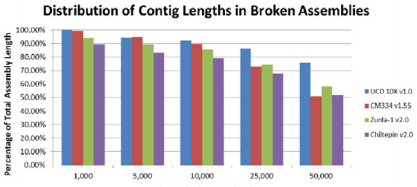


图 1.6 不同版本辣椒基因组 Contig 序列长度分布

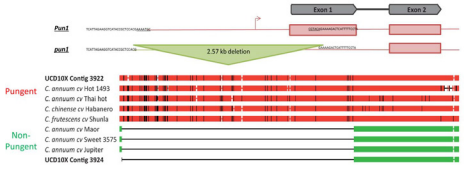


图 1.7 辣椒 *pun1* 基因序列比对

2 解决方案二：10X Genomics 全基因组重测序

基于 10X Genomics 平台建库的全基因组重测序，依托其超长 Linked-Reads 的优势，可以精确获得基因组结构变异信息，包括 InDel、SV、CNV 和 Fusion Gene，通过 10X Genomics 平台独特的 Phasing 功能，轻松获得样本的单体型信息，更精准地解读动植物基因组的遗传信息，助力动植物基因定位、分子系统进化及功能基因验证等研究。

方案优势

- 之前由于受到测序成本的限制，对于 SV 的研究较少，但目前已经到了从 SNP 到 SV 研究的转型时代
- 对于 SV 检测的准确率提高了 30% 以上，可实现染色体 Phasing，获得精确的单体型信息，对动植物分子辅助育种具有重要意义
- 技术新颖，结果靠谱度高，所有结果可视化，发表文章较容易
- 博奥晶典项目经验丰富，已经建库 300+ 个，可提供稳定可靠的结果

方案技术流程及信息分析内容

方案案例

- **博奥晶典 10X Genomics 全基因组重测序对某作物 BSA 定位结果区域进行精细定位**

  研究对象：某作物突变型个体和野生型个体各 1 个

  测序策略：≥ 40X/ 个样本

  精细定位结果：对于某作物前期的研究利用 BSA 定位区间为 55Mb, 博奥晶典通过 10X Genomics 大片段建库进行重测序分析 , 分别对两个材料检测 SNP、InDel、SV, 比较两材料之间 SNP、InDel 和 SV 的差异 , 根据突变位点所在的基因组位置 , 筛选突变基因。最终 , 在 5 号染色体上定位到性状关联的基因 , 并获得此基因的序列信息 , 进一步缩小了定位的区间 , 实现了精细定位的目的 , 并且结果采用全可视化展示方式 , 更加清晰方便。

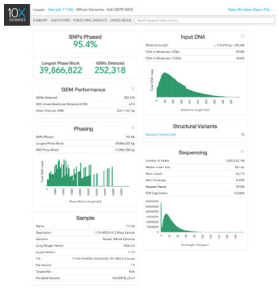


图 2.1 统计结果可视化

图 2.2 结构变异结果可视化说明

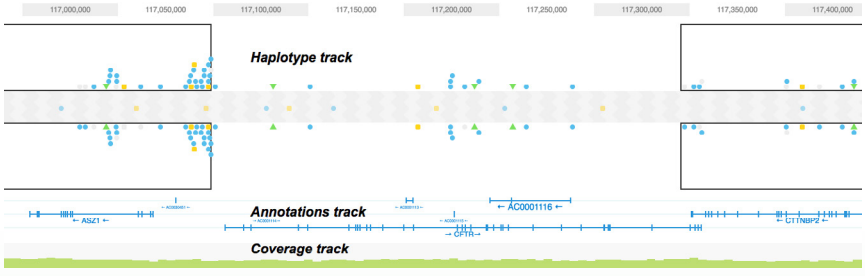


图 2.3 Haplotype 可视化说明

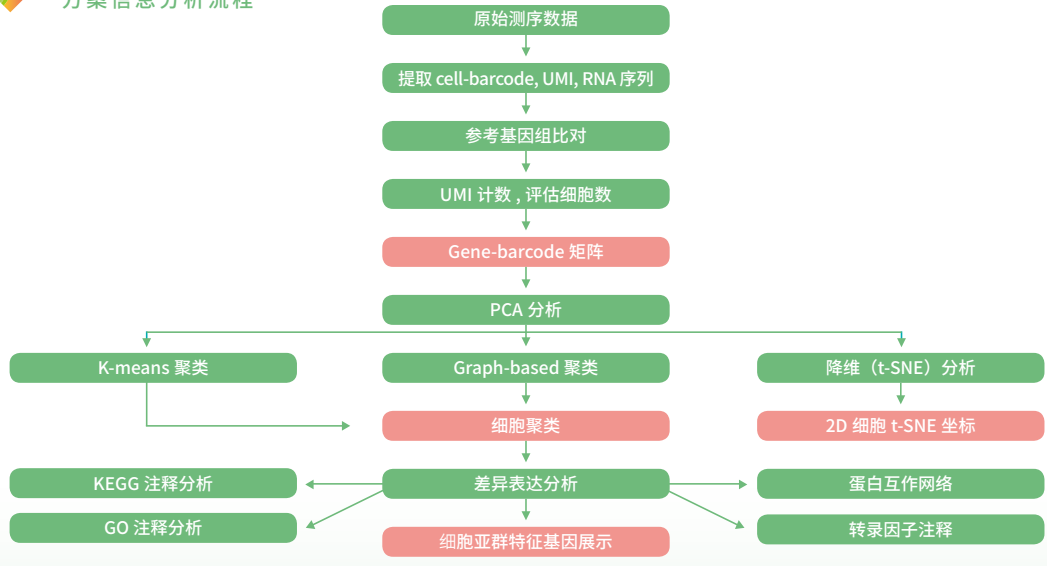
3 解决方案三：10X Genomics 单细胞转录组测序

Single Cell 3' Solution 是一种在单个细胞水平对 mRNA 进行高通量测序的新技术，为单细胞转录组研究提供更具扩展性的平台。其基于 10X Genomics 平台，一次性能够分离 500~10,000 个单细胞，并能够将分离的单个细胞中的微量 mRNA 通过高效扩增后再进行高通量测序，有效地解决了组织样本无法破解的细胞异质性难题，有助于发现新的细胞类型；并基于大量的单细胞基因表达数据，可进行细胞表达特征聚类、亚群表达特征、标志物筛选等方面的分析，是一种高效的细胞基因表达水平的检测技术。

方案优势

- **细胞选取灵活：** 每个样本细胞数可达 500-10,000 个，细胞选取灵活；
- **项目周期短：** 10 min 内实现对上万个细胞的封装，2 天内即可完成细胞悬液制备、单细胞捕获、扩增及建库，且可同时获得上万个细胞的分析结果，大大缩短项目周期；
- **细胞捕获效率高：** 单个细胞捕获率高达 65%，高效获取每个细胞中的基因表达；
- **真正大规模单细胞测序：** 单个液滴捕获到多个细胞的概率极低，0.9%/1,000 cells，严格质控，利用 barcode 信息及 UMI 去除含多个细胞数据，实现真正的大规模单细胞测序；
- **低成本：** 相较于其他单细胞平台，如 Fluidigm C1 平台、已推的 Illumina Bio-Rad Single Cell Solution 平台或传统人工操作扩增方法等，成本大幅降低；
- **应用范围广：** 对细胞类型无限制，已成功应用于识别细胞周期、肿瘤细胞异质性、鉴定罕见细胞类型、免疫细胞分型、疾病分型等领域；
- **高性价比：** 与传统手工细胞分离测序方法相比，价格优惠了数十倍；
- **丰富的项目经验：** 作为国内第一家引进 10X Genomics 平台的公司，博奥晶典率先推出了 Chromium™ Single Cell 3' Solution 平台服务，并完成了人和小鼠的肝脏、肺脏、皮肤、心脏、脑、肾脏、血液、卵巢，以及水稻、拟南芥、白菜、油菜等植物原生质体的单细胞测序。
- **文章优势：** 支持客户发表国内首篇文章 PNAS（IF=9.423）和国内最高分值文章 Cell Stem Cell（IF=23.290）。
- **北上广三城实验室及必要情况下的仪器上门服务，全程保障您的样品安全和实验稳定性**

方案信息分析流程



方案应用方向

1. 干细胞或者原代细胞的性质
2. 植物细胞对环境的适应性
3. 特定部位的细胞对于不同环境的反应程度
4. 植物发育

方案案例

- 细胞系单细胞转录组测序结果评估

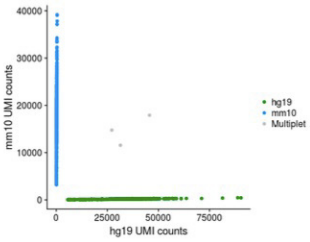


图 3.1 293T 和 3T3 细胞混合后检出散点图

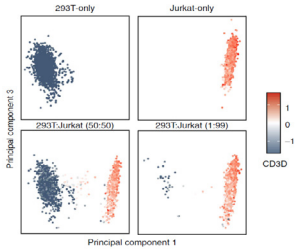


图 3.2 主成分分析展示

- 细胞分群

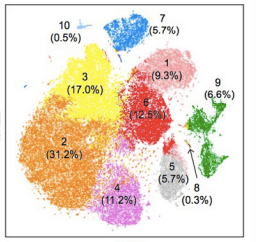


图 3.3 细胞分群二维展示图

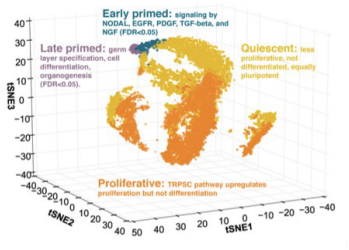


图 3.4 细胞分群三维展示图

- 差异表达分析、鉴定亚群特征基因

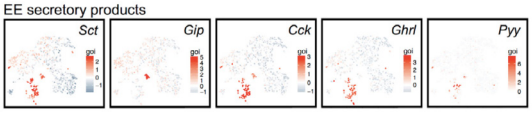


图 3.5 细胞亚群特定基因表达热图

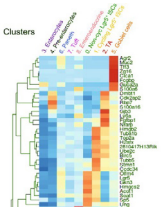


图 3.6 细胞亚群特征基因聚类图

- 差异基因功能注释和富集分析

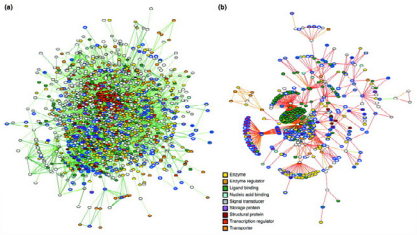


图 3.7 蛋白互作网络分析

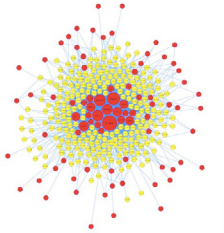


图 3.8 转录因子注释分析

方案发表文献

截止到 2018 年 07 月，Chromium™ Single Cell 3' Solution 平台已经发表 97 篇高水平文章（Nature、Cell、Nature Methods 等），总影响因子超过 1851 分，已经广泛应用于方法学研究、细胞发育、免疫反应、干细胞分化、肿瘤异质性等研究领域。

表 3.110X Genomics 单细胞转录组测序部分已发表文章展示（红色表示博奥支持文章）

发表时间	发表期刊	影响因子	研究方向
2018.06	***Cell Stem Cell***	23.290	大规模单细胞转录组测序揭示化合物诱导体细胞重编程早期胚胎样过程
2018.04	*Cell*	28.710	群体大规模宏观单细胞分析映射人造血分化的连续调节景观
2018.04	*Nature biotechnology*	43.113	对 CRISPR-Cas9 诱导的遗传印记同时进行谱系追踪和细胞类型鉴定
2018.04	*Cell Stem Cell*	23.290	粘膜下层肌上皮细胞可作为气道损伤再生的储备干细胞
2018.04	*Science*	34.166	单细胞转录组测序揭示小鼠肾脏疾病的潜在细胞靶点
2018.04	*Nature Genetics*	31.616	单细胞 RNA 测序鉴定细胞类型特异性顺式 eQTLs 和共表达 QTLs
2018.03	*Nature*	40.137	皮质抑制性中间神经元的发育多样化
2018.03	***PNAS***	9.423	I 型肺泡在个体分化发育中的命运
2018.02	*Cell*	28.71	BAFME 的发病机制导致负责肌阵挛性震颤和癫痫的神经元功能障碍
2018.02	*Nature Immunology*	21.056	单细胞转录组表征肿瘤中的内皮细胞异质性和抗血管生成治疗后的变化
2018.02	*Nature Immunology*	21.056	单细胞基因表达揭示了由 TCR 形成的调节性 T 细胞表型的景观
2018.01	*Nature Neuroscience*	16.724	单细胞 RNA 测序揭示出生后发育中齿状回神经发生的保守性质
2018.01	*Nature Cell Biology*	18.699	小鼠器官发育阶段白三烯途径在调节血液祖细胞形成中的作用
2017.11	*Nature*	38.138	在单细胞水平研究小肠上皮细胞
2017.10	*Nature biotechnology*	43.113	单细胞转录组测序 + 单细胞蛋白
2017.09	*Nature*	38.138	肺炎机制 -- 神经肽 NMU 放大了 ILC2 驱动的过敏性肺炎
2017.05	*Nature*	40.648	哺乳动物肠道干细胞的分化 Wnt 和 RSPO 调节机制
2016.12	*Cell*	31.356	基因修饰对细胞造成的影响及这些基因如何相互作用和依赖

参考文献

1. Wang X, Xu Y, Zhang S, et al. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction[J]. Nature Genetics, 2017, 49(5):765-772.

2. Maretty L, Jensen J M, Petersen B, et al. Sequencing and de novo assembly of 150 genomes from Denmark as a population reference.[J]. Nature, 2017, 548(7665):87.

3. Zhao Q, Feng Q, Lu H, et al. Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice[J]. Nature Genetics, 2018, 50(2):278.

4. Agaba M, Ishengoma E, Miller W C, et al. Giraffe genome sequence reveals clues to its unique morphology and physiology[J]. Nature Communications, 2016, 7:11519.

5. Weisenfeld N I, Kumar V, Shah P, et al. Direct determination of diploid genome sequences.[J]. Genome Research, 2017, 27(5):757.

6. Mohr D W, Naguib A, Weisenfeld N, et al. Improved de novo Genome Assembly: Linked-Read Sequencing Combined with Optical Mapping Produce a High Quality Mammalian Genome at Relatively Low Cost[J]. bioRxiv, 2017: 128348.

7. Hulse-Kemp A M, Maheshwari S, Stoffel K, et al. Reference Quality Assembly of the 3.5 Gb genome of Capsicum annuum from a Single Linked-Read Library[J]. bioRxiv, 2017: 152777.

©2018 CapitalBio Technology. All rights reserved.

KYFW-20180003 V1.0

博奥晶典控股成员

博奥检验 CapitalBio MedLab　博奥木华基因 CapitalBio Genomics　博奥新景 CapitalBioDx　博奥颐和 CapitalBio eHealth

北京博奥晶典生物技术有限公司

地址：北京市经济技术开发区科创六街生物医药园 C 座　邮编：101111　电话：010-69002900　传真：010-69002988　客服：400-9968-999
http://www.capitalbiotech.com