

生命科学服务 动植物研究产品手册



公司简介

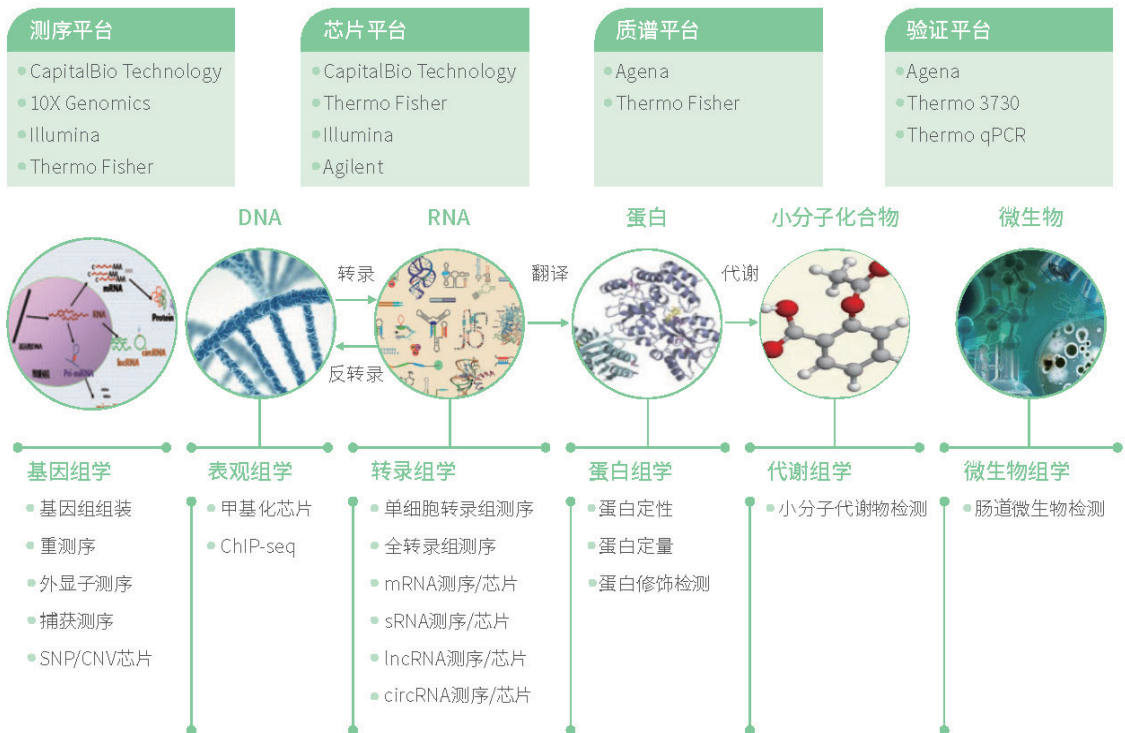
北京博奥晶典生物技术有限公司(以下简称“博奥晶典”)是博奥生物集团有限公司暨生物芯片北京国家工程研究中心的产业化平台。博奥晶典是亚洲唯一一家入围 MAQC 的机构,被 Nature 列为全球知名微阵列产品和服务供应商,最早通过 17025 国家合格实验室认可认证的科研服务机构。

博奥晶典生物检测实验室实力

- 国内首家引进 10X Genomics 平台的服务机构
- 国内最大的 Affymetrix 和 Agilent 芯片认证服务商
- 国内最大 Agena SNP 质谱检测服务机构
- RayBiotech 公司蛋白芯片认证服务商
- 国内首批部署单机通量最高的 NovaSeq 测序服务机构
- 国内唯一提供高通量特征基因表达谱 HTS² 技术服务机构
- 国内首家完成蛋白全谱检测服务机构
- 芯片、测序和蛋白质谱检测平台通过CNAS认证

博奥晶典 科研服务 农业研究解决方案

致力于打造中国最完善的高通量生物技术服务平台



至今,共支持客户发表 SCI 文章 1500 余篇,影响因子累计超过 6700 分,影响因子>5 的文章超过 360 篇,影响因子>10 的文章超过 70 篇,影响因子>20 的文章超过 32 篇。

01 基因组学

- 01 基因组调研图
- 02 10X Genomics 基因组 *de novo* 测序
- 06 全基因组重测序
- 12 10X Genomics 重测序
- 14 农业 SNP 芯片分型
- 19 Agena MassARRAY® SNP 服务平台
- 20 IMAP 检测服务

02 转录组学

- 21 10X Genomics 单细胞转录组测序
- 25 mRNA 测序
- 27 lncRNA 测序
- 29 Small RNA 测序
- 32 全转录组测序

03 表观组学和蛋白组学

- 35 Agena MassARRAY® 甲基化检测
- 36 蛋白组研究

1 基因组调研图 (Genome Survey)

高通量测序技术迅猛发展,极大推动了基因组的 *de novo* 组装,选择合适的建库测序策略对于一个基因组高质量组装是十分必要的。基因组特征的准确评估对于选择更好的测序方式组装出高质量基因组十分重要,并且基因组大小的准确评估还可以预估测序所需的深度,更好地避免测序时数据量不足或超出太多等问题。

技术流程



图 1.1 项目流程图

产品应用

针对基因组的复杂程度,一般可分为简单基因组和复杂基因组。通过构建基因组调研图,快速获得基因组大小、杂合率、重复率等基本信息,为制定该物种的全基因组组装测序策略提供有效依据。

案例解析

● 鱼类调研图构建

通过基因组 k-mer 分布图,可初步判断基因组基本特征,通过进一步公式计算基因组大小、杂合率、重复率。

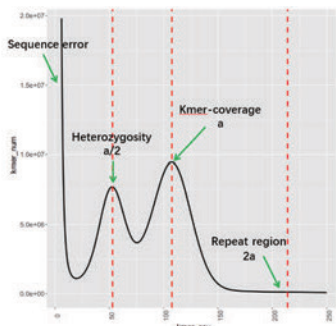


图 1.2 kmer 频率分布图

注 x轴为 kmer coverage, y轴为 kmer 频数。x, y(90, 10⁶) 表示有 10⁶ 个不同的 kmer 均出现 90 次

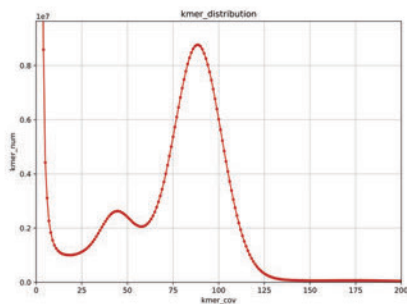


图 1.3 kmer 频率分布图

注: 横坐标为 kmer coverage 即 kmer 覆盖度, 纵坐标为 kmer 频数

x 轴最左端出现的峰为测序错误峰，可用于评估测序错误率；k-mer coverage 即 x 轴为 a 处有一个主峰，对应的 a 值可用于评估基因组大小；在 $x=a/2$ 处同样出现了一个峰，称之为杂合峰，用于评估基因组杂合度；基因组重复序列 k-mer 出现会成倍增加，因而在 x 轴为 2a 后的区域主要用于评估基因组重复。若是出现两个主峰，峰高相差不大，两峰横坐标又是 2 倍关系，说明该个体为高杂合或高重复。

如对某鱼类进行调研图的构建（图 1.3），k-mer 分析评估得到物种基因组大小为 378.68M，杂合为 0.52%，重复序列比例为 11.27%。并且对提取 DNA 的外源污染情况和基因组 GC 含量分布进行分析。

2 10X Genomics 基因组 *de novo* 测序

利用 10X Genomics 的 barcode 序列信息，采用 Supernova™ Assembler 软件，可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发现、分子系统进化研究、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。

技术路线

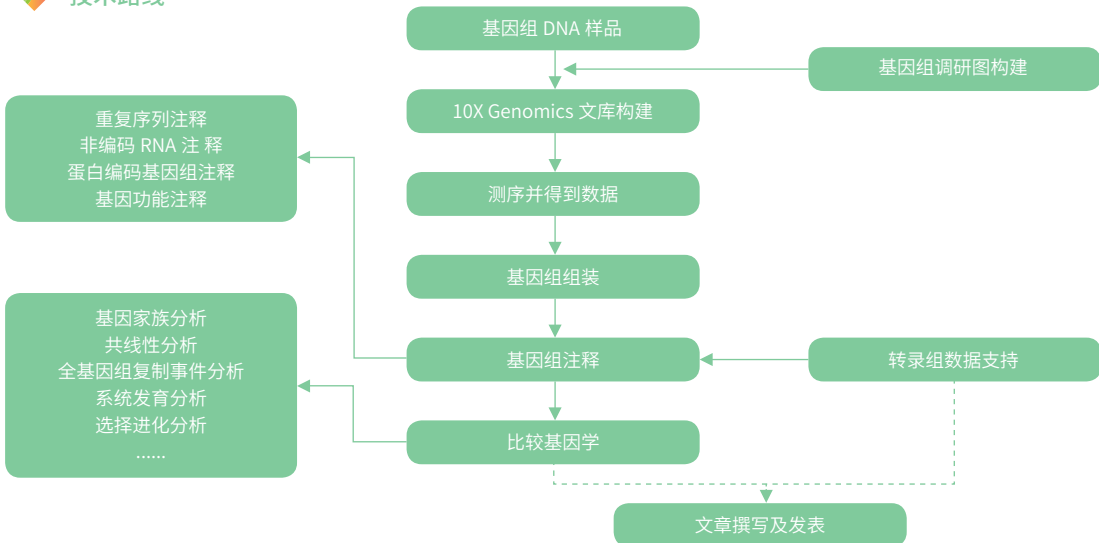


图 2.1 项目流程图

产品优势

- 1 大跨度 Linked-reads 为 Contig 准确组装成 Scaffold 提供充分支持
- 2 仅需构建一个 10X Genomics 文库，成本超低
- 3 组装算法的优化大幅缩短了组装周期
- 4 作为国内第一家引进 10X Genomics 平台的服务商，博奥晶典已成功建库 300+ 个，具有丰富的项目经验，可提供稳定可靠的结果
- 5 对已有的基因组序列，能够有效延长 Scaffold 长度
- 6 具有丰富的信息分析经验，可提供高质量组装结果

 产品参数

建库策略	测序深度	组装软件	承诺指标
10X Genomics 长片段文库	基因组 ≤ 1.6G, 测序 120G 基因组 > 1.6G, 测序 70X	Supernova 2.0	鸟类和哺乳动物: Contig N50 ≥ 40Kb Scaffold N50 ≥ 4Mb

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。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 2.1 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 2.2 博奥晶典 10X Genomics 基因组组装案例 (部分)

物种	Size (Gb)	Contig N50 (Kb)	Scaffold N50 (Mb)	组装软件
水稻	0.38	56	1.45	Supernova 1.2
某果树 1	0.78	45	1.15	Supernova 1.2
某果树 2	0.3	57.23	2.27	Supernova 2.0
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

 产品应用

由于 10X Genomics *de novo* 具有组装质量高、成本低、周期短等优势,因此其应用于物种泛基因组构建有突出的优势。

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,挖掘其中的特有基因和共有基因信息,及其变异信息(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。

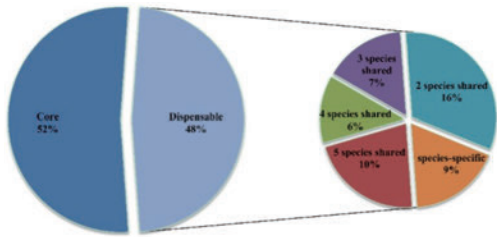


图 2.2 不同柑橘中基因组中特异基因和共有基因占比^[1]

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。

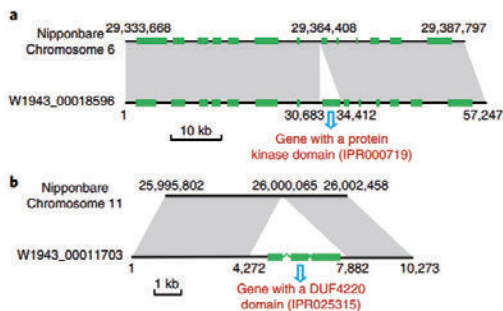


图 2.4 基于序列组装的水稻 SV 检测^[3]

2018 年发表在 Nature Genetics 上的 66 个水稻泛基因组,基于组装的基因组序列充分挖掘了水稻属的变异位点(尤其是大片段缺失、插入等 SV 信息)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。

案例解析

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 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相,Phase block N50: 2.7~9.3Mb。

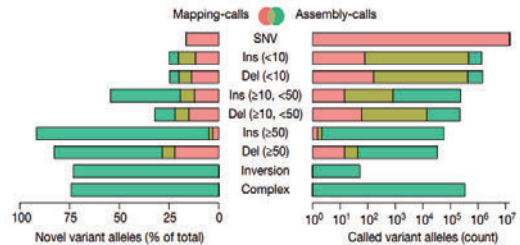


图 2.3 基于 mapping 和基于 *de novo* 组装的变异检测效果对比^[2]

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。

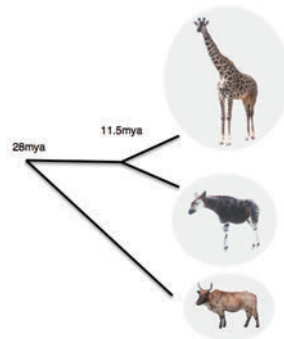


图 2.5 长颈鹿和霍加狓系统进化及分化时间分析^[4]

2016 年发表在 Nature Communications 上的长颈鹿基因组研究中,通过构建长颈鹿及其近缘种霍加狓基因组,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。

表 2.3 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 10x library	56	115	114.6	8.0	18.7	2.1
B	NA19240	Yoruban	F	One 10x library	56	125	118.8	9.3	16.4	2.3
C	HG00733	Puerto Rican	F	One 10x library	56	106	123.6	3.4	17.8	2.0
D	HG00512	Chinese	M	One 10x library	56	102	113.2	2.7	15.4	2.2
E	NA24385	Ashkenazi	M	One 10x library	56	120	106.4	4.2	15.1	2.6
F	HGP	European	M	One 10x library	56	139	120.2	4.5	18.6	2.5
G	NA12878	European	F	One 10x 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 2.4)。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 2.6)。

表 2.4 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	DISCOVER
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

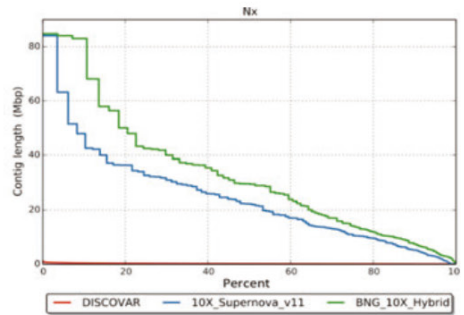


图 2.6 僧海豹基因组三种组装策略序列长度统计

● 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,而利用 10X Genomics 构建的辣椒基因组 Contig N50 为 123Kb, Scaffold N50 为 3.69Mb,图 2.7 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,导致辣椒辣味的丧失(图 2.8)。

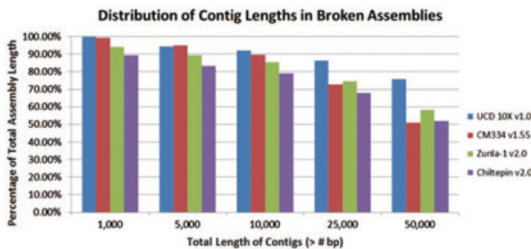


图 2.7 不同版本辣椒基因组 Contig 序列长度分布

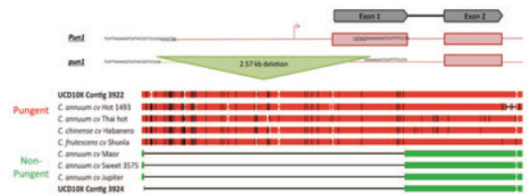


图 2.8 辣椒 *pun1* 基因序列比对

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.Marett 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):757-767.
- 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.Hulsekemp 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]. Hortic Res, 2017, 5(1):4.

3 全基因组重测序 (WGS)

普通重测序 (WGS)

全基因组重测序即对基因组序列已知物种的样品进行测序，并在个体或群体水平进行全基因组的差异性分析，通过序列比对，可全面挖掘基因序列的差异和结构变异，包括 SNP、InDel、SV、CNV 等，进一步能够实现遗传进化分析及重要性状候选基因的预测。目前该技术已广泛应用于变异检测、功能基因挖掘、遗传图谱构建、全基因组关联分析、群体进化等研究，具有重大的科研和产业价值，成为动植物基因挖掘和研究群体进化的最为迅速、有效的方法之一。

● 重测序技术路线

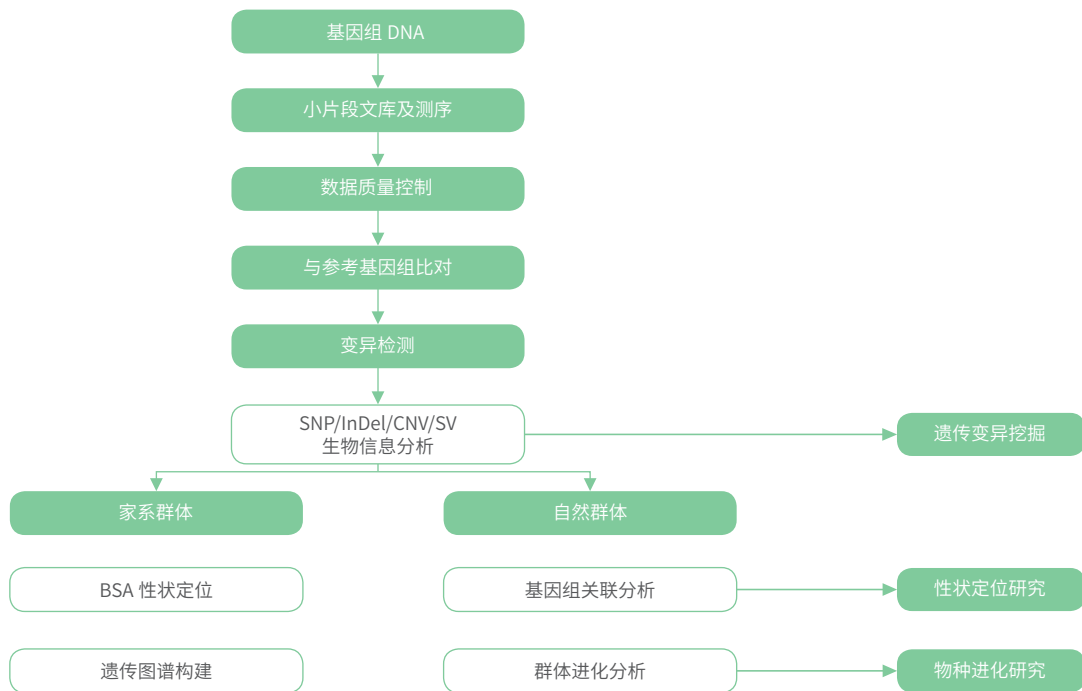


图 3.1 项目流程图

 产品分析内容

产品分类		分析内容
基础信息分析		原始数据质控 参考基因组比对 SNV 检测、注释及统计 InDel 检测、注释及统计 Circos 全局展示
高级分析	变异检测	SV 检测、注释及统计 CNV 检测、注释及统计 Circos 全局展示 变异信息差异分析 有意义高频突变基因检测及功能分析
	BSA 性状定位	SNP 频率差异分析 目标性状关联区域定位 关联区域功能注释
	遗传图谱构建	遗传标记筛选 高分辨率遗传图谱构建 结合表型的 QTL 定位分析
	基因组关联分析	基因型填补 群体分层分析（群体主成分分析、构建系统进化树） 连锁不平衡分析 全基因组关联分析（单位点关联分析、多重检验校正） 显著位点注释及其附近单体型构建
	群体遗传分析	群体遗传多样性分析（群体主成分分析、群体进化树分析、群体结构分析） 连锁不平衡分析 选择性清除分析 选择区域注释及富集分析 种群动态分析

 产品优势

平台优势	博奥晶典拥有Novaseq测序平台，重测序性价比更高
分辨率高	技术简单、稳定性好，能够达到单碱基分辨率
发现新变异	与芯片分型相比较，测序可以检测到新的变异
覆盖率高	全基因组范围扫描变异，对低频位点也有很好效果
经验丰富	具有18年的基于SNP的生物信息分析经验

技术参数

建库策略	参考基因组	分析类型	测序深度
350bp 文库	该物种已有参考基因组	变异检测	SNP、InDel $\geq 10X$, SV $\geq 20X$, CNV $\geq 30X$
		BSA 性状定位	$\geq 10X$ / 亲本个体; $\geq 20X$ / 每个子代池
		遗传图谱	10~30X/ 亲本个体; 4~5X/ 子代个体
		全基因组关联分析	$\geq 10X$ / 个体
		群体进化	$\geq 10X$ / 个体

产品应用

- 1 物种起源与进化历史研究
- 2 目标性状功能基因挖掘及优良性状基因定位
- 3 分子标记开发及辅助品质改良和品种选育
- 4 性状遗传机制解析
- 5 辅助种质创新与鉴定

案例解析

1. 变异检测

基于全基因组重测序可全面挖掘 SNP、InDel、CNV、SV 等遗传变异信息并进行分析, 构建该物种的遗传变异信息库, 对于促进种质资源保存、交流、利用以及基因库管理等具有重要的学术和实用意义。

● 四种栽培葡萄的序列多样性和结构变异检测^[1]

对 4 种萨丁地区的栽培葡萄进行测序, 并在公共数据库下载了 3 种葡萄的序列, 与葡萄参考基因组 PN40024 进行比对, 结果表明每个样品中 SNP 的密度从 5508 个 /M 到 8522.1 个 /M; InDel 的密度从 213.4 个 /M 到 728.4 个 /M。

经鉴定, 6526 个遗传区域存在拷贝数的不同, 其中 81 个区域是四种栽培葡萄共有的, 特有区域达到 4953 个区域。另外鉴定了栽培葡萄中大的结构缺失, 大约 1.1M 的大结构缺失是 4 种栽培葡萄共有的, 共涉及 44 个蛋白编码基因。

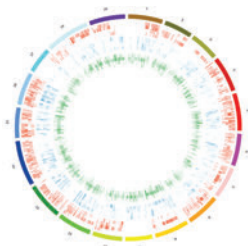


图 3.2 拷贝数变异 (CNV) 在葡萄 19 条染色体上的分布

表 3.1 不同葡萄中大片段缺失变异统计

Cultivar	#Large deletions	Large deletions (bp)	Private (bp)	%Private
Bovale	412	26858904	7513499	28.0
Cannoa	1990	156925773	85551167	54.5
Carignano	529	29768056	6171910	20.7
Vermantino	50	3779734	640538	16.9
Common	14	1195168		

对鉴定到的变异位点进行功能分析, 其功能主要集中在防御应答和细胞凋亡进程中。CNV 获得区域的基因主要富集在类黄酮合成和其他辅助代谢过程, 而 CNV 缺失区域的基因主要与抗性相关, 此外, 大结构缺失区域的基因也主要与生物 / 非生物胁迫响应相关。

2. BSA 性状定位

BSA (Bulked Segregant Analysis, BSA) 为集群分离分析法或混合分组分析法, 利用该技术结合全基因组重测序可对物种特异性状相关基因进行快速定位。与传统的育种方法时间长、精力耗费大、费用高的特点相比, BSA 性状定位具有时间短、成本低、方便快速的特点。

● BSA 定位分析狗尾草稀疏花序基因^[4]

狗尾草和玉米有广泛的基因共线性, 相似的结构和相同的习性, 都显示了它作为可转移模型的潜力。本研究用亚硝酸甲基脲 (NMU) 诱导狗尾草突变体, 并从中筛选出花序稀疏 (*spp*) 的品系, 利用 BSA 法构建突变型和野生型杂交后代的混池, 对来自 3 个杂交系的 5 个 F₂ DNA 池进行测序, 发现在 5 号染色体末端大约 1Mb 的区域可能与 *spp* 有关。

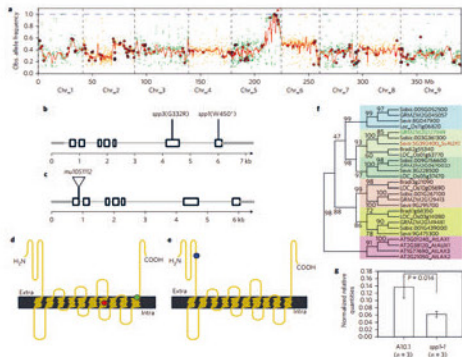


图 3.3 BSA 作图和 *SvAUX1* 和 *ZmAUX1* 等位基因的变异

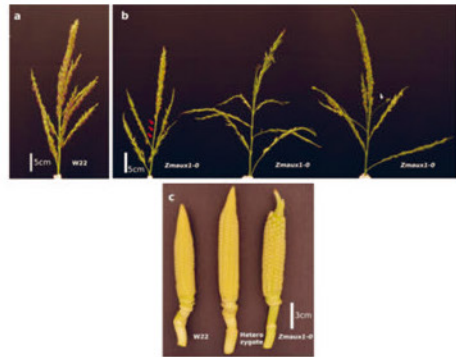


图 3.4 *Zmaux1-0* 的表型变异

研究团队通过同源搜索和分析在玉米中找到了同源基因 *ZmAUX1*, 为研究其功能, 获得了该基因突变体。插入的突变序列会打断第一个内含子的剪切信号, 导致与狗尾草中 *spp* 相似的功能缺失。玉米 *ZmAUX1* 突变体表现出混乱的花序分枝, 研究组将突变体与野生型 W22 回交, 纯合体相比较野生型和杂合体, 分枝数、穗数、初级分枝长等均显著减少, 且每行的穗更少。

3. 遗传图谱

利用重测序技术, 可快速便捷地对群体多态性进行鉴定并完成遗传图谱构建, 其常用于数量性状位点 (QTL) 定位; 获得与目的性状连锁的基因组片段, 开展基因的图位克隆或辅助基因组组装。另外, 通过图谱比较, 还可以提供基因组进化信息, 进而探索物种的遗传演化进程。

● 水稻遗传图谱构建及产量相关 QTL 定位^[3]

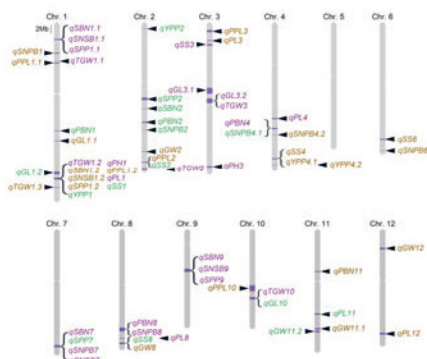


图 3.5 产量相关 QTL 位点在染色体上的分布

以栽培水稻野生近缘种 W2014 和优良籼稻品种 93-11 以及由两者繁衍的 131 个渐渗系为材料。双亲筛选到 181,957 个高质量 SNP, 平均每 10Kb 有 4.9 个 SNP, 92,127 个 SNP 位于基因区。

构建了包含 1,070 个 bin 标记的高密度遗传连锁图谱。调查了 13 个产量相关性状, 所有环境条件下, YPP 与 PPL、PL、SBN、SNSB 和 SPP 均存在极显著正相关, PPL 和 PL、PPL 和 TGW 以及 PL 和 GW 之间均存在负相关。为了评估图谱的质量及精确度, 调查了落粒性 (SH) 和谷壳颜色 (HC) 2 个高遗传率性状并进行 QTL 分析。结果表明本研究所构建的 Bin map 具有较高的质量及精确度。

4. 全基因组关联分析

基于全基因组重测序的 GWAS 是研究动植物复杂性状的重要手段。对遗传多样性丰富的自然群体的每个个体进行全基因组重测序，结合目标性状的表型数据，基于一定的统计方法进行全基因组关联分析，可以快速获得影响目标性状表型变异的染色体区段或基因位点。

● GWAS 解析大豆复杂农艺性状间的调控网络^[4]

共对 809 份大豆进行了重测序 (8.3×) 分析，深入解析了大豆 84 个农艺性状间的遗传调控网络，共鉴定出 245 个显著关联位点，发现其中 95 个关联位点和其它位点存在上位性效应。

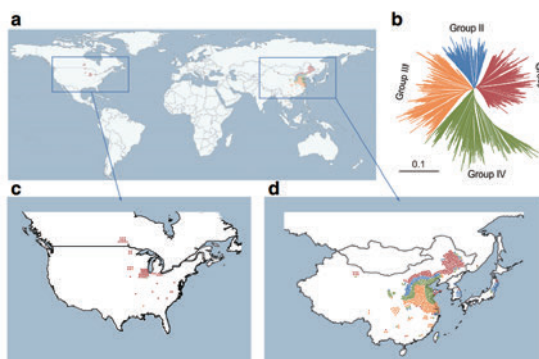


图 3.6 809 份大豆取样地理位置分布及大豆群体结构分析

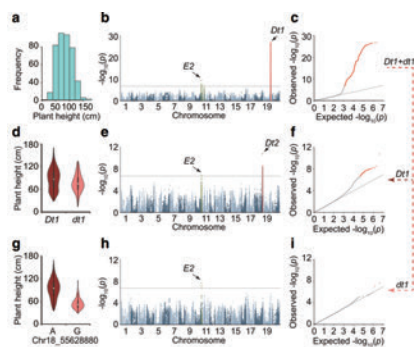


图 3.7 大豆植株高度性状的 GWAS 分析

例如，对于油含量相关性状，共鉴定到 24 个脂肪酸代谢相关和 21 个脂代谢相关的基因。深入分析发现，这些基因是通过加性效应共同调控多个大豆油脂性状的形成。

这些关联位点揭示了不同性状间相互耦合的遗传基础。根据连锁不平衡分析，发现 115 个关联位点可相互连锁，并与所观测的 51 个性状联系起来，形成复杂的多性状多位点调控网络，该遗传调控网络很好地解释了不同性状间的耦合关系。研究还发现其中 23 个关联位点，包括 16 个新鉴定的位点，对不同性状的形成起到关键调控作用。

5. 群体进化

由于生物生存环境的不同，同一物种会由于自然选择、遗传漂变、人工驯化等因素形成不同的亚种或亚群，群体进化研究就是用来追溯和揭示这个进化过程的。利用全基因组重测序可获得同一物种的不同亚群，或不同地理分布的品种之间的大量 SNP、InDel、SV 和 CNV 等变异信息，基于此，进一步进行群体的遗传结构、群体主成分、连锁不平衡和选择性消除等群体遗传学分析，从而在基因组水平揭示群体遗传结构、基因交流情况、群体进化动态、物种形成机制、环境适应机制等生物学问题。

● 虎鲸群体进化研究^[5]

本研究对 5 个生态型群体共 48 只虎鲸进行重测序。得到高质量的测序数据，进行群体遗传多样性分析，选择消除分析，功能富集分析等。

群体遗传多样性分析表明，虎鲸包括 10 个北太平洋海域的 transient 生态型和 resident 生态型；7 个来自南极海域的以大型哺乳动物为食的 B1 生态型，11 个以企鹅为食的 B2 生态型，10 个以鱼为食的 C 生态型。

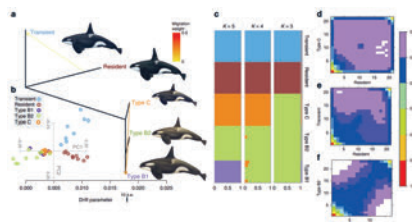


图 3.8 虎鲸生态类型进化关系展示

使用 PSMC 方法，分析有效种群大小 (N_e) 和进化率之间的关系，以重建虎鲸的种群历史。结果表明三种生态型都表现出先下降至瓶颈期后扩增的模式 (transient 型、C 型、resident 型)，正如在冰川周期中应对全球环境压力所预料的那样，瓶颈期是顺序发生的。

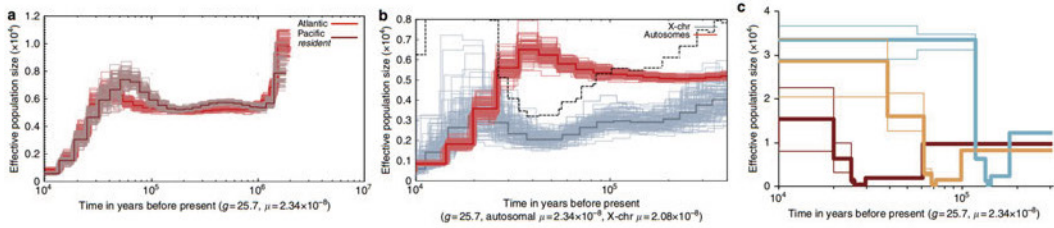


图 3.9 虎鲸生态型群体历史重建

比较北太平洋温带地区和南极地区的虎鲸生态型，脂肪组织发育相关基因发生了显著富集，这证明富含脂肪的饮食在快速适应寒冷气候中起到重要作用。例如 *FAM83H* 基因和 *CES2* 基因分别参与皮肤发育和脂肪代谢中。

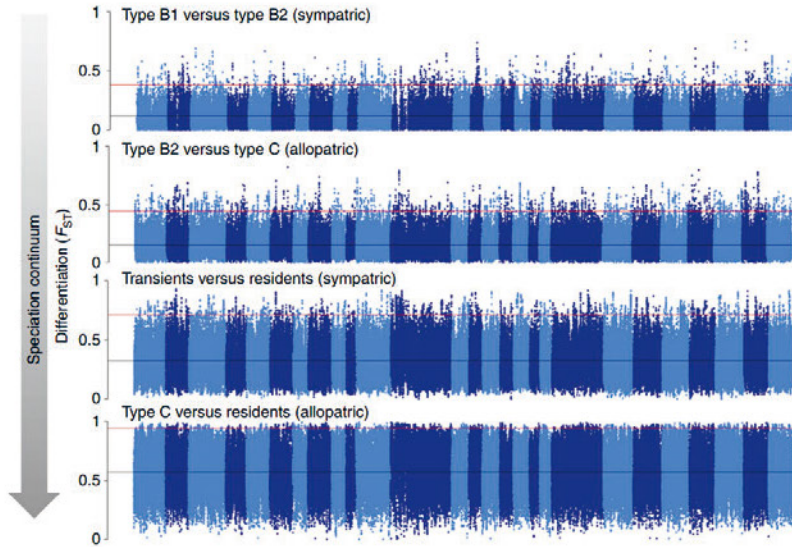


图 3.10 虎鲸全基因组分化



参考文献

1. Mercenaro L, Nieddu G, Porceddu A, et al. Sequence Polymorphisms and Structural Variations among Four Grapevine (*Vitis vinifera* L.) Cultivars Representing Sardinian Agriculture. [J]. *Frontiers in Plant Science*, 2017, 8.
2. Huang P, Jiang H, Zhu C, et al. Sparse panicle1 is required for inflorescence development in *Setaria viridis* and maize. [J]. *Nature Plants*, 2017, 3(5):17054.
3. Ma X, Fu Y, Zhao X, et al. Genomic structure analysis of a set of *Oryza nivara* introgression lines and identification of yield-associated QTLs using whole-genome resequencing. [J]. *Scientific reports*, 2016, 6: 27425.
4. Fang C, Ma Y, Wu S, et al. Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. [J]. *Genome biology*, 2017, 18(1): 161.
5. Foote A D, Vijay N, Ávila-Arcos M C, et al. Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. [J]. *Nature communications*, 2016, 7: 11693.

4 10X Genomics 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 SNP、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。

技术流程

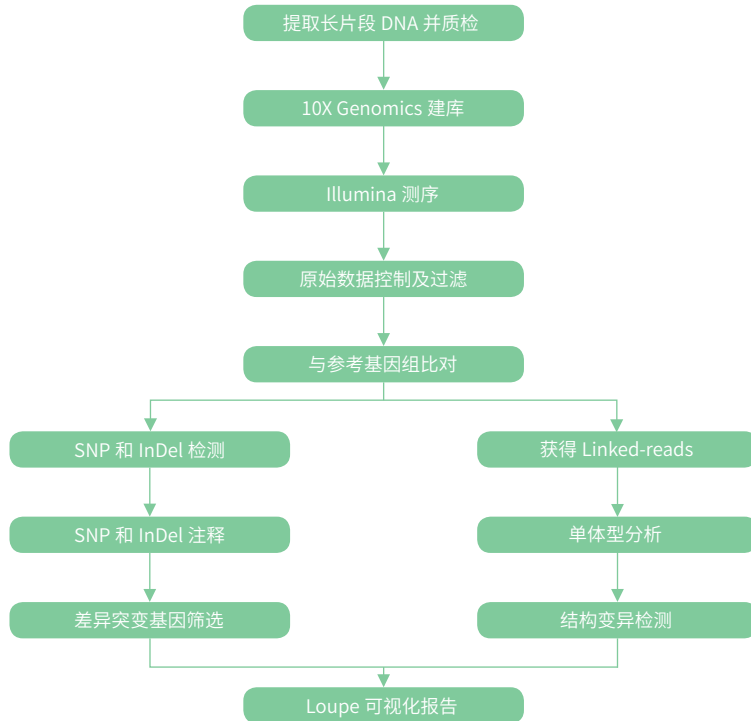


图 4.1 项目流程图

产品优势

- 1 对于 SV 检测的准确率提高 30% 以上，可实现染色体 Phasing，获得精确的单体型信息
- 2 技术新颖，结果可靠性高，所有结果可视化，发表文章较容易
- 3 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

产品应用

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；对动植物分子辅助育种具有重要意义。

技术参数

建库策略	参考基因组	测序深度
10X Genomics长片段文库	该物种已有参考基因组	≥40X

案例解析

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，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 4.2 统计结果可视化



图 4.3 结构变异结果可视化说明

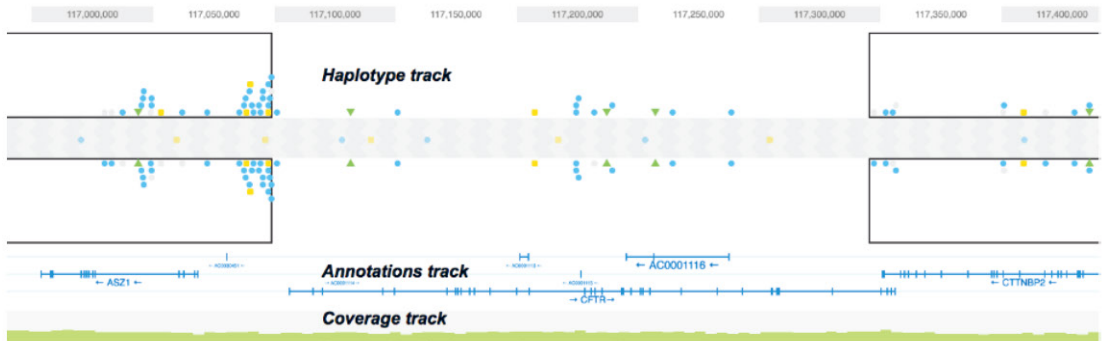


图 4.4 Haplotype 可视化说明

5 农业 SNP 芯片分型

SNP 作为第三代分子标记，在分子育种相关过程中的许多方面具有重要的作用。对已有目录芯片的物种可直接进行 SNP 检测。但如果研究者研究的物种还没有目录化 SNP 芯片，或者目录芯片 SNP 位点不能有效代表研究者关注的特有品种或性状的信息时，需要先进行 SNP 芯片的定制。通过 SNP 分型，一方面可利用 QTL 定位、GWAS、选择信号分析等策略找到与性状相关的 SNP 标记位点，用于分子标记辅助育种，另一方面在全基因组水平采用基因组选择方法进行分子育种。

技术流程

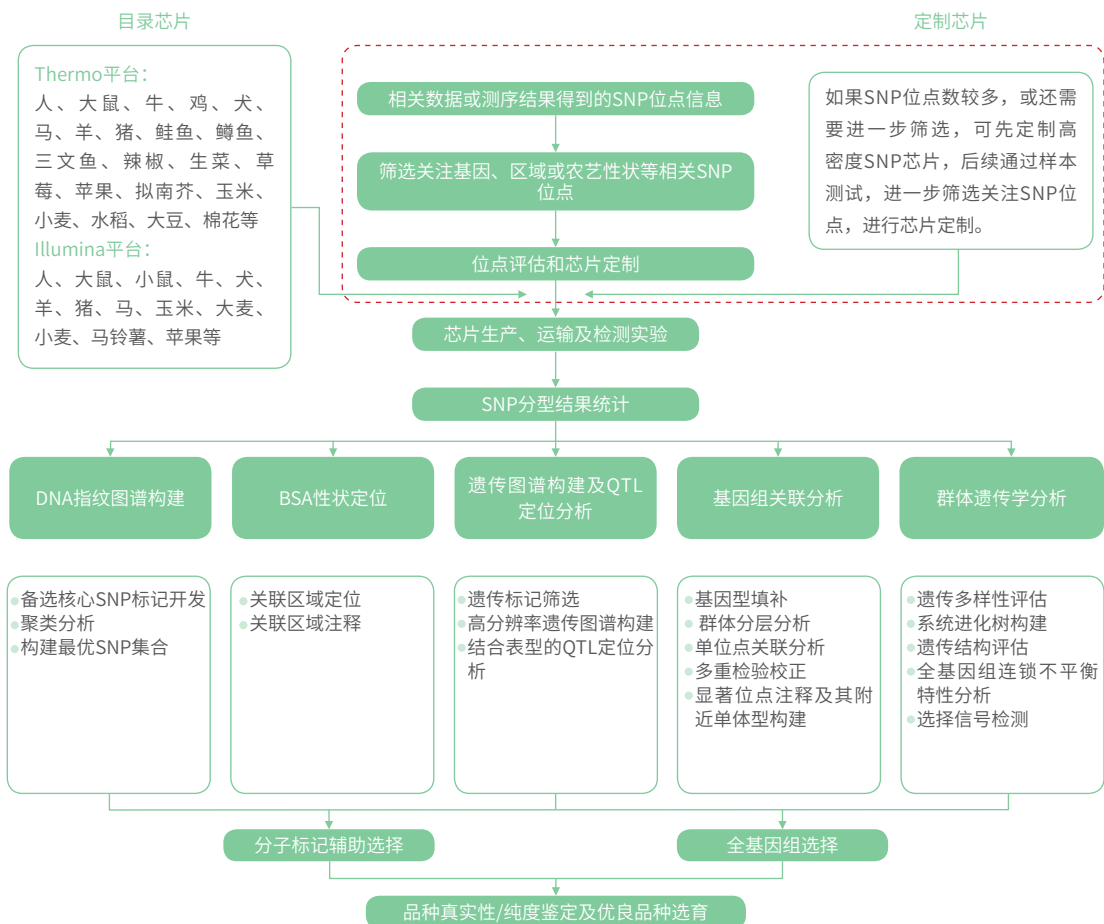


图5.1 项目流程图

产品优势

数据质量高	性能稳定，检测准确，性价比高，适合大量样品研究
定制灵活	按需灵活定制，支持高低密度的SNP定制，一张芯片可定制多物种
应用广泛	可广泛用于各物种的基因组研究和分子育种
经验丰富	认证芯片实验体系，18年累计芯片经验，能够提供稳定可靠的结果

表5.1 博奥晶典已完成的SNP定制芯片项目（部分）

单位	物种	芯片密度
中国农业科学院	小麦	660K/55K/50K
中国农业科学院	玉米	55K
中国科学院遗传所	大豆	660K
南京农业大学	大豆	355K
北京农林科学院	桃	200K
郑州烟草研究所	烟草	300K
江西农业大学	猪	1.2M

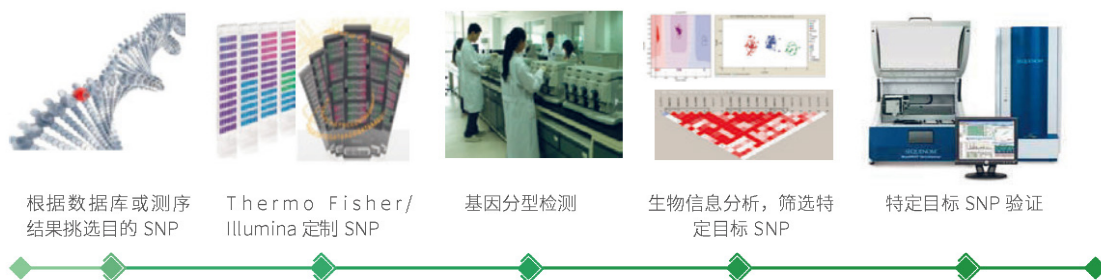


图 5.2 SNP 芯片定制—检测—分析—验证一体化流程

产品应用

- 1 分子标记辅助选择及基因组选择
- 2 GWAS、遗传图谱构建及QTL定位
- 3 群体基因组学、选择性清除及遗传多样性研究
- 4 品种资源鉴定和物种指纹图谱构建

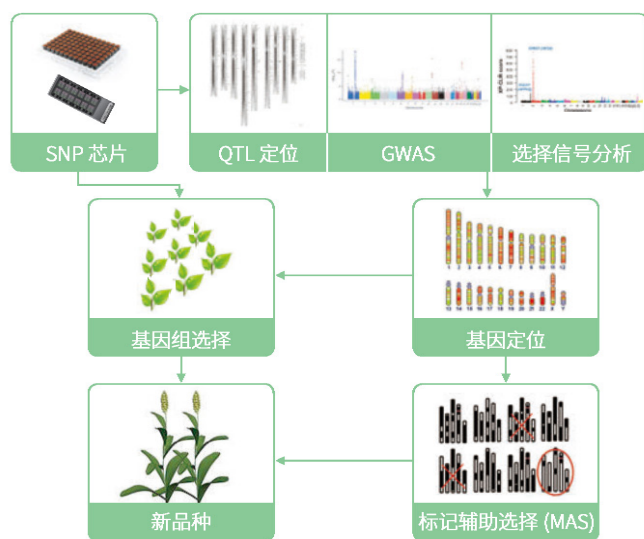


图 5.3 SNP 定制芯片分子育种流程

案例解析

1. 小麦 660K SNP 定制芯片

贾继增老师课题组综合运用从头测序、重测序等策略对 111 个普通小麦、2 个四倍体小麦以及 3 个 A 基因组与 2 个 D 基因组共计 118 个小麦及其近缘属种进行高通量测序，获得 6Tb 数据，用中国春 survey sequence 为参考序列采用严格标准经过生物信息学分析获得 5,138 万个 SNP，与北京博奥晶典生物技术有限公司合作，经过生物信息学筛选和芯片预实验，最后形成小麦 660K SNP 芯片。

a. 小麦 660K SNP 芯片优势

多态性高

该芯片高多态性且高分辨率标记占 58.9%，在 66 份普通小麦的自然群体群体中，PolyHighResolution、NoNinorHom 和 OTV 三类 SNP 占 65.35%

通量高

可检测 >60 万个 SNP 位点，六倍体普通小麦中 Call rate >95%，是目前有效检测位点最多的一款小麦 SNP 芯片

准确性高

标记经过初筛，淘汰了分辨不清的标记

注释清晰

绝大多数 SNP 探针物理位置和遗传位置信息齐全，为进一步深入分析研究奠定了基础

性价比高

在现有小麦 SNP 标记检测中最为便宜

b. 小麦黑点病 GWAS 分析^[1]

本研究利用小麦 660K 和 90K 高分辨率 SNP 芯片，对 166 份黄淮海区重要推广品种和高代品系进行基因型检测，并开展群体结构分析和全基因组关联分析 (GWAS)。

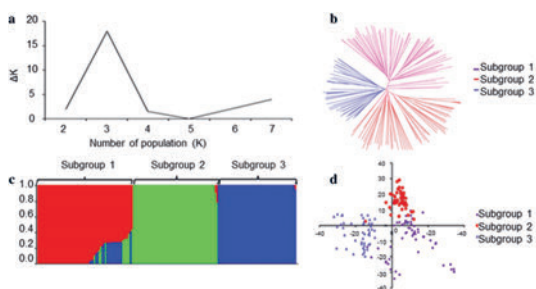


图 5.4 166 个小麦群体结构分析

结合 5 个环境下的田间表型数据，采用经典的 Tassel MLM 模型和较为严谨的 FarmCPU 共同进行 GWAS 分析发掘黑胚病显著关联位点，选择在两种模型中均存在的位点作为最终结果。共计检测到 25 个位点，位于 2A、2B、3A、3B (2)、3D、4B (2)、5A (3)、5B (3)、6A、6B、6D、7A (5)、7B 和 7D (2) 染色体上，分别解释 7.9% 到 18.0% 的表型变异。在发现的位点中，13 个位点在多个环境下存在，均有稳定的遗传效应，7 个位点与前人研究结果一致，18 个位点为全新位点。

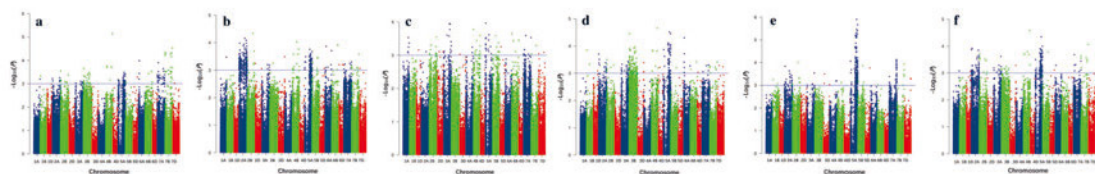


图 5.6 利用 Tassel v5.0 软件的混合线性模型 (MLM) 对 166 个小麦的黑点病抗性进行全基因组关联分析 (GWAS)

该研究不仅为小麦黑胚病抗性育种提供了 SNP 分子标记和抗性等位基因，而且为黑胚病抗性反应的遗传机制提供了新的线索。

2. 玉米 55K SNP 定制芯片

中国农业科学院作物科学研究所玉米分子育种技术和应用创新团队与博奥晶典生物技术有限公司、南通新禾生物技术有限公司合作，运用 Thermo Fisher Axiom 平台联合开发了玉米 55K SNP 芯片。该芯片具有高密度、低成本、全覆盖、多用途的特点，可以用于玉米遗传作图、种质资源鉴定、基因定位、分子标记辅助育种等各项研究中。

a. 玉米 55K SNP 芯片优势

- | | | | |
|---|-----------------------------------|---|----------------------|
| 1 | 基因组覆盖率高，检测位点分布均匀 | 2 | 更低的位点缺失率和杂合率 |
| 3 | 能够清晰地划分中国的玉米杂种优势群 | 4 | 在热带和温带玉米群体间不具有明显的偏向性 |
| 5 | 包含了影响玉米重要农艺性状的候选基因，加入了已公开转基因事件标签位 | | |
| 6 | 表现稳定、分型结果高度可重复性 | | |

b. 玉米开花相关性状 GWAS 分析^[4]

研究群体由 724 个杂交株系组成，来源于 28 个温带玉米自交系和 23 个热带玉米自交系。基于该群体，利用 55K SNP 芯片产生的基因型数据和开花性状 (播种 - 抽雄的天数、播种至吐丝的天数、播种至开花的天数、开花 - 吐丝间隔期) 进行 GWAS 分析。通过分析，鉴定出与开花性状有关的 5 个基因组区域，并用于候选基因和数量性状核苷酸的挖掘。此外，对 51 个玉米自交系进行深度重测序，获得了 790 万个 SNP 标记，为更高分辨率的 GWAS 打下了基础。

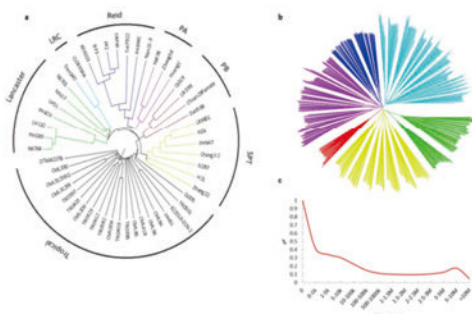


图 5.7 51 个亲本自交系和 724 个多系杂交群体 NJ 树图及整个基因组的 LD 衰减图

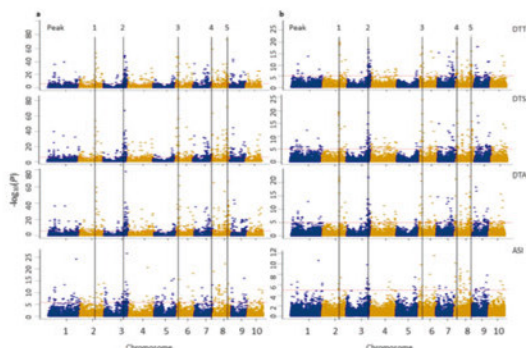


图 5.8 对 MHP 群体的 4 个开花性状进行 GWAS 分析，通过 GWAS 分析，在基因组上得到 5 个显著相关的区域。

2012 年至今，博奥晶典支持客户发表农林牧渔及微生物领域的英文文章 420 余篇。

表5.2 博奥晶典支持客户客户发表文章列表（部分）

研究内容	文章名称	期刊
连锁图谱构建	Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea.	BMC Genomics 11 :468
QTL定位	Genetic architecture of aluminum tolerance in rice (<i>Oryza sativa</i>) determined through genome-wide association analysis and QTL mapping.	PLoS Genet 7(8): e1002221.
复杂性状关联分析	Genome-wide association studies of 14 agronomic traits in rice landraces.	Nature Genet 42(11): 961-7.
关联分析	Highly effective SNP-based association mapping and management of recessive defects in livestock.	Nature Genet 40, 449-54.
遗传进化	The genome sequence of taurine cattle a window to ruminant biology and evolution.	Science 324(5926): 522-8.
遗传进化	Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds.	Science 324(528): 528- 32.
品种资源鉴定	Cereal variety identification using MALDI-TOF mass spectrometry.	J Cereal Sci 52: 356-61.
遗传变异及驯化相关基因研究	Development and application of a novel genome-wide SNP array reveals domestication history in soybean.	Scientific reports 6 (2016).
QTL定位	Genome-wide linkage mapping of QTL for black point reaction in bread wheat (<i>Triticum aestivum</i> L.).	Theoretical and Applied Genetics (2016): 1-12.
关联分析	An R2R3-type MYB transcription factor, GmMYB29, regulates isoflavone biosynthesis in soybean.	Plos Genetics , 2017 , 13 (5) :e1006770.
关联分析	Development of a multiple-hybrid population for genome-wide association studies: theoretical consideration and genetic mapping of flowering traits in maize.	Scientific Reports 2017,7: 40239.

参考文献

- 1.Liu J, He Z, Rasheed A, et al. Genome-wide association mapping of black point reaction in common wheat (*Triticum aestivum* L.).[J]. BMC plant biology, 2017, 17(1): 220.
- 2.Wang H, Xu C, Liu X, et al. Development of a multiple-hybrid population for genome-wide association studies: theoretical consideration and genetic mapping of flowering traits in maize.[J]. Scientific reports, 2017, 7: 40239.

6 Agena MassARRAY® SNP 服务平台

对于在全基因组水平发现的突变位点，通常需要进行独立大样本量验证工作。Agena MassARRAY® 质谱作为高通量检测平台，是 300 个以内 SNP 位点分型检测质量与效率均佳的首选检测方法。

技术流程

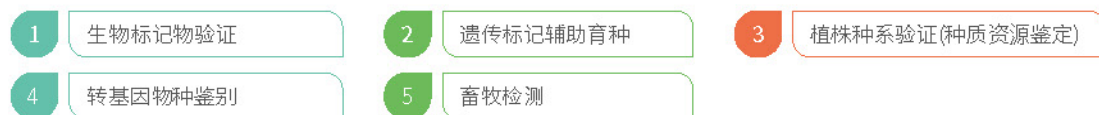


图6.1 项目流程图

产品优势

灵活	可以自由选择感兴趣的SNP位点和需要检测的样品
准确	直接检测待测物分子量，具有极高的准确度，可有效区分三等位基因
高通量	每反应孔可实现最多40重反应；每张芯片可完成384个样品的多重IPLEX分型实验
高性价比	无需荧光标记，降低每次基因型分析的成本；适合于SNP位点小于300，样本量大于500份的项目
便捷	可对95%以上已被证实的SNP进行实验设计；专业的引物设计和基因分型软件，自动化判读分型实验是否成功

产品应用



案例解析

牛 *CLECT7A* 基因遗传变异与副结核病 ELISA 血清阳性的关联性研究^[1]

牛的副结核病感染每年都会给对乳制品及牛肉行业带来巨大的经济损失。为了研究 SNP 与牛副结核病感染易感性之间的关联性，研究人员利用一代测序方法对 *CLECT7A* 基因启动子、编码区、5' UTR 及 3' UTR 区域的 SNP 进行鉴定，对已鉴定的 9 个 SNP 位点，利用 Agena MassARRAY® 质谱平台对感染组和非感染组牛进行检测，最终发现 *CLECT7A* 基因上的一个错义突变与牛副结核病感染易感性之间存在强关联。

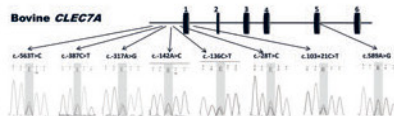


图6.2 *CLECT7A*基因SNP位点位置及分型

参考文献

1. Pant S D, Verschoor C P, Schenkel F S, et al. Bovine *CLECT7A* genetic variants and their association with seropositivity in Johne's disease ELISA. [J]. Gene, 2014, 537(2):302.

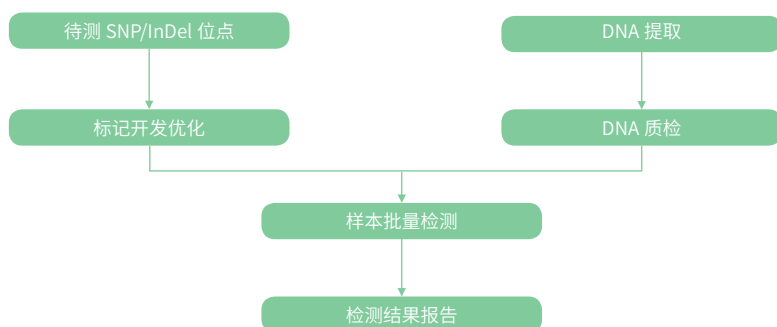
7 IMAP 检测服务

随着 GWAS 等研究工作的广泛开展，有大量与性状相关的位点被挖掘出来，选择什么方法对这些位点进行进一步检测分析，也成为研究者关注的内容。

博奥晶典可以根据用户的实际情况提供高性价比且灵活的检测方法（包括博奥晶典自主研发的微流控芯片检测系统——IMAP），实现准确、灵活、经济和高效的 SNP/InDel 检测，服务于广大用户。

服务内容包括从上游的标记开发到下游的批量样本的分型检测 / 验证，适用于 GWAS 后续验证、精细定位、标记辅助选择育种、转基因检测、种子真实性检测、种子纯度检测等。同时还可将验证后的位点开发成芯片产品。

技术流程



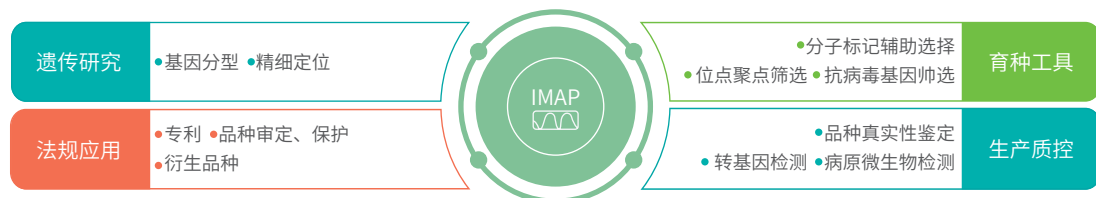
检测优势

成功率高	对位点优化设计有丰富经验，可提高检测的成功率
灵活性高	位点增减灵活，可独立检测，无需样本拼凑
性价比高	采用通用荧光探针，减少探针合成费用；微量反应体系，降低试剂成本
技术成熟	已完成多个物种的分型检测，如玉米、水稻、大豆、苹果、绵羊、奶牛、鸡、狗、人等

样本要求

基因组 DNA 的电泳条带清晰无明显弥散，无 RNA 污染，260/280 在 1.7~2.1 之间，总量 $\geq 1.0 \mu\text{g}$ 。

应用方向



转录组学

1 10X Genomics 单细胞转录组测序

Chromium™ 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 500~10000 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。

技术流程

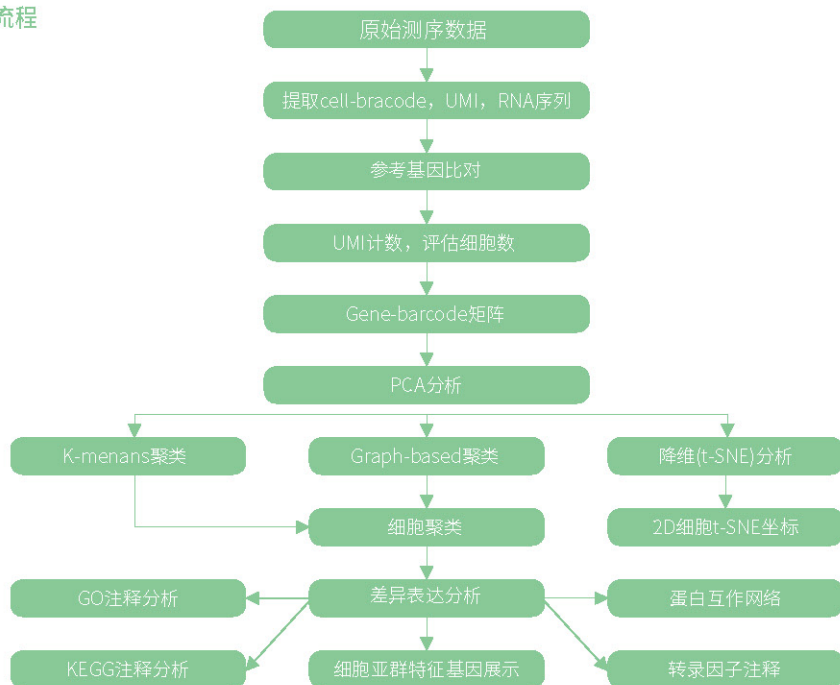


图1.1 项目流程图

产品优势

高通量	一次实验可以同时获取500-10,000个细胞的转录组和免疫组数据
丰富的项目经验	10+物种, 50+样品类型, 700+样品
低价格	开启单个细胞转录组测序低成本时代的黑马技术
首家引进	博奥晶典是国内首家引进10X Genomics平台的服务商, 具有丰富的平台经验
一站式服务	从样品处理到数据分析、结果可视化展示
首篇科研论文	博奥晶典支持发表10X Genomics单细胞研究国内首篇科研论文PNAS (IF=9.423)
首篇CNS级科研论文	博奥晶典联合北京大学合作在干细胞领域顶级期刊《Cell Stem Cell》(IF=23.290) 发表科研论文
上门服务	北上广三城实验室及全国范围的仪器上门服务, 全程保障您的样品安全和实验稳定性

技术参数

建库策略	参考基因组	建议数据量	承诺指标
10X Genomics单细胞文库	有参	110G或220G	过滤后Q30≥80%

产品应用

1

干细胞或者原代细胞的性质研究

2

植物对环境的适应性

3

特定部位的细胞对于不同环境的反应程度

4

植物发育调控机制研究

案例解析

肺泡 I 型细胞在肺泡发育和肺泡再生过程中异质性的首次阐述^[1]

研究背景:

肺泡 I 型 (AT1) 细胞对肺的气体交换功能至关重要, 同时在损伤诱导的肺泡再生期间具有肺泡可塑性, 其在肺泡发育和再生过程中存在着异质性。将单细胞测序, 小鼠遗传学和肺泡小体器官培养等技术相结合, 第一次探究了肺泡上皮细胞中 AT1 细胞在转录组水平上发育和再生过程中的变化, 发现了一个新的 AT1 细胞遗传分子标记物, 即胰岛素样生长因子结合蛋白 2 (Igfbp2)。鉴定出生后的 AT1 细胞群实际上有两个 AT1 细胞亚型 ($Hopx^+Igfbp2^+$ 和 $Hopx^+Igfbp2^-$ AT1 细胞), 在肺泡再生过程中具有不同的细胞命运。

(该文章10X Genomics单细胞转录组测序实验部分在博奥晶典完成)

研究路线:

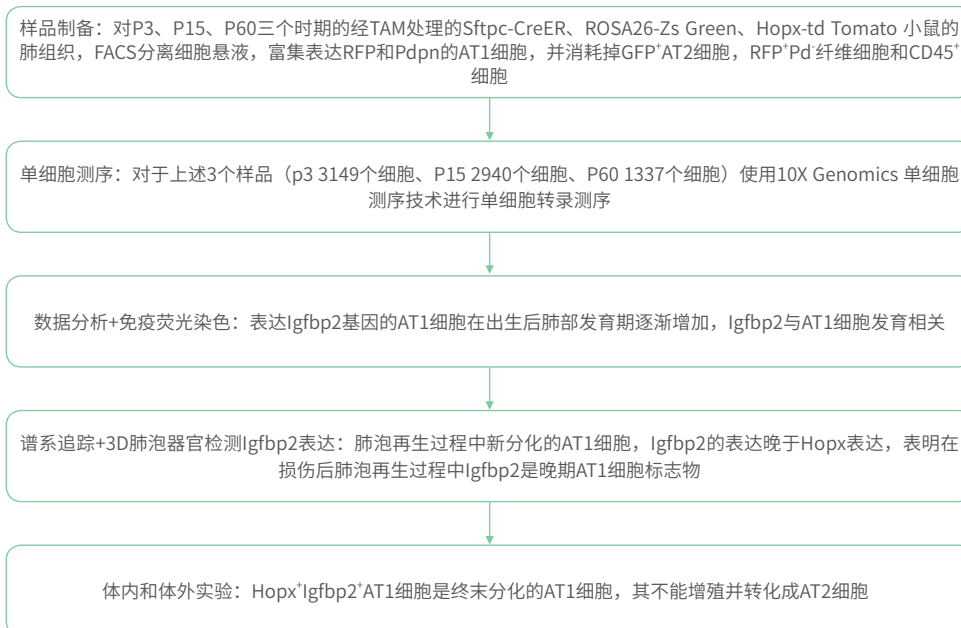


图1.2 文章研究思路总结

结果展示:

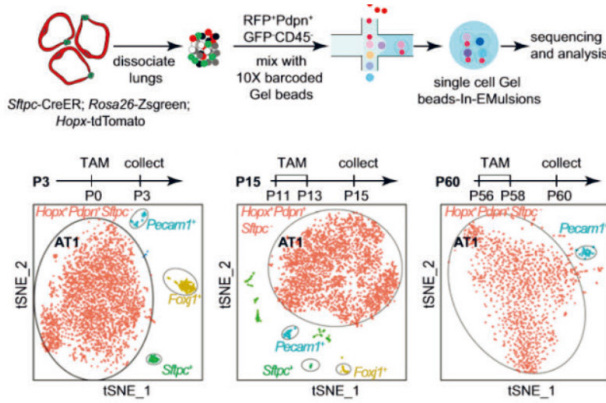


图1.3 实验流程和单细胞测序分群结果

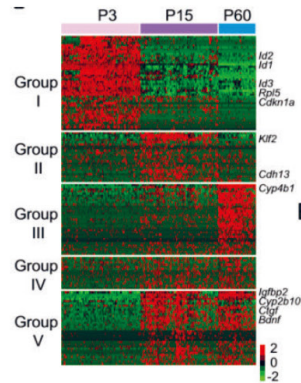


图1.4 从P3、P15和P60中随机挑选1000个AT1细胞中基因聚类图

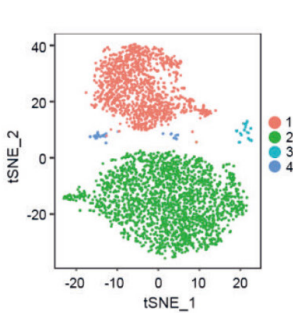


图1.5 scRNA-seq分析AT1和AT2细胞的分群图

Genes upregulated in AT1 cells as compared with AT2 cells

Gene Ontology	KEGG pathway enrichment
-Regulation of cell shape	-Focal adhesion
-Cell adhesion	-Tight junction
-Positive regulation of cell migration	-ECM-receptor interaction
-Extracellular matrix organization	-Hippo signaling pathway
-Actin cytoskeleton organization	-PI3K-Akt signaling pathway
-Endocytosis	-Rap1 signaling pathway
-Regulation of cell growth	-Leukocyte transendothelial migration
-Angiogenesis	-Regulation of actin cytoskeleton
-Regulation of Rho protein signal transduction	-Endocytosis
-Caveola assembly	-Amoebiasis

图1.6 GO、KEGG富集图

研究结论:

1. Igfbp2 是一种高度特异性的 AT1 细胞终末分化的遗传标记。
2. 成熟 AT1 细胞群包含两种不同类型的细胞：95% $Hopx^+Igfbp2^+$ 和 5% $Hopx^+Igfbp^-$ AT1 细胞。
3. $Hopx^+$ AT1 细胞终末分化为 $Hopx^+Igfbp2^+$ AT1 细胞亚型；但它们不能转分化成 AT2 细胞，并且在肺泡再生过程中不能增殖，占 95%。
4. $Hopx^+Igfbp2^+$ AT1 细胞，占成熟的 $Hopx^+$ AT1 细胞 5% 以下，可分化成 AT2 细胞，从而有助于损伤下、肺泡再生过程中修复组织的完整性。

至今，科研人员利用 10X Genomics 平台已发表 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1.1 10X Genomics 单细胞转录组测序部分已发表文章展示(红色表示博奥晶典支持文章)

发表时间	发表期刊	影响因子	研究方向
2018.07	Nature Medicine	30.357	肺肿瘤微环境中基质细胞的表型鉴定
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为早期胚胎干细胞过程
2018.05	Immunity	24.082	淋巴结基质细胞的单细胞RNA测序揭示其生态位相关的异质性
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	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
2018.3	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
2018.01	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	Science	34.661	CD8+T细胞规律性呈递干细胞表观遗传性质
2017.12	Genome Biology	11.313	单细胞分析人脑胶质瘤从而揭示了肿瘤微环境中巨噬细胞个体发育差异
2017.11	Nature communications	12.124	单细胞水平构建小鼠乳腺发育谱系
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+单细胞蛋白联合研究
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
2017.09	Cell	28.71	不同的肺间质细胞促进肺部上皮的自我更新和瘢痕组织形成机制
2017.08	Nature	38.138	成年小鼠Müller神经胶质细胞功能性神经元再生刺激研究
2017.05	Nature	40.648	Wnt 和 Rspo对哺乳动物肠道干细胞的分化和再生调控机制
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 Y, Zan T, Huang H, et al. Pulmonary alveolar type I cell population consists of two distinct subtypes that differ in cell fate[J]. Proc Natl Acad Sci U S A, 2018, 115(10):2407.

2 mRNA 测序

转录组可以随外部环境条件而有所转变，这是与基因组不同的。由于转录组包括了所有在细胞里的 mRNA 的转录，除却异常的 mRNA 降解现象（例如转录衰减）以外，转录组反映了在任何给定时间内活跃表达的基因。转录组测序的研究对象为特定细胞在某一功能状态下所转录出来的所有 mRNA 的总和。转录组研究是基因功能及结构研究的基础和出发点，通过新一代高通量测序技术，能够全面快速的获得某一物种特定组织或器官在某一状态下 mRNA 表达的信息。目前已经被广泛应用于转录本结构（基因边界鉴定、可变剪切研究等），转录本变异研究（基因融合、编码区 SNP 研究等），基因表达水平研究以及全新转录本发现。

技术流程

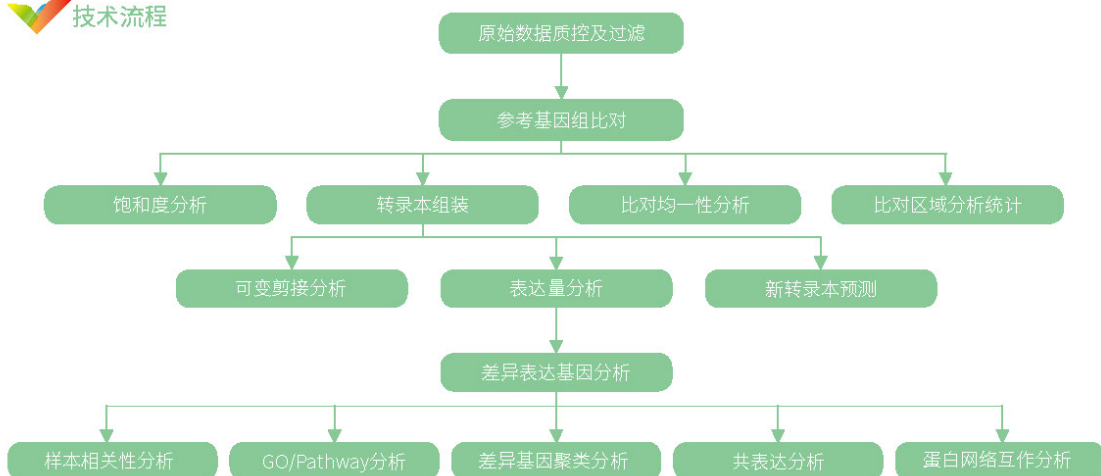


图2.1 项目流程图

产品优势

- 1 最高通量测序平台Novaseq，保证测序质量和周期
- 2 丰富的项目经验，建库0失败率，保证检测稳定性
- 3 全面的数据分析，为数据挖掘保驾护航

技术参数

类型	参考基因组	建议数据量
真核生物	有参/无参	6G
原核生物	有参	2G

产品应用

- 1 动植物的发育机制研究
- 2 品种间差异研究
- 3 物种间保守型研究
- 4 生物与非生物胁迫的应答机制研究
- 5 中药的药用价值研究

耐旱水稻渗入系及其亲本在干旱胁迫下的比较转录组测序^[1]

研究背景:

水稻 (*Oryza sativa* L.) 对于干旱胁迫比其他谷类更敏感, 栽培水稻基因库与其野生近缘种中存在大量的耐旱性基因型差异 (DT)。干旱敏感的原种 (DS), 和有选择性渗入 DT 片段的子一代回交是提高水稻干旱胁迫耐受性的有效方法。为了剖析水稻中 DT 的分子机制, 使用转录组测序来研究 DT 渗入系 H471、DT 供体 P28 和干旱敏感的原种回交亲本 HHZ 在干旱胁迫下的转录组差异。

本研究由博奥晶典提供mRNA测序支持及数据分析 (IF=4.041)

研究路线:

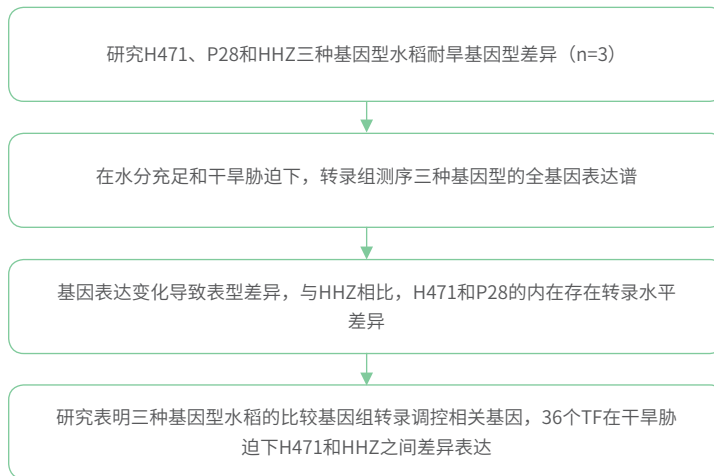


图2.2 文章研究思路总结

部分研究结果:

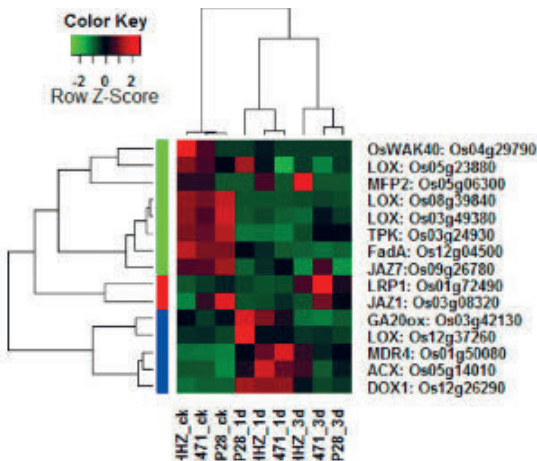


图2.3 JA相关基因和一些共同调节基因的分层聚类图

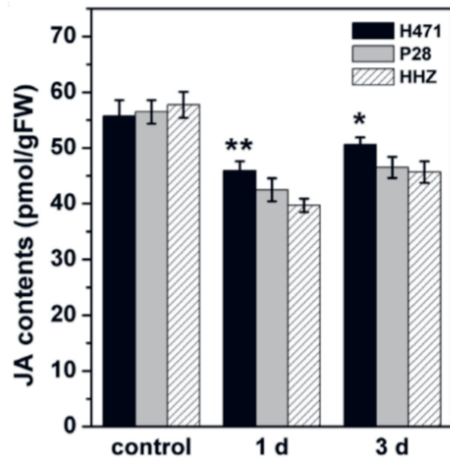


图2.4 H471、HHZ和P28在干旱条件下的JA含量

1.Huang L, Zhang F, Zhang F, et al. Comparative transcriptome sequencing of tolerant rice introgression line and its parents in response to drought stress.[J]. Bmc Genomics, 2014, 15(1):1026.

3 lncRNA 测序

长链非编码 RNA (long noncoding RNA, lncRNA) 是一类长度超过 200 nt 的不编码蛋白的 RNA, 它通过与 DNA、RNA 或蛋白质结合, 来调控基因表达。采用 Illumina 平台进行测序, 针对有参考基因组物种开展 lncRNA 分析, 同时提供针对测序数据中 mRNA 的分析。

技术流程

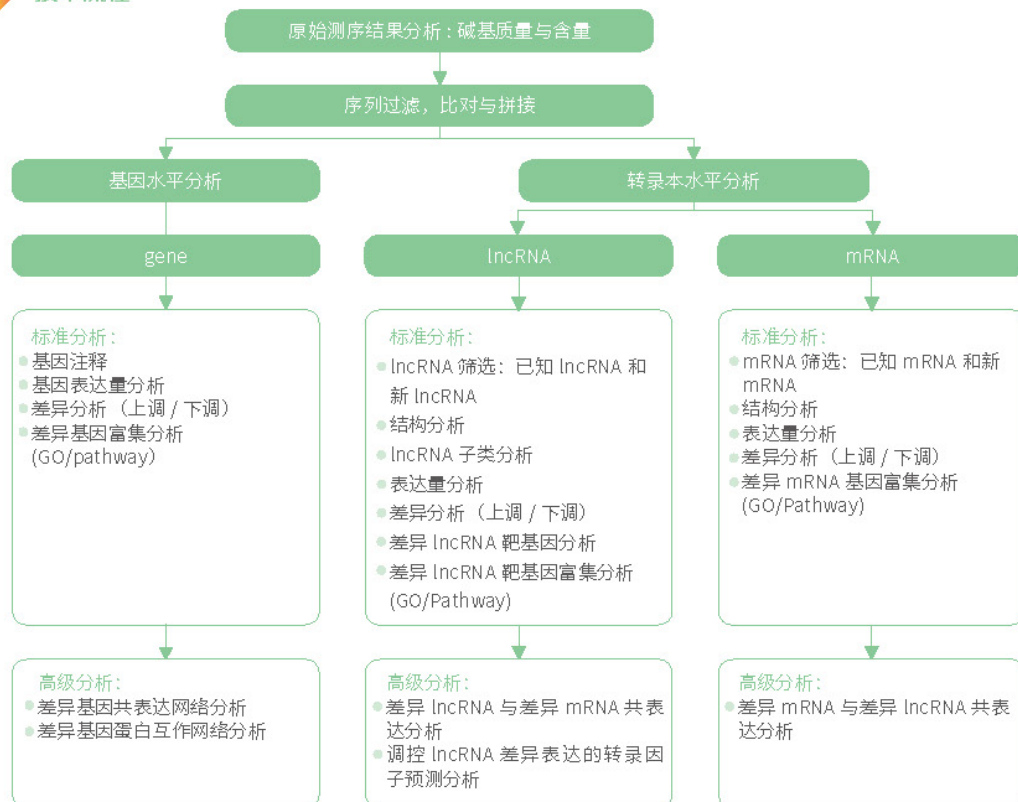


图3.1 项目流程图

产品优势

- 1 采用去除rRNA的策略, 同时检测带polyA尾和不带polyA尾的lncRNA; 链特异性建库则保留来源于两条链的转录本信息, 可以分析反义lncRNA
- 2 全面的数据分析, 同时分析lncRNA和mRNA, 包括已知lncRNA和新的lncRNA
- 3 深入的数据挖掘, 构建lncRNA和mRNA共表达网络, lncRNA转录因子调控网络, 基因蛋白互作网络等

技术参数

建库策略	参考基因组	建议数据量
去 rRNA 和链特异性建库	有参且具有完整的注释	≥10G

产品应用

- 1 动植物的发育机制研究
- 2 品种间差异研究
- 3 重要性状研究
- 4 生物与非生物胁迫的应答机制研究

案例解析

耐旱玉米品种 lncRNA 的鉴定及表达^[1]

研究背景:

最近很多报道证明 lncRNA 在植物胁迫响应机制中起重要作用。本研究使用深度 RNA-seq 对来自智利阿塔卡马沙漠的耐旱玉米品种 Luteño 的盐度和过量硼诱导的综合胁迫进行综合鉴定，研究共发现 48,345 种不同的 lncRNA，其中 28,012 个在玉米品种中是保守的，其他为潜在的 Luteño 特有 lncRNA 转录物。根据 B73 玉米参考基因组序列，发现大多数 Luteño 玉米 lncRNA 对应于基因间转录物。综上，本研究确定了 1,710 个可能对盐和硼暴露胁迫条件响应的转录本。并通过 RT-qPCR 进行验证，发现大量存在于玉米根和叶中对盐和硼胁迫响应的 lncRNA，本研究为第一篇对极端生长环境下的玉米品种进行 lncRNA 鉴定的报道。

研究路线:

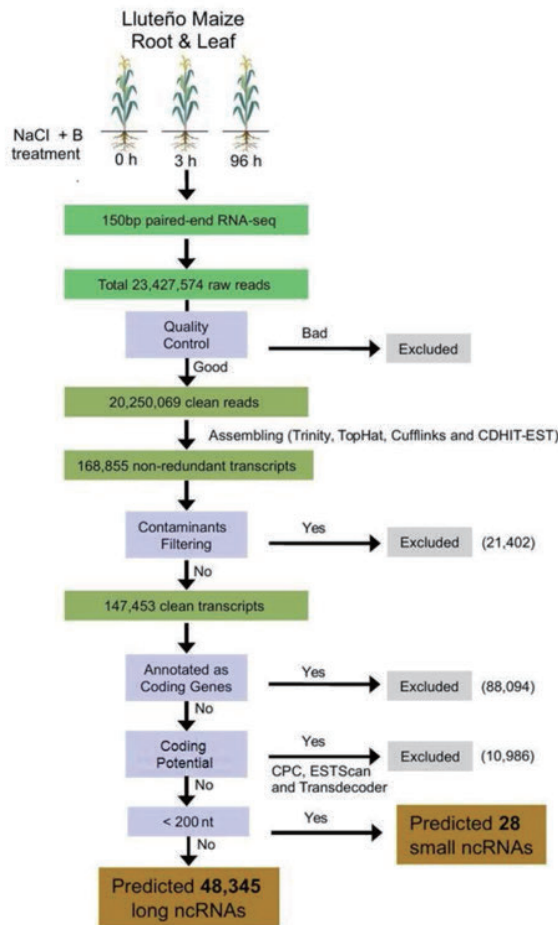


图3.2 文章研究思路总结

部分研究结果:

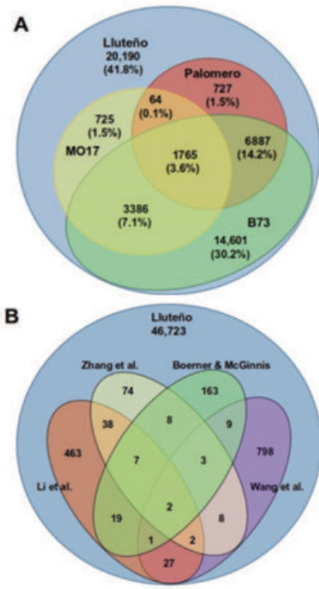


图3.3 比较Luteño玉米和其他玉米中转录组及lncRNA的差异

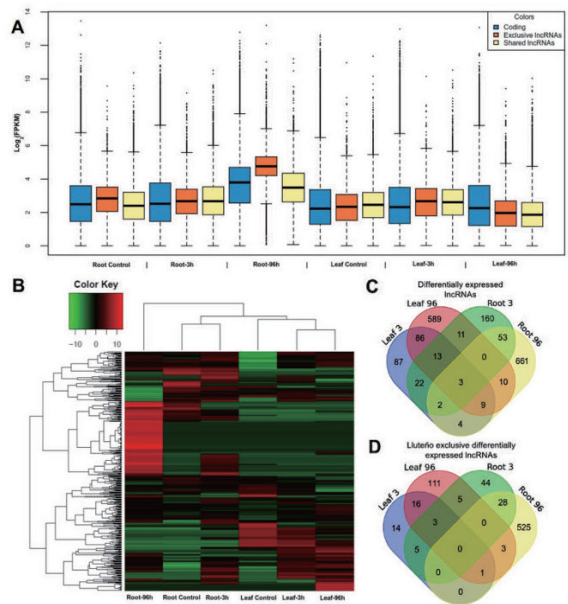


图3.4 Luteño玉米中lncRNA的差异表达

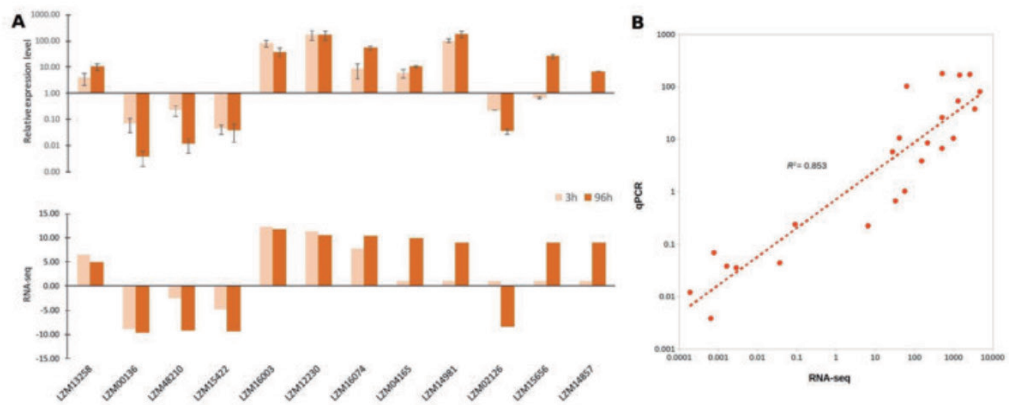


图3.5 RT-qPCR验证测序鉴定到的盐和硼响应相关lncRNA的差异表达

参考文献

1. Huancamamani W, Ariascarrasco R, Cárdenasinasivincha S, et al. Long Non-Coding RNAs Responsive to Salt and Boron Stress in the Hyper-Arid Luteño Maize from Atacama Desert. [J]. Genes, 2018, 9(3):170.

4 Small RNA 测序

小RNA (small RNA) 是指长度在 18-30 nt 的一类非编码 RNA。种类繁多, 包含 miRNA (微小 RNA), piRNA (蛋白互作小 RNA), endo-siRNA (内源干扰小 RNA), snoRNA (核仁小 RNA) 等。其中研究最普遍的为 miRNA, 其长度主要集中在 20-24 nt, 普遍存在于动植物中。在动物中, miRNA 主要作用是转录后抑制 mRNA 的翻译。在植物中, miRNA 主要作用是直接降解靶 mRNA。采用 Illumina 平台进行测序, 针对有参考基因组样本开展 miRNA 分析。

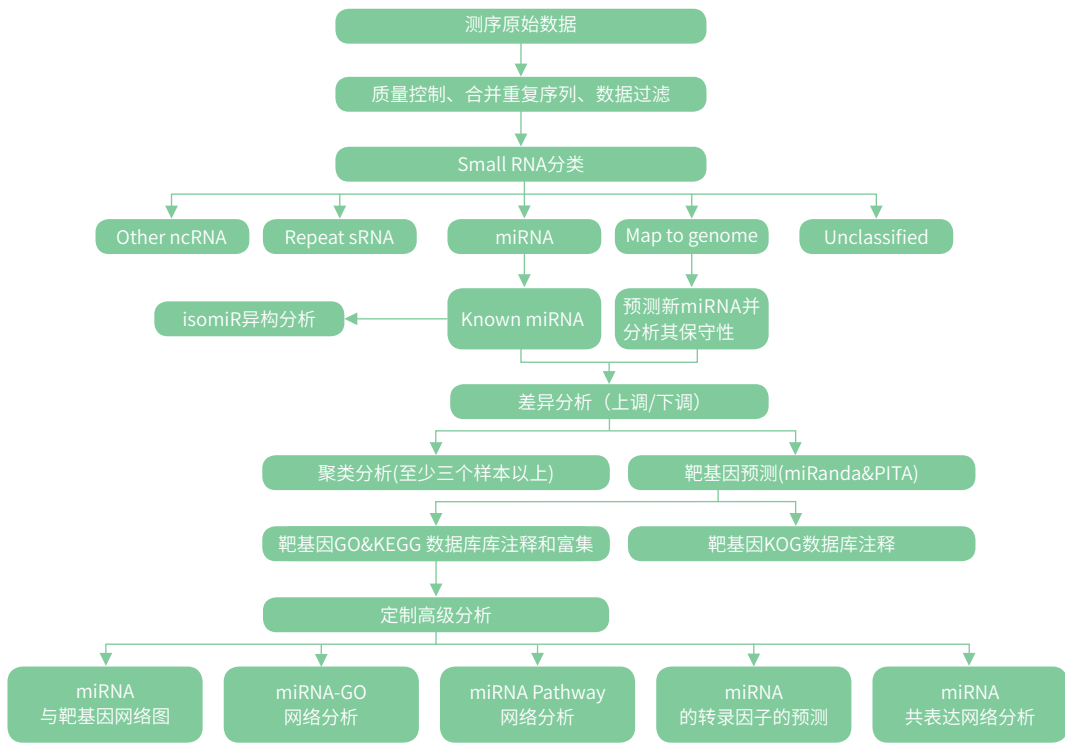


图5.1 项目流程图

产品优势

- 1 项目经验丰富，主流测序方法，严谨实验过程
- 2 分析采用miRBase数据库最新版本22.0，获得最新的microRNA信息
- 3 专业的数据分析，鉴定已知和新的miRNA，分析表达差异，预测靶基因等

技术参数

建库策略	参考基因组	测序策略	建议数据量
Small RNA 文库	有参且具有完整的注释	SR75	10~20M reads

产品应用

- 1 动植物的发育机制研究
- 2 品种间差异研究
- 3 重要性状研究
- 4 生物与非生物胁迫的应答机制研究
- 5 与转录组联合，研究RNA调控网络

案例解析

瓜尔豆组织特异性 miRNA 研究^[1]

研究背景

瓜尔豆包含 80%可溶性膳食纤维，高蛋白质含量和高价值黏胶，因此最近已成为重要的豆科经济植物。发育中的瓜尔豆种子在胚乳中积累了 90%的半乳甘露聚糖，因此可以作为模式作物来研究半乳甘露聚糖的生物合成及其调控原理。miRNA 是其相应靶基因的微小主调控因子，能够调控靶基因终产物的量发生变化。因此本研究从基因型 RGC-936 的瓜尔豆的 5 种组织中制备小 RNA 文库并测序，鉴定已知 miRNA 和新 miRNA，结果共发现 187 个已知 miRNA 和 171 个新 miRNA，选择 10 个 miRNA 进行验证，验证结果与测序结果一致。靶基因的 GO 注释结果显示：50.10% 的基因参与碳水化合物代谢，20.75% 的基因与激酶或其它酶相关，10.20% 的与转录因子相关，8.35% 与转运蛋白相关。2 个新的 unigenes 被注释为 ManS（甘露糖基转移酶 / 甘露聚糖合成酶）和 UGE（UDP-D- 葡萄糖 4- 差向异构酶），并且是 3 个新 miRNA（即 Ct-miR3130，Ct-miR3135 和 Ct-miR3157）的靶基因。该研究结果表明，这些新 miRNA 可能在调控半乳甘露聚糖途径中发挥重要作用。

研究路线

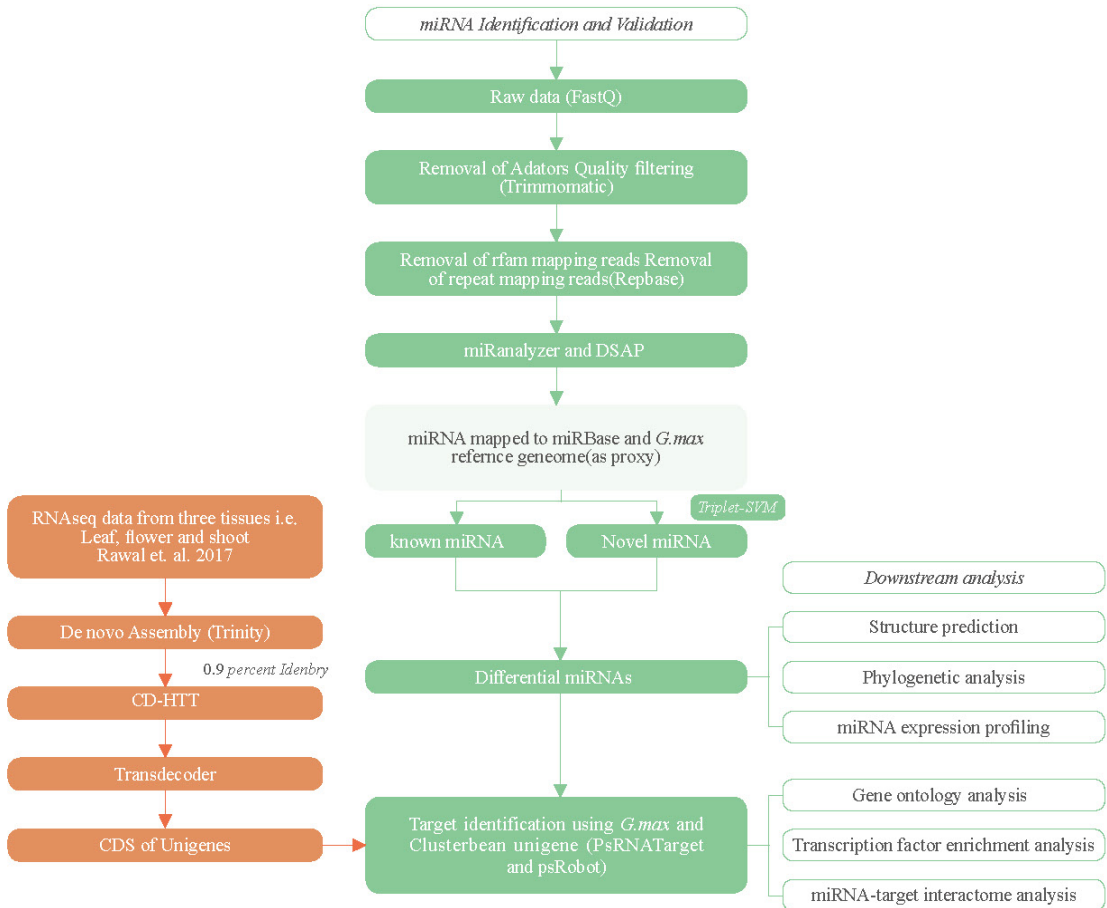


图5.2 文章研究数据分析思路总结

部分研究结果:

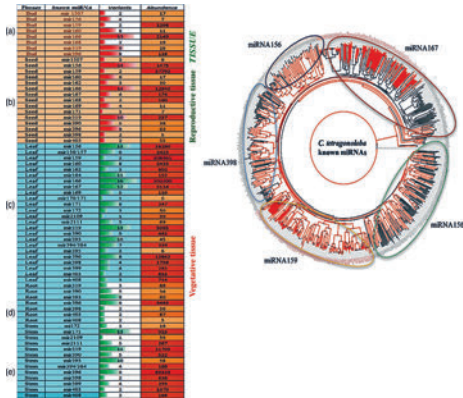


图5.3 定性和定量的分析已知miRNA家族

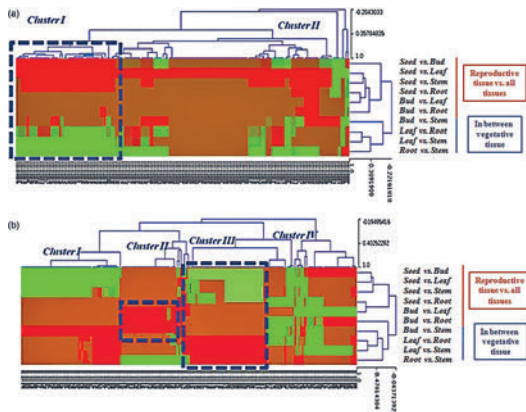


图5.4 热图和聚类分析171个新的和187个已知的miRNA在不同组织中的表达

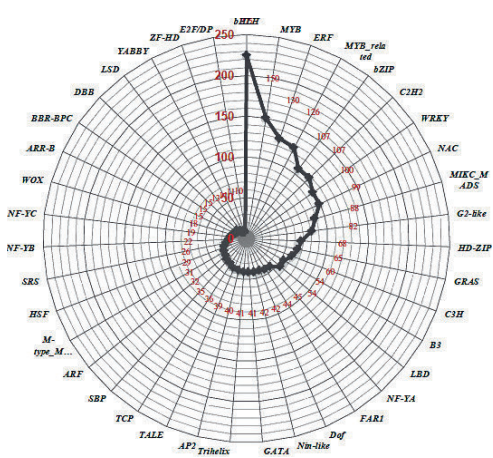


图5.5 作为miRNA靶基因的39个转录因子家族聚类分析

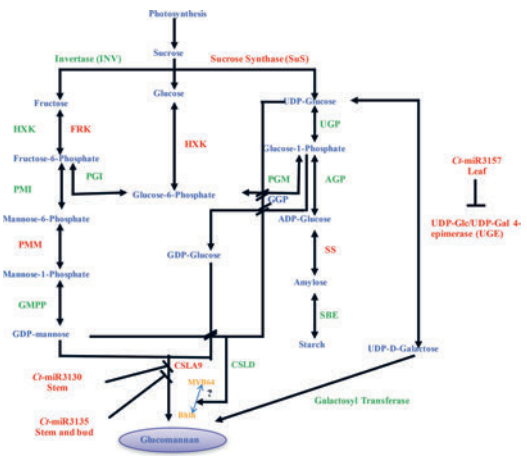


图5.6 半乳甘露聚糖合成代谢途径展示

参考文献

1. Tyagi A, Nigam D, S. V. A M, et al. Genome - wide discovery of tissue - specific miRNAs in clusterbean (*Cyamopsis tetragonoloba*) indicates their association with galactomannan biosynthesis[J]. Plant Biotechnology Journal, 2017.

5 全转录组测序

全转录组是指细胞所能转录出来的所有类型 RNA 的总和，包括 mRNA 和非编码 RNA (non-coding RNA, ncRNA)。目前针对非编码 RNA 的研究主要集中在 small RNA, lncRNA 和 circRNA。对特定组织或器官在特性状态下的全转录组研究，可以从 RNA 层面系统地揭示生物学背后的转录调控规律。

技术流程

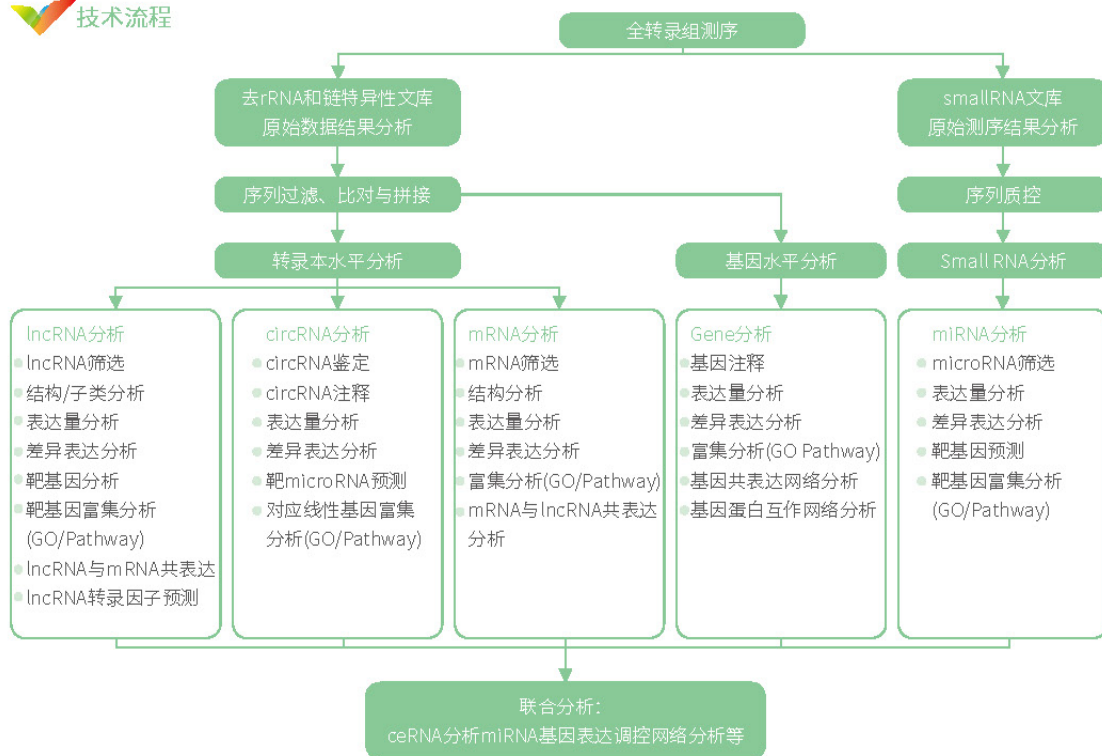


图6.1 项目流程图

产品优势

- 1 突破单一分子限制，实现RNA的全面研究，获取特定组织或器官在特定状态下的转录本集合信息
- 2 开展一次实验，采用一份样本，构建两种文库，可同时检测四种分子
- 3 丰富数据分析，包括各种RNA的表达差异和靶基因等；还可进行多维度的关联分析
- 4 紧扣RNA研究热点，构建ceRNA调控网络

技术参数

RNA 类型	建库策略	参考基因组	测序策略	建议数据量
mRNA	去 rRNA 和链特异性文库	有参且具有完整的注释	PE 150	12G/样本
lncRNA				
circRNA				
Small RNA	Small RNA文库		SR75	10M reads/样本

产品应用

- 1 动植物的发育机制研究
- 2 品种间差异研究
- 3 重要性状研究
- 4 生物与非生物胁迫的应答机制研究
- 5 中药的药用价值研究

研究背景

珠心胚 (NE) 是柑橘无融合生殖常见的形式, 其中无性胚胎直接从珠心细胞周围胚囊开始。NE 和脱毒砧木无性系的后代会产生理想的农艺性状, 却阻碍了杂交育种的发展。尽管 NE 在柑橘育种和商业生产中具有重要意义, 但对其潜在的分子机制却知之甚少。本研究通过对多胚柑桔进行组织学观察确定了其珠心胚的起始阶段 (NEI)。为探索 NEI 阶段涉及的基因和调控通路, 我们对两对品种 NEI 前及 NEI 期间的胚珠组织进行了 RNA-Seq 和 sRNA-Seq 分析。共鉴定得到差异表达基因 305 个。GO 分析显示, 基因显著注释到氧化应激反应中。本研究筛选到近 150 个 miRNA, 其中 90 个为已知的, 60 个新的 miRNAs。这种集转录和转录后调节为一体的研究为多胚柑桔研究提供了新的见解, 这将有助于揭示植物无融合生殖的调控机制。

本研究由博奥晶典提供RNA-seq和sRNA-seq实验技术服务。

研究路线:

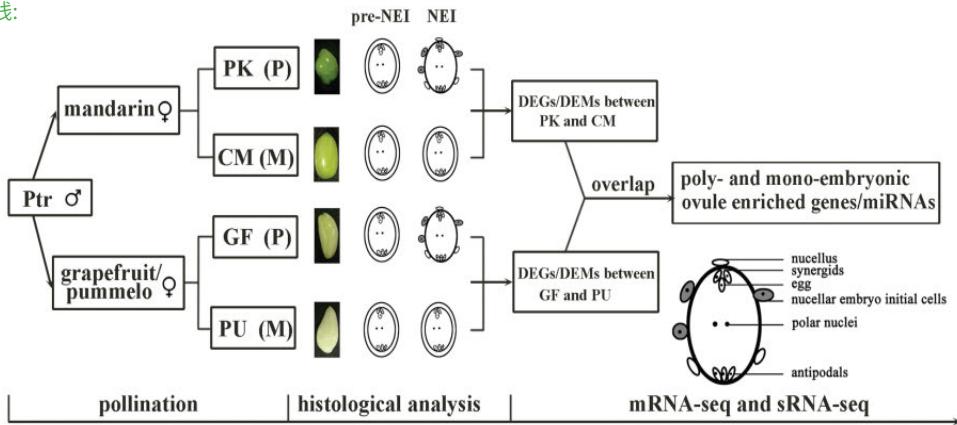


图6.2 文章研究思路总结

部分研究结果:

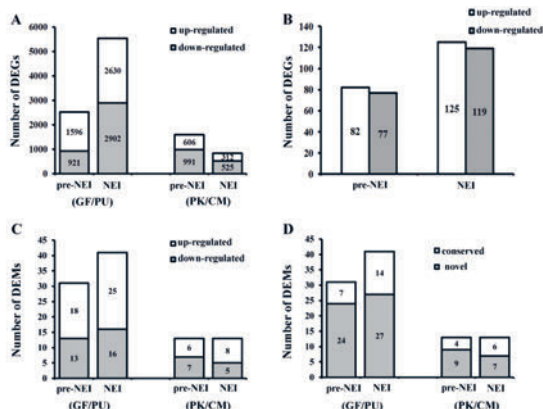


图6.3 珠心的胚胎发生时mRNA和miRNA的差异表达

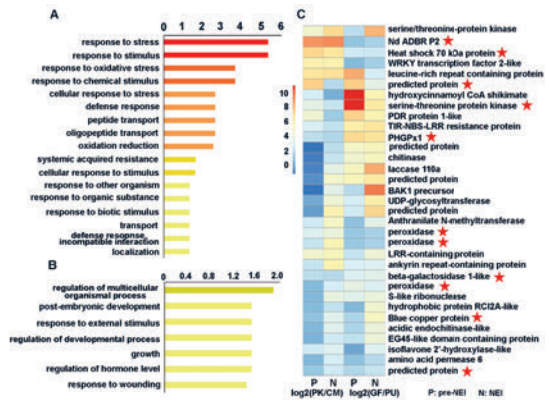


图6.4 GO富集分析生物学进程中上调 (A) 和下调 (B) 的基因及热图展示抗性响应中上调的基因 (C)

参考文献

1. Long J M, Liu Z, Wu X M, et al. Genome-scale mRNA and small RNA transcriptomic insights into initiation of citrus apomixis. [J]. Journal of Experimental Botany, 2016, 67(19):5743-5756.

表观组学和蛋白组学

1 Agena MassARRAY® 甲基化检测

针对全基因组甲基化测序或其他检测方法发现的差异甲基化位点，或者针对候选基因的特殊区域，通常需要进行大样本的定量检测工作。基于 Agena MassARRAY® 质谱平台，博奥晶典可提供候选区域甲基化位点定量检测服务，一次反应最多可检测长达 600bp DNA 片段上各 CpG 位点的甲基化程度。

技术流程

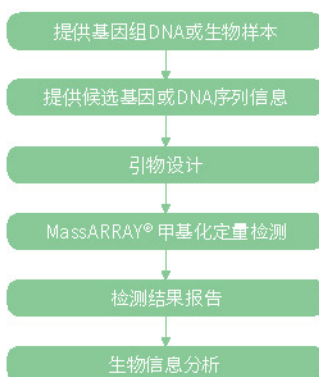


图1.1 项目流程图

产品优势

- 1 定性定量分析一步到位
- 2 可对200-500 bp的靶区域进行相对甲基化程度检测
- 3 可对每个扩增子内的多个CpG位点进行准确定量
- 4 可视化的数据产出，检测结果一目了然
- 5 适合于全基因组甲基化测序结果的验证及特定DNA片段的甲基化研究

产品应用

- 1 动植物个体发育与细胞分化
- 2 物种及种群进化
- 3 动植物育种

案例解析

猪脂肪和肌肉组织的 DNA 甲基化图谱构建^[1]

来自四川农业大学和北京大学的研究人员以猪为动物模型，选择长白猪，藏猪和荣昌猪三类猪种，对 180 份脂肪和肌肉进行甲基化免疫沉淀文库测序和质谱验证，同时进行组织基因表达谱检测，构建猪脂肪和肌肉全基因组 DNA 甲基化图谱和基因表达图谱（其中涉及的 Agena 质谱甲基化检测工作由博奥生物有限公司完成）。甲基化图谱显示了不同猪种、性别、解剖部位甲基化模式的共性及特性。其中特异性甲基化区域与已知肥胖相关基因或新基因的启动子区高度相关。

这项研究首次构建猪不同部位脂肪和肌肉组织的 DNA 甲基化图谱，为预防人类肥胖疾病的发生和促进猪肌肉生长及脂肪沉积性状的研究，奠定了重要的表观遗传学基础。

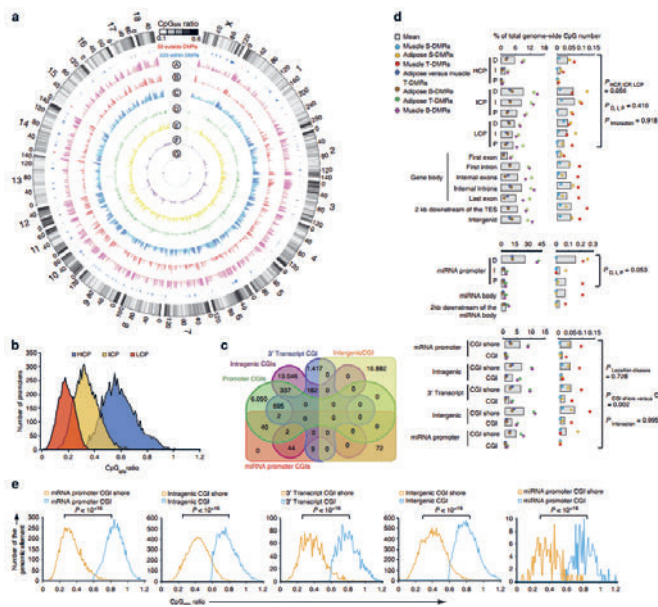


图1.2 猪全基因组差异化甲基化区域分布

参考文献

1.Li M, Wu H, Luo Z, et al. An atlas of DNA methylomes in porcine adipose and muscle tissues.[J]. Nature communications, 2012, 3: 850.

2 蛋白组研究

在蛋白组学的检测，博奥晶典提供基于 RayBiotech 平台的蛋白芯片检测和基于赛默飞平台的蛋白质组学检测。

博奥晶典是 RayBiotech 授权和认证的蛋白芯片服务提供商，RayBiotech 是世界上第一家专业从事蛋白芯片研发和生产的公司，也是国际上公认的蛋白芯片的先驱企业，博奥晶典可以提供 RayBiotech 全部玻璃基片的双抗夹心法抗体检测服务，如细胞因子、血管生成、细胞凋亡、信号转导等定量和半定量抗体芯片检测服务。

蛋白质组水平的定量检测方法主要是基于质谱检测的定量方法。随着高通量质谱定量技术的成熟，基于高分辨率质谱平台的蛋白质组学检测服务也已经逐渐成为研究的主流。

技术流程

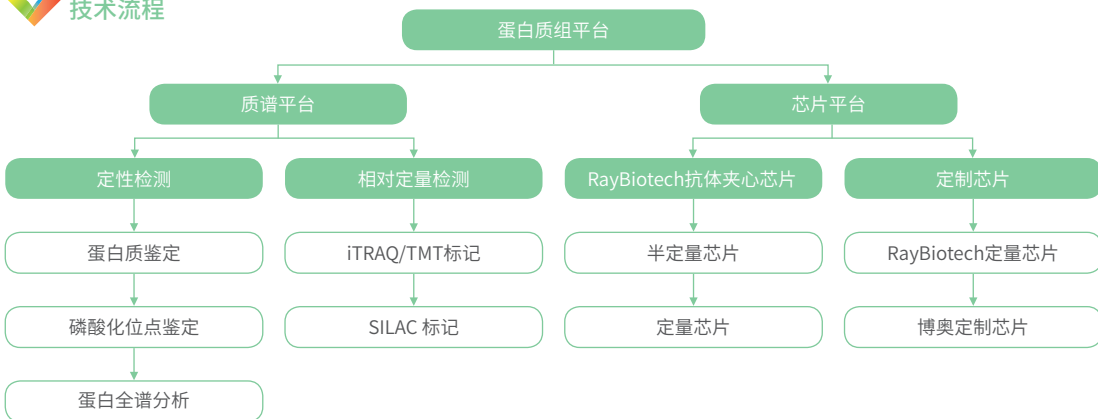


图2.1 项目流程图

产品优势

- 1 完整技术平台，可多平台联合分析
- 2 高效率、高准确性，国内首家完成单物种蛋白全谱分析的科研服务单位
- 3 专业技术支持，数据结果一对一解答
- 4 丰富的服务内容及项目经验，质量严格把控

技术参数

定性鉴定	IP、CO-IP、蛋白复合物中鉴定蛋白质（后续western blot验证）
相对定量	化学标记法（TMT标记定量：6或10个标记、iTRAQ标记定量：8个标记）

产品应用

- 1 疾病研究 生理或病理过程、发现新的疾病分子标志物、疾病分型、预后效果、蛋白质翻译后修饰功能研究等
- 2 药物研究 药物作用靶点和机制等
- 3 食品科学 成分鉴定、过敏原检测等
- 4 微生物学 病原微生物致病机制、耐药机理等
- 5 农业领域 如育种、转基因、抗逆研究等

案例解析

蛋白质谱在作物雄性不育机制研究中的应用^[1]

研究背景:

大豆是提供蛋白质和油的重要全球作物。杂种优势是增加大豆产量的一个显著潜力的方法。细胞质雄性不育 (CMS) 在杂交种子的生产中起着至关重要的作用。然而，大豆雄性不育的遗传和分子机制仍然需要进一步阐明。在本研究中，使用蛋白质组学技术进行分析，结果表明，一些关键的蛋白质参与大豆雄性不育相关联的途径。这项工作为了解大豆 CMS 的遗传和分子机制提供了新的视角。

研究路线:



图2.2 文章研究思路总结

部分研究结果:

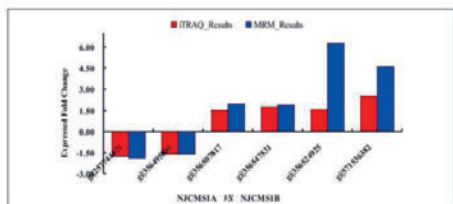


图2.3 蛋白质谱结果MRM验证 (选择10个, 成功检测6个)

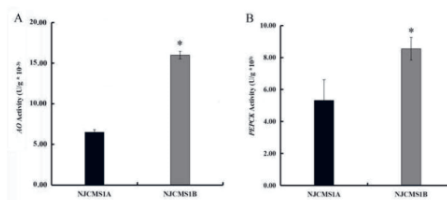


图2.4 雄性不育相关蛋白的酶活验证

参考文献

1.Li J, Ding X, Han S, et al. Differential proteomics analysis to identify proteins and pathways associated with male sterility of soybean using iTRAQ-based strategy.[J]. Journal of Proteomics, 2016, 138:72-82.

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
CapitalBio Dx



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>

