

系统的生物学研究 全面解决方案

2015 年 3 月



目 录

第一章 公司简介.....	4
第二章 平台简介.....	9
微阵列芯片服务平台.....	10
新一代测序服务平台.....	14
分子检测平台.....	18
生物标志物平台.....	19
生物信息分析平台.....	21
样品处理平台.....	23
第三章 微阵列芯片服务.....	24
(一) SNP 芯片服务	25
(二) CytoScan™ HD 芯片服务	38
(三) CGH 芯片服务	41
(四) 表达谱芯片服务.....	45
(五) miRNA 芯片服务	57
(六) DNA 甲基化芯片服务	64
(七) lncRNA 芯片服务	70
(八) 定制芯片服务.....	76
(九) 在线服务.....	77
(十) 蛋白质抗体芯片.....	78
第四章 测序服务.....	86
DNA 水平研究	87
(一) 全基因组重测序服务	87
(二) 全外显子组测序服务	89
(三) 目标区域捕获测序服务	96
(四) RAD-Seq 服务.....	100
RNA 水平研究	103
(一) 转录组测序服务	103
(二) miRNA 测序服务.....	118
(三) CircRNA-Seq	126
(四) RIP-seq 服务.....	128
表观遗传学水平.....	129
(一) 全基因组甲基化测序服务	129
(二) 甲基化 DNA 免疫共沉淀测序服务	135
(三) 染色质免疫共沉淀测序服务	141
(四) NimbleGen SeqCap Epi 甲基化序列富集方案.....	145



(五) Agilent sureselect 甲基化序列富集方案.....	146
(六) 焦磷酸测序法检测甲基化	148
(七) PacBio RSII 测序法检测甲基化	150
第五章 生物标志物验证及分子检测服务.....	156
(一) 中低通量核酸检测.....	158
(二) 蛋白标志物的免疫检测技术.....	167
(三) 标志物功能筛选及验证.....	171
(四) 基因组编辑技术-CRISPR/Cas9 技术	173
(五) 基因克隆蛋白表达纯化平台.....	173
第六章 生物信息分析服务.....	177
第七章 试剂耗材服务.....	183
(一) RNAsafety 组织 RNA 保存液	184
(二) Agilent PLUF 系列试剂耗材产品	185
(三) DNase I footprinting assay 服务	188
2012-2014 年度经典论文	191
经典服务案例.....	195

第一章 公司简介



上海伯豪生物技术有限公司
SHANGHAI BIOTECHNOLOGY CORPORATION

上海伯豪生物技术有限公司（以下简称上海伯豪）是上海生物芯片有限公司/生物芯片上海国家工程研究中心根据国内外研发外包发展的需要，整合旗下系统化的分子生物学技术平台、商业化服务体系、高素质服务团队等资源成立的致力于研发外包服务公司。



大股东为国家级生物芯片产业化基地—上海生物芯片公司
/生物芯片上海国家工程研究中心

公司于2011年12月6日获得高新技术企业认定。2012年3月获批浦东新区企业研发机构。



2011 年上海伯豪被批准为上海市高新技术企业



2013 年获批浦东新区研发机构

上海伯豪拥有六大技术平台：生物样品分析平台、微阵列芯片平台、新一代测序平台、生物标志物验证平台、分子检测平台、生物信息平台，凭借先进的技术平台，多样化的服务和严格的质量控制等竞争优势，公司向国内外企业和相关单位提供系统的生物学研究全面解决方案。目前正在为多达18家跨国制药企业（包括排名前10位的跨国制药企业）和超过1800家的国内科研机构、医院等提供基因表达谱、基因分型、比较基因组学、DNA甲基化、miRNA、生物标志物筛选及确认、生物信息等技术服务。

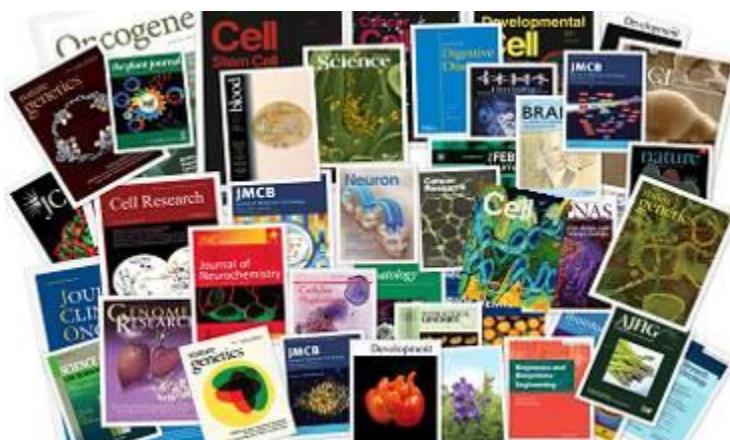
公司拥有一支以上海为中心，辐射全国的强大市场营销队伍和销售网络，已设立华北、华南、华东三个大区十五个办事处，推广公司的主导技术服务、代理产品，快速提升公司品牌的知名度，扩大影响力。历年来承接项目数超过4000个，用户单位1800家以上，客户发表SCI论文600余篇（影响因子2500余分），全面推动了中国基因组学服务产业的发展。



国际客户分布图



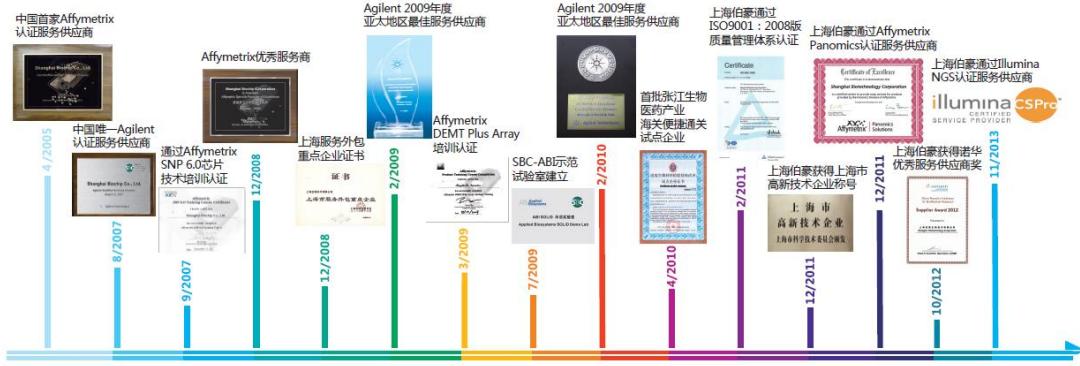
国内客户分布图



部分利用公司服务发表 SCI 论文封面展示

公司建立了规模化的质量控制体系，通过了ISO9001: 2008质量技术体系的认证，并参照GLP-L的标准，通过128个SOP文件对项目、样品、实验过程、生物信息分析过程、数据信息管理、保密等进行严格质控。尤其是数据管理，做到异地备份，保密管理。为用户的数据真实性、可靠性提供了保证。

公司的系统化技术平台和高质量的服务体系得到了广大客户认可和合作伙伴的高度评价。公司平台陆续获得了国际知名高通量技术厂商Illumina、Affymetrix、Agilent、ABI的技术服务中心认证。公司于2012年获得诺华（CNIBR）优秀供应商奖（Supplier Award 2012）。公司是Affymetrix公司在中国第一家认证的服务提供商，是Agilent公司在国内唯一认证的服务提供商，先后被授予“Agilent公司亚太区最佳服务供应商”和“Affymetrix公司优秀服务商”称号。



公司各类认证展示

2013年起，公司加强兼并和收购，先后参股成立第三方检测专业公司“上海伯豪医学检验所有限公司”，控股医药销售专业公司“上海伯豪医药科技有限公司”。

公司在今后的发展中，将一如既往利用高端的现代化仪器设备、多项成熟完善的技术平台、高素质的技术团队、国际水平的商业化服务流程、积极进取的经营理念，为广大客户提供持续、高效和稳定的技术支持和服务。

伯豪公司园区图片展示





上海伯豪生物技术有限公司
SHANGHAI BIOTECHNOLOGY CORPORATION



伯豪临检所图片展示



地址：上海市张江高科路444号15幢1层(201205)

电话：021-51520288 传真：021-51520266

8 / 201

技术服务热线：800-820-5086 / 400-880-5086

技术服务网站：<http://www.ebioservice.com>

第二章 平台简介



微阵列芯片服务平台

上海伯豪生物技术有限公司作为国内较早从事生物芯片服务的团队，具有多年的研究和服务经验，已建立了 SBC 点制芯片、Affymetrix 芯片、Agilent 芯片、Illumina 芯片等微阵列技术服务，为国内外数千家科研单位和企业提供技术服务和技术支撑。

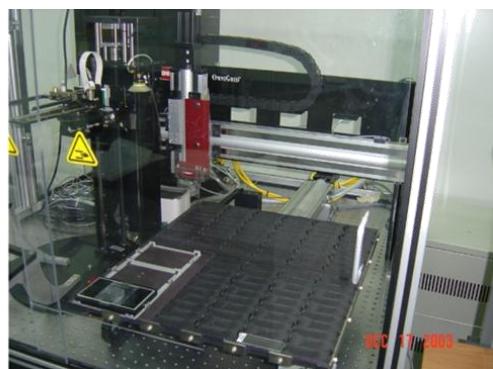
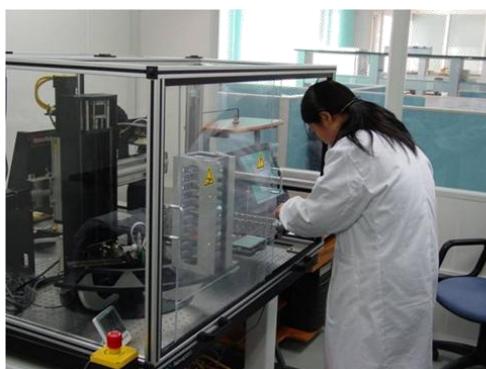
上海伯豪微阵列芯片服务平台总览

研究类型	服务平台	用途	服务种类
基因组层次 研究服务	SNP 芯片服务平台	用于研究各种生物 DNA 序列的差异和单核苷酸多态性，同时研究这些差异对疾病、诊断、预后的影响。	Affymetrix SNP 系列芯片服务 Illumina SNP 系列芯片服务
	CNV 芯片	用于研究基因拷贝数的变化，从而明确疾病遗传基因，阐明其致病机制。	Agilent CGH 系列芯片服务 Affymetrix 系列芯片服务
转录组层次 研究服务	基因表达谱芯片服务平台	用于大规模分析一定生物对象在特定生物工程中基因表达变化的全面信息。	Affymetrix 基因表达谱系列芯片服务 Agilent 基因表达谱系列芯片服务 Illumina 基因表达谱系列芯片服务
	miRNA 芯片服务平台	用于研究 miRNA 的功能及其在动植物的生长发育、细胞分化和凋亡以及人类疾病的发生等过程中的作用。	Agilent miRNA 系列芯片服务 Affymetrix miRNA 系列芯片服务
	DNA 甲基化芯片服务平台	在全基因组水平研究 DNA 甲基化在发育、X 染色体失活、衰老以及人类疾病中的变化。	Illumina DNA 甲基化芯片系列服务

SBC 芯片服务平台

SBC 芯片服务平台经过十年的发展已经成为国内最为系统稳定的高通量基因组研究技术平台，从样品制备、芯片制作、芯片杂交、数据扫描到后期的数据管理，储存以及深度数据挖掘都有了标准化的流程、坚实的理论和实验的支持。

- 芯片点样

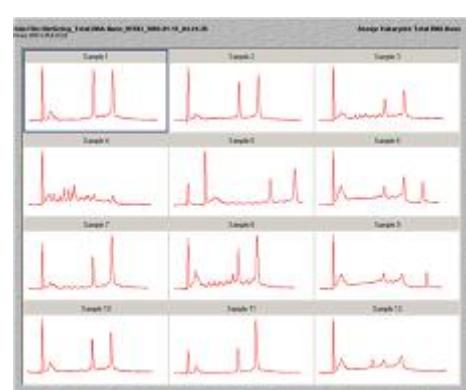


GeneMachine OmniGrid System 100 基因芯片点样系统

- 样品检测



Agilent 2100 分析仪

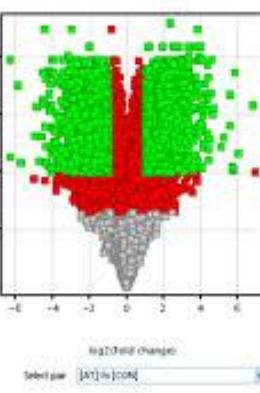
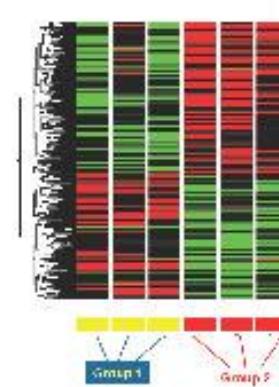


Agilent 2100 实验结果图

- 扫描及数据分析



芯片扫描及数据分析



生物信息分析结果图



Affymetrix 微阵列芯片平台

美国 Affymetrix 公司开发的寡核苷酸原位光刻合成专利技术(light-controlled in situ synthesis of DNA microarrays)，是目前最高密度的芯片制备技术。Affymetrix GeneChip 生物芯片检测系统是世界上第一种经欧盟和美国 FDA 批准的可作为体外诊断的芯片系统。Affymetrix 在表达谱方面，有 3'IVT 和全转录本表达谱芯片，全转录组表达谱芯片克服了传统 3'IVT 芯片只检测 3' 端的不足；在基因分型方面，有 SNP 6.0 芯片，Axiom[®] Genome-Wide CHB 1 & 2 Array Plate, Axiom[®] Exome 319 Array Plate；在细胞遗传学方面，有 CytoScanTM HD 芯片，该芯片解决了传统核型分析、荧光原位杂交(FISH) 等技术分辨率低、只能定性分析的问题。

作为中国首家通过 Affymetrix 认证的技术服务供应商以及 Affymetrix 优秀服务商，本平台可为广大科研工作者提供 mRNA 表达谱、SNP 基因分型、miRNA 表达谱，细胞遗传学等一系列生物芯片服务解决方案。



Affymetrix 芯片系统



Affymetrix GeneAtlas 系统

Agilent 微阵列芯片平台

美国 Agilent 公司采用专利的非接触式工业化喷墨印制方法，依据设计序列，将 60-mer 长度的寡核苷酸单体均一地置于特殊处理的载玻片上进行原位合成印制。这一技术无需接触载玻片表面，也不会引入表面接触的不均匀性，最终实现样点的一致性及可重现性。第三代 SurePrint G3 微阵列芯片产品延续了安捷伦传统以来在检测低表达水平基因方面的高灵敏度和检测结果的高准确度，确保用户获得高质量的芯片数据。

作为中国首家通过 Agilent 认证的技术服务供应商以及 Agilent 亚太地区最佳服务供应商，该平台可为广大科研工作者提供 mRNA 和 lncRNA 表达谱、miRNA 表达谱、CGH 等一系列生物芯片服务解决方案。同时利用 Agilent eArray 芯片探针设计软件，为客户提供简便经济的定制芯片服务。



Agilent 2100 分析仪



Agilent 2200 分析仪



杂交炉



Agilent 扫描仪



Illumina 微阵列芯片平台

Illumina 公司以测序业务闻名，而其在生物芯片方面的业务也具有不可忽视的优势和竞争力。激光共聚焦微珠芯片技术是公司的核心技术之一，也是目前最前沿的芯片技术之一。

2012 年，上海伯豪生物技术有限公司整合上海南方基因科技有限公司平台资源，为国内外科研用户整合推出“基于 Illumina Beadlab System 的 Illumina 生物芯片全面解决方案”。该方案的推出，将为广大科研工作者提供基于 Illumina Beadlab System 的 SNP 基因分型和拷贝数变异（CNV）检测、DNA 甲基化芯片和 mRNA 表达谱芯片等一系列生物芯片服务解决方案。



Illumina 芯片操作系统



新一代测序服务平台

新一代测序技术作为对传统测序的变革，近几年迅速发展。目前，新一代测序技术已广泛应用于生物学研究的各个领域，很多生物学问题都可以借助新一代测序技术予以解决。上海伯豪生物技术有限公司自 2009 年起先后建立了 ABI SOLiD 新一代测序服务平台和 Illumina 新一代测序服务平台，拥有 Illumina HiSeq2500 V4（1 台），HiSeq2500 V3（1 台），Illumina GAIIx（1 台），ABI Ion PGM 1 台，ABI SOLiD5500（2 台）和 ABI SOLiD4（2 台）。

2010 年起，上海伯豪生物技术有限公司正式对外推出“SBC 新一代测序系统解决方案”，将新一代测序技术与生物芯片技术相结合，协助广大科研工作者以更全面、更深入和更低廉的价格获得基因组、转录组及修饰组的各项数据。力争尽快将新一代测序技术及生物芯片技术相结合，更加系统地分析基因组的结构和表达的变化。

服务种类一览

基因组测序（DNA-Seq）	转录组测序（RNA-Seq）	表观基因组测序
<i>de novo</i> 测序	转录组测序（无参考基因组）	染色质免疫沉淀测序（ChIP-Seq）
全基因组重测序	转录组测序（有参考基因组）	甲基化 DNA 免疫共沉淀测序（MeDIP-Seq）
全外显子组测序	small RNA 测序	全基因组 Bisulfite 甲基化测序
DNA 目标序列捕获重测序	微量样本转录组测序	
RAD-Seq	定向转录本测序	
单细胞全基因组测序	宏转录组测序	
单细胞外显子捕获测序	RIP-Seq	
16S rDNA 分子生态测序		
宏基因组测序		

Illumina HiSeq2500 V4 新一代测序服务平台

Illumina HiSeq 测序系统是一种基于荧光标记的可逆终止化学反应原理实现的单分子簇边合成边测序技术。由于全部四种可逆终止 dNTP 都会在同一测序循环中出现，其中的自然竞争过程可使合成偏差降至最低，所以对于同聚物的检测也不会有太大问题，这种可逆终止的化学方法严格确保了逐个碱基的测序。目前，illumina solexa GAIIx 测序系统双端读长可达 300nt，HiSeq2500 测序系统高通量模块双端读长可达到 250nt，单次运行可产生 1000G 的原始数据。快速模块单次运行仅需 27 小时，可产生 120G 的原始数据（以 2x100 计算）。此外，快速模块读长最长可达到双端 500nt，仅需 60 小时，即可产生 300G 的原始数据。



Illumina Hiseq2500 V4 系统



Illumina solexa GAIIx 系统



cBot Cluster Generation

Illumina Miseq 平台

Miseq 平台是 Illumina 以久经考验的边合成边测序技术为基础，当单个碱基结合 DNA 链时，通过专利的可逆终止子方法对数百万到数千万个片段进行大规模并行测序。新的 MiSeq 试剂 以 25 M 测序 reads 和 2x300 bp 读长实现了高达 15 Gb 的产量。作为一台在单个仪器上整合了扩增、测序及数据分析的测序仪器，MiSeq 以其高效、灵活的优势广泛应用于小基因组（细菌、线粒体或叶绿体基因组）de novo 测序及重测序、PCR 扩增产物 Barcoded Sequencing、转录组测序、Small RNA 测序、ChIP-Seq 等。



Illumina Miseq 系统

ABI SOLiD 新一代测序服务平台

SOLiD 系统包含测序组件、化学组件、计算集群和数据存储组件。这个平台基于通过寡核苷酸连接和检测进行测序。与聚合酶测序方法不同的是，SOLiD 系统利用专利的逐步连接（stepwise ligation）技术来产生高质量的数据。目前，SOLiD 系统单次运行最大可产生 200G 的原始数据（以 2x60nt 计算）。



ABI SOLiD 平台一览

Pacbio RS II 测序技术平台

PacBio RS 系统专利的 SMRT Cells，其含有数以万计纳米级的零模波导孔（zero-mode waveguides, ZMWs）。测序反应是在专利的 SMRT cell 中进行的，每个 SMRT cell 中有 150,000 个 ZMW。每个 ZMW 都能够包含一个 DNA 聚合酶及一条 DNA 样品链。这样，SMRT Cell 能够平行检测大约 75,000 个单分子测序反应。

地址：上海市张江高科技园区李冰路 151 号（201203） 电话：021-51320288 传真：021-51320266 15 / 201

技术服务热线：800-820-5086 / 400-880-5086

技术服务网站：<http://www.ebioservice.com>



当测序进行时，专利的包被技术保证 DNA 聚合酶和模板形成的复合体被锚定在 ZMW 的底部，反应溶液中带有标记着不同荧光磷酸基团的高浓度核苷酸可有效保证聚合酶合成的速度、精确性和持续合成的能力，检测装置则透过 ZMW 底部的基层来实时观察 DNA 的合成过程：当 DNA 聚合酶检测到正确的核苷酸并将其插入模板时，这些碱基会在插入位置保留几十毫秒，而游离状态的碱基扩散的速度大约是几十微秒。在这个时间差中，检测器就可以检测到插入碱基的荧光信号，并将其转换成相对应的碱基类型。然后，DNA 聚合酶将碱基所带的标记着荧光集团的磷酸基团剪掉以形成天然 DNA 链。此后，荧光信号迅速衰减至基线并开始下一轮的合成。PacBio RS 使用一套优化的算法，将光学系统所捕获的信息翻译成对应的 ACGT 碱基信息。一旦测序开始，实时的数据就被传送到初步分析系统中，以实时生成碱基组成和质量值等信息。



Pacbio RS II 测序仪

Ion Torrent 测序技术平台

Ion Torrent 是最新一代的测序技术，它的问世给测序技术的应用带来了激动人心的进展。它采用了半导体技术和简单的化学试剂进行 DNA 测序，而不是使用光 作为媒介。在半导体芯片的微孔中固定 DNA 链，随后依次掺入 ATCG。随着每个碱基的掺入，释放出氢离子，在它们穿过每个孔底部时能被检测到，通过对 H⁺ 的检测，实时判读碱基。



Ion Torrent PGM 测序仪

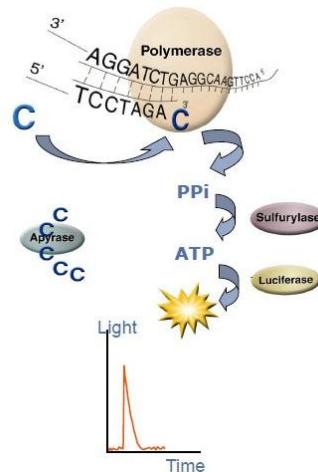
焦磷酸测序技术平台

焦磷酸测序技术（Pyrosequencing）作为一种新的序列分析技术，能够快速地检测甲基化的频率，对样品中的甲基化位点进行定性及定量检测，为甲基化研究提供了新的途径。上海伯豪为您提供基于 QIAGEN 公司焦磷酸测序平台的服务。

焦磷酸测序技术原理

焦磷酸测序技术是由 4 种酶（DNA 聚合酶、ATP 硫酸化酶、荧光素酶和三磷酸腺苷双磷酸酶）催化的同一反应体系中的酶级联化学发光反应。具体过程如下：

1. 仪器向待检测样品中每次只加入一个核苷酸碱基(dNTP)
2. 互补的碱基结合在模板上时，会释放出焦磷酸基团 Pyrophosphate (PPi)
3. PPi 在 ATP 硫酸化酶(ATP Sulfurylase)作用下，合成 ATP，ATP 作为能量分子，驱动荧光素在荧光素酶(Luciferase)的催化下转化为氧化荧光素，并发出与 ATP 能量成正比的荧光信号
4. 信号被 CCD 数码相机捕捉，形成峰形 PyrogramTM
5. 未与模板结合的碱基将被三磷酸腺苷双磷酸酶(Apyrase)降解
6. 加入下一种 dNTP



Qiagen Q96 ID



分子检测平台

上海伯豪生物技术有限公司分子检测平台建于 2013 年，主要进行中低通量的核酸检测服务，主要包括 Sanger 测序技术，焦磷酸测序技术，荧光定量 PCR 技术，PCR array 等服务。本平台主要检测仪器有 ABI 7900HT，Qiagen 96Q ID，ABI9700，Agilent Mx3005P 等。分子检测平台主要进行中低通量的核酸检测服务，其中 Sanger 测序，焦磷酸测序和 Massarray 技术主要应用有肿瘤相关基因的突变检测，肿瘤易感基因的检测，基因的 cDNA 测序，甲基化位点检测等；其中荧光定量 PCR 技术和 PCR array 技术主要应用于 mRNA、microRNA、lncRNA 的定量检测，基因拷贝数检测以及基因分型的检测。于此同时，分子检测平台运用以上检测技术

针对本公司微阵列芯片平台和新一代测序平台的结果，提供相应的验证服务。



Qiagen Q96 ID



PCR 仪



ABI 7900HT



Agilent 3005



生物标志物平台

核酸标志物检测及验证平台

运用低密度芯片、液相芯片，以及荧光定量 PCR 等技术，对多物种的各类核酸标志物分子进行定量及定性分析的技术平台。

基于 branched DNA 技术开发的多基因定量检测技术，Affymetrix Quantigene MultiPlex Assay 无需分离核酸，直接裂解后可在单管内同时检测 3-80 个基因的表达水平，可检测低至 1,000 个拷贝的核酸分子，具有很高的灵敏度；特别适用于血液样本和保存多年后 mRNA 高度降解的 FFPE 样本。

基于 branched DNA 技术开发的单核酸分子原位检测技术，Affymetrix QuantiGene ViewRNA ISH 结果为图像，可以定性和半定量分析目标基因的表达及分布情况。在单细胞水平检测低至一个拷贝的 mRNA，在定量与结构生物学水平原位检测多基因的 mRNA，在抗体无法获得的情况下，对免疫细胞化学研究是一种极好的补充。



BioPlex 系统



AB 7900 荧光定量 PCR 仪



AB3730 测序仪



Affymetrix-Panomics 服务认证



常规测序仪标准认证



荧光定量 PCR 仪标准认证

推荐服务：

	服务名称	应用
低通量多基因验证技术	Qiagen RT ² Profiler PCR Array AB TaqMan® Low Density Arrays	30-360 个基因的通量化基因表达水平、SNP、CNV、miRNA 等检测和验证。
	Affymetrix QuantiGene 2.0 MultiPlex Assay	3-80 个多核酸分子 (mRNA、miRNA、LncRNA、融合基因检测、DNA 拷贝数等) 通量化检测和验证，尤其适用于 FFPE 样本及血液样本。
荧光定量 PCR 技术	Taqman 探针法	进行 SNP 位点、miRNA、mRNA、SNP、CNV 等精确定量检测和验证。
	Sybgreen 法	进行多个或单个基因 (mRNA、miRNA、)



		LncRNA 等) 表达水平、CNV 拷贝数变化等快捷低成本检测。
常规测序技术	ABI 的 3730 测序仪、pyrosequencing 测序仪	采用 PCR 产物测序或克隆测序，进行 DNA 的甲基化、SNP 和突变验证检测。
单核酸分子的原位检测技术	DNA FISH	对 DNA 拷贝数及融合基因等进行检测和验证。
	Affymetrix QuantiGene ViewRNA ISH	对各类物种的组织切片及组织芯片进行 RNA 定位、分布、迁移、相互作用等方面的检测验证。

蛋白标志物检测平台

采用基于免疫反应检测的技术，主要进行蛋白表达水平的检测及定量定性分析。

Fullmoon 蛋白芯片：所有抗体都通过专利的 3-D 聚合材料固定在高质量的玻片表面，从而确保抗体芯片产品高度的有效性和特异性。为了增大数据的可信度，每个抗体都进行了六次重复。

液相芯片基于 xMAP 技术原理，整合了荧光编码微球、激光检测、应用流体学、最新的高速数字信号和计算机运算法则等多项技术，每个微球作为单独的检测体，可同时进行大量的生物检测，只需要 10~20 μl 的样本量就可以一次检测多达 100 个指标（FLEXMAP 3DTM 可多达 500 个指标），最快可达 10000 次测试/小时，真正实现了“高通量”与“高速度”。检测低限可达 0.01 pg/Ml, 检测范围达 3.5-6 个数量级，与 ELISA 和质谱分析具有很强的一致性。

推荐服务：

服务名称	应用
Fullmoon 蛋白芯片	广谱磷酸化蛋白芯片，广谱及功能特异的信号传导通路蛋白芯片，肿瘤标志物蛋白芯片等的通量化检测
液相芯片	Panomics 的 Cytokine 细胞因子检测、Bio-Rad 的 Cytokine 细胞因子、磷酸化多重蛋白因子检测试剂盒，Millipore 疾病相关因子 3-60 重的多因子检测。

CRISPR/Cas9 技术服务平台

CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats) 是最新出现的一种由 RNA 指导的 Cas9 核酸酶对靶向基因进行编辑的技术。CRISPR/Cas9 是细菌和古细菌为应对病毒和质粒不断攻击而演化来的获得性免疫防御机制。在这一系统中，crRNA (CRISPR-derived RNA) 通过碱基配对与 tracrRNA (trans-activating RNA) 结合形成双链 RNA，此 tracrRNA/crRNA 二元复合体指导 Cas9 蛋白在 crRNA 引导序列靶定位点切断双链 DNA(图 2)。在基因组 编辑过程中，tracrRNA 和 crRNA 可以融合成为 1 条 RNA(sgRNA) 表达同样可以起到靶向剪切的作用。CRISPR/Cas9 的优点是操作简单，对基因组的效率高。需要对某一个靶位点编辑的时候，只需要表达相应的 sgRNA 即可，不需要对 Cas9 核酸酶进行改造。它可对任何物种的基因组进行高 效率的定向编辑。基因组编辑技术的一个缺点是脱靶效应，可能在靶点以外的地方切断 DNA，产生 indel。本公司用一个突变的 Cas9 酶 Cas9n 克服了这个局限，，它只切断 DNA 的一个链，这个单链缺口会促进同源重组，但不会产生脱靶突变。



生物信息分析平台

随着微阵列芯片和高通量测序等高通量基因组学技术的发展，实验中产生了海量数据和信息，而这些数据的存储、整理和分析都成为目前生命科学研究中极具挑战性的课题。

上海伯豪生物技术有限公司作为生物芯片上海国家工程研究中心致力于研发外包服务的子公司，利用中心多年生物信息分析研发积累，目前正在为多达 18 家跨国制药企业（包括排名前 10 位的跨国制药企业）和超过 1100 家的国内科研机构、医院等提供各层次生物信息分析解决方案。

1. 硬件设备 — “高性能计算节点+高速网络+高效能存储”的整套解决方案

上海伯豪生物技术有限公司根据自身基因组学服务的需要定制了天梭 TS10000 高性能集群，共由 20 个节点和 1 个存储构成，储存能力达到 150TB，计算能力超过 1.8 亿次每秒以上。该集群在设计上专注于提高 HPC 应用的运行效率，增强系统的可扩展性，致力于提升系统的总体计算能力、加速节点间的数据交换和系统的 I/O 带宽。同时，上海伯豪生物技术有限公司还根据实际运算需求，开发了不同级别的软件应用，促进系统内部服务器、存储、应用之间最大限度趋近协同运算。

整体硬件设计采用了“高性能计算节点+高速网络+高效能存储”的整套定制方案。计算节点采用的是基于 Intel®至强®5500 处理器的浪潮 NF5220 和四路六核服务器 NF520D2。该集群不仅计算性能高，还具有高扩展、低功耗的特点，由于采用了智能功耗控制技术，该集群整体功耗比同类产品至少降低 10%，在任务较少时，还可以通过浪潮天梭监管软件远程关闭没有工作的节点，进一步节能减耗。



数据备份系统
China Telecom in WuXi, Jiangsu)

2. 软件设备 — 商业化和自主开发相结合的微阵列芯片生物信息分析综合解决方案

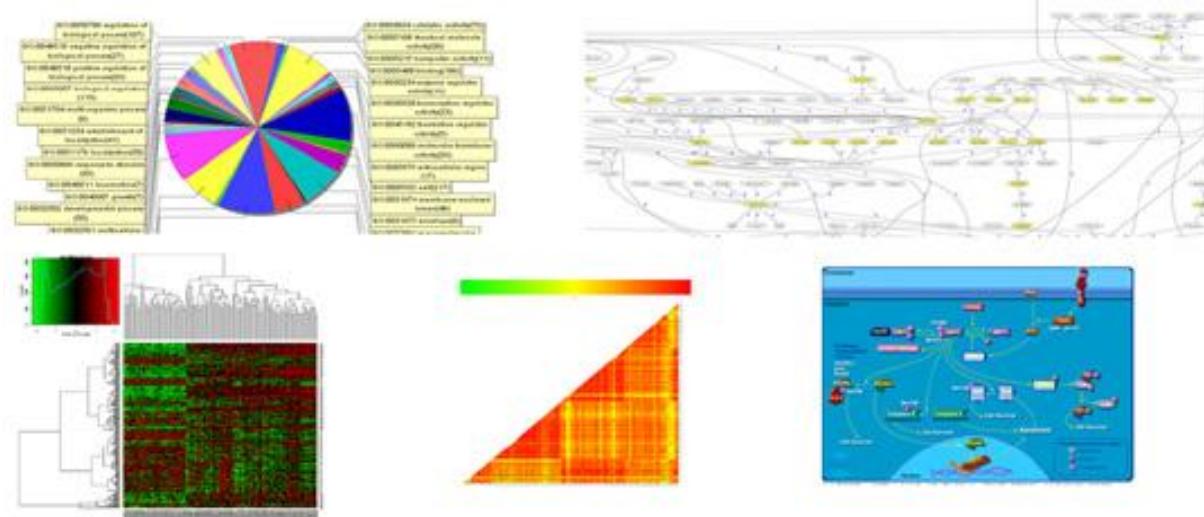
上海伯豪生物信息平台的软件系统采用商业化和自主开发相结合的方式建立了微阵列芯片数据分析综合解决方案，可针对表达谱芯片、miRNA 芯片、SNP 芯片、DNA 甲基化芯片、ChIP-on-Chip、CGH 芯片等提供各种个性化服务。

同时，上海伯豪还针对广大科研工作者设计工作需要开发向伯豪用户开放的个性化实时在线

分析系统—SBC Analysis System, SAS 系统。

该系统核心算法基于 R 语言(R-software)程序包，可以完成大部分的表达谱分析运算，系统本地整合 8 大公共数据库，实现对多个数据库的交叉查询，帮助您发现实验揭示的生物学意义。

该在线分析系统将帮助广大科研工作者表达谱芯片进行各类统计学分析和功能注释工作。这些分析包括，差异倍数计算(Fold Change), t 检验 (t-Test), SAM 检验(Significant Analysis of Microarray, SAM), 一维方差分析(one-way ANOVA), 二维方差分析(two-way ANOVA), 相关性分析(Correlation analysis), 聚类分析(Unsupervised hierarchical clustering), 主成份分析(Principal Component Analysis, PCA), GO 富集分析 (GO enrichment analysis), pathway 富集分析(Pathway enrichment analysis), microRNA 检索, 蛋白质相互作用检索(protein interaction)等。



个性化实时在线分析系统——SBC Analysis System, SAS 系统

生物信息分析服务一览

生物芯片分析服务			
基因组研究层次	转录组研究层次	表观遗传研究层次	SBC 特色分析服务
SNP 芯片分析服务	表达谱芯片分析服务	DNA 甲基化芯片分析服务	各类芯片联合分析服务
CNV 芯片分析服务	miRNA 芯片分析服务		SBC Analysis System 分析服务
	lncRNA 芯片分析服务		GEO 上传服务
新一代测序分析服务			
基因组研究层次	转录组研究层次	表观遗传研究层次	SBC 特色分析服务
<i>de novo</i> 测序分析服务	RNA 测序分析服务	全基因组甲基化测序分析服务	16S r DNA 测序分析服务
DNA 重测序分析服务	Small RNA 测序分析服务	服务	宏基因组测序分析服务
外显子组测序分析服务	LncRNA 测序分析服务	MeDIP-Seq 分析服务	宏转录组测序分析服务
目标区段捕获测序分析服务		ChIP-Seq 分析服务	PacBio RSII 测序分析服务
		RIP-Seq 分析服务	

样品处理平台

完整 RNA 或 DNA 样品的提取和纯化，是进行后续实验服务的前提。无论是微阵列芯片、新一代测序实验服务，还是文库构建、Real-Time PCR 实验服务，都需要由高质量的 RNA 或 DNA 样品来完成。

而对于 RNA 的处理更是我进行表达谱芯片实验服务的关键，由于 RNA 酶无处不在，致使 RNA 随时都有可能降解，因此严格规范的实验操作以及良好的实验环境就成为决定 RNA 质量的关键因素。

SBC 每年都要完成数以万计的样本抽提，其中包含各个种类的样本：如植物的组织样本，动物的组织、细胞、血液样本，微生物样本等等。十年来的实践操作为 SBC 积累了丰富的核酸抽提及质检经验。



为了降低导致样品污染、降解的可能性，上海伯豪对样本处理的每个环节都按照标准严格执行。SBC 从美国引进了无 RNA 酶污染的组织粉碎系统，每份要粉碎的样本均使用一次性的 Ttbag，并配套以一次性无酶的匀浆头和匀浆管，一系列高新技术产品确保了样本的安全。

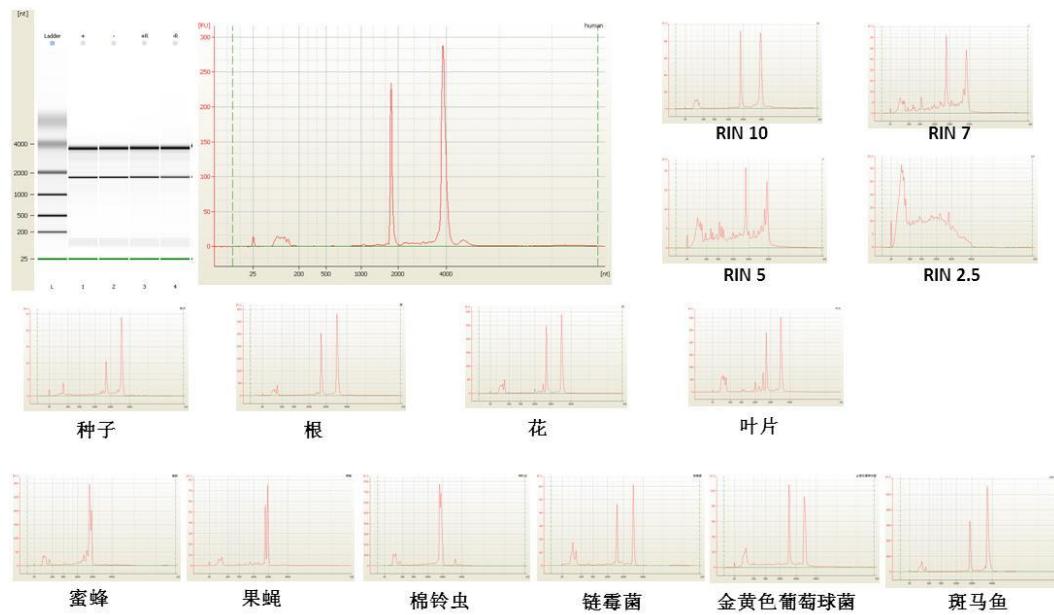
同时，平台采用了 Agilent 2100 Lab-on-a-chip 核酸质检系统进行样品质检工作。该系统仅需微量的 RNA 样品，就能将样品的浓度、质量等信息精确测定，更能将其所测得的数据信息模拟成凝胶电泳图，给人以清晰直观的效果。SBC 的所有样品均运用 Agilent2100 系统质检，所有数据信息均由仪器测得计算后得出，相比起通过肉眼判断的凝胶电泳图片结果更具有科学说服力。



Agilent 2100 分析仪



Nanodrop® 1000 紫外分光光度



RNA 质检部分实例

第三章 微阵列芯片服务



上海伯豪微阵列芯片服务一览

DNA		RNA		表观遗传学
SNP 芯片	CNV 芯片	表达谱芯片	miRNA 芯片	DNA 甲基化芯片
用于研究各种生物 DNA 序列的差异和单核苷酸多态性，同时研究这些差异对疾病、诊断、预后的影响	用于研究基因拷贝数目的变化，从而明确疾病相关遗传基因、阐明其致病机制	用于大规模分析一定的生物对象在特定生物过程中基因表达变化的全面信息	用于研究小 RNA 的功能及其在动植物的生长发育、细胞分化和凋亡以及人类疾病的发生等过程中作用	在全基因组水平研究 DNA 甲基化在发育、X 染色体失活、衰老及人类疾病中的变化
Affymetrix SNP 芯片	Agilent CGH 芯片	Affymetrix 基因表达谱芯片	Agilent miRNA 芯片	Illumina 甲基化芯片
Illumina SNP 芯片	Affymetrix 芯片	Agilent 基因表达谱芯片	Affymetrix miRNA 芯片	
		Illumina 基因表达谱芯片		

(一) SNP 芯片服务

人类基因组由 30 亿个核苷酸组成，它们携带着人类的遗传信息并决定人体的生理特征。人类基因组序列的 0.1%-0.2% 在人种、人群和个体之间存在 DNA 序列的差异，即单核苷酸多态性 (SNP, single nucleotide polymorphism)。这些单核苷酸多态性可引起不同的遗传性状，如 ABO 血型位点标记，白细胞 HLA 位点标记和个体药物代谢差异等。了解这些 DNA 序列的差异和单核苷酸多态性以及这些差异所表现的意义将有助于探索动植物群体遗传结构及遗传多态性，并且在医学领域，将会给疾病的预测、诊断、预后和预防带来革命性的变化。

SBC 利用平台优势，为广大科研工作者提供基于 Affymetrix SNP 芯片平台和 Illumina 微珠基因分型分析系统的 SNP 芯片技术服务。

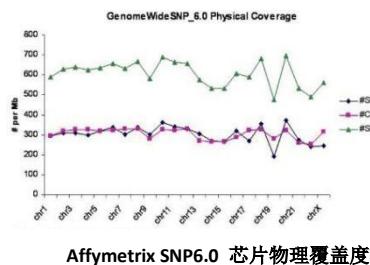
Affymetrix SNP 芯片服务

芯片推荐：Genome-Wide Human SNP Array 6.0

芯片介绍：Affymetrix Genome-Wide Human SNP 6.0 芯片产品涵盖超过 1,800,000 个遗传变异标志物：包括超过 906,600 个 SNP 和超过 946,000 个用于检测拷贝数变化 (CNV, Copy Number Variation) 的探针：

- ✓ 482,000 个 SNP 来自于前代产品 500K 和 SNP5.0 芯片；
- ✓ 424,000 个 SNP 包括国际 HapMap 计划中的标签 SNP, X, Y 染色体和线粒体上具有代表性的 SNP，以及来自于重组热点区域和 500K 芯片设计完成后新加入 dbSNP 数据库的 SNP；
- ✓ 202,000 个用于检测 5,677 个已知拷贝数变异区域的探针，这些区域来源于多基因组变异数数据库。该数据库中 3,182 个非重叠片断区域平均用 61 个探针来检测；

- ✓ 744,000 个探针平均分配到整个基因组上，用来发现未知的拷贝数变异区域。



芯片推荐：Affymetrix Mouse Diversity Genotyping Array

芯片介绍：这款芯片与 Human SNP 6.0 芯片类似，既包含 SNP 探针，也包含 CNV 探针。一共涵盖 623,000 个 SNP 和 916,000 个 CNV 位点。

SBC 也可为您提供以下物种的 Affymetrix SNP 芯片服务：犬、稻、生菜、胡椒、拟南芥、玫瑰、草莓、鳟鱼

基于 GeneTitan 平台的 Axiom 基因分型芯片

Affymetrix Axiom 平台支持全基因组关联和候选基因关联相关研究。研究者将不再受困于固定内容的芯片，全新的基因分型平台采用酶介导、单碱基测序步骤来确保实验的低成本、分型结果的高重复性和流程的全自动化。该产品参照 HapMap 和千人基因组计划(1KGP)数据库设计探针，与 GeneTitan™ 配套使用。

芯片推荐：Axiom® Biobank Genotyping Arrays

芯片介绍：能经济地对大样品集合进行基因分型，如生物样本库、基因组中心及核心实验室筛查的样品。

芯片有目录版本和定制版本可供选择。

- ✓ 功能强大、适合主要人类族群的高覆盖度分析的全基因组关联研究（GWAS）分型模块；
- ✓ 适合外显子组分析 Axiom® Exome 外显子分型模块；
- ✓ 吸收、分布、代谢和排泄(absorption, distribution, metabolism, and excretion ,ADME)及其他临床相关的标记；
- ✓ 新发现的功能丧失变异，包括最近外显子组测序项目中发现的序列插入和缺失。

芯片推荐：Axiom® miRNA Target Site Genotyping Arrays

芯片介绍：一款人类全基因组范围评估 miRNA 靶基因位点的高密度基因分型工具。该芯片可检测在 miRNA 级联放大反应中影响基因功能的 238,000 个 SNP 和插入缺失位点，包括 miRNA、它们的 mRNA 靶基因位点以及沉默机制相关的位点。其中 80% 的位点是在其他相关基因分型芯片上所没有的。因此也可以用它来补充一些 GWAS 研究和基因调控相关的研究

芯片推荐：Axiom® Genome-Wide Population-Optimized Human Arrays

芯片介绍：最大限度覆盖不同人种的稀有和常见变异，关注与疾病密切相关以及功能相关的 SNP，包含 Axiom® 中国人基因分型芯片(CHB 1 和 CHB 2)。

芯片推荐：人 Axiom® Exome Genotyping Arrays

芯片介绍：最完整的外显子组基因分型芯片，可筛查功能变异，有助于更好的发现新的人类编码区 SNP 以及插入缺失。



芯片推荐：牛 Axiom® Genome-Wide BOS 1 Bovine Array

芯片介绍：此款芯片是与牛基因分型领域的重要科学家合作开发的，实现了准确的遗传价值评估、全基因组关联研究(GWAS)、生物多样性和连锁不平衡(LD)研究。此芯片旨在最大限度提高Bos taurus和Bos indicus品种的遗传和物理SNP覆盖度。芯片超过640,000个SNP标记，代表了来自Affymetrix牛基因组数据库的大约300万个SNP的遗传多样性。

芯片推荐：水牛 Axiom® Buffalo Genotyping Array

芯片介绍：Axiom® Buffalo Genotyping Array带来了多个物种和品系的水牛（*Bubalus bubalis bubalis*和*Bubalus bubalis carabanensis*）中多态性SNP的最高全基因组范围覆盖，这些水牛是意大利、巴西和南亚最重要的肉牛和奶牛之一。此芯片是目前市场上唯一一款高密度的水牛基因分型工具，含有90,000个常见和稀有标记。。

芯片推荐：鸡 Axiom® Genome-Wide Chicken Genotyping Array

芯片介绍：此芯片是第一款商业化的高密度鸡基因分型芯片。此芯片包含 580,000 常见和稀有的变异标记，这些标记存在于商业的蛋鸡和肉鸡以及非商业的远交群中。它适用于蛋鸡和肉鸡的基因组选择研究、全基因组关联研究、高分辨率遗传图谱、孟德尔性状作图及选择进化研究。

芯片推荐：玉米 Axiom® Maize Genotyping Array

芯片介绍：玉米是一个二倍体物种（n=10），其基因组大小为 2.5Gb。不同品种的玉米基因组表现出重复DNA 和高的基因组与结构多样性。玉米还表现出 50-300bp 的极低连锁不平衡(LD)，从而需要非常高密度的芯片来研究单倍型结构和基因组多样性。Axiom Maize Genotyping Array 带来了极高密度的（600K）全基因组 SNP 覆盖。这些多态性 SNP 在 285 个硬粒型(Flint) 和马齿型(Dent) 玉米中检测到，代表了 161 个来自世界各地的不同品种。

芯片推荐：三文鱼 Axiom® Salmon Genotyping Array

芯片介绍：此款芯片由 Affymetrix 联合苏格兰的科学家共同开发，包含 130,000 个 SNP 标记。芯片上的高密度标记确保了三文鱼基因组的广泛覆盖，又避免了重要区域的代表性不足。此芯片适用于三文鱼复杂性状研究、分子育种、选择进化研究。

芯片推荐：大豆 Axiom® Soybean Genotyping Array

芯片介绍：此款芯片也称为 Axiom® SoyaSNP 96-Array Plate，共包括 180,961 个标记，适用于对栽培和野生的大豆品种进行基因分型。它是全基因组关联研究和高密度遗传作图的理想选择，可确定野生及栽培大豆种质中的重要适应性基因。

芯片推荐：小麦 Axiom Wheat HD Genotyping Array & Breeder's Genotyping Array

芯片介绍：小麦的基因分型芯片是通过 Affymetrix 专家设计项目与英国布里斯托大学合同设计的。芯片有两种：高密度(HD)的芯片包含 817,000 个 SNP 位点，而 Breeder's 基因分型芯片包含 35,000 个 SNP 位点。这些芯片上的标记都可以在 CerealsDB 网站 (<http://www.cerealsdb.uk.net/>) 上查询。

SBC 也可协助研究者根据实验需求，定制 Axiom® myDesign™ Genotyping Arrays。



Illumina SNP 芯片服务

Omni 家族芯片

Omni 芯片家族自 2009 年上市以来，已经提供了一系列灵活的芯片，它们带来了精选自国际 HapMap 计划和千人基因组计划的常见和稀有变异。有了 Omni 芯片家族，研究人员能够立即开始新一代 GWAS、基因分型和比较基因组研究。

芯片名称	样本/ 芯片	位点数	自定义 位点*	位点选择
Human OmniZhongHua-8 (推荐!)	8	SNP 位点超过 90 万	N/A	HapMap Phase 1-3 和 1KGP, 特别覆盖中国人特有的常见变 异和稀有变异
HumanOmni5-Quad (推荐!)	4	SNP 位点约 430 万	50 万	HapMap Phase 1-3 和 1KGP(MAF > 1%)
HumanOmni2.5-8	8	SNP 位点约 250 万	20 万	1KGP (MAF > 2.5%)
HumanOmni2.5S	8	SNP 位点约 250 万	50 万	1KGP (MAF > 1%)
HumanOmniExpress-24	24	SNP 位点超过 70 万	5 万	HapMap Phase 1-3(MAF > 5%)
HumanOmniExpress Exome BeadChip	8	SNP 位点超过 90 万 (MAF > 5%), 外显子 序列超过 24 万	3 万	Illumina 人外显子组芯片在原有 Omni 家族芯片基础上，增加了 精选自 12,000 多个个体的外显 子组和全基因组序列。外显子内 容包含了约 250,000 个标志物， 它们代表了不同群体，包括欧 洲、非洲、中国和西班牙个体以 及多种常见病，如 II 型糖尿病、 癌症、代谢及精神疾病等，以支 持各种旨在发现功能相关的疾 病关联的研究类型。此外，研究 人员还可在 Exome BeadChip 芯片上添加自定义的标志物，可 以方便更高密度靶定基因组的 特定区域，聚焦目标人群，或掺 入精选的疾病相关变异。
HumanOmni2.5Exome BeadChip	8	SNP 位点超过 250 万 (MAF > 2.5%), 外显 子序列超 24 万	N/A	
HumanOmni5Exome BeadChip (推荐!)	4	SNP 位点超过 450 万 (MAF > 1%) 外显子序 列超过 24 万	20 万	

注：Omni 家族芯片同款一般有两种类型：一种是芯片的 SNP 位点固定，另一种是可自定义添加 SNP 位点。自定义位点是在芯片 SNP 位点基础上，可自定义添加最大位点数，N/A 表示不能添加。

HumanOmni ZhongHua-8 BeadChip 芯片服务（推荐!）

特别覆盖中国人特有的常见和稀有变异

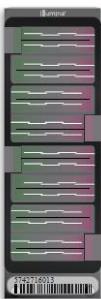


- ✓ 第一款人类种群特异的全基因组芯片，经过优化的标签 SNP 内容来自 HapMap 所有三个阶段以及千人基因组计划 (1kGP)，可用于在中国人种群中探索全新的疾病和性状关联。
- ✓ 特别覆盖中国人 81% 的常见变异 (MAF > 5%) 和 60% 的稀有变异 (MAF > 2.5%)，适合全基因组关联研究 (GWAS)。
- ✓ 采用 Illumina 专利的 BeadArray™ 技术，可提供非常高的数据质量，平均检出率 > 99%，

重复率> 99.9%。

HumanOmni5-Quad 芯片服务 (推荐!)

人类变异的高密集覆盖



- ✓ 4 个样品的 Omni5 芯片特有超过 430 万个标记物/样品。
- ✓ 特有来自千人基因组计划的新颖、常见和稀有变异。除了提供人类变异 (MAF 低至 1%) 的密集覆盖，覆盖了基因组中与疾病关联的高价值区域，包括基因区域、MHC 区域和非同义 SNP。
- ✓ 特有来自千人基因组计划的近 5000 个插入缺失和多碱基替换标记物，能帮助研究人员阐明结构变异如何影响性状和疾病。
- ✓ 可自定义的 500K 位点让研究变得更具灵活性，可以方便更高密度靶定基因组的特定区域。

动植物全基因组分型芯片解决方案

芯片名称/物种	样本/ 芯片	位点数	位点来源
牛 BovineHD	24	> 77 万	与 USDA-ARS、University of Missouri 和 University of Alberta 合作开发
牛 BovineSNP50 v2	24	> 5 万	
犬 CanineHD	12	> 17 万	与 LUPA 合作开发，信息来自于 CanFam2.0 参考序列
羊 OvineSNP50	12	> 5 万	作为国际羊基因组计划的一部分，与 AgResearch, Baylor UCSC, CSIRO 和 USDA 合作开发
猪 PorcineSNP60	12	> 6 万	与包括爱荷华州立大学，伊利诺斯大学，剑桥大学和瓦格宁根大学等领先的猪类研究所合作开发，平均每隔 40Kb 就有一个 SNP 位点，可用于不同种类猪的 SNP 分型研究。
马 EquineSNP50	12	> 5 万	与国际 Equine Genome Mapping Workshop 和 Morris Animal Foundation's Equine Genome Consortium 合作开发，平均每隔 43.2Kb 就有一个 SNP 位点。
玉米 MaizeSNP50	24	> 5 万	位点选自玉米 B73 系基因组序列，与 TraitGenetics、INRA 和 Syngenta 合作开发，平均每隔 40Kb 就有一个 SNP 位点。

SBC 同时可提供 Illumina 合作者基于 Infinium 技术研发的苹果, 梨, 樱桃, 番茄, 马铃薯, 山羊 SNP 芯片等。

Illumina 定制 SNP 芯片

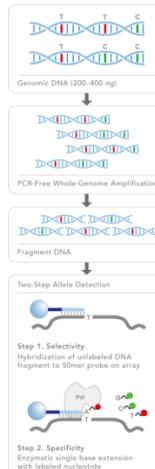
iSelect 定制芯片平台

Infinium 采用了 Illumina 专利的 InfiniumHD 检测技术，为每一个 SNP 位点设计一个微珠，反应通过 50 个碱基的探针和样品杂交，并通过末端单碱基延伸完成，末端双色标记(Cy3,Cy5)来检测该位点的等位基因情况。充分保证了检测的高灵敏性和特异性以及准确性。反应无需 PCR，不依赖于限制性内切酶，无需电

泳，保证了高质量的数据。

- ✓ 高通量的 Infinium 检测技术，检出率高；
- ✓ 除了 SNP 外，还可以针对拷贝数变异（CNV）进行设计；
- ✓ 任意物种、任意位点随意组合。

定制产品	位点数量	样品/芯片
iSelect 3–90k BeadChip	3–90k	24
iSelect 90–250k BeadChip	90–250k	12
iSelect 250–1M BeadChip	250–1M	4

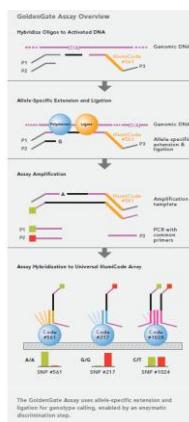


iSelect 技术原理

GoldenGate 定制芯片平台

Illumina 专利的 GoldenGate 技术是理想的中等通量自定义芯片技术，不仅可以定制 SNP 基因分型和拷贝数变异（CNV）检测芯片，还可以定制 DNA 甲基化芯片。

- ✓ 低成本的中等通量基因分型解决方案
- ✓ 方法经典可靠，为 HapMap 计划主要应用平台
- ✓ 数据质量高，平均响应率 > 99%
- ✓ 定制灵活



GoldenGate 技术原理



科研工作者除可利用 HapMap, dbSNP 等数据库挑选 SNP 位点外, Illumina 可免费对客户提供的 rs 号, 基因名称, 序列或要研究的区域的位置进行生物信息学分析, 挑选 SNP 位点并进行位点实验成功率评分, 评分结果返回用户处后和用户沟通一起来确定要进行研究的 SNP 位点, 以合成探针组 (OPA), 进行基因分型的研究。一个样品可以同时进行 96, 384, 768、1536 或 3072 个 SNP 位点的研究。

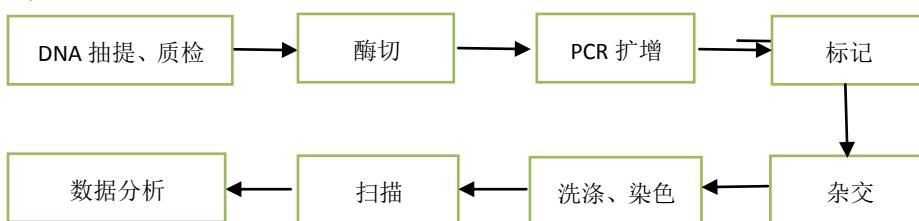
目前, SBC 已经完成的 GoldenGate 定制 SNP 芯片项目物种有: 人, 大豆、鱼等。

服务相关(以 Affymetrix SNP6.0 芯片为例)

样品要求

1. 样品纯度: OD 260/280 值应在 1.7~2.0 之间, A260/A230 > 1.5; RNA 应该去除干净; 不含有其它个体或其它物种的 DNA 污染。
2. 样品浓度: 浓度不低于 55ng/ μ l;
3. 样品总量: 每个样品总量不少于 2 μ g 。
4. 样品溶剂: 溶解在 Reduced TE (10mM Tris, pH 8.0, 0.1mM EDTA) 中。
5. 样品运输: DNA 低温运输 (-20°C); 在运输过程中请用 parafilm 将管口密封好, 以防出现污染;

服务流程



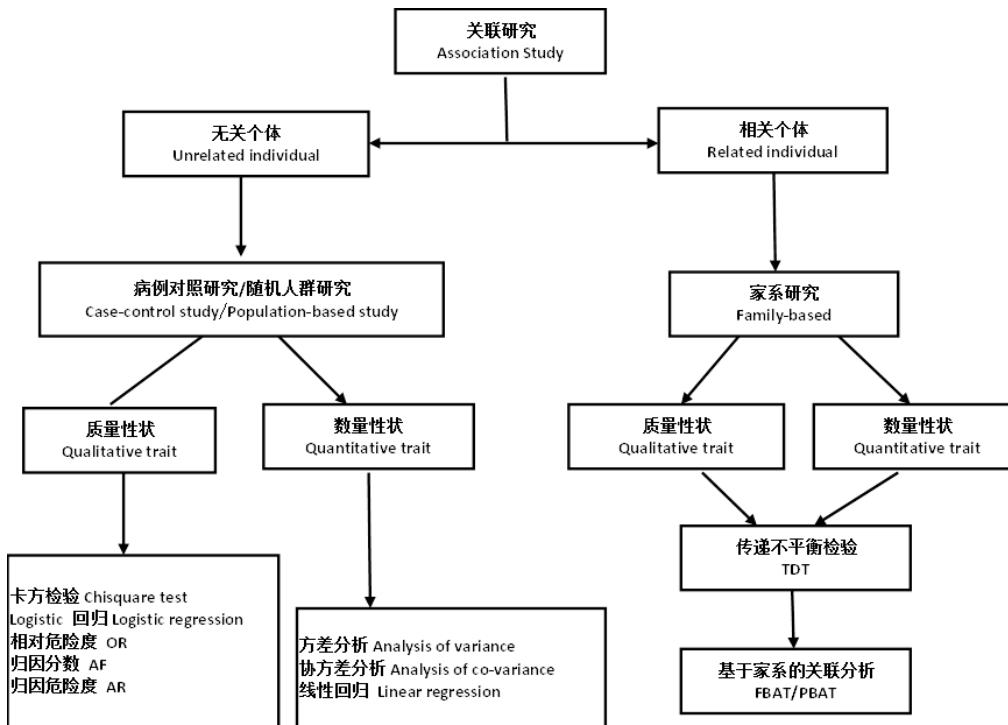
数据分析

遗传遗传图谱构建	全基因关联分析	家系连锁分析	群体进化研究
BinMap/遗传图谱构建	群体结构分析	连锁分析	群体进化树
QTL 定位	连锁不平衡分析	单倍型分析	群体结构分析
	全基因组关联分析		连锁不平衡分析
			多态性分析
			选择分析

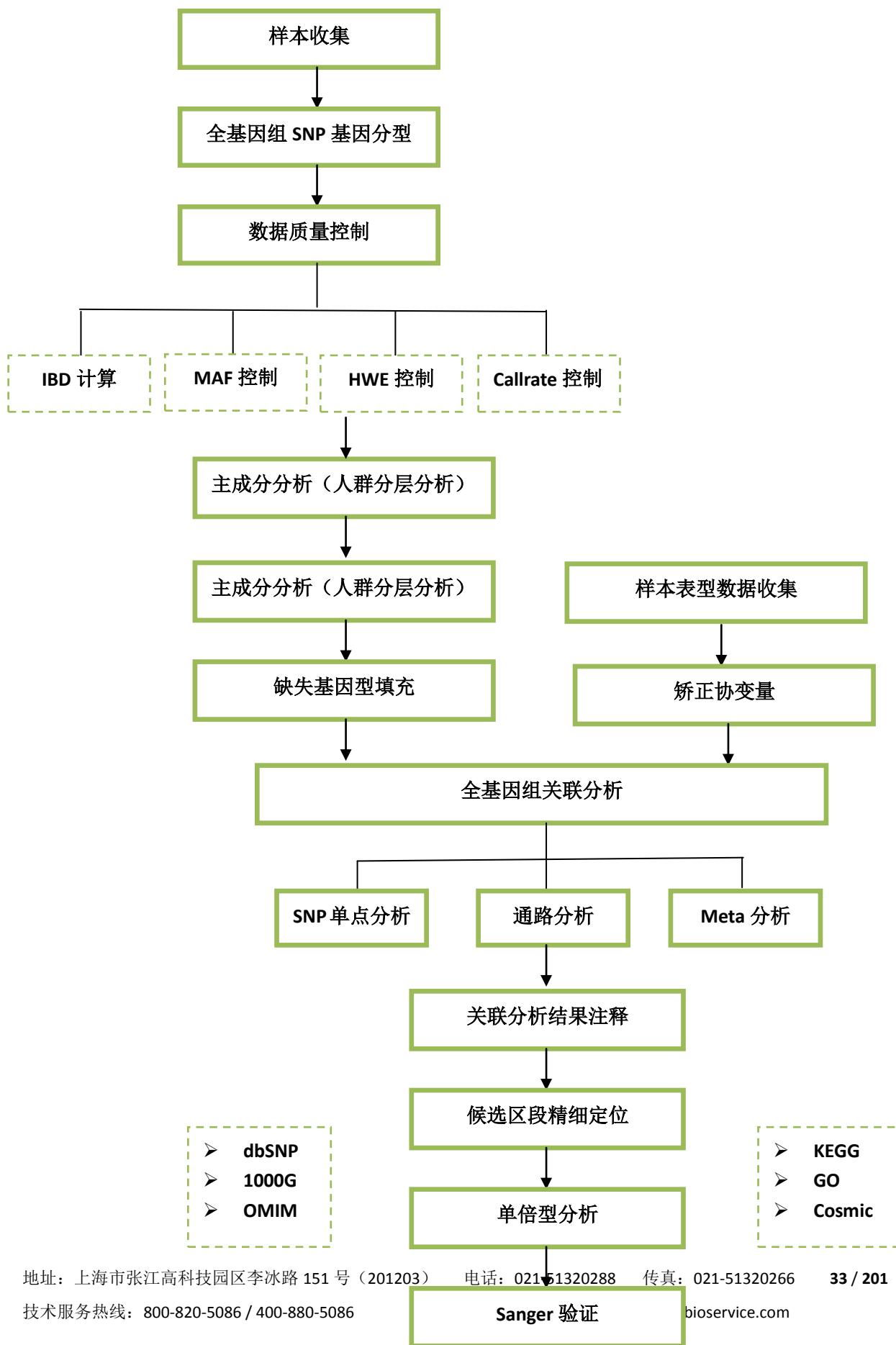
全基因组关联分析

全基因组关联研究（Genome-wide association study, GWAS）是用来检测全基因组范围的遗传变异与可观测的性状之间的遗传关联的一种策略。基于高密度 SNP (Single nucleotide polymorphism) 标记的全基因组关联分析是近几年提出的复杂性状功能基因鉴定的新策略，其基本思想是基于连锁不平衡的原理，直接检测基因本身或基因附件的微小区域 (<0.1 cM) 的 SNP 标记与复杂性状表型信息的关联来实现目标性状或疾病的精细定位。

GWAS 的统计分析依据研究设计不同可采用不同的分析方法，目前常规的分析方法如下图所示：



技术路线



1. 数据质量控制

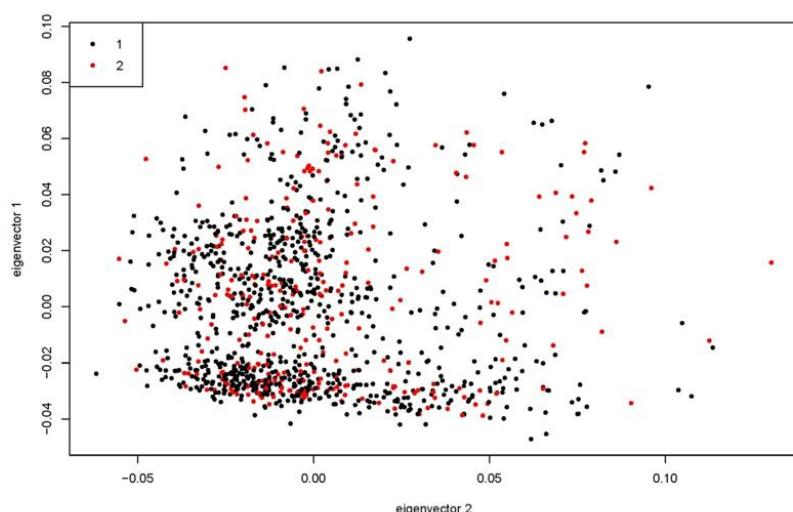
从质量方面对数据进行预处理，保证分析数据的准确性及结果可靠性。

- ✓ 去除最小基因频率（Minor allele frequency, MAF）小于 0.05 的 SNP 位点
- ✓ 去除检出率（Callrate）小于 0.90 的 SNP 位点
- ✓ 去除哈温检测中 P 值小于 10E-3 的 SNP 位点
- ✓ 去除检出率小于 0.90 的个体
- ✓ 去除重复样本以及有一级亲缘关系的样本

2. 主成分分析

主成分分析（Principal Component Analysis, PCA）的原理是找到数据方差最大的两个或者三个主成分(就是向量)，将数据投影在这些主成分上，以达到降维的目的，通过图像上的点之间的相互距离来显示样品之间的相似度。考察样品的分布情况，验证实验设计的合理性，生物学重复样品的均一性（至少 2 组数据）。

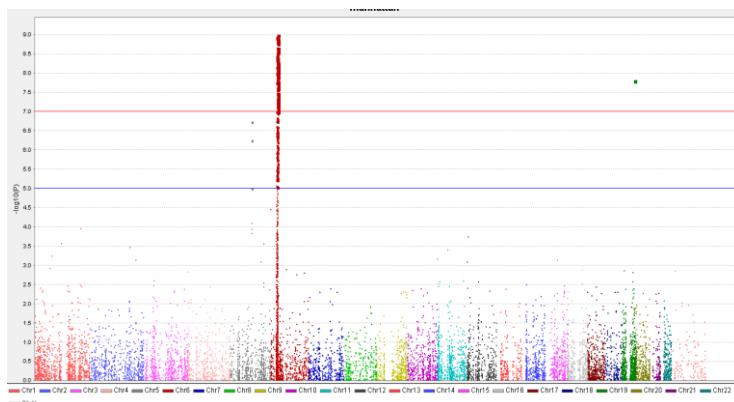
将不同区域的样本进行 PCA 主成分分析，观察样本的聚类情况，探索地域差异与遗传差异对群体差异的影响。选择使用合适方法对关联分析的结果进行校正，减少人群分层带来的假阳性



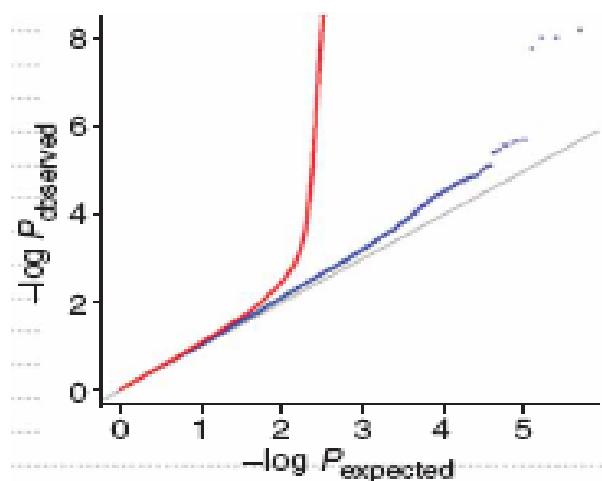
PCA 分析示意图

3. 关联分析

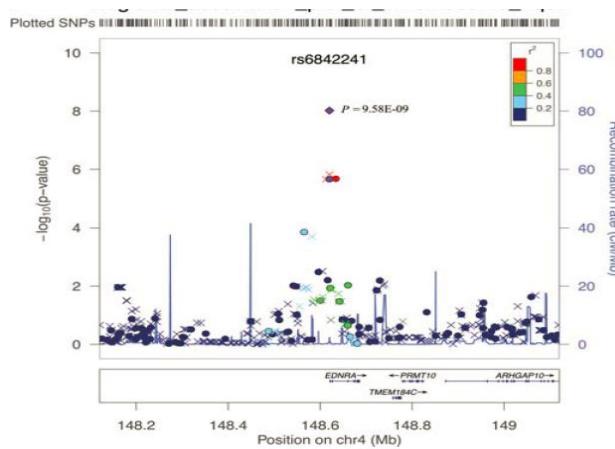
针对不同的疾病分型或性状表型，采用合适的模型进行关联分析，分析方法包括：卡方检验、Fisher 精确性检验、Cochran-Armitage 趋势检验及传递不平衡检验。

Manhattan plot

X-轴为基因组坐标，Y-轴为每个单核苷酸多态性的关联 P-值的负对数。

QQ plot**4. 区域关联分析图**

以 4 号染色体上的某一区段为例



5. 显著区段连锁不平衡分析

连锁不平衡 (linkage disequilibrium) 是指基因组中不同基因座间存在的非随机关联，即不同基因座的非等位基因间的非随机组合。LD Plot 表示该基因所有 snp 的连锁情况，各个方块的颜色由浅至深（白-红），表示连锁程度由低到高，深红色表示完全连锁。



6. 显著位点注释

基于各类公共数据库（1000G、dbSNP、Cosmic、OMIM、KEGG/GO）对显著位点进行注释，综合注释信息及关联分析结果，在全基因组范围内筛选与疾病关联的位点，并计算这些位点等位基因频数，危险等位基因频数，OR 值以及矫正后的 p-value 值。

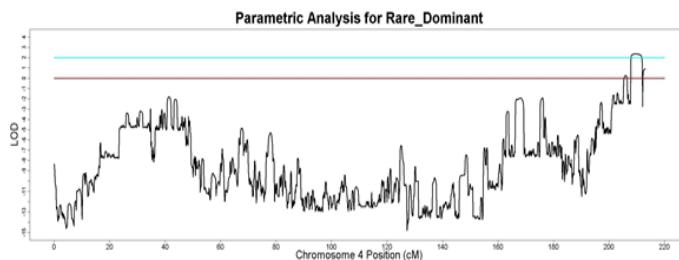
家系连锁分析

在家系中，位于同一条染色体上的两个基因座（QTL 与遗传标记）在减数分裂的过程中会发生交换和重组，染色体上的两个基因组相距越远，发生重组的几率越高，两个基因座在一起传给后代的机会越少。因此，由标记与 QTL 间的重组率可估算出两者间的距离及连锁程度。家系连锁分析就是通过寻找与 QTL 紧密连锁的某一标记，从而确定该基因在染色体上的粗略定位。

进行连锁分析所需要的条件包括，完整且正确的家系系谱信息及明确的患病与否，高密度的遗传标记。

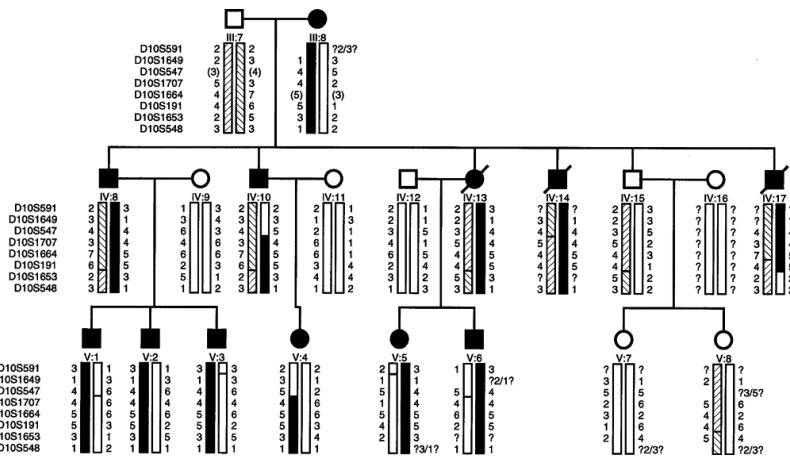
LOD plot

LOD(log odds score)，优势对数记分法。是根据遗传标志与致病基因的连锁，和在家系中的重组值，即两者之间的遗传距离，得出两者连锁的似然性比例。Lod 值为 0，意味着连锁假设与不连锁假设的可能性相等；Lod 值为正值，有利于连锁；Lod 值为负值，表示有一定重组率的连锁。显著的域值是 +3 和 +2。Lod=+3 时，连锁的概率为 95%。



单倍型分析

单倍型又称单体型，是 tagSNP 的 call 在染色单体上的线性排列，单倍型分析的目的是看是否所有患病个体都继承了同样的单体型。在下图中，患病个体画框的里面单体型是一样的。



伯豪案例

应用案例一：Affymetrix SNP 6.0 芯片揭示先天多毛症遗传缺失

先天性全身多毛症（congenital generalized hypertrichosis, CGHT）是一种极其罕见的体毛生长异常。中国医学科学院-北京协和医学院张学教授领导的联合团队首先收集了3个常染色体显性遗传的中国人CGHT家系和1名患重度CGHT伴牙龈增生的散发个体；继之通过全基因组扫描将CGHT的遗传位点定位到染色体17q24.2-q24.3区域；进一步应用全基因组拷贝数检测、荧光定量PCR和荧光原位杂交（FISH）等多项先进技术，确定染色体17q24.2-q24.3拷贝数突变对无论有无牙龈增生的CGHT都起到决定性的作用。该课题芯片相关实验部分在生物芯片上海国家工程研究中心完成。



原文出处：Copy-number mutations on chromosome 17q24.2-q24.3 in congenital generalized hypertrichosis terminalis with or without gingival hyperplasia. Am J Hum Genet. 2009; 84(6):807-13.
(IF12.303) (客户发表文章) 使用服务种类：Affymetrix Genome-Wide Human SNP Array 6.0 服务

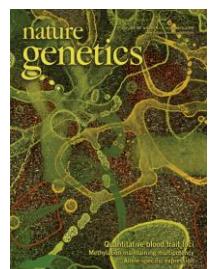
应用案例二：Affymetrix GeneChip Human Mapping 100K 芯片揭示亚洲人群迁徙规律由人类基因组机构（Human Genome Organization, HUGO）泛亚太地区 SNP 联盟（Pan-Asian SNP Consortium）发起的一项大型研究首次通过 SNP 分析获得了亚洲人群 SNP 遗传多元化图谱，并以此分析了亚洲人群的历史移民路线。生物芯片上海国家工程研究中心作为作者单位之一承担了中国样品的 SNP 实验。研究表明，亚洲大陆的人口主要是通过某单一的迁徙活动而形成的。



原文出处：Mapping Human Genetic Diversity in Asia. Science. 2009, 326(5959):1541-5. (IF29.747) 使用服务种类：Affymetrix GeneChip Human Mapping100KArray 服务

应用案例三：Illumina SNP 芯片发现汉族人群红斑狼疮易感基因

安徽医科大学第一附属医院张学军教授领衔的研究团队通过对12000多例中国汉族红斑狼疮患者和健康对照样本进行研究，发现了5个与汉族人群发病密切相关的易感基因ETS1、IKZF1、RASGRP3、SLC15A4 和 TNIP1，并确定了4个新的易感位点；研究同时验证了在欧洲人中发现的7个易感基因在汉族人中同样存在。



原文出处：Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nature Genetics. 2009, 41(11):1234-7. (IF34.284) 使用服务种类：Illumina Human 610-Quad BeadChips 服务



(二) CytoScan™ HD 芯片服务

染色体水平的缺失、重复等变化是细胞遗传学研究的热点。以往科学家运用核型分析、荧光原位杂交(FISH)等技术已获得染色体结构变异的许多可贵进展，但这些传统技术分辨率低、且只能定性分析。Affymetrix推出CytoScan™ HD细胞遗传学芯片解决了上述问题，能够在全基因组范围内高分辨率的检测DNA拷贝数的变异。

产品亮点

- ✓ 目前密度最高的细胞遗传学芯片，含有超过270万种探针，同时具备CNV探针和SNP探针，包括75万个SNP探针和195万个拷贝数探针。CNV探针平均覆盖全基因组，对于>400kb的基因组结构变异检出率达99%。SNP探针进行SNP分型的准确率高于99%。
- ✓ 不但能检测基因缺失、重复，还能检测杂合性缺失(LOH, loss of heterozygosity)和单亲源二体(UPD, uniparental isodisomy)
- ✓ 基因覆盖全面。百分百覆盖ISCA constitutional genes、Sanger cancer genes，同时覆盖96%的RefSeq genes。

芯片参数

Markers for copy number analysis	数目
CNV探针数	2,696,550
non-polymorphic探针数	1,953,246
SNP探针数	743,304
“genotype-able” SNP探针数	749,157
常染色体探针数	2,491,915
拟常染色体探针数	4,624
基因内探针数	1,410,535
基因间探针数	1,286,015

基因覆盖情况

数据库	基因覆盖情况
ISCA constitutional genes (340)	100%
Cancer genes (526)	100%
OMIM Morbid genes (2,640)	98%
X chromosome OMIM Morbid genes (177)	100%
RefSeq genes (36,121)	96%

服务相关

样品要求

1. 样品纯度：OD 260/280值应在1.7~2.0之间，A260/A230>1.5；RNA应该去除干净；不含有其它个体或其它物种的DNA污染。
2. 样品浓度：浓度不低于55ng/ μ l；
3. 样品总量：每个样品总量不少于1 μ g。
4. 样品溶剂：溶解在Reduced TE(10mM Tris, pH 8.0, 0.1mM EDTA)中。

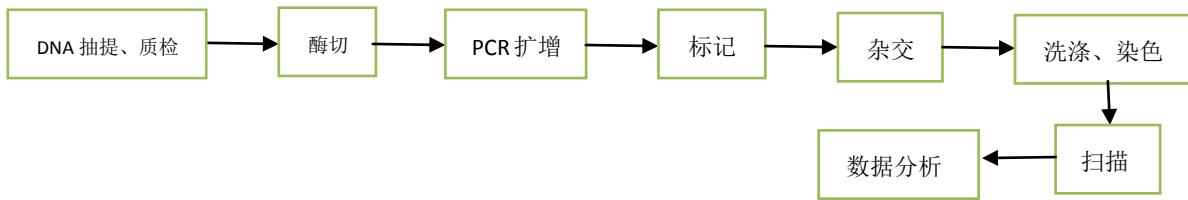
地址：上海市张江高科技园区李冰路151号（201203） 电话：021-51320288 传真：021-51320266 38 / 201

技术服务热线：800-820-5086 / 400-880-5086

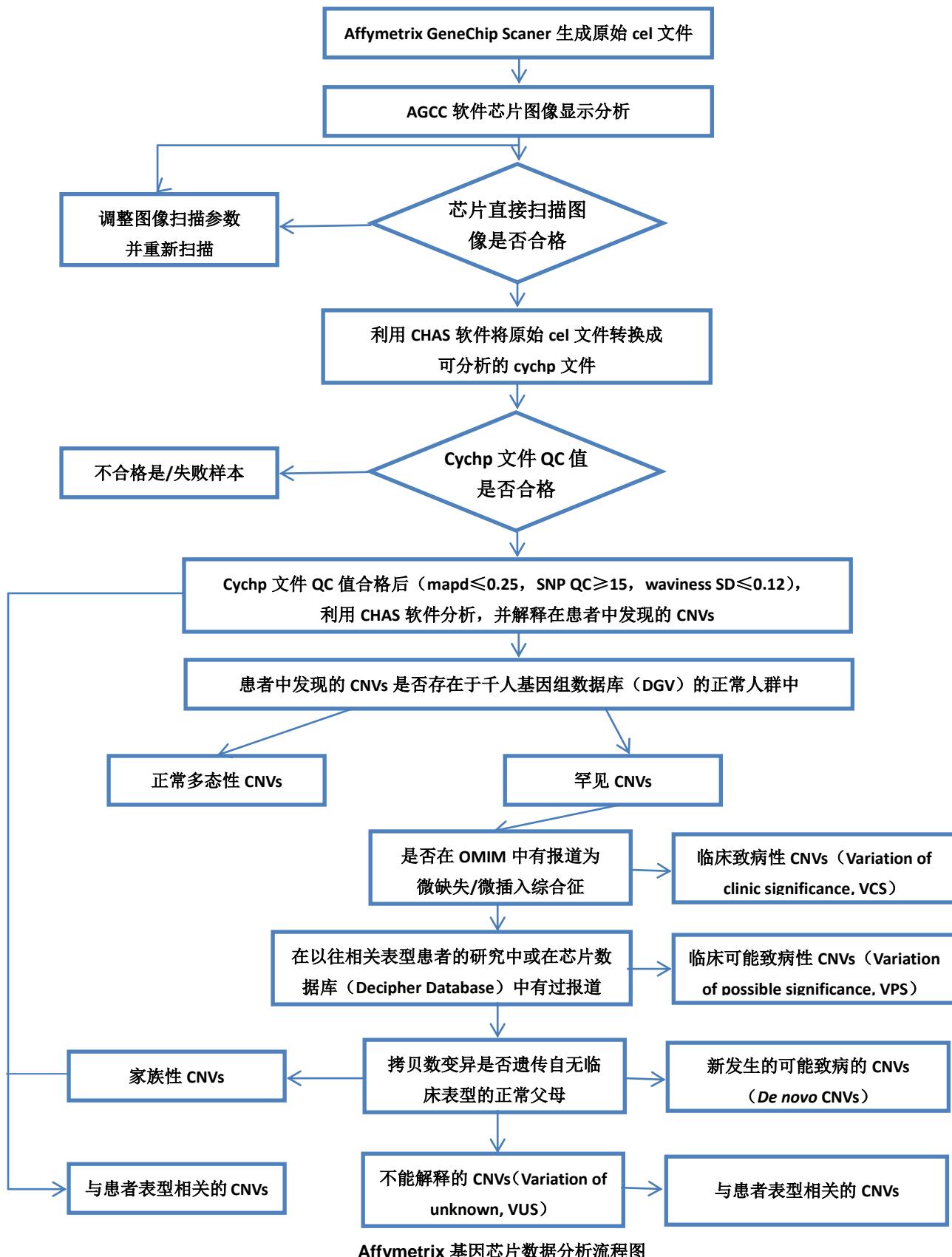
技术服务网站：<http://www.ebioservice.com>

5. 样品运输：DNA 低温运输 (-20°C)；在运输过程中请用 parafilm 将管口密封好，以防出现污染；

服务流程



数据分析流程





数据分析说明:

Affymetrix Genechip Scaner 产生的原始数据 Cel 文件，经过 Affymetrix 自带软件 AGCC(Affymetrix GeneChip Convert Console)将 Cel 文件转换成 Cychp 文件，Cychp 文件用 Affymetrix 自带软件 CHAS(Chromosome Analysis Suite)分析。通过和各种数据库结合，每个患者 CNV 的性质根据一系列完整的流程最终分成四类：临床致病性 CNV (Variation of clinic significance, VCS)、临床可能致病性 (Variation of possible significance, VPS)、不能解释(Variation of unknown,VUS)以及良性 CNV (Benign CNV)。

应用案例

基因芯片技术揭示慢性淋巴细胞白血病与 10q24.32 复发性等位基因缺失有关

2012 年 3 月美国遗传医学学院会议上，犹他大学与 ARUP 公司诊断实验室做了联合使用 Affymetrix 基因芯片分析揭示慢性淋巴细胞白血病与 10q24.32 复发性等位基因缺失有关的报告。慢性淋巴细胞白血病具有非常多样化的临床表型。而通过染色体的畸变类型可以判断病人属于哪种 CLL 亚型，评估预后以及检测残留病灶，所以正确识别相关的畸变对于病人管理具有非常重要的意义。此项研究中，10 例 CLL 样本全部由 CytoScan HD (Affymetrix) 芯片检测 CNVs。其中两个独立的样本中都发现了 10q24.32 的微观间质性缺失。这两个样本中检测到的缺失片段大小均在 670kb 左右。缺失区域包含 TMEM180, ACTR1A, SUFU, TRIM8, ARL3, SFXN2, C10orf26, CYP17A1, C10orf32, AS3MT 及 CNNM2 基因。而另一个案例检测到的 10q 末端缺失伴有 10q24.1 的断裂点。在这些基因当中，SUFU 是肿瘤抑制基因，TRIM8 是生长抑制基因。它们可能是该临床显著缺失区域的候选基因。

原文出处：Genomic microarray analysis of chronic lymphocytic leukemia reveals a recurrent monoallelic deletion of 10q24.32. American College of Medical Genetics (ACMG). 2012, 29 March .

(三) CGH 芯片服务

作为疾病的一项生物标志，染色体水平的缺失、扩增等变化已成为科学家研究的热点。以往科学家运用传统的染色体核型分析方法(如染色体区带染色分析, FISH 等)已获得许多染色体结构变异的可贵进展。但传统的方法实验操作繁琐，分辨率低，不能覆盖全基因组，难以提供染色体变异位点的精确定位。基于基因芯片的比较基因组杂交技术, aCGH (array-based Comparative Genomic Hybridization) 平台的出现解决了上述问题。这种平台能够支持研究人员通过微阵列准确研究与疾病有关的染色体的变化，大大加速了人类对基因组 DNA 结构变异在遗传疾病中意义的理解。

目前，SBC 可为广大科研工作者提供基于 Agilent 芯片平台以及 Affymetrix 芯片平台的服务。Agilent aCGH 的探针覆盖全基因组，在包括基因区域（内含子区、外显子区）、基因间区域以及对疾病研究极其重要的亚端粒区域（重复序列除外）的分布大致相同。具有敏感度高、精确度高、分辨率高的优点。

CGH 芯片服务

服务名称		芯片介绍
人 CGH 芯片服务	Agilent SurePrint G3 Human CGH Microarray 1x1M 服务	芯片格式为 1x1M。包含 963,029 个探针，其中重复探针(5 倍)为 1,000 个，内部质控探针为 6,685 个。探针设计参照数据库为 UCSC hg18(NCBI Build 36, March 2006)。探针间距为 2.1KB。 (该服务其他格式推荐：2x400K、4x180K、8x60k)
	Agilent Human Genome CGH Microarray 4x44K 服务	芯片格式为 4x44K，包含 42,494 个探针，其中重复探针(5 倍)为 301，内部质控探针为 2,118 个。探针设计参照数据库为 UCSC hg18 (NCBI Build 36, March 2006)。探针间距为 43 KB。(该服务其他格式推荐：



		1 x 244K、2 x 105K)
小鼠 CGH 芯片服务	Agilent Unrestricted HD-CGH Microarray ISCA v2, 2x105K 服务	芯片格式为 2 x 105K, 包含 19,647 个探针在 ISCA 区域 以及 80,757 个 backbone 探针, 内部质控探针为 4,626 个。探针设计参照数据库为 UCSC hg19 (NCBI Build 37, February 2009)。探针间距为 35KB。(该服务其他格式推荐: 4 x 44K、4 x 180K)
	Agilent SurePrint G3 Unrestricted CGH ISCA v2, 8x60K 服务	芯片格式为 8x60K, 包含 18,851 个探针在 ISCA 区域以及 40,208 个 backbone 探针, 内部质控探针为 3,886 个。探针设计参照数据库为 UCSC hg19 (NCBI Build 37, February 2009)。探针间距为 60KB。
大鼠 CGH 芯片服务	Agilent SurePrint G3 Mouse Genome CGH Microarray 1x1M 服务	芯片格式为 1x1M。包含 963,261 个探针, 其中重复探针(5 倍)为 1,000 个, 内部质控探针为 6,745 个。探针设计参照数据库为 UCSC mm9(NCBI Build 37), July 2007。探针间距为 1.8KB。(该服务其他格式推荐: 4 x 180K)
	Agilent Mouse Genome CGH Microarray Kit, 1x244K 服务	芯片格式为 1 x 244K。包含 235,402 个探针, 其中重复探针(5 倍)为 1,400 个, 内部质控探针为 5,099 个。探针设计参照数据库为 UCSC mm8 (NCBI Build 36), February 2006。探针间距为 7.8 KB。(该服务其他格式推荐: 2 x 105K)
	Agilent SurePrint G3 Rat Genome CGH Microarray 1x1M 服务	芯片格式为 1x1M。包含 962,967 个探针, 其中重复探针(5 倍)为 1,000 个, 内部质控探针为 7,038 个。探针设计参照数据库为 UCSC rn4(Baylor HGSC Build 3.4), November 2004。探针间距为 1.7KB。(该服务其他格式推荐: 4 x 180K)
	Agilent Rat Genome CGH Microarray Kit, 1x244K 服务	芯片格式为 1 x 244K。包含 235,974 个探针, 其中重复探针(5 倍)为 1,000 个, 内部质控探针为 5,492 个。探针设计参照数据库为 UCSC rn4 (Baylor HGSC Build 3.4), November 2004。探针间距为 4.9KB。(该服务其他格式推荐: 2 x 105K)

✓ Agilent 还提供其它物种的 CGH 芯片: 牛、鸡、犬、黑猩猩、恒河猴、水稻以及定制 aCGH 芯片。

CGH+SNP 芯片服务

Agilent 推出的 CGH+SNP 芯片服务在原有的 Agilent CGH 芯片的基础上, 增加了能检测 SNP 的探针, 从而在一张芯片上能同时检测 CNV (拷贝数差异) 和 Copy-neutral LOH/UPD (拷贝数平衡的杂合性缺失, 单亲源二体)。

服务名称	芯片介绍
SurePrint G3 Human Genome CGH+SNP Microarray 2x400K	芯片格式为 2x400K, 包含 292,097 个 CGH 探针, 119,091 个 SNP 探针, 其中重复探针(5 倍)为 600 个, 内部质控探针为 6,700 个, 探针设计参照数据库为 UCSC hg19(NCBI Build 37, February 2009)。探针间距为 7.2kb。
SurePrint G3 ISCA CGH+SNP Microarray Kit, 4x180K	芯片格式为 4x180K, 包含 110,712 个 CGH 探针, 59,647 个 SNP 探针, 其中重复探针(5 倍)为 600 个, 内部质控探针为 8,121 个, 探针设计参照数据库为 UCSC hg19(NCBI Build 37, February 2009)。探针间距为 25.3kb。



SurePrint G3 Cancer CGH+SNP Microarray Kit, 4x180K

芯片格式为 4x180K, 包含 110,712 个 CGH 探针, 59,647 个 SNP 探针, 其中重复探针(5 倍)为 600 个, 内部质控探针为 8,121 个, 探针设计参照数据库为 UCSC hg19(NCBI Build 37, February 2009)。探针间距为 25KB。

服务相关

样品要求

1. 样品纯度: OD 260/280 值应在 1.7~2.0 之间, A260/A230 > 1.5; RNA 应该去除干净; 不含有其它个体或其它物种的 DNA 污染;
2. 样品浓度: 浓度不低于 75ng/ μ l;
3. 样品总量: 每个样品总量不少于 3 μ g;
4. 样品溶剂: 溶解在水或 TE 中;
5. 样品运输: DNA 低温运输 (-20°C); 在运输过程中请用 parafilm 将管口密封好, 以防出现污染。

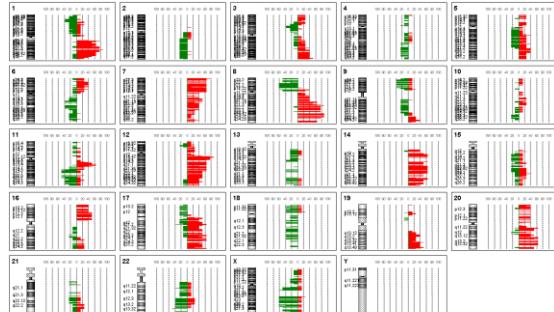
服务流程



数据分析

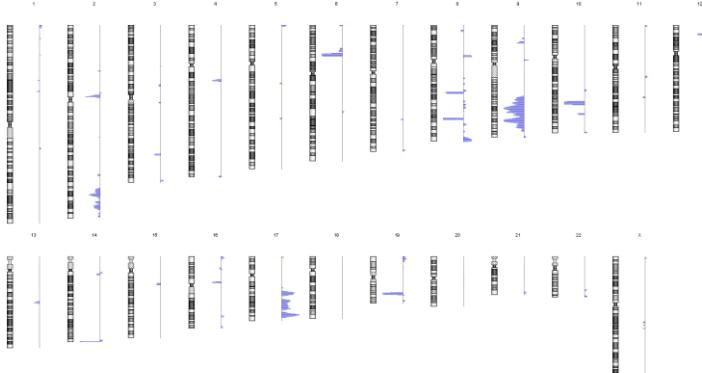
作为疾病的一项生物标志, 染色体水平的缺失、扩增等变化已成为科学家研究的热点。基于基因芯片的比较基因组杂交技术 (array-based Comparative Genomic Hybridization, aCGH), 它通过单一的一次杂交可对某一疾病有关基因组的染色体拷贝数量的变化进行检查, 大大加速了人类对基因组 DNA 结构变异在遗传疾病中意义的理解。

探针改变百分比

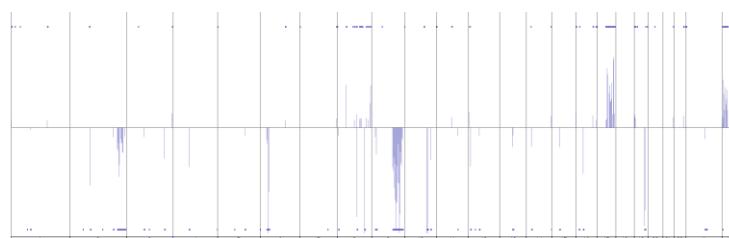


基于探针信号的 log 比值, 显示探针改变的百分比情况

变异区段作图



染色体分开的变异区段作图



染色体连在一起的变异区的作图。

应用领域

- 研究肿瘤发生过程中伴随的基因变化，如染色体的畸变等；
- 通过研究基因拷贝数目的变化，寻找新的原癌基因或抑癌基因；
- 揭示生物体抗药性的机制及代谢过程中可能发生的基因组变化；
- 发现新的生物标记，可用于临床诊断、疾病的分型等。

应用案例

应用案例一：利用 CGH 芯片寻找结肠癌基因组结构变异的共性

通过 15 例结肠癌样本与正常组织之间的比较基因组杂交芯片分析，作者发现了美国非洲裔人群结肠癌患者在基因组结构变异上的一些共性： \times 染色体一般扩增；4、8、18 号染色体一般发生缺失；ExOC4, EPHB6, GNAS, MLL3 和 TBx22 基因一般缺失；HAPLN1, ADAM29, SMAD2 和 SMAD4 基因一般扩增。

原文出处：Distinct Genetic Alterations in Colorectal Cancer. PLoS ONE. January 2010 (5) 1.

应用案例二：表达谱芯片与 CGH 芯片的联合运用

该文章对 199 例肺癌样品同时进行了基因表达谱分析 (Affymetrix U133A array) 和基因拷贝数差异分析 (Human Genome CGH Microarray 4 \times 44K)。结果显示，肺癌样品中 EGFR 的突变往往伴随着 DUSP4 表达量的降低，而 DUSP4 表达量的降低则是由 8 号染色体上的一段缺失引起的。

原文出处：An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. Oncogene. 2009 August 6; 28(31): 2773–2783.

应用案例三： aCGH 数据与第二代测序全基因组数据结合用于绘制韩国人全基因组图谱

该文章将 aCGH 实验数据和第二代测序得到的全基因组 DNA 序列数据结合在一起，从而得到更加准确的 CNV 注释。在这项研究中，首尔大学的研究团队同时采用了几种相辅相成的方法来检测 CNV，其中包含由安捷伦生产的含 2400 万个寡核苷酸探针的高分辨定制 CNV 芯片。该团队最初鉴定出 1237 个 CNV



区域, 经过进一步的数据分析, 使它降至 238 个缺失区域, 从 277 个碱基到 196900 个碱基, 总共 2.4 MB, 以及 77 个拷贝数增加区域, 共 7 MB。在这些 CNV 中, 有 148 个缺失和 33 个拷贝数增加是新的发现, 不在已公布的基因组变异数据库中(DGV)。

原文出处: A highly annotated whole-genome sequence of a Korean individual. Nature. 2009 Aug 20;460(7258):1011-5.

FFPE 样品检测

芯片推荐: Affymetrix OncoScan™ FFPE Assay Kit **NEW!**

芯片介绍: Affymetrix 推出全新的 OncoScan™ FFPE 芯片, 一种适用于高度降解的 FFPE 实体瘤样本的全基因组拷贝数分析方案。全新的 OncoScan FFPE 芯片利用 Affymetrix 独特的分子倒置探针 (MIP) 技术, 能够快速经济地分析来自 FFPE 样本的少量的高度降解的 DNA, 让实体瘤癌症分析向前迈进了一大步。该产品在一次实验中即可提供全基因组范围的拷贝数数据, 且在全基因组内大约 900 个已知的癌基因、杂合性缺失 (LOH) 上的分辨率极高, 同时可以提供临床相关的获得性突变的数据。在分析软件方面, 研发人员升级了早一代的 BioDiscovery Nexus 软件, 增加了许多转为癌症分析设计的功能, 命名为 OncoScan™ Nexus Express 软件, 它实现了几分钟内数百个样本的拷贝数分析能力, 也将随产品免费提供。

(四) 表达谱芯片服务

基因表达谱芯片可使科研工作者实现在 mRNA 水平上同时平行研究成百上千乃至上万条基因的表达关系。它与传统的研究基因表达的方法 (如差异 cDNA 文库筛选、Northern blot 和 PCR) 相比较, 可为使用者节省大量的研究经费和时间, 并获得范围更广、更具有关联性的研究结果。它的主要用途是用于大规模分析一定的生物对象在特定生物过程中 (如疾病、发育、分化、凋亡等) 基因表达变化的全面信息。目前已经广泛应用于基因表达检测、基因功能研究、新基因识别、药物靶点筛选、药物作用机理和毒理、疾病相关基因和生物标记物等众多研究领域。

SBC 可为客户提供 Affymetrix、Agilent、Illumina 等基因表达谱芯片服务。

服务推荐

芯片名称		芯片介绍
人表达谱 芯片服务	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	此款芯片包含了 47,000 个转录本, 代表了 38,500 个明晰的人类基因。数据库来源于 GeneBank、dbEST、RefSeq、UniGene database(Build 159 January 25 2003)、Washington University EST trace repository、NCBI human genome assembly(Build 31)。
	Affymetrix GeneChip Human Transcriptome Array 2.0 NEW!	最高分辨率的基因芯片, 支持所有转录本异构体的全转录组分析, 包含超过六百万条不同探针, 覆盖编码和非编码转录本。该芯片为每个外显子设计约 10 条探针, 每个外显子之间的剪接点有 4 条探针, 确保每个实验都能获得完整、准确和可重复的数据。重复性高 (批内相关系数 ≥ 0.99), 所有检测过的组织观察到的变异系数 (CV) $< 6.5\%$ 。样品类型既包括新鲜冻存样品, 也包括 FFPE 样本。
	Affymetrix GeneChip PrimeView™ Human Gene Expression Array	新版人 3' IVT 表达谱卡式芯片, 利用 GeneChip® Scanner 3000 7G system 平台, 据 2011 年 6 月份的数据库设计, 覆盖人



		类基因组中确认的注解明晰的 20000 多个基因的 36000 多个转录本，具有超强灵敏度和高度的可重复性，低 CV 值，完全覆盖带注解基因组。
	Affymetrix GeneChip Human Gene 2.0ST Array	Affymetrix 新一代全转录本芯片，在一张芯片上可以同时检测 mRNA 和 lincRNA，在其上一代 1.0 ST Array 的基础上，该芯片内容进一步丰富并更新数据库。2.0 ST Array 有更多的优点：全面覆盖转录组，可以同时检测>30,000 个编码的转录本和>11,000 个长链非编码转录本；探针设计最大程度覆盖所有外显子，可检测选择性拼接/转录本变体。每个转录本覆盖特异探针的数目中值达到 21 种，每种探针的长度为 25bp，这意味着每个转录本将被检测的碱基数中值高达 525 个；重复性高；灵敏度高。
	Agilent Whole Human Genome Oligo Microarray 4x44K Kit	Agilent 的人全基因组表达谱芯片汇总目前所了解的人类全基因组序列信息，代表了超过 41,000 人类基因和转录本。设计该产品所用的序列信息源于对 RefSeq、Goldenpath、Ensembl 和 Unigene 等知名数据库的深入研究，并与人类基因组拼接进行比对。绝大多数探针经过 Agilent 专利的实验验证程序的检验和优化。
小鼠表达谱芯片服务	Agilent SurePrint G3 Human Gene Expression 8x60K v2 Microarray Kit	新一代的安捷伦人类表达谱芯片，包含了最新公布的人类 lincRNA 以及 TUCP 目录，据 RefSeq 数据库更新了 mRNA，其探针覆盖 49,000 个转录本。
	Agilent SurePrint G3 Human Gene Expression 8x60K Kit	安捷伦人类表达谱芯片，基于 G3 平台。除了涵盖 27,958 条 Entrez Gene RNAs 外，更涵盖 7,419 条 lincRNA。不仅能检测蛋白编码 RNA 表达量的变化，也能检测非编码 RNA 的表达量变化。探针设计参照的数据库为：RefSeq Build 36.3; Ensemble Release 52; Unigene Build 216; GenBank (April 2009)。非编码 RNA 探针是和 John Rinn 实验室（隶属于麻省理工学院-哈佛大学博德研究所）共同设计的。
	Illumina HumanHT-12 v4 Expression BeadChip	该全基因组表达谱芯片包括超过 47,000 个探针，覆盖 31,000 多个注释基因，可同时检测 12 个样本。信息来源 RefSeq 数据库 (Release38), UniGene 数据库 (Build 199)。
小鼠表达谱芯片服务	Affymetrix GeneChip Mouse Genome 430 2.0 Array	Affymetrix 小鼠基因组 430 2.0 芯片涵盖了 39,000 个转录本，代表 34,000 个明晰的小鼠基因。序列信息基于 GeneBank、dbEST、RefSeq、UniGene database。
	Affymetrix GeneChip Mouse Transcriptome Assay 1.0 NEW!	最高分辨率的小鼠基因芯片，支持所有转录本异构体的全转录组分析，覆盖超过 23,000 个编码基因(其中包括超过 114,000 个转录本; 332,000 个外显子以及 215,000 个外显子簇)和超过 55,000 个非编码基因 (其中包括 101,000 个转录本; 222,000 个外显子以及 162,000 个外显子簇)。该芯片为每个外显子设计约 10 条探针，每个外显子之间的剪接点有 4 条探针，确保每个实验都能获得完整、准确和可重复的数据。重复性高，批内相关系数 ≥ 0.99 。样品类型既包括新鲜冻存样品，也包括 FFPE 样本。这种芯片的应用将加快人类疾病的临床前研究，提高了从小鼠模型研究转化到人类



		疾病研究的成功率。
	Affymetrix GeneChip Mouse Gene 2.0 ST Array	小鼠全转录本芯片，在一张芯片上可以同时检测 mRNA 和 lincRNA，可同时检测>35,000 个编码的和非编码的转录本及>2,000 个长链基因间非编码转录本
	Agilent Whole Mouse Genome Oligo Microarray, 4x44K	Agilent 小鼠全基因组表达谱芯片真正代表小鼠基因组中所有已知基因及其产生的转录本，代表了超过 41,174 个小鼠基因和转录本。设计该产品所用的序列信息源于对 UCSC、NIA、RefSeq、Ensembl、Unigene 和 RIKEN 等知名数据库的深入研究，而且绝大多数探针经过 Agilent 专利的实验验证程序的检验和优化。
	Agilent SurePrint G3 Mouse Gene Expression 8x60K Kit	安捷伦最新设计的小鼠表达谱芯片，基于 G3 平台。除了涵盖 39,430 条 Entrez Gene RNAs 外，更涵盖 16,251 条 lincRNA。不仅能检测蛋白编码 RNA 表达量的变化，也能检测非编码 RNA 的表达量变化。探针设计参照的数据库为：RefSeq Build 37；Ensembl Release 55；Unigene Build 176；GenBank (April 2009)；RIKEN 3。
	Illumina MouseRef-8 v2 Expression BeadChip	该全基因组表达谱芯片包括约 25,600 个转录本，参考数据库为：NCBI RefSeq (Build 36, Release 22)。
	Illumina MouseWG-6 v2 Expression BeadChip	该全基因组表达谱芯片包括约 45,200 个转录本，参考数据库为：NCBI RefSeq (Build 36, Release 22); RIKEN FANTOM2; MEEBO。
大鼠表达谱芯片服务	Affymetrix GeneChip Rat Genome 230 2.0 Array	Affymetrix 大鼠基因组 230 2.0 芯片包含有 31,000 个 probe sets，代表 28,000 个明晰的大鼠基因。序列信息基于 GenBank , dbEST, RefSeq 等权威数据库。
	Affymetrix GeneChip Rat Gene 2.0 ST Array	大鼠全转录本芯片，在一张芯片上可以同时检测 mRNA 和 lincRNA，可同时检测>27,000 个蛋白编码转录本及>23,500 个 Entrez 基因
	Agilent Whole Rat Genome Oligo Microarray, 4x44K	Agilent 大鼠全基因组表达谱芯片真正代表大鼠基因组中所有已知基因及其产生的转录本，代表了超过 41,012 个大鼠基因和转录本。设计该产品所用的序列信息源于对 RefSeq、Ensembl、Unigene 和 GenBank 等知名数据库的深入研究，而且绝大多数探针经过 Agilent 专利的实验验证程序的检验和优化。
	Agilent SurePrint G3 Rat Gene Expression 8x60K Kit	安捷伦最新设计的大鼠表达谱芯片，基于 G3 平台。涵盖 30,003 条 Entrez Gene RNAs。探针设计参照的数据库为：RefSeq Build 36.2; Ensembl Release 55; Unigene Build 177; GenBank (January 2009)。



SBC 提供其它物种的表达谱芯片：

Affymetrix 3' IVT Expression Array: 拟南芥、苜蓿、稻、大豆、大麦、小麦、柑橘、棉花、玉米、白杨、甘蔗、西红柿、葡萄、酵母、果蝇、秀丽隐杆线虫、牛、犬、鸡、猪、恒河猴、非洲爪蟾、热带爪蟾、斑马鱼。

Affymetrix Gene 1.0 ST Array : 拟南芥、苜蓿、稻、大豆、果蝇、秀丽隐杆线虫、牛、犬、鸡、猪、恒河猕猴、食蟹猕猴、马、猫、豚鼠、狨猴、羊、兔、恒河猴、斑胸草雀、斑马鱼、番茄。

Agilent 3' IVT Expression Array: 拟南芥、苜蓿、稻、大麦、棉花、芸苔属植物、烟草、西红柿、小麦、酵母、大肠杆菌、稻瘟菌、果蝇、秀丽隐杆线虫、蚊子、牛、犬、鸡、猪、马、恒河猴、兔、三文鱼、羊、非洲爪蟾、斑马鱼。

FFPE 样品的表达谱芯片服务

芯片推荐： Illumina Human Whole-Genome DASL HT Assay BeadChip

芯片介绍：Human Whole-Genome DASL (cDNA 介导的退火、选择、延伸和连接) Assay 是唯一一个能从低丰度或部分降解的人 RNA 样本中产生全基因组表达图谱的优化整合系统，特别适用于那些来自甲醛固定石蜡包埋(FFPE)组织的 RNA 样本。芯片覆盖约 29,000 个转录本。信息来源 RefSeq 数据库 (Build 36.2, Release 38)。

芯片推荐： Affymetrix Almac Xcel™ Array

芯片介绍：Almac Xcel™ 芯片是目前市场上唯一一款为福尔马林石蜡包埋(FFPE)样本设计和优化的 3' 端表达谱芯片。芯片包含了针对 97,000 多个转录本的探针组，其中许多是其他芯片没有的。芯片约 70% 内容与 GeneChip Human Genome U133 Plus 2.0 芯片的内容相同。还有 8% 是 U133 plus 2.0 芯片未包含的经过验证的 RefSeq 更新信息，其余 22% 来自高质量的 Almac 公司内部测序数据和验证过滤后的公开数据。

Affymetrix GeneChip Human Transcriptome 2.0 Array 和

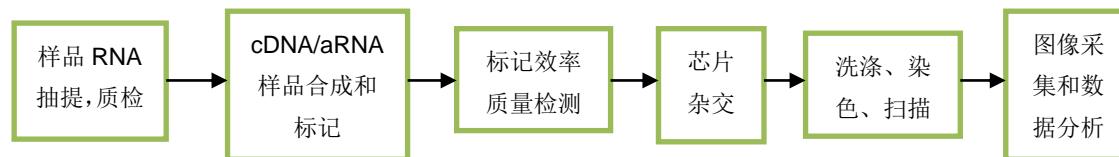
Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 这两款芯片在我们已经服务过的 FFPE 样品表达谱检测应用中也同样表现优异。

服务相关

样品要求（非 FFPE 样本）

1. 样品纯度：OD 260/280 值应在 1.9~2.2 之间；DNA 应该去除干净。
2. 样本完整性：Agilent 2100 RIN \geq 7，且 28S/18S \geq 0.7。
3. 样品浓度：最低浓度不低于 100ng/ μ l。
4. 样品总量：每个样品总量不少于 2 μ g。
5. 样品溶剂：要溶解在 RNase-free H₂O 或 RNase-free TE (pH 8.0) 中。
6. 样品运输：RNA 用冻存管保存，并用干冰或液氮运输；

服务流程

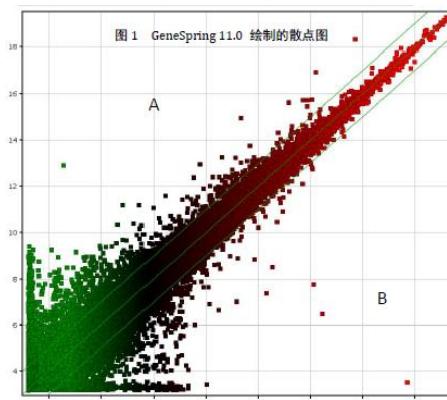


数据分析

归一化

归一化的目的是消除芯片实验过程中系统变异对基因表达水平所带来的影响，使得各个样本的数据处于相同的水平，为后续分析的有效性奠定基础。一般采用的归一化的方法为quantile normalization。

单荧光芯片的原始数据经过标准化处理，转化为 \log_2 的对数后，在一个二维直角坐标系平面中，绘制散点图(scatter plot)。芯片数据的散点图常用于评估两组数据总体分布集中趋势。散点图中每一个点代表芯片上的探针点，该点在二维平面中的位置由其X轴坐标和Y轴坐标确定。



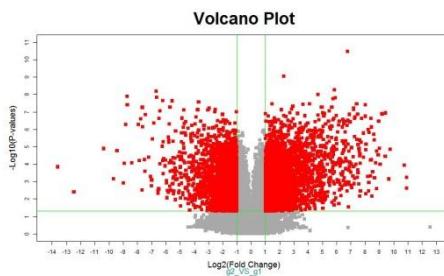
差异基因筛选

差异基因的筛选是微阵列实验数据分析的最关键一步，后续的功能分析和功能预测都是基于这些筛选出来的差异基因，因此可以毫不夸张地说，目的明确的实验设计配合合理的差异基因筛选方法，是整个实验成功的关键。我们根据您的实验设计选择针对的差异基因筛选方法，主要方法有：

- 1) 单因素两组数据统计分析, t-test
- 2) 单因素多组数据统计分析, one-way anova
- 3) 多因素数据统计分析 two(N)-way anova
- 4) SAM 分析 (R-software)
- 5) 倍数差异

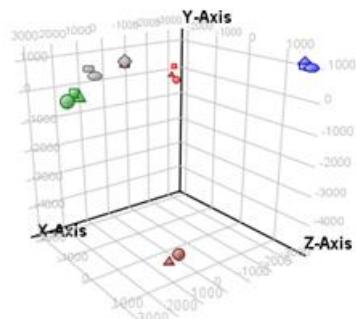
火山图

单荧光芯片数据组间T检验分析得到p-value值与Fold change值两个因素共同绘制火山图(Volcano plot)，用于显示两组样品数据的显著性差异。在火山图里，其中一个坐标显示着由t-test演算出来的负 \log of p-values，另外一个坐标则显示在两个条件比较下 \log_2 转换后的改变值。



主成分分析

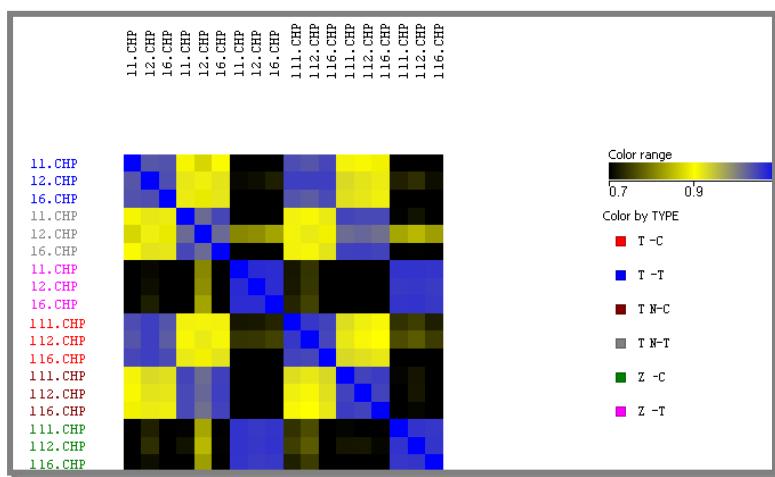
PCA (Principal Component Analysis) 的原理是找到数据方差最大的两个或者三个主成分(就是向量)，将数据投影在这些主成分上，以达到降维的目的，通过图像上的点之间的相互距离来显示样品之间的相似度。考察样品的分布情况，验证实验设计的合理性，生物学重复样品的均一性（至少2组数据）。



样本相关性 (correlation)

目的：检测各个样品间的相关性，验证生物学重复之间的均一性。

要求：设有生物学重复实验的样品组数据。

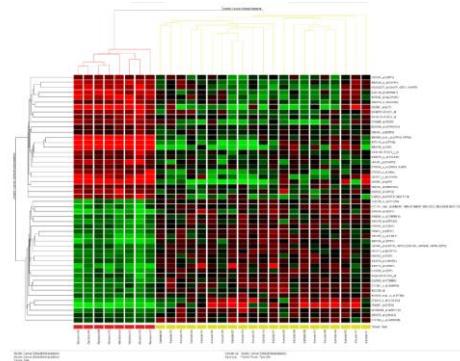


聚类分析

聚类分析包括层次聚类和表达模式聚类两种。

1 层次聚类

为了全面的直观的展示样品之间的关系及差异情况，将表达基因做聚类分析（见下图）。用挑选的差异基因的表达情况来计算样品直接的相关性。一般来说，同一类样品能通过聚类出现在同一个簇（cluster）中，聚在同一个簇的基因可能具有类似的生物学功能。

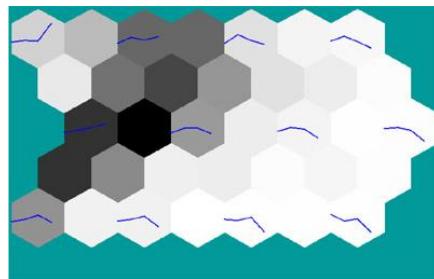


2 表达模式聚类

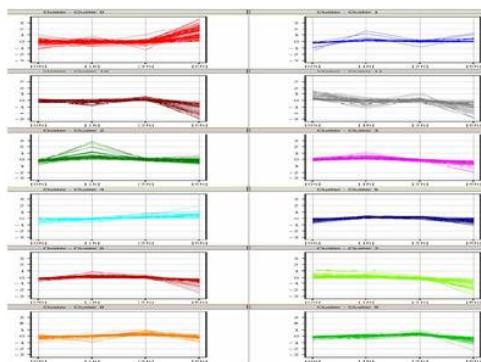
在研究按照处理时间、剂量、疾病恶化程度顺序设计的序列实验时，为了得到差异基因当中与样本变

化相关的基因群，可以对基因进行表达趋势聚类，并对趋势进行显著性分析，可以精确、直观地筛选出随样本顺序变化影响最显著的基因群。常规的是趋势聚类如 SOM 或者 k-means，可以将差异基因所参与的主流表达趋势科学、准确的筛选出来。

要求：三个或三个时间点以上的数据，或者是具有逻辑关系顺序的数据，如：正常-疾病-治疗。

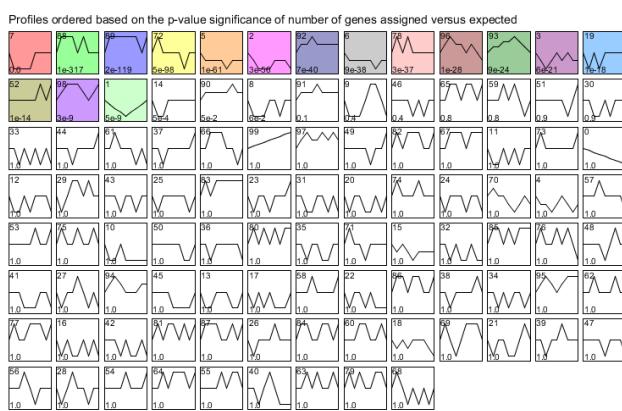


自组织映射聚类 SOM 结果图。每个趋势用蓝色的折线表示。每个趋势可以通过一个六边形同另外一个趋势连接。六边形的灰度表示相关性系数 (correlation) 的大小，白色: correlation=1，黑色: correlation=0，灰度表示不同的相关性系数。通过连接相邻两个趋势的六边形的灰度可以了解它们之间的相关性系数。

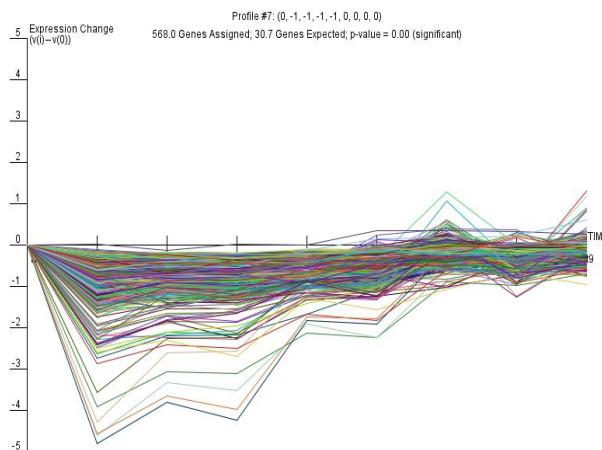


3 趋势显著性分析

趋势总图，每个小图表示一种表达趋势，折线代表这类表达模式的趋势类型。图中标颜色的为显著性的趋势，代表实验中基因主要的表达变化趋势。



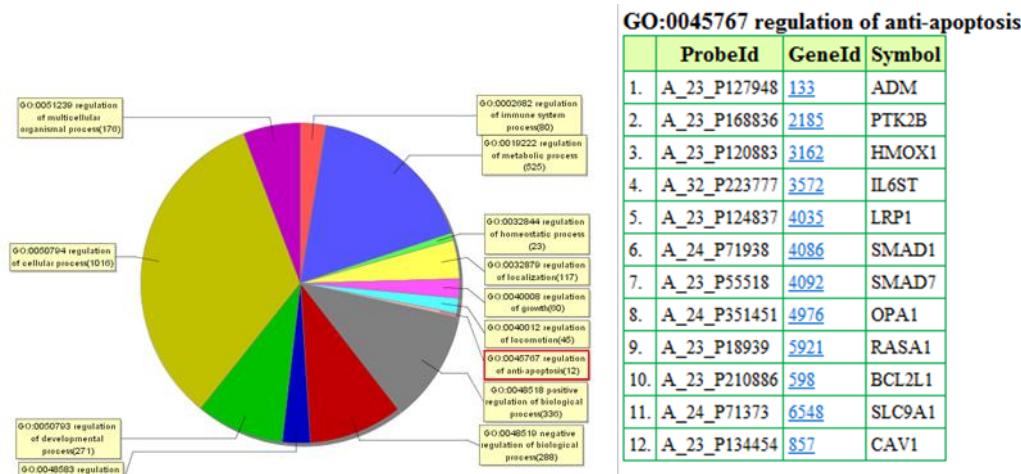
显著性趋势图，图中 x 轴为时间点，y 轴为基因的标准化的表达值，图中的每条线为一个基因的变化趋势。



基因功能分析 (GO analysis)

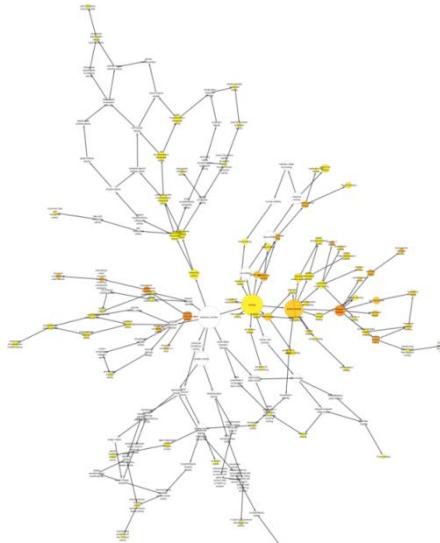
Gene Ontology(简称 GO)是基因功能国际标准分类体系。GO 可分为分子功能(Molecular Function), 生物过程(biological process) 和细胞组成(cellular component) 三个部分。GO-Analysis 对差异基因等按 GO 分类，并对分类结果进行基于离散分布的显著性分析、误判率分析、富集度分析，得出与实验目的有显著联系的、低误判率的、靶向性的基因功能分类，该分类即导致样本性状差异的最重要的功能差别。通过该分析有可能找到导致性状变化的重要功能，并且找到该功能所对应的基因。

GOid	Name	Hits	Total	Percent	Enrichment test pvalue	qvalue
GO:0050793	regulation of developmental process	271	1280	21.17%	0.0	0.0025
GO:0050793	regulation of developmental process	271	1280	21.17%	0.0	0.0025
GO:0051094	positive regulation of developmental process	145	648	22.38%	2.0E-4	0.0182
GO:0051094	positive regulation of developmental process	145	648	22.38%	2.0E-4	0.0182
GO:0002520	immune system development	88	358	24.58%	3.0E-4	0.0224
GO:0001775	cell activation	88	366	24.04%	5.0E-4	0.0299
GO:0051093	negative regulation of developmental process	125	562	22.24%	6.0E-4	0.0299
GO:0051093	negative regulation of developmental process	125	562	22.24%	6.0E-4	0.0299
GO:0040011	locomotion	115	525	21.9%	0.0015	0.0677
GO:0006928	cell motion	120	557	21.54%	0.0020	0.0727
GO:0051674	localization of cell	120	557	21.54%	0.0020	0.0727



功能层次网络构建 (GO-Network)

功能层次网络构建 (GO-Network) 是基于 GO 的层次结构, 将所有差异基因同时参与的显著性 GO 及其相互从属关系构建功能网络, 从全局角度, 系统地概括功能间相互作用关系及所属分层关系。

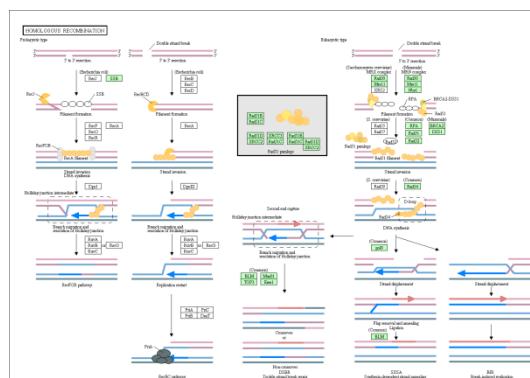
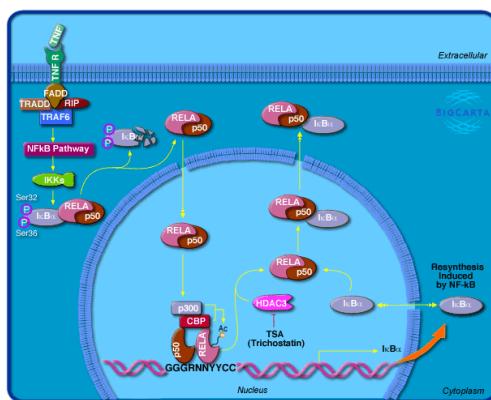


圈的大小代表 GO 的层级, 越大越高, 黄色的深浅代表富集度的高低, 越深越显著

信号通路分析 (Pathway analysis)

信号通路 (Pathway) 是多个蛋白质间相互作用, 共同调节细胞功能和代谢活动的过程。而信号通路分析是通过对差异基因按照 Pathway 的主要公共数据库 KEGG 和 Biocarta 来进行分类, 对 Pathway 中的基因进行基于离散分布的显著性分析, 得到与实验目的有显著联系的 Pathway 分类, 该分类即导致样本性状差异的最重要 Pathway。

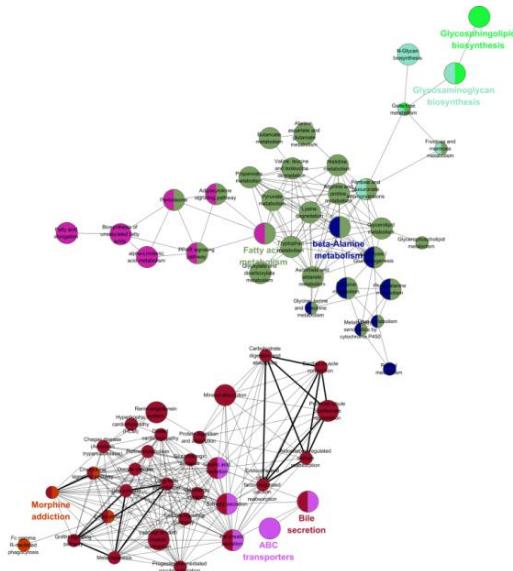
PathwayDB	Name	Hits	Total	Percent	Enrichment test pvalue	qvalue
Biocarta	?arrestins in GPCR Desensitization	5	11	45.45%	0.0013	1.0E-4
Biocarta	Acetylation and Deacetylation of RelA in The Nucleus	5	13	38.46%	0.0024	1.0E-4
Biocarta	Actions of Nitric Oxide in the Heart	4	19	21.05%	0.0355	8.0E-4
Biocarta	Activation of cAMP-dependent protein kinase, PKA	3	6	50.0%	0.011	3.0E-4
Biocarta	Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor	7	19	36.84%	4.0E-4	0.0
Biocarta	Activation of PKC through G protein coupled receptor	1	5	20.0%	0.289	0.0044
Biocarta	Activation of Src by Protein-tyrosine phosphatase alpha	2	8	25.0%	0.1023	0.0019
Biocarta	Acute Myocardial Infarction	4	13	30.77%	0.0124	3.0E-4
Biocarta	Adhesion and Diapedesis of Granulocytes	9	17	52.94%	0.0	0.0
Biocarta	Adhesion and Diapedesis of Lymphocytes	6	14	42.86%	6.0E-4	0.0
Biocarta	Adhesion Molecules on Lymphocyte	6	9	66.67%	1.0E-4	0.0
Biocarta	ADP-Ribosylation Factor	1	17	5.88%	0.6406	0.0090



信号通路调控网络构建 (Pathway-Network)

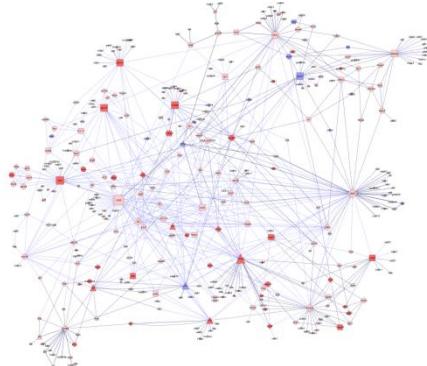


信号通路调控网络构建（Pathway-Network）是根据所有差异基因同时参与的 Pathway 之间的相互调控关系构建信号通路调控网络，从系统的角度研究各个信号通路间的信号传导和调控过程，在多个显著性 Pathway 中发现受实验影响的核心 Pathway，以及实验影响的信号通路之间的调控机理。

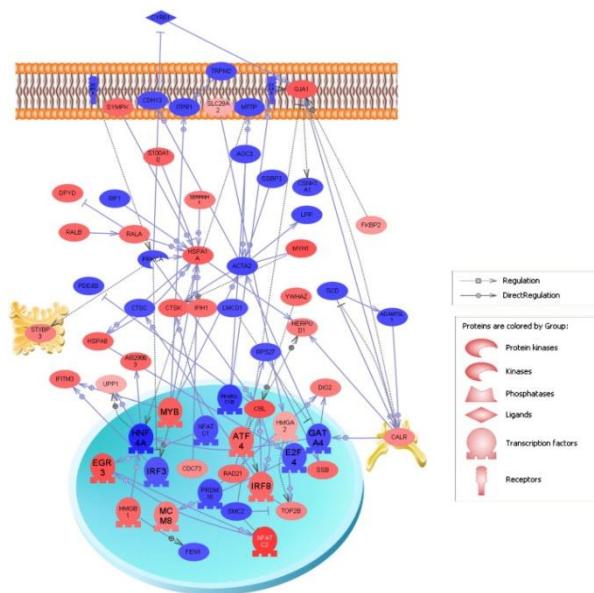


基因相互作用网络图

基因网络是一组基因的集合体，这些基因通过相互协作来控制生物体重要的生命过程。我们通过筛选的差异基因进行基于实验证，其他亲缘物种数据库检索、文献挖掘等多种信息进行基因互作网络的构建。



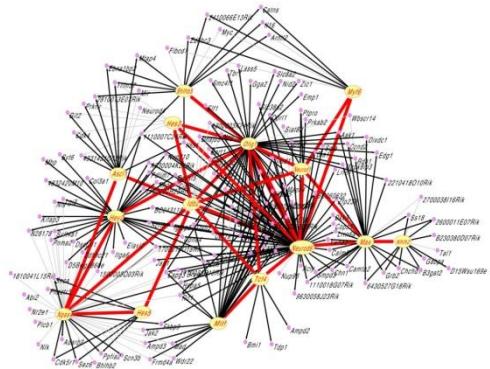
图中节点的大小表示的是 degree，节点的形状表示不同的功能类型（炎症，代谢，其他等），颜色代表上下调的信息，边的颜色或者线型代表是相互作用的来源）。



细胞外的刺激物质找到它的配体，利用数据库对基因的关联以及定位进行搜索，找到差异基因表达的 pathway，而且能够找到 pathway 中起调控作用的关键蛋白。

转录因子分析及网络构建

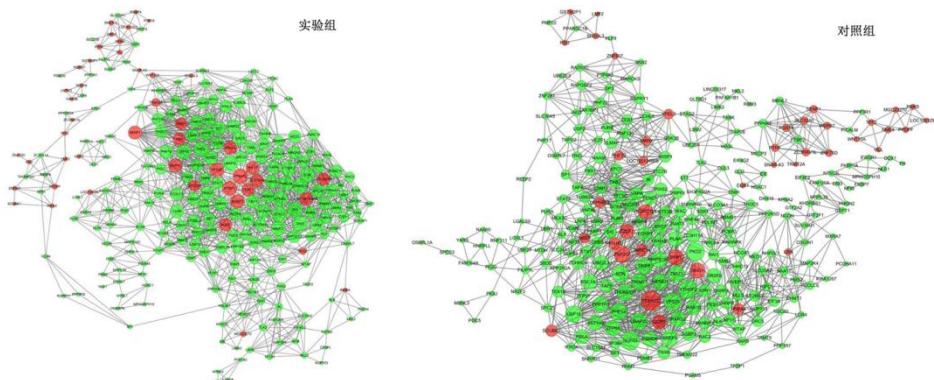
转录因子与基因表达调控区域的特异性结合是基因表达的重要调控方式。针对特定转录因子的实验，可对差异基因进行转录因子结合位点的 motif 分析，从而推测差异基因中有哪些基因可能直接受该转录因子调节。最后结果会用统计学方法进行检验，找到调控目标生物性状，统计学上有显著差别的转录因子。找到差异的转录因子及其调控的靶基因，我们可以构建 **TF-network**，从中得出关键的转录因子及被关键转录因子转录调控的基因，系统的研究转录因子与基因的调控关系。



共表达网络构建(Coexpression Network)

共表达网络图（Coexpression Network）是根据基因表达信号值的动态变化，计算基因间的共表达关系，得到基因间的表达调控关系及调控方向，从而构建基因的表达调控网络。利用共表达网络图，研究者可通过分析基因调控能力，获得样本随实验变化的核心调控基因。

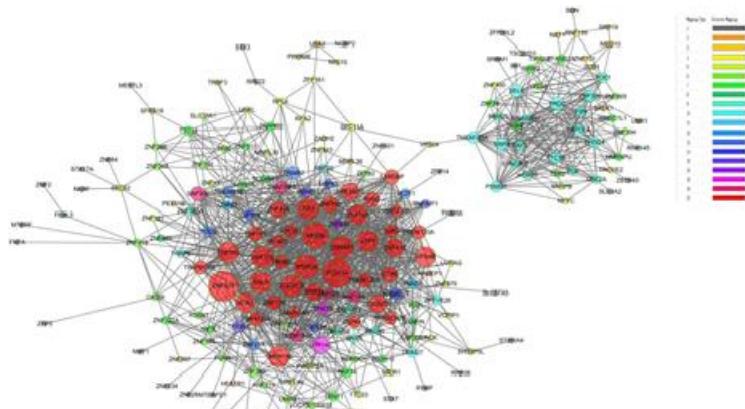
要求：每组数据 3 个或 3 个以上生物学重复



图中红色为上调基因，绿色为下调基因

时间序列的共表达网络

根据基因表达信号值随着时间序列的变化，计算基因间的共表达关系，得到基因间的表达调控关系及调控方向，从而构建基因的表达调控网络。利用此网络图，研究者可通过分析基因调控能力，获得样本在时间过程中起调控作用的基因。

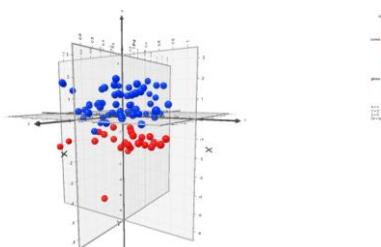


图中颜色为对网络的分层，其中红色区域的基因处于网络的核心地位，起着调控网路的核心作用

分子建模预测(MAKER 预测)

采用模式识别与数据挖掘技术有效进行预测模型的构建，将部分芯片数据拿来做预测模型，然后部分芯片数据作为测试数据集（独立样本）来验证模型的准确性。目的在于利用实验数据来筛选出一批靶标基因，并以此构建模型，以进行早期诊断、疾病预测和预后分析。在微阵列运用于疾病预测方面，有各种分类方法，我们的常用方法是以下几种：SVM、Compound covariate predictor、Diagonal linear discriminant analysis、Nearest centroid predictor、Bayesian compound covariate predictor

并利用了 Leave-one-out cross-validation (LOOCV) 以及 cross-validated misclassification error rate 的筛选策略找到最优 MAKER。





伯豪案例

应用案例一：表达谱芯片助力脑干胶质瘤研究

首都医科大学和杜克大学医学院的研究人员合作，发现了脑干胶质瘤（brainstem glioma, GSB）的新的遗传病因。脑干胶质瘤是一种致命的肿瘤，患者多为儿童及青少年。他们通过外显子组测序分析，发现 PPM1D 基因第 6 外显子的截短突变可引起脑干胶质瘤。该基因的这类突变是功能获得突变，它们不仅促进癌细胞的生长，而且能阻止癌细胞的灭亡。突变也导致放射线疗法对部分脑干胶质瘤患者不再有效。为了进一步研究脑干胶质瘤发病的遗传机制，作者对 21 例胶质瘤样本进行了表达谱芯片检测，并进行了基因富集分析 (Gene Set Enrichment Analysis, GSEA)。表达谱芯片检测由上海伯豪生物技术有限公司完成。

原文出处: Exome sequencing identifies somatic gain-of-function PPM1D mutations in brainstem gliomas. Nat Genet. 2014 Jul;46(7):726-30 (IF35.209) 使用服务种类: Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 服务

应用案例二：在细胞重编程研究中表达谱芯片的应用

中国科学院上海生命科学研究院生物化学与细胞生物学研究所惠利健研究组通过表达三个肝脏转录因子，成功地将人体皮肤细胞转变为肝细胞。该研究组在将小鼠皮肤细胞转变为肝细胞研究结果的基础上，经过重新筛选和优化条件，成功建立了诱导人成纤维细胞重编程为肝细胞 (hiHep 细胞) 的方法。hiHep 细胞可表达肝脏基因，并具有肝细胞的许多功能。为了证明诱导得到的 hiHep 细胞具有类似于人肝细胞的基因表达特征，他们对 hiHep 细胞和人肝细胞以及其它对照细胞进行了全基因表达谱检测并做了聚类分析。检测和分析结果从一个方面显示他们成功诱导了细胞重编程。表达谱芯片服务由上海伯豪生物技术有限公司提供。

原文出处: Direct Reprogramming of Human Fibroblasts to Functional and Expandable Hepatocytes. Cell Stem Cell, 2014 Mar 6;14(3):370-84. (IF25.315) 使用服务种类: Agilent Whole Human Genome Oligo Microarray 服务

应用案例三：表达谱芯片筛选发现肝癌扩散转移相关基因

该研究使用 Agilent 人全基因组表达谱芯片对 15 个经过 DAC+TSA 处理和没有经过处理的肝癌细胞系进行了表达谱分析，发现 SCARA5 基因沉默可导致肝癌癌细胞扩散转移。J Clin Invest 评审专家认为，这不仅是发现了一个与癌症相关的重要基因，为肝癌提供了潜在新型诊断和治疗方法，还采用了新的研究方法，拓展了研究思路。该课题相关的芯片技术服务在生物芯片上海国家工程研究中心完成。本文第一作者黄建博士这样评价我们的工作：“全面的解决方案；先进的技术平台；优质的服务水平。”

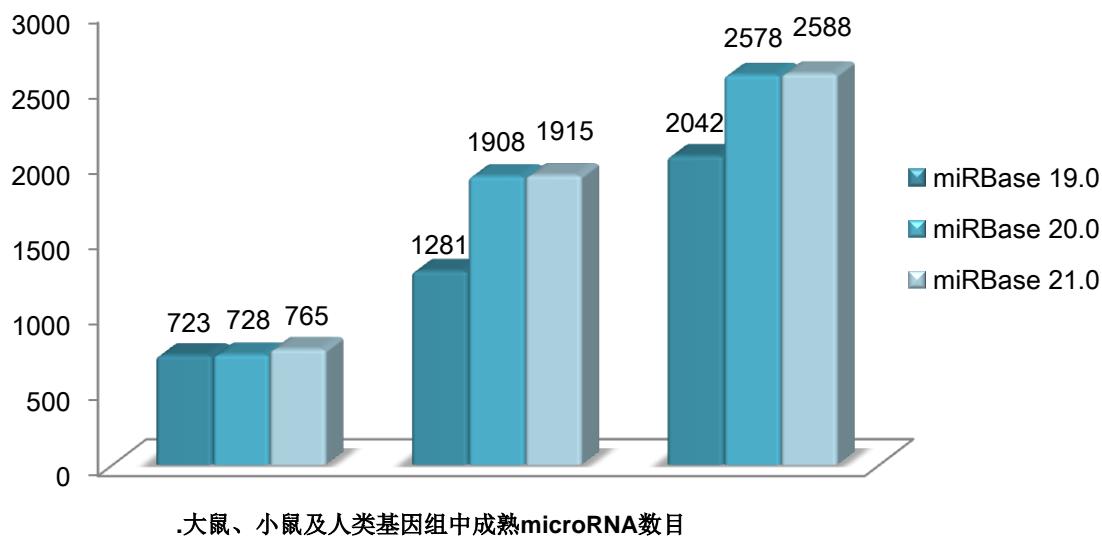


原文出处: Genetic and epigenetic silencing of SCARA5 may contribute to human hepatocellular carcinoma by activating FAK signaling. J Clin Invest. 2010, 120(1):223-41. (IF15.387) 使用服务种类: Agilent Whole Human Genome Oligo Microarray 服务

(五) miRNA 芯片服务

microRNA(microRNA)是一种小的内源性非编码 RNA 分子，大约由 21-25 个核苷酸组成。通过与靶基因的 3'UTR 配对，促进 mRNA 的降解或抑制 mRNA 的翻译，从而抑制靶基因的表达。

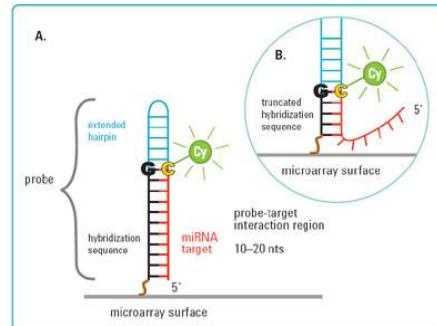
现在,miRBase(1)共收录 223 个物种的 28645 个 microRNA 前体,35828 个成熟 microRNA (miRBase, v21, 2014 年 6 月).



上海伯豪生物技术有限公司在 miRNA 芯片领域建立起 Agilent miRNA 芯片平台(针对动物 miRNA, 已经进行了超过 11000 份份样品的 miRNA 芯片检测服务) 和 Affymetrix miRNA 芯片平台 (覆盖 203 个物种), 为客户提供最具针对性的高端技术服务。

Agilent miRNA 芯片服务

安捷伦科技有限公司 (Agilent Technologies Inc.) 作为国际领先的生物芯片制造企业, 利用其先进的芯片生产工艺, 开发出具有独特设计的 miRNA 检测芯片, 能特异地检测出成熟的 miRNA 并能很好地区分高度同源的 miRNA 分子。该平台还具有特异性高、灵敏度高等特点。可以广泛地应用于检测和比较各类样品的 miRNA 表达状态, 并通过分析其调控的靶基因来深入分析其调控机制或生物学功能。



Agilent microRNA 芯片探针设计图

技术特点

- 唯一能通过特殊的探针设计区分成熟 miRNA 和 miRNA 前体的芯片平台;
- 无需分离 miRNA, 直接用 Total RNA 进行实验, 只需 100ng 的 Total RNA 即可用来进行标记实验 (不包括质检所消耗的量), 避免了分离、富集、放大等过程带来的影响;
- 样品适用范围广: 组织、血清、血浆、FFPE、冰冻切片组织、脱落细胞等均可进行实验;
- 高特异性: 特殊的专利探针设计能区别一个碱基的差异;
- 高灵敏度: 检测丰度跨 4 个 log, 最高可达 6 个 log, 只要存在 6000 个拷贝的分子就可以检出;
- 数据来源于 Sanger microRNA 数据库 (miRBase), 方便及时更新;
- 人、大鼠、小鼠之外的动物物种, 可以方便的通过 earray 进行个性化定制。



服务推荐

物种	芯片名称	P/N	Design ID	Format
人	SurePrint G3 Human miRNA Microarray, Release 21	G4872A	070156	8 x 60K
小鼠	SurePrint G3 Mouse miRNA Microarray, Release 21	G4872A	070155	8 x 60K
大鼠	SurePrint Rat miRNA Microarray, Release 21	G4471A	070154	8 x 15K
其他物种	Agilent Custom microRNA Microarray	定制	定制	定制

服务内容:

总 RNA 抽提、样品质检、microRNA 标记、杂交、洗脱、图像扫描、数据分析。

Affymetrix miRNA 芯片服务

上海伯豪生物技术有限公司作为 Affymetrix 认证服务供应商，利用 Affymetrix 开发的 GeneChip® microRNA 4.0 Array 为您提供优质的服务。

Affymetrix GeneChip miRNA 4.0 芯片服务特点:

- 高覆盖度:** 涵盖 miRBase Release 20.0 中所有成熟的 microRNA 序列;
- 新型的分析:** 采用同样的芯片可以检测人、小鼠、大鼠或其他物种的 microRNA（同时检测 203 个物种的 microRNA）;
- 信息量大:** 芯片结果包括: host gene ID、microRNA target genes、clustered microRNA 等;
- 样本量低:** 仅需要 130ng 的 total RNA;
- 高灵敏度:** 检测丰度跨 4 个 log，可以检测到 1.0 attomol 的 microRNA 含量。

芯片内容:

Affymetrix microRNA 芯片 4.0 版本，
100% 覆盖 miRBase v20 数据库；
包含 203 个物种的 30424 个成熟 microRNA；
5214 个人、小鼠、大鼠 microRNA；
1996 个人类 snoRNA, scaRNA；
3770 个探针专门针对人类、小鼠、大鼠前体发夹状序列设计。

总探针数	人			小鼠		大鼠	
	成熟 miRNA	Pre-microRNA	snoRNA, scaRNA	成熟 miRNA	Pre-microRNA	成熟 miRNA	Pre-microRNA
30,424	2,578	2,025	1,996	1,908	1,225	728	490

Affymetrix microRNA 4.0 芯片探针数目:(人、大鼠、小鼠)



服务相关

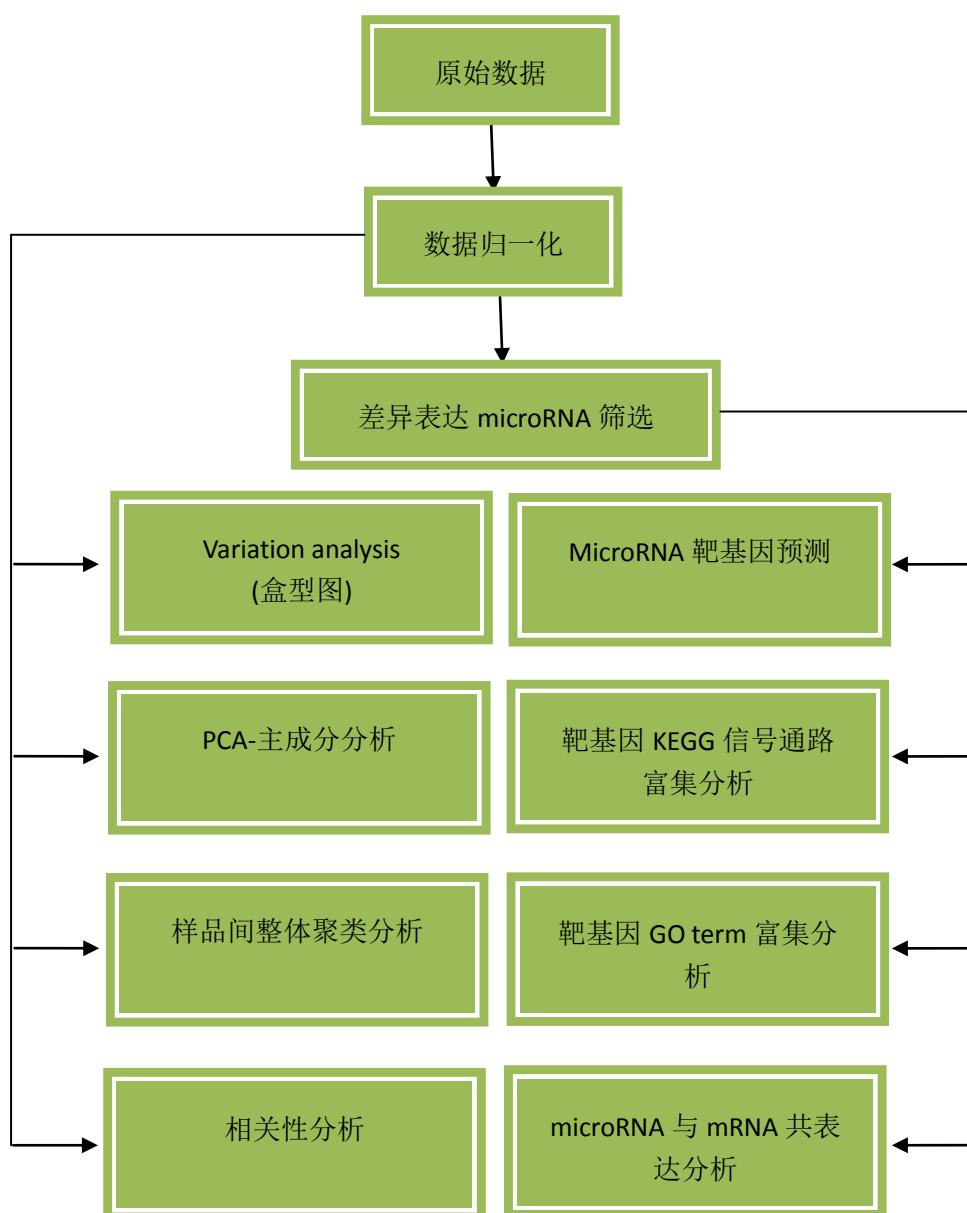
样品要求

1. 样品纯度: OD 260/280 值应在 1.8~2.1 之间; DNA 污染应该去除干净;
2. 样品浓度: 浓度不低于 20ng/ μ l (细胞来源样本);
3. 样品总量: 每个样品总量不少于 0.5 μ g ;
4. 样品溶剂: 溶解在水中;
5. 样品运输: RNA 干冰低温运输; 在运输过程中请用 parafilm 将管口密封好, 以防出现污染。

服务流程



数据分析



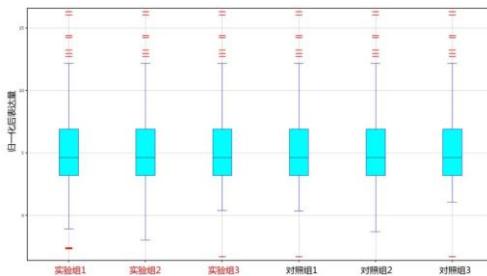
microRNA芯片的数据分析包括数据预处理、microRNA表达分析、差异基因分析、差异基因GO/KEGG富

集分析。

MicroRNA芯片数据分析内容（以Agilent microRNA芯片为例）

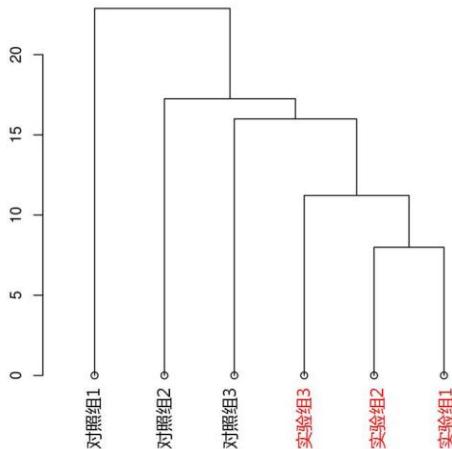
归一化

将原始数据归一化后，对归一化后的数据用Box Plot图来观察各样品数据分布的整体特征。microRNA芯片采用的归一化的方法为quantile normalization, loess normalization。



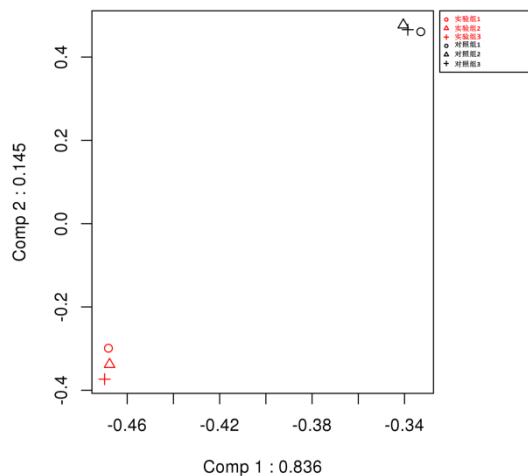
聚类

为了全面和直观展示样品之间的关系及差异情况，将表达基因或者差异基因数据做聚类分析（见下图）。同一类样品聚类后出现在同一个簇（cluster）中，聚在同一个簇的基因可能具有类似的生物学功能。



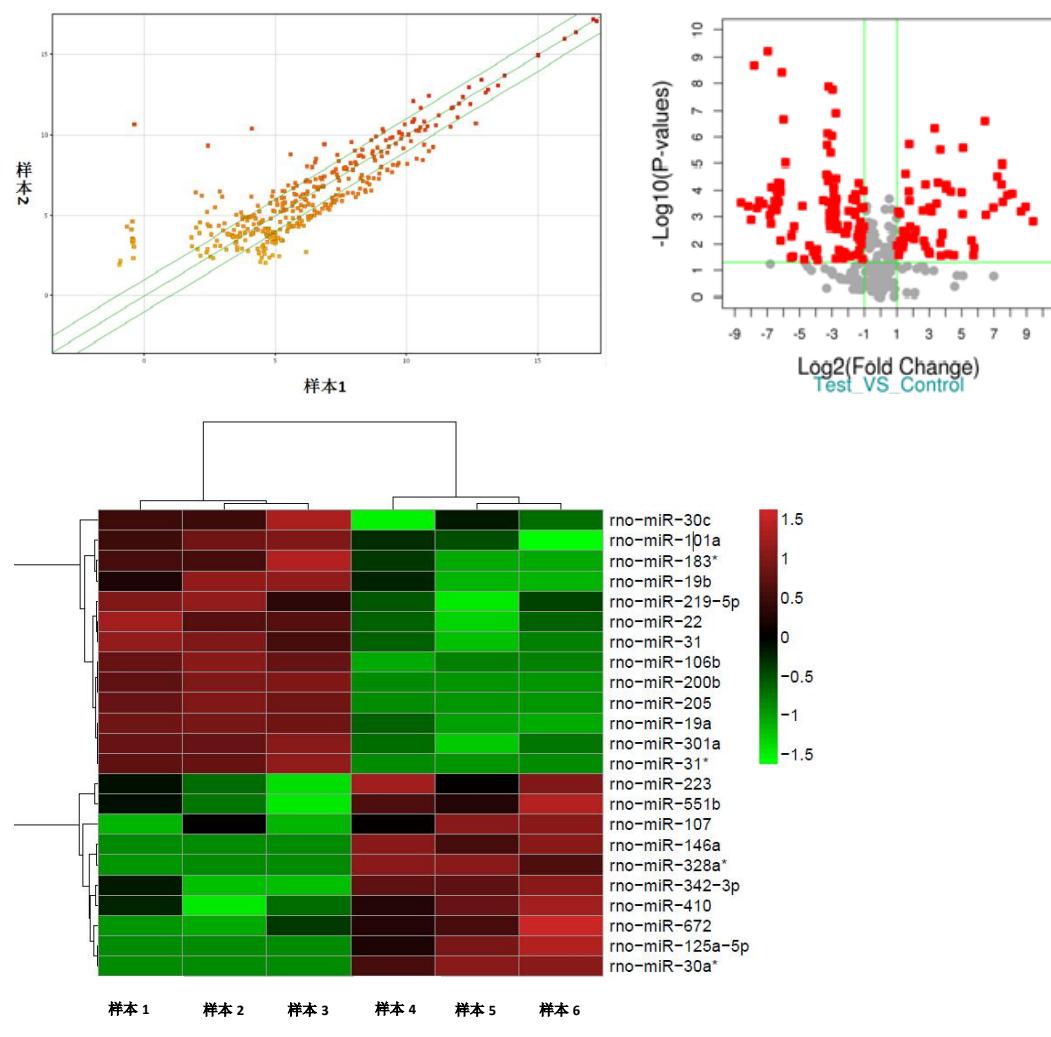
主成分分析

PCA (Principal Component Analysis) 的原理是找到数据方差最大的两个或者三个主成分(就是向量)，将数据投影在这些主成分上，以达到降维的目的，通过图像上的点之间的相互距离来显示样品之间的相似度。PCA 通常用来评价样品的分组情况，验证实验设计的合理性，生物学重复样品的均一性等（至少 2 组数据）。



差异 microRNA 筛选

原始数据归一化后。采用 Fold-change(表达差异倍数)以及 T-test 等检验统计学对差异基因进行筛选。将筛选得到的不同组间的差异基因，通过散点图、火山图及热图等方法，进行直观的图形展示。



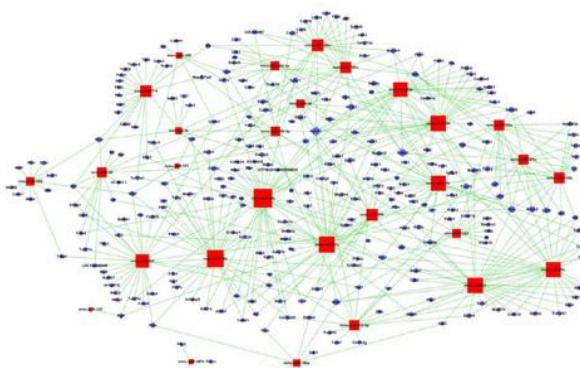
microRNA靶基因预测

应用 TargetScan 等数据库的软件预测 microRNA 靶基因。

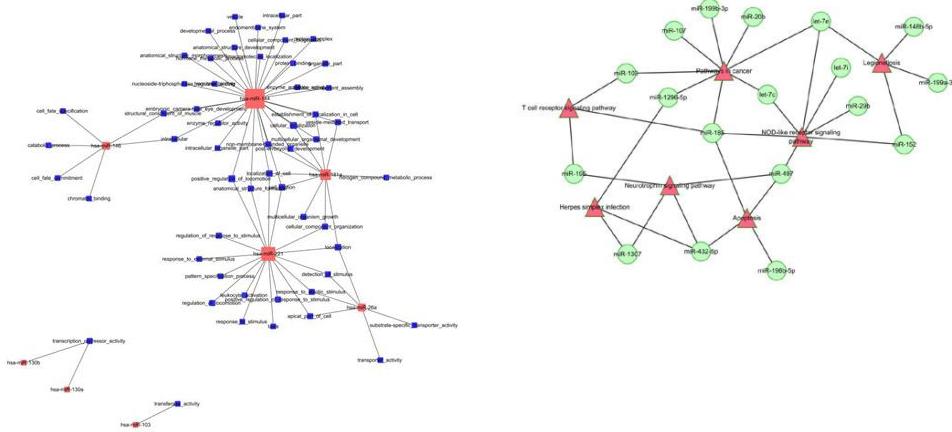
miRNA	Symbol	description	Entrez Gene
rno-miR-150 rno-miR-383 rno-miR-140 rno-miR-495 rno-miR-31	Eif4g2 Pmfbp1	translation initiation fac ulated factor 1 bindin	171362 171414
rno-miR-20a rno-miR-93 rno-miR-130	Ubb	ubiquitin B	192255
rno-miR-130a rno-miR-378 rno-miR-1	C6	complement component 6	24237
rno-let-7b rno-miR-185 rno-miR-33	Cyp11b8	P450, subfamily XIB 8	24295
miRNA	map_location	dbNires	type_of_gene
rno-miR-150 rno-miR-383 rno-miR-140 rno-miR-495 rno-miR-31	9q21 19q12	821477 377 Ensembl: ENSRNOG000000000000	pseudo protein-coding
rno-miR-20a rno-miR-93 rno-miR-130	10q23	562 Ensembl: ENSRNOG000000000000	protein-coding
rno-miR-130a rno-miR-378 rno-miR-1	2q16	MAF: 125 Ensembl: ENSRNOG000000000000	protein-coding
rno-let-7b rno-miR-185 rno-miR-337 rno-let-7i	B6D: 2455 RATMAP: 32967		other

microRNA与靶基因网络图

将显著性功能与显著性Pathway所包含的靶基因取交集后与microRNA构建基因与microRNA调控网络(待确认),可以在全局的水平上直观的反应基因之间的相互关系,同时反映了基因调控网络的稳定性。根据网络中microRNA的位置函数计算出microRNA在网络中的关系强度,即microRNA的网络特征值。特征值最高microRNA处于网络的枢纽性地位,该microRNA调控能力最强,对网络结构和样本性状有重要的调控价值,同时从网络中也可以得到被microRNA调控的关键靶基因。

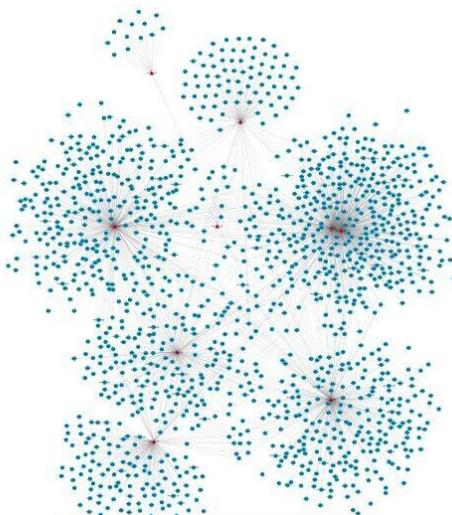


microRNA靶基因富集分析



microRNA与表达谱芯片平台的联合分析

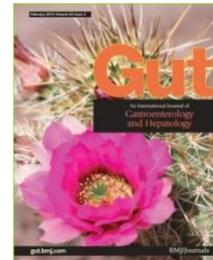
microRNA 与表达谱芯片平台的联合分析除了和甲基化与表达谱芯片分析类似外，还可以将 **microRNA** 预测的靶基因和表达谱芯片的差异基因取交集，进行 **microRNA-DiffGene-Network**。



伯豪案例

应用案例一：基因芯片筛选发现大肠癌分子标志物

该研究课题首先收集了 133 例大肠癌样品，利用激光共聚焦捕获技术（laser capture microdissection, LCM）切取组织样品中的上皮细胞，随后提取总 RNA，利用 Agilent microRNA 芯片筛选，RT-PCR 验证等方法找到了 3 个 microRNA 的组合，最后，利用这 3 个 microRNA 的组合来对 58 个 FFPE 样本进行早期诊断，发现准确率达到 89%（AUC=0.918）。这一研究成果对于大肠癌的及时发现和治疗有重大意义。其中 microRNA 基因芯片技术服务由上海伯豪完成。



原文出处：A microRNA panel to discriminate carcinomas from high-grade intraepithelial neoplasms in colonoscopy biopsy tissue. Gut. 2013, 62(2): 280-289. (IF10.111) 使用服务种类：Agilent Human miRNA 芯片服务

应用案例二：基因芯片筛选发现乙型肝癌血清血浆 microRNA 生物标志物

复旦大学附属中山医院、复旦大学肝癌研究所樊嘉教授领衔的课题组利用 microRNA 芯片对比肝癌患者 VS 正常人群，肝癌患者 VS 慢性乙型肝炎（chronic hepatitis B, CHB）患者，肝癌患者 VS 肝硬化（cirrhosis）患者血清血浆 microRNA 的表达谱，筛选出 15 个明显差异表达的 microRNA。随后，通过 RT-PCR 验证等方法寻找到 7 个 microRNA 的组合。最后，科学家们又利用 390 份样品对建立好的诊断模型进行验证，该模型区分肝癌患者和正常人群的正确率能达到 94.1%。这项研究成果对于肝癌的早期诊断和治疗有重大意义，microRNA 生物芯片技术服务由上海伯豪提供。



原文出处：Plasma MicroRNA Panel to Diagnose Hepatitis B Virus–Related Hepatocellular Carcinoma. J Clin Oncol. 2011. 29(36):2697. (IF18.970) 使用服务种类：Agilent Human miRNA 芯片服务

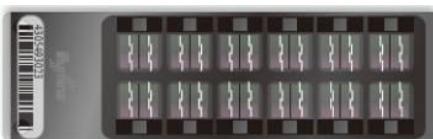
（六）DNA 甲基化芯片服务

DNA 甲基化对生命过程非常重要，是表观遗传学的重要调控方式，也是基因精确调控的方法之一，通过对位于基因启动子及第一外显子区的 CpG 岛的甲基化而抑制基因的表达。它在细胞正常发育、基因表达

模式以及基因组稳定性中起着至关重要的作用。目前，SBC 可利用 Illumina 服务平台，为您提供 DNA 甲基化芯片服务。

Illumina Infinium Human 甲基化 450K 芯片 HOT!

Illumine Human Methylation 450K BeadChip 芯片可检测人全基因组 48 万个甲基化位点，全面覆盖了 96% 的 CpG 岛，并根据需求加入了 CpG 岛以外的 CpG 位点、人类干细胞非 CpG 甲基化位点、正常组织与肿瘤（多种癌症）组织差异甲基化位点、编码区以外的 CpG 岛、miRNA 启动子区域和已通过 GWAS 的疾病相关区域的位点，同时覆盖了 Human Methylation27 BeadChip 的 94% 的位点。每张芯片可平行进行 12 个样本的检测。集合了全面、高样品通量，以及价格实惠等优点，广泛应用于干细胞研究、肿瘤和其他疾病研究，是唯一一款适合表观基因组全关联分析研究的全基因组 DNA 甲基化芯片。



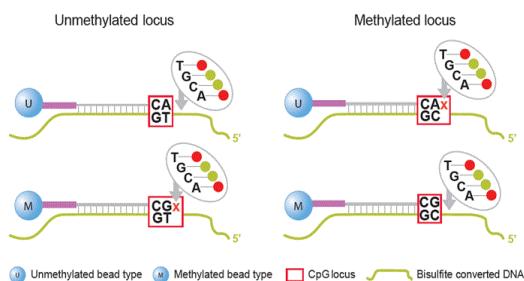
- ✓ 覆盖全面：485,553 多个甲基化位点；
- ✓ 无需“甲基化 DNA 免疫共沉淀”，亚硫酸氢盐处理基因组 DNA 即可进行芯片实验；Infinium 探针设计，直接识别甲基化位点；
- ✓ 技术重复相关系数 $R^2 = 0.9969$ ；
- ✓ 与“全基因组亚硫酸氢盐-测序”相关系数 $R^2 > 0.959$ ；
- ✓ 可检测 FFPE 样品甲基化分析。

芯片原理

Illumina Human 甲基化 450K 芯片采用 Illumina BeadArray™ Infinium 技术，特异性的识别甲基化位点。在甲基化 450K 芯片中，同时采用了 Infinium I 和 Infinium II 探针设计，从而使得检测范围最大化。

探针类型	探针数/位点数	反应类型	延伸所用的核苷酸	区分方式
Infinium I	两根探针检测一个位点	延伸反应	dNTP-BioT	看信号强弱
Infinium II	一个探针检测一个位点	单碱基延伸	ddNTP-BioT ddNTP-DNP	荧光颜色及强弱区分

(1) Infinium I 探针设计

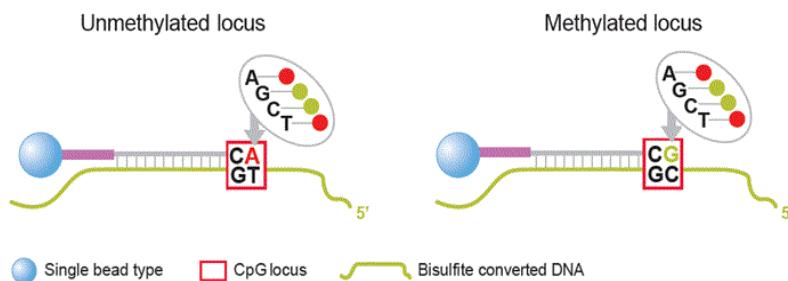


对于每个甲基化位点，都对应设计有两种探针：M型磁珠、U型磁珠。M型磁珠尾部为G，用来检测甲基化位点。U型磁珠尾部为A，用来检测未甲基化位点。



基因组上的某一位点，如果被甲基化了，那么在亚硫酸氢盐的处理下，GC 仍为 GC，与 M 型磁珠配对，荧光标记的核苷酸掺入后能被检测到荧光信号。M 型磁珠发光。反之，如果没有被甲基化，那么在亚硫酸氢盐的处理下，GC 变为 GT，与 U 型磁珠配对，延伸后 U 型磁珠发光。

(2) Infinium II 探针设计



Infinium II 探针只使用一种磁珠，探针末端为 C。配对后只掺入单个碱基（ddNTP-BioT, ddNTP-DNP）。根据荧光类型判断掺入的碱基类型，从而判断是否被甲基化。

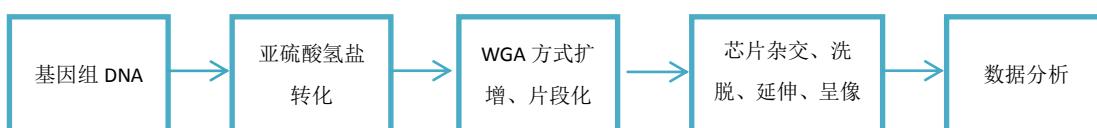
服务相关

样品类型：组织、细胞、基因组 DNA (> 2μg)

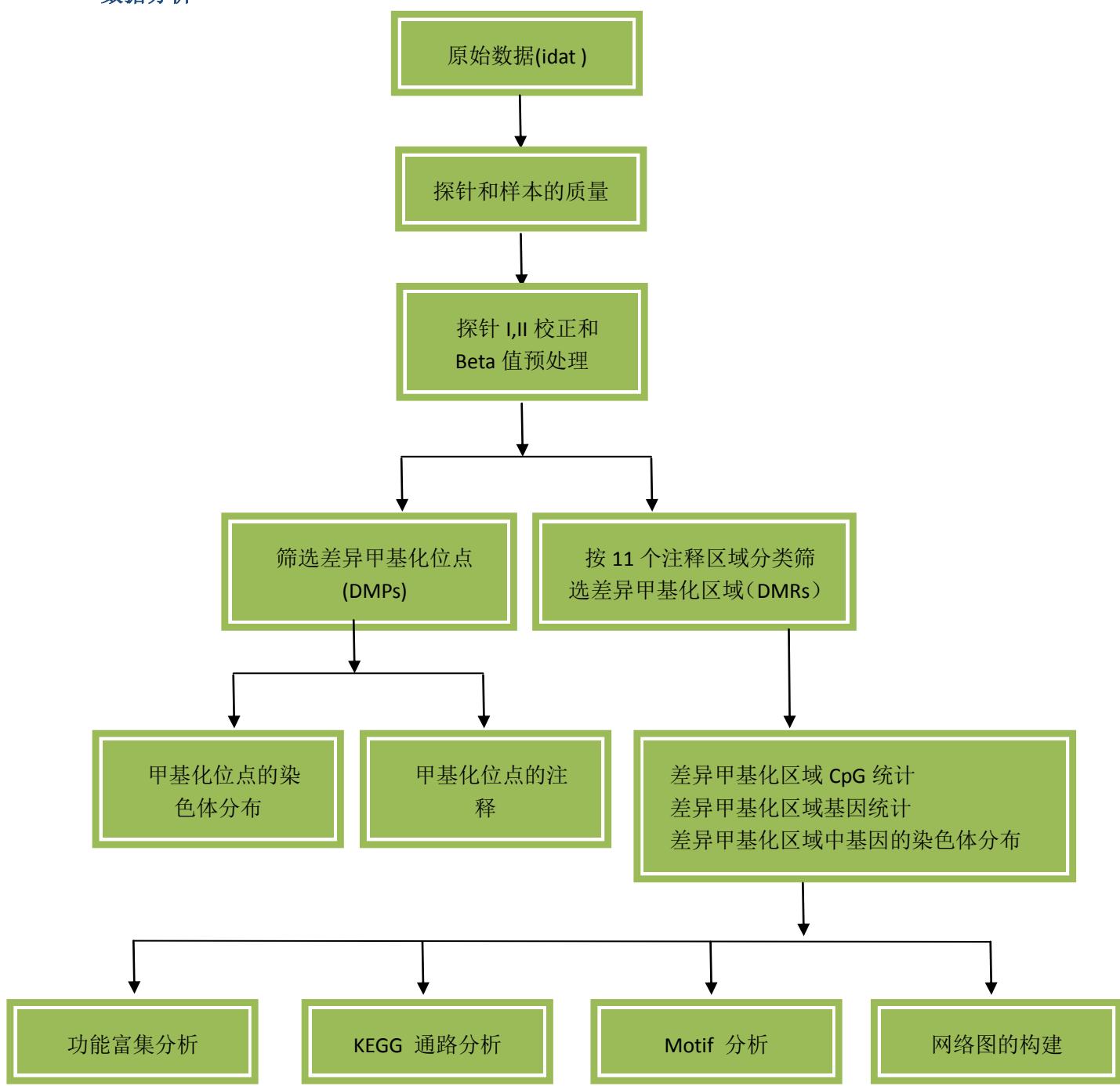
样品要求：

1. 样品纯度：OD 260/280 值应在 1.7~2.0 之间；RNA 应该去除干净；不得有其它个体或其它物种的 DNA 污染。
2. 样品浓度：浓度不低于 55ng/μl；
3. 样品总量：每个样品总量不少于 2μg 。
4. 样品溶剂：溶解在 TE 中。
5. 样品运输：DNA 低温运输 (-20°C)；在运输过程中请用 parafilm 将管口密封好，以防出现污染。

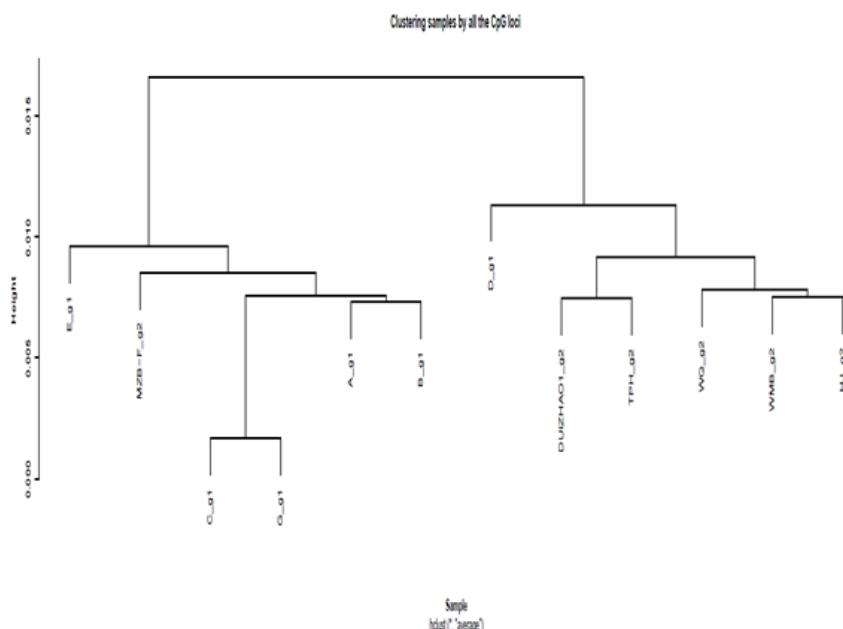
实验流程：



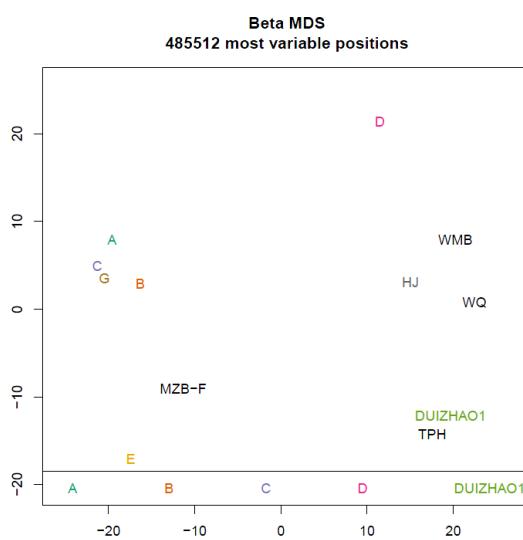
数据分析



基础分析

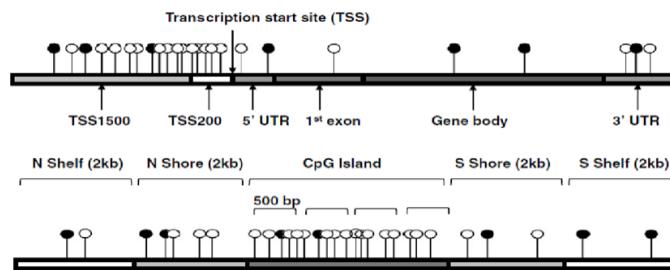


应用探针位点的 beta 值做样本层次聚类。聚类分析后相似的样本出现在同一个簇（cluster）中，聚在同一个簇的基因可能具有类似的生物学功能。通过聚类图可以证实样本分组是否符合实验设计需求。

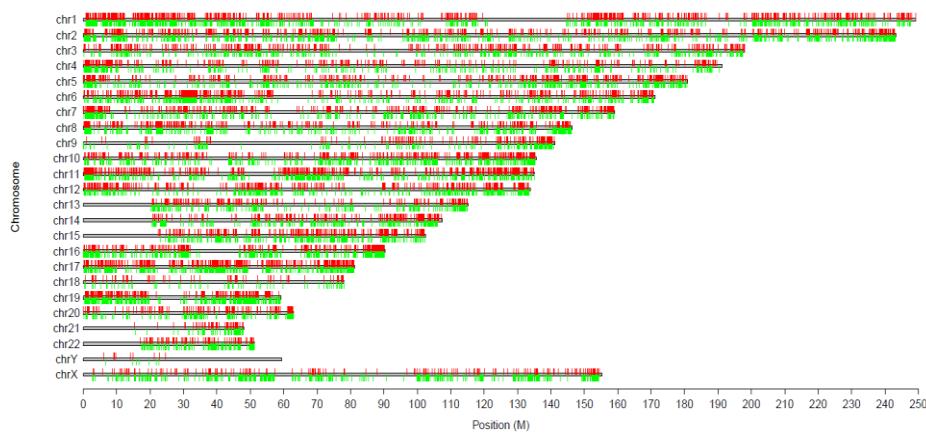


采用多维尺度分析（MultiDimensional Scaling）统计方法计算样本之间的距离，图中的点（对象）的距离反应了它们的相似性或差异性（不相似性）。

Gene/CpG island probe distribution

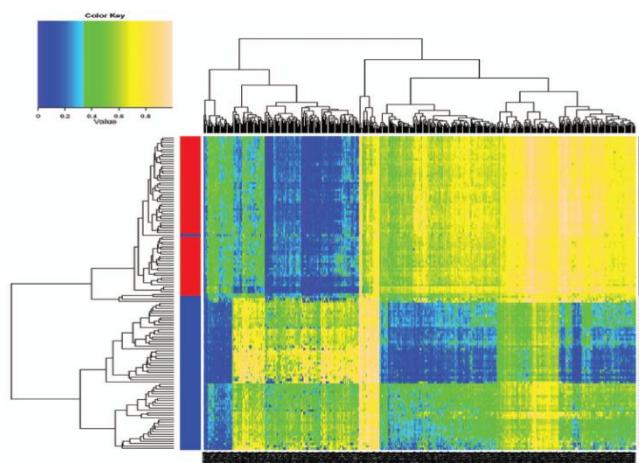


按基因注释分类和 CpG 岛注释分类分别筛选样本组间的差异甲基化区域。

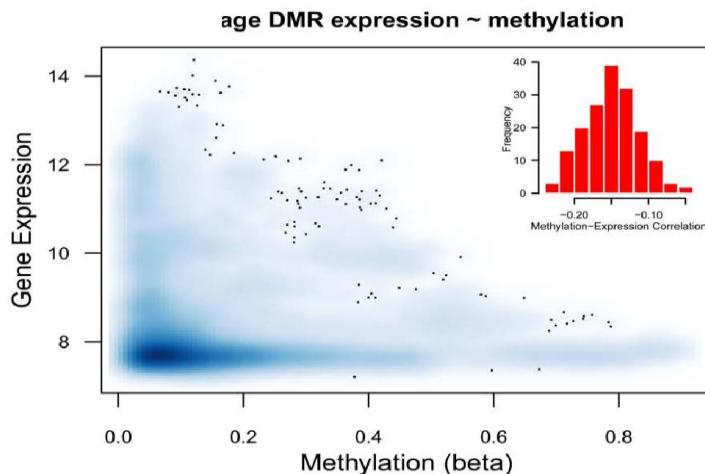


差异甲基化位点定位在染色体，高甲基化或低甲基化的基因是否有染色体的偏向性或成簇分布的特点。其中红色代表实验组高甲基化位点，蓝色代表实验组低甲基化基因位点，灰色表示甲基化位点在组间没有变化。

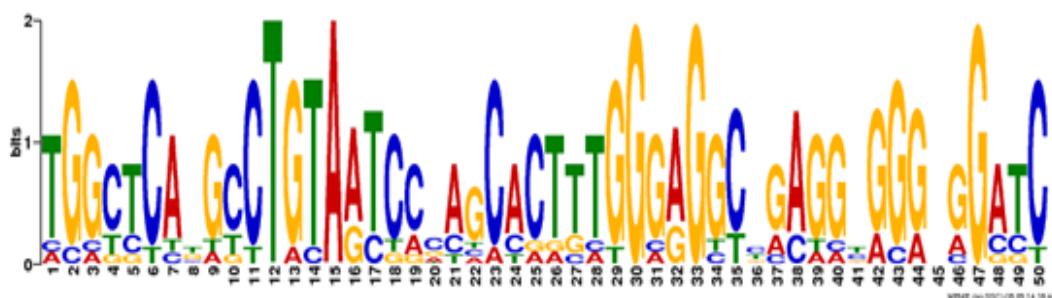
高级分析



挑选显著差异的 CpG 甲基化位点作为样本分类的标志物。



基因的甲基化程度和表达值之间的关系图。基因的甲基化程度越高，基因的表达值反而下降。



筛选差异甲基化区域在 TSS 200-TSS1500 范围内的基因，根据序列预测相应的 motif。

伯豪案例

应用案例：miR-146a/b 通过调控 UHRF1 基因来调节胃癌的入侵及迁移

研究者通过分析 miR-146a/b 以及 UHRF1 基因在人类胃癌组织中的含量发现 miR-146a/b 与 UHRF1 基因的表达呈负相关。该发现描绘了一幅 UHRF1 基因调控的新机制及胃癌组织中异常 DNA 甲基化的蓝图，新发现的 miR-146a/b/UHRF1 路线为胃癌迁移灶发生的研究提供了新思路，同时也为阻止胃癌迁移提供了新的治疗方式。

原文出处: Regulation of UHRF1 by miR-146a/b modulates gastric cancer invasion and metastasis.

FASEB J. 2013 Aug 27. (IF5.704) 使用服务类型: Infinium HumanMethylation450 BeadArray

(七) lncRNA 芯片服务

长链非编码 RNA(long noncoding RNA, lncRNA) 指的是长度在 200-100000 nt 之间的 RNA 分子。它们不编码蛋白，但参与细胞内多种过程调控。近年来的研究表明，lncRNA 参与了 X 染色体沉默，基因组印记以及染色质修饰，转录激活，转录干扰，核内运输等多种重要的调控过程，lncRNA 的这些调控作用也开始引起人们广泛的关注。lncRNA 的种类远远超过编码 RNA，哺乳动物基因组序列中 4%~9% 的序列产生的转录本是 lncRNA (相应的蛋白编码 RNA 的比例是 1%-5%)。lncRNA 已经成为非编码 RNA 研究领域的一个热点。

上海伯豪针对广大客户的需求，开发出全新的 lncRNA 芯片，并利用 Agilent 公司专利的 SurePrint 技



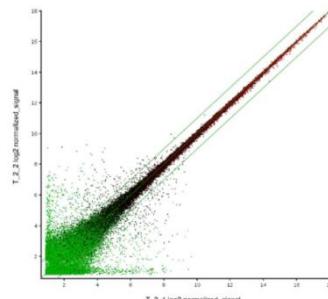
术制造。为您提供完善的基因芯片技术服务，并完成数据分析，提供完整的实验报告。

服务推荐

	人 Human(V5.0) New!	大鼠 Rat	小鼠 Mouse
格式	4×180K	4×44K	4×180K
lncRNA 数目	63,431	9,843	33,231
mRNA 数目	39,887	30,451	39,430
探针数目	175,906	40,344	151,045
探针长度	60 nt	60 nt	60 nt
lncRNA 数据库	GENCODE v17 (7,346) LNCipedia v2.1 (19,324) Lncrnadb (10) Noncoder v3 (14,382) Ensembl (2,275) NCBI (4,866) UCSC (4,029) HTA2.0 (403) Agilent G3 v2 (10,796)	NCBI RefSeq (660) UCSC (6067) UCR (462) Homolog (2755)	NCBI RefSeq (1,394) UCSC (3,616) ENSEMBL (1,055) FANTOM (26,171) UCR (910) LNCRNA-DB (85)
mRNA 数据库	SurePrint G3 Human Gene Expression 8×60K 中的 mRNA 探针	Agilent 4×44K Whole Rat Genome Oligo Microarray 中的 mRNA 探针	SurePrint G3 Mouse Gene Expression 8×60K 中的 mRNA 探针

技术特点

- ✓ 目前市面上探针密度最高的一款 lncRNA 芯片产品；
- ✓ 芯片采用 Agilent eArray 平台设计，使用 Agilent SurePrint 技术合成，技术重复性高于 99%；
- ✓ 覆盖主流数据库所有 lncRNA，可同时检测 lncRNA 和 mRNA，并挖掘两者之间的关联；
- ✓ 检出率高：随机引物扩增方式，可同时扩增带 Poly-A 和不带 Poly-A 的 lncRNA，确保更多的 lncRNA 被检测出来；
- ✓ 提供完整 lncRNA 芯片数据分析解决方案



技术重复散点图

服务相关

样品要求

1. 样品纯度：OD 260/280 值应在 1.8~2.1 之间；DNA 污染应该去除干净；
2. 样品浓度：浓度不低于 40ng/ μ l；
3. 样品总量：每个样品总量不少于 2 μ g；
4. 样品溶剂：溶解在水中；
5. 样品运输：RNA 干冰低温运输；在运输过程中请用 parafilm 将管口密封好，以防出现污染。

服务流程



数据分析

归一化

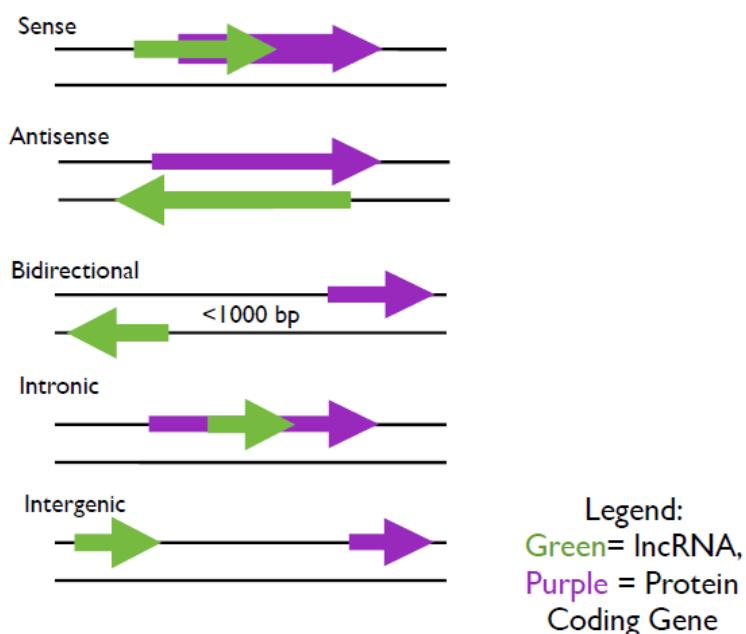
lncRNA 芯片采用的归一化的方法为 quantile normalization。

差异 lncRNA 的筛选

lncRNA 芯片中既有 lncRNA 的探针又有 mRNA 的探针，分别做差异基因的筛选，筛选方法同表达谱的筛选方法是一致的，参见表达谱的差异基因筛选。

差异 lncRNA 分类

lncRNA 与蛋白编码基因的染色体位置关系，往往预示了 lncRNA 的功能。根据 lncRNA 与相邻编码蛋白基因的位置关系进行分类。可以分为：1) Sense-overlapping：与 mRNA 外显子有重叠区，转录方向一致；2) Antisense-overlapping：与 mRNA 外显子有重叠区，转录方向相反；3) Bidirectional：与 mRNA 转录方向相反，转录起始位置距离在 1kb 内；4) Intronic：lncRNA 完全落编码蛋白基因的内含子中；5) Intergenic：lncRNA 位于编码蛋白基因的基因间区。



差异 lncRNA 靶基因的预测

lncRNA 可能通过调控 mRNA 发挥功能，因此有必要预测 lncRNA 的靶基因。

lncRNA 的靶基因预测分为 cis 预测与 trans 预测。Cis 预测的原理是找出 lncRNA 所处染色体位置上下游 10 kb 范围内的蛋白编码基因。Trans 预测先采用 blast 选出序列上与 lncRNA 互补 mRNA 序列，再利用 RNAPlex 计算两序列之间的互补能量，筛选出可能稳定结合的 mRNA。

差异 lncRNA 的重注释

lncRNA 芯片注释不完善，因此需要将筛选出来的 lncRNA 进行重注释。将差异 lncRNA 在基因组上位置上下游延伸，以寻找 lncRNA 附近的有功能的基因。

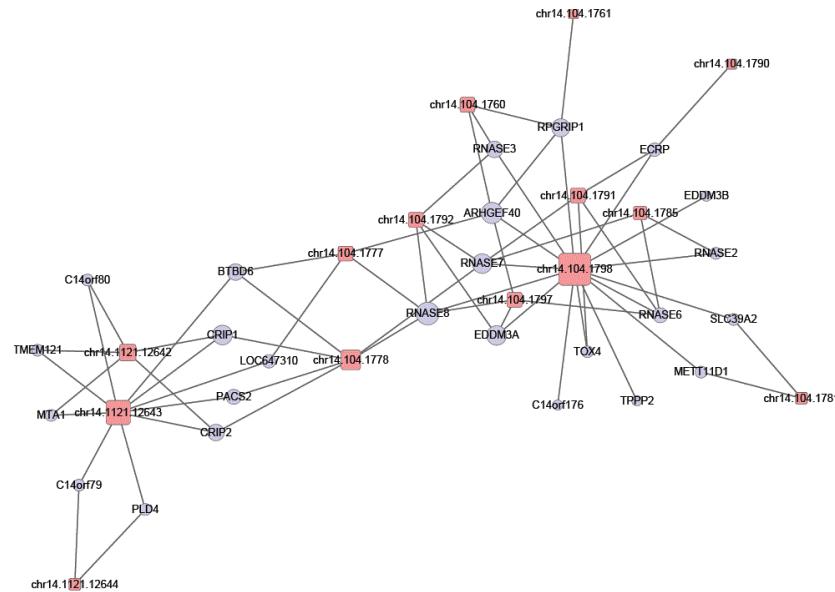
差异 lncRNA 靶基因的预测

lncRNA 可能通过调控相应的 mRNA 发挥功能，因此有必要预测 lncRNA 的靶基因。我们提取差异 lncRNA 和 mRNA 的序列，首先用 blast 进行初筛，之后用 RNAPlex 进行进一步筛选，以预测 lncRNA 可能调控的 mRNA。

差异 lncRNA 与靶基因共表达网络

预测出 lncRNA 的靶基因后，并可进一步在 mRNA 的数据中探寻该 mRNA 是否发生表达量的变化。由此构建差异 lncRNA 与靶基因相互作用网络图。

差异 lncRNA 与靶基因相互作用网络图。方框代表 lncRNA，圆形代表 mRNA。连线表示可能的调控关系。节点面积越大，表示调控的 mRNA 越多，预示该 lncRNA 在调控网络中所起的作用可能越大。

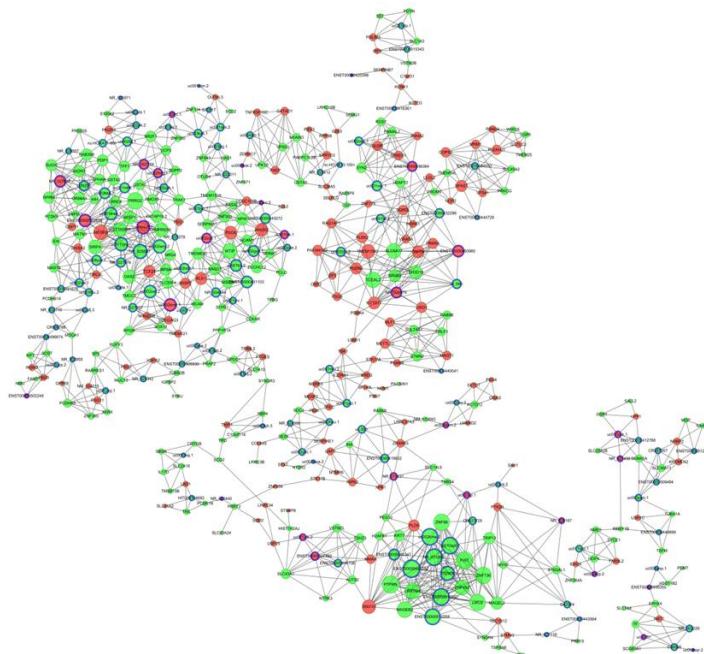


差异 lncRNA 与差异 mRNA 的共表达分析

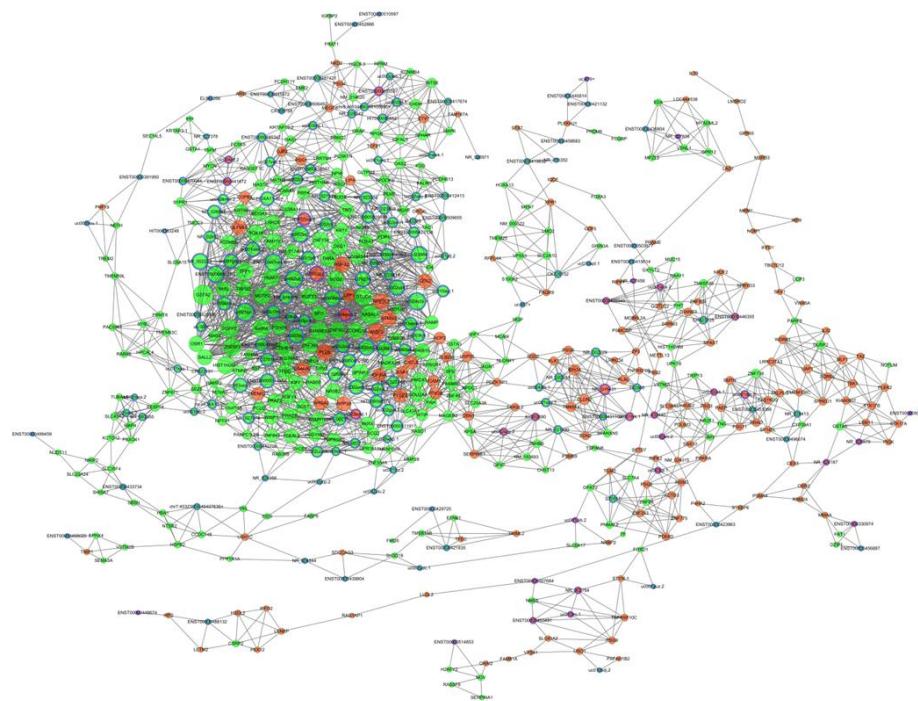
SBC Human lncRNA 芯片能同时检测出差异表达的 lncRNA 和 mRNA。我们将差异 lncRNA 和差异 mRNA 在一组样品中进行共表达分析，可以发现与某个 lncRNA 具有相同表达模式的 mRNA。

要求：每组数据 3 个或 3 个以上生物学重复

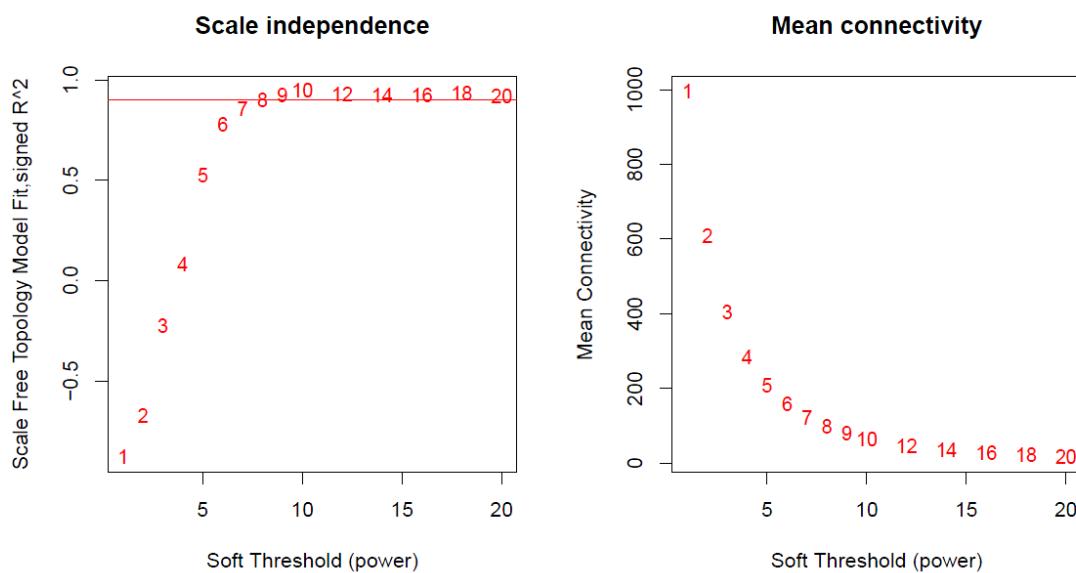
实验组：

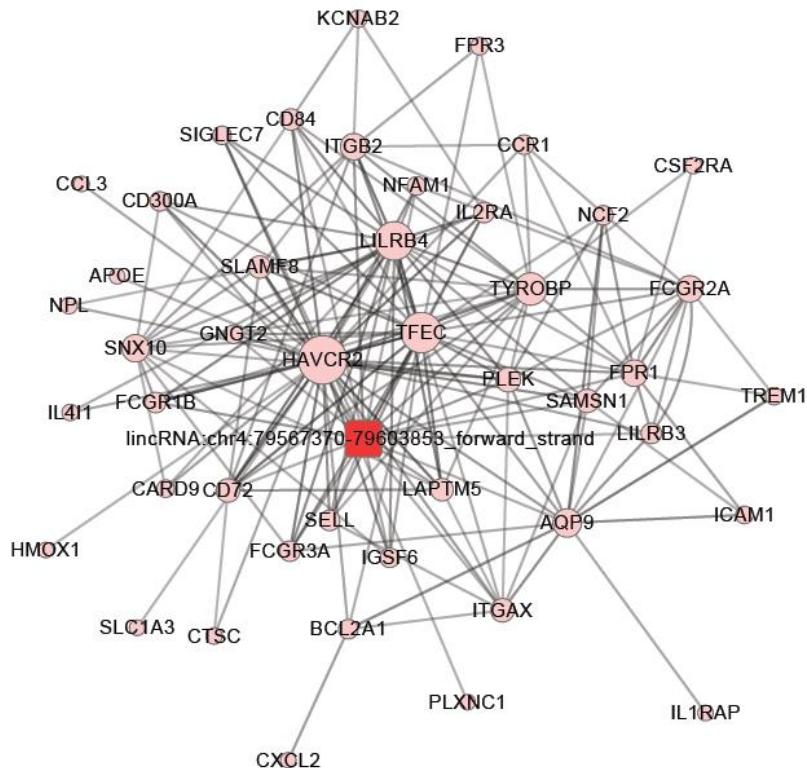


对照组：



lncRNA 与 mRNA 共表达分析作用图，圆形带圈代表 lncRNA，圆形代表 mRNA。红色为上调基因，绿色为下调基因。





差异 lncRNA 靶基因的 GO analysis

对 lncRNA 的靶基因进行 GO Ontology 的生物学的分类,根据 Fisher's Exact Test,得到 p-value,得到 lncRNA 靶基因对应的显著性功能,从而了解 lncRNA 的功能。

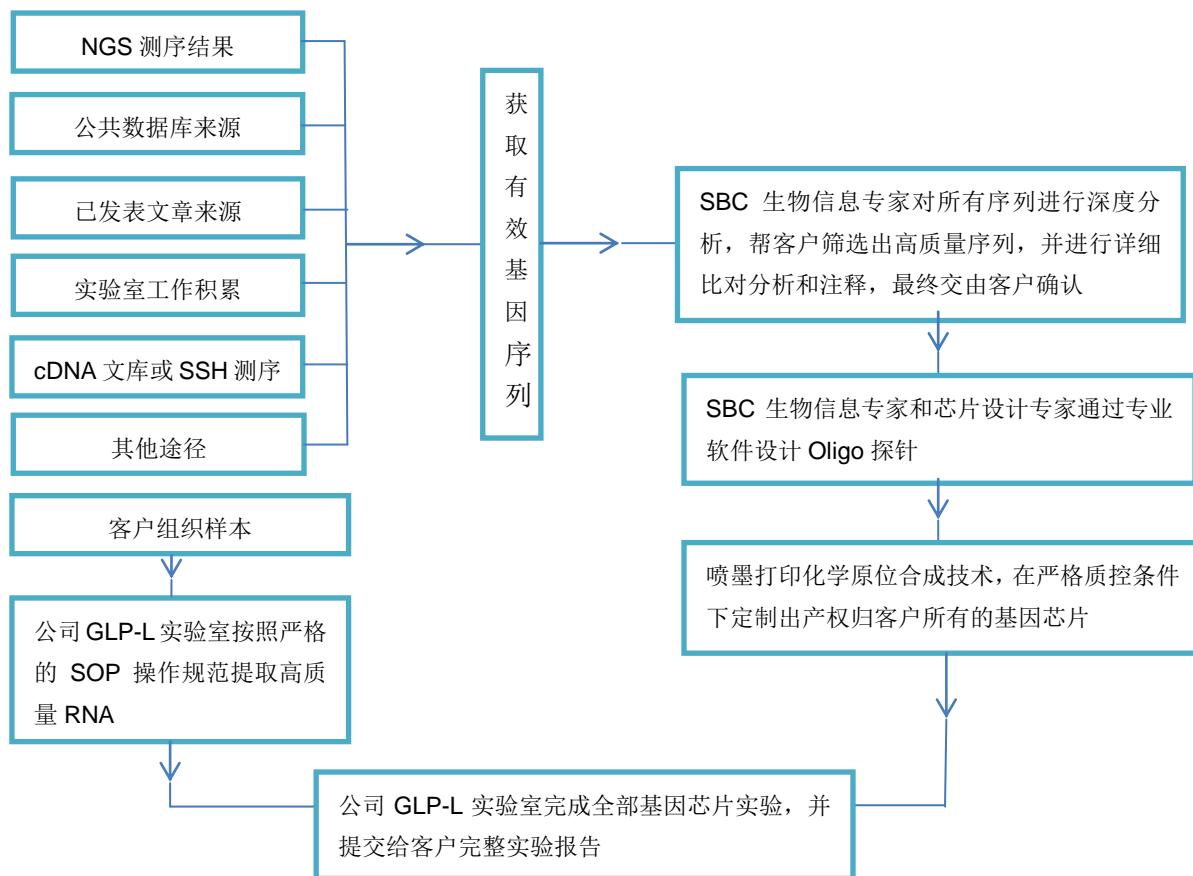
差异 lncRNA 靶基因的 pathway analysis

对 lncRNA 靶基因按照 Pathway 的主要公共数据库 KEGG 和 Biocarta 来进行分类,对 Pathway 中的基因进行基于离散分布的显著性分析,得到与实验目的有显著联系的 Pathway 分类,由这些 pathway 对应相应的靶基因,从而获得该分类即导致 lncRNA 差异的最重要 Pathway。

差异 lncRNA 的转录因子的预测

提取 lncRNA TSS 的上游 2000bp, 下游 500bp, 利用 HMM 的算法根据 TRANSFAC8.1 数据库预测其转录因子。

(八) 定制芯片服务



服务特色及服务流程

✓ 强大的 e-array 软件辅助:

由 e-array 软件支持, SBC 生物信息专家的帮助, 可以对任何已知基因组序列或 EST 序列的物种进行精准探针设计。

✓ 国际化制作标准:

利用原位合成技术进行芯片制作, 确保芯片质量严格遵循 Agilent 公司芯片生产质量标准, 确保与商业化芯片完全一致的品质。

✓ 流程简化、成本降低:

无需预先合成探针, 无需花费成本进行繁琐的摇菌、抽质粒、PCR、纯化等步骤。

更加自由: 不受克隆限制, 没有克隆, 照样制作。

✓ 不受物种限制:

可以对任何有背景信息的物种设计制作芯片并完成服务。

✓ 种类繁多:

可定制表达谱芯片、CGH 芯片、ChIP-on Chip 芯片等。

✓ 多种芯片格式可供选择:

HD 平台 (1x244K、2x105K、4x44K、8x15K) 和 G3 平台 (1x1M、2x400K、4x180K、8x60K)



SBC 已经完成的 e-array 物种清单：

动物：羊、鸡、猪、壁虎、家蚕、对虾、清鳞鱼、天牛、线虫、血吸虫、涡虫、朱砂叶螨、白符跳、淡水珍珠蚌、马氏珠母贝.....

植物：小麦、水稻、西瓜、柑橘、月季、茄科、玉米、甘蔗、牡丹、短柄三叶草、红花、芦笋、麻风树、山核桃、刺五加、百脉根、烟草、草菇、当归、雨生红球藻.....

微生物：烟曲霉、乳酸杆菌、丙酮丁醇梭菌、猪链球菌、马尔尼菲青霉菌、布鲁氏菌、拜氏梭菌、胸膜肺炎放线杆菌、克雷伯氏肺炎菌、沙雷菌、希瓦氏菌、耻垢分枝杆菌、副猪嗜血杆菌、谷氨酸棒杆菌、禾谷镰刀菌、结核杆菌、绿脓杆菌、沙门氏杆菌、水稻纹枯菌、苏云金芽孢杆菌、铜绿假单孢菌、弯曲杆菌、阿维链霉菌、天蓝色链霉菌、林可链霉菌、始旋链霉菌、集胞藻.....

(九) 在线服务

向伯豪用户开放的个性化实时在线 表达谱芯片分析系统

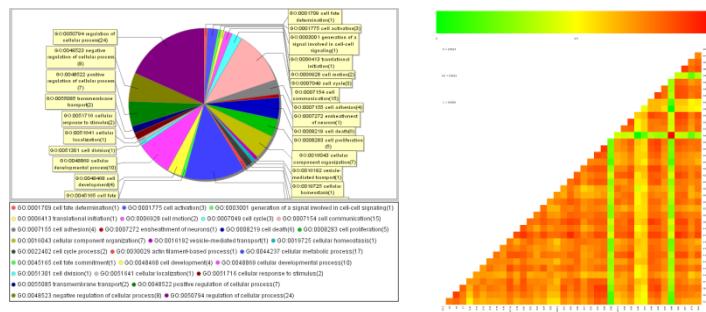
SBC Analysis System (SAS 系统) 是上海伯豪生物技术有限公司推出的一款在线的表达谱芯片数据分析软件，旨在为客户提供更快速便捷的数据分析工具，您在装有 IE 浏览器的计算机上即可通过客户端进行数据分析。

SAS 系统的统计计算与数据库查询均在服务器上完成，且运行结果都保存在伯豪公司服务器上，不会占用用户计算机的空间和内存。用户可以随时上网查看自己的数据，也可以将数据下载到自己的计算机上保存。

该系统核心算法基于 R 语言(R-software)程序包，可以完成大部分的表达谱分析运算，系统本地整合 8 大公共数据库，实现对多个数据库的交叉查询，帮助您发现实验揭示的生物学意义。

SAS 系统即将升级至 2.0 版本，该在线分析系统将帮助广大科研工作者表达谱芯片进行以下各类数据管理、统计学分析和功能注释工作：

数据管理	数据分析		功能注释
自行上传数据	计算	差异基因筛选 T-test 分析 方差分析	基因信息检索 染色体定位图
文件夹及文件编辑		散点图分析 火山图分析 主成分作图分析 相关性作图分析 样本单项聚类图分析 聚类热图分析 SOM 图分析 GO 聚类图分析 柱状图分析 韦恩图分析	动植物 GO 分类检索 Pathway 检索 蛋白相互关系图 MicroRNA 靶基因预测 启动子预测 转录因子预测 变异检索 (SNP、CNV、InDel) 蛋白信息检索
初始数据筛选	作图		
结果数据精选			

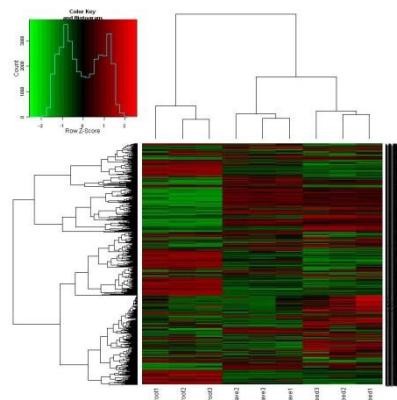


GO 富集分析

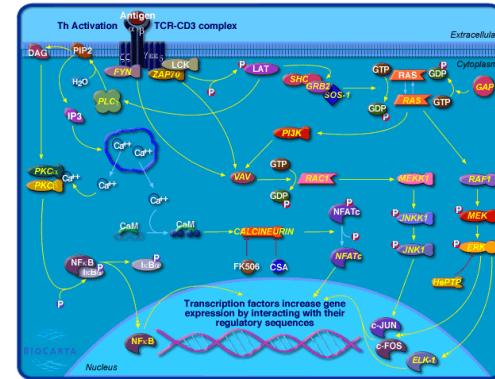
相关性分析

<i>p</i> 值	差异倍数	<i>q</i> 值	均值	均值	中位数	中位数	标准差	标准差	
probeset	pvalues	foldchange	qvalues	mean1	mean2	median1	median2	stdvar1	stdvar2
A_24_P627	0.013	3.144	0.008	12.932	14.584	12.651	14.654	0.494	0.438
A_32_P778	0.001	16.674	0.000	4.531	8.591	4.359	8.583	0.467	0.206
A_24_P9373	0.044	2.582	0.014	7.127	8.495	7.250	8.444	0.594	0.228
A_24_P841	0.009	2.267	0.014	7.759	8.941	7.800	8.885	0.183	0.314
A_32_P893	0.002	0.518	0.024	7.171	6.221	7.207	6.146	0.109	0.156
A_23_P662	0.093	0.332	0.008	5.394	3.804	5.319	3.968	0.144	0.925
A_22_P169	0.069	2.022	0.024	3.224	5.279	4.002	4.922	0.744	0.806

上海伯豪生物技术有限公司的在线分析系统 DiffGene 分析运算的结果表格



层次聚类



Pathway 分析

个性化实时在线分析系统—SBC Analysis System (SAS 系统)

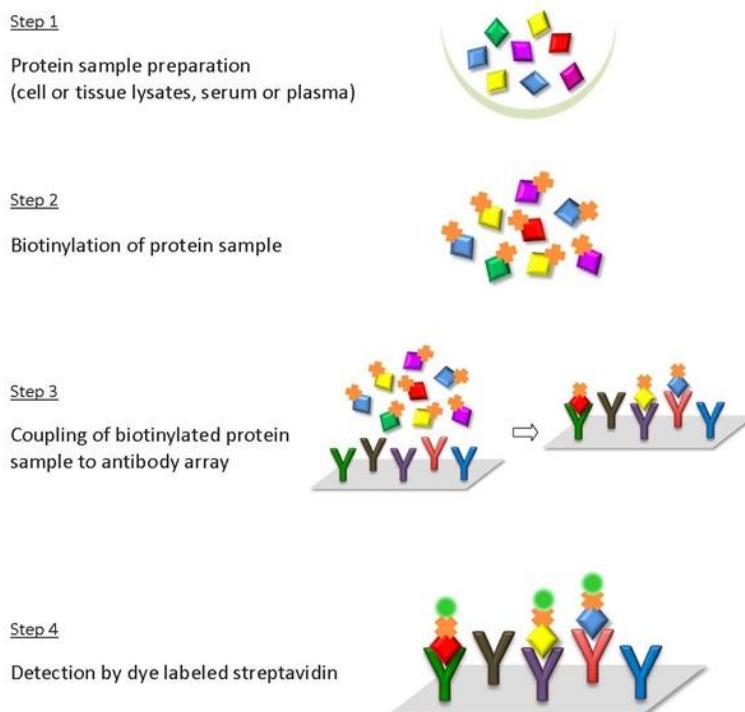
(十) 蛋白质抗体芯片

蛋白质芯片技术是继基因芯片技术后发展起来的生物检测技术,是蛋白质组学研究中除了酵母双杂交、双向电泳技术、质谱技术等之外的一种重要的工具。蛋白质芯片技术具有平行、快速、自动化的优点,并且检测到的平行数据误差更小,更准确,这对于高通量的研究具有非常重要的意义。它的出现对于生物学、临床检验医学、遗传学、肿瘤学等多学科的进步具有极大的推动作用。蛋白质芯片技术根据芯片上固定的物质不同又分为抗体芯片与蛋白芯片。

近几年来，随着抗体技术的发展，抗体库的建设速度越来越快，生产了大量优质的抗体，从而推动了抗体芯片技术的发展与应用。抗体芯片技术在医学领域的应用，带动了相关科技论文的发表，数量与质量均呈上升趋势。

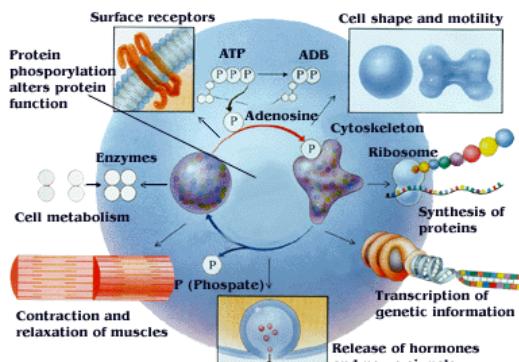
上海伯豪生物技术有限公司引进国外知名抗体芯片技术公司(Full Moon Biosystems)公司的技术和产品，建立了丰富的抗体芯片技术平台，推出的43余种抗体芯片可以用于检测生理或病理过程相关蛋白的表达丰度、筛选疾病标志物、探索药物作用机制、检测信号通路中成员的激活状态。使用的芯片产品已经在Nature Medicine, Journal of Clinical Investigation, Cancer Research, PLoS One, Journal of Biological Chemistry等专业学术杂志发表了高质量的论文。可以检测的样本包括培养的细胞、细胞培养上清、血清、血浆、冰冻组织、以及福尔马林浸泡或者Trizol浸泡的组织，石蜡包埋组织和切片也可以进行应用。

Full Moon 蛋白芯片检测原理



服务内容一：信号通路磷酸化水平检测

可逆的蛋白磷酸化反应调节着细胞的大部分功能。蛋白质的可逆磷酸化是细胞中普遍存在的一类酶促共价键化学修饰反应，通过可逆的磷酸化反应，导致蛋白质活性和功能的改变，从而影响和调节着细胞的各种功能，如能量储存、细胞形态变化、蛋白质的合成、基因表达、信号因子的释放、肌肉收缩以及细胞代谢等。



可提供技术服务芯片

产品目录号 (Cat.No.)	Full Moon 磷酸化水平检测抗体芯片	抗体数目
PEX100	Phospho Explorer Antibody Array (磷酸化探索者)	1318
PAA137	AKT Phospho Antibody Array	137
PAB216	AKT/PKB Phospho Antibody Array	216
PAM174	AMPK Signaling Phospho Antibody Array	174
PAP247	Apoptosis Phospho Antibody Array (细胞凋亡磷酸化)	247
PAC155	Cancer Apoptosis Phospho Antibody Array (肿瘤相关细胞凋亡磷酸化)	155
PCS248	Cancer Signaling Phospho Antibody Array (肿瘤相关信号通路)	269



PCC238	Cell Cycle Control Phospho Antibody Array (细胞周期调控因子磷酸化)	238
PCC076	Cell Cycle Phospho Antibody Array (细胞周期磷酸化)	76
PCT173	Chromatin/Transcription Phospho Antibody Array (染色质、转录调控磷酸化)	173
PCR174	CREB Phospho Antibody Array	174
PCP141	Cytoskeleton Phospho Antibody Array (细胞骨架磷酸化)	141
PEG214	EGF Pathway Phospho Antibody Array	214
PER239	ErbB/Her Signaling Phospho Antibody Array	239
PEK208	ERK Signaling Phospho Antibody Array	208
PGF169	FGF Phospho Antibody Array	169
PGP193	GPCR Signaling To MAPK/ERK Phospho Antibody Array	193
PIR245	IGF1R Phospho Antibody Array	245
PIG219	Insulin Receptor Phospho Antibody Array (胰岛素受体磷酸化抗体芯片)	219
PJS042	Jak/Stat Phospho Antibody Array	42
PJS202	Jak/Stat II Phospho Antibody Array	202
PMK185	MAPK Phospho Antibody Array	185
PMT138	mTOR Phospho Antibody Array	138
PNS170	Neuroscience Phospho Antibody Array (神经生物学磷酸化水平)	167
PNK215	NFkB Phospho Antibody Array	215
PNR052	Nuclear/Membrane Receptor Phospho Antibody Array (核、膜受体磷酸化检测)	52
PFT196	p53 Signaling Phospho Antibody Array	196
PDG195	PDGF Phospho Antibody Array	195
PTC188	T-Cell Receptor Phospho Antibody Array	188
PTG176	TGF-beta Phospho Antibody Array	176
PTK098	Tyrosine Kinase Adaptor Phospho Antibody Array	98
PST228	Tyrosine Phosphorylation ProArray	228
PVE185	VEGF Pathway Phospho Antibody Array	185
PNT227	Wnt Signaling Phospho Antibody Array	227

服务内容二：蛋白质表达水平检测

可提供技术服务芯片（相对定量）

产品目录号 (Cat.No.)	Full Moon 蛋白表达水平检测抗体芯片	抗体数目
APP069	Apoptosis Antibody Array (细胞凋亡)	73
SCB200	Cancer BioMarker Antibody Array (肿瘤标记物)	247
ACC058	Cell Cycle Antibody Array (细胞周期)	60



SCK100	Cytokine Profiling Antibody Array (细胞因子检测)	310
ASB600	Explorer Antibody Array (探索者芯片)	656
AVK276	Kinase Antibody Array (激酶)	276
SMK060	MAPK Signaling Antibody Array (MAPK 信号通路)	63
AST160	Signal Transduction Antibody Array (信号转导)	165
SET100	Signaling Explorer Antibody Array (信号因子探索者)	1358

服务相关

样品要求

◆ 细胞

约 1×10^7 个贴壁细胞去培养基后，常温 PBS 洗涤 3 次，直接加入裂解缓冲液，推荐使用 Full Moon 的 Protein Extraction Buffer 或者 RIPA Lysis Buffer，根据培养皿大小加入适量缓冲液，反复吹打，待细胞裂解后，收集溶液，4℃离心后收集上清，置-70℃冰箱保存，干冰运输。

悬浮培养细胞直接收集，收集约 1×10^7 个细胞，常温 PBS 洗涤 3 次，低速离心沉淀细胞，弃去 PBS (尽量去除干净)，直接加入裂解缓冲液，推荐使用 Full Moon 的 Protein Extraction Buffer 或者 RIPA Lysis Buffer，根据细胞数量加入适量缓冲液，按操作流程裂解后，4℃离心后收集上清，置-70℃冰箱保存，干冰运输。

◆ 组织

采集新鲜组织 (注意：生物体血液限流后 20 分钟内取材料并保存好)，组织块以 PBS 或生理盐水清洗干净 (尽量洗净血液)，切为小块，放入液氮速冻后-70℃保存，干冰运输，或福尔马林/TriZol 中 4℃保存，一周内，常温运输。

◆ 血清和血浆

采集新鲜血液，切忌溶血。全血分离血清，或 EDTA 抗凝血分离血浆，取 100-500 微升置-20℃或-70℃冰箱保存，干冰运输。

服务流程



伯豪案例

CXCL5(epithelial neutrophil-activating peptide-78)促进中性粒细胞浸润并可作为肝癌预后的判断指标

复旦大学中山医院肝癌研究所癌变与侵袭教育部重点实验室，复旦大学生物医学研究院的研究人员证实趋化因子 CXCL5 在肝癌的肿瘤生长、侵袭及预后中发挥重要的作用。研究人员通过比较 CXCL5 在几组具有不同转移性能的肝癌细胞系和 919 名肝癌患者体内 mRNA 和蛋白水平的变化，结合磷酸化抗体芯片这种高通量的实验平台发现 CXCL5 主要通过激活 PI3K-Akt 和 ERK1/2 两个信号通路促进肝癌的增殖、转移和侵袭。另一方面，在体外实验发现 CXCL5 对中性粒细胞具有直接趋化作用，而体内实验也证明肝癌细胞中的 CXCL5 促进肿瘤的生长、肺转移和瘤内中性粒细胞浸润。免疫组化分析表明，肝癌中 CXCL5 的过



表达与中性粒细胞的浸润、患者总生存期短以及肿瘤复发相关。而多变量分析发现，单独的 CXCL5 的过表达或其与瘤内中性粒细胞的存在状态相结合，都可以作为一个新的预后衡量指标，同时 CXCL5 也可以作为肝癌一个新的治疗靶点。

原文出处： Overexpression of CXCL5 Mediates Neutrophil Infiltration and Indicates Poor Prognosis for Hepatocellular Carcinoma. Hepatology. 2012, 56(6) :2242-54. (IF10.8) 使用服务种类：蛋白芯片

微阵列芯片服务应用文章精选 (2010-2014)

- [1] Liu X, Sun H, Qi J, Wang L, He S, Liu J, Feng C, Chen C, Li W, Guo Y,)Qin D, Pan G, Chen J, Pei D, Zheng H. Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT-MET mechanism for optimal reprogramming. Nat Cell Biol. 2013, 15(7):829-38. (IF20.761) 使用服务种类：Agilent 小鼠表达谱芯片服务
- [2] Han W, Xin Z, Zhao Z, Bao W, Lin X, Yin B, Zhao J, Yuan J, Qiang B, Peng X. RNA-binding protein PCBP2 modulates glioma growth by regulating FHL3. J Clin Invest. 2013, 123(5):2103-18. (IF13.069) 使用服务种类：Agilent 人表达谱芯片服务
- [3] Cai J, Guan H, Fang L, Yang Y, Zhu X, Yuan J, Wu J, Li M. MicroRNA-374a activates Wnt/β-catenin signaling to promote breast cancer metastasis. J Clin Invest. 2013 Jan 16. (IF13.069) 使用服务种类：Agilent Human miRNA 芯片服务
- [4] Yan XL, Jia YL, Chen L, Zeng Q, Zhou JN, Fu CJ, Chen HX, Yuan HF, Li ZW, Shi L, Xu YC, Wang JX, Zhang XM, He LJ, Zhai C, Yue W, Pei XT. Hepatocellular carcinoma-associated mesenchymal stem cells promote hepatocarcinoma progression: Role of the S100A4-miR155-SOCS1-MMP9 axis. Hepatology. 2013 May 1. (IF11.665) 使用服务种类：Agilent 人表达谱芯片服务
- [5] Wang S, Wang L, Bayaxi N, Li J, Verhaegh W, Janevski A, Varadan V, Ren Y, Merkle D, Meng X, Gao X, Wang H, Ren J, Kuo WP, Dimitrova N, Wu Y, Zhu H. A microRNA panel to discriminate carcinomas from high-grade intraepithelial neoplasms in colonoscopy biopsy tissue. Gut. 2013, 62(2):280-9. (IF 10.111) 使用服务类型：Agilent Human miRNA 芯片服务
- [6] Li W, Shuai L, Wan H, Dong M, Wang M, Sang L, Feng C, Luo GZ, Li T, Li X, Wang L, Zheng QY, Sheng C, Wu HJ, Liu Z, Liu L, Wang L, Wang XJ, Zhao XY, Zhou Q. Androgenetic haploid embryonic stem cells produce live transgenic mice. Nature. 2012, 490(7420):407-11. (IF36.280) 使用服务种类：Affymetrix GeneChip Mouse Genome 430 2.0 芯片服务
- [7] Huang W, Hu J, Yang DW, Fan XT, Jin Y, Hou YY, Wang JP, Yuan YF, Tan YS, Zhu XZ, Bai CX, Wu Y, Zhu HG, Lu SH. Two MicroRNA Panels to Discriminate Three Subtypes of Lung Carcinoma in Bronchial Brushing Specimens. Am J Respir Crit Care Med. 2012 Oct 4. (IF11.080) 使用服务种类：Agilent Human miRNA 芯片服务
- [8] Zhou SL,Dai Z, Zhou ZJ, Wang XY, Yang GH, Wang Z,Huang XW , Fan J,Zhou J. Overexpression of CXCL5 Mediates Neutrophil Infiltration and Indicates Poor Prognosis for Hepatocellular Carcinoma. Hepatology. 2012 :56(6):2242-54.(IF 11.19) 使用服务种类：蛋白芯片
- [9] Zhu S, Pan W, Song X, Liu Y, Shao X, Tang Y, Liang D, He D, Wang H, Liu W, Shi Y, Harley JB, Shen N, Qian Y. The microRNA miR-23b suppresses IL-17-associated autoimmune inflammation by targeting TAB2, TAB3 and IKK-α. Nat Med. 2012 Jun 3. (IF22.464) 使用服务种类：Affymetrix GeneChip Human Genome U133 Plus 2.0 芯片服务
- [10] Yang H, Shi LY, Wang BA, Liang D, Zhong CQ, Liu W, Nie YZ, Liu J, Zhao J, Gao X, Li DS, Xu GL,



- Li JS. Generation of Genetically Modified Mice by Oocyte Injection of Androgenetic Haploid Embryonic Stem Cells. *Cell.* 2012, 149(3):605-617. (IF 32.403) 使用服务种类: Affymetrix Mouse Genome 430 2.0 Array 服务
- [11] Huang Y, Li W, Yao X, Lin QJ, Yin JW, Liang Y, Heiner M, Tian B, Hui J, Wang G. Mediator Complex Regulates Alternative mRNA Processing via the MED23 Subunit. *Mol Cell.* 2012 Jan 18. (IF14.148) 使用服务种类: Affymetrix GeneChip Human Exon 1.0 ST Array 服务
- [12] Jiang L, Lin C, Song L, Wu J, Chen B, Ying Z, Fang L, Yan X, He M, Li J, Li M. MicroRNA-30e* promotes human glioma cell invasiveness in an orthotopic xenotransplantation model by disrupting the NF- κ B/I κ B α negative feedback loop. *J Clin Invest.* 2012, 122(1):33-47. (IF13.069) 使用服务种类: Agilent Human miRNA 芯片服务
- [13] Song L, Dai T, Xie Y, Wang C, Lin C, Wu Z, Ying Z, Wu J, Li M, Li J. Upregulation of miR-1245 by c-myc targets BRCA2 and impairs DNA repair. *J Mol Cell Biol.* 2012, 4(1):108-17. (IF13.400) 使用服务种类: Agilent 人 CGH 芯片 (4×180K) 服务
- [14] Zhou J, Yu L, Gao X, Hu J, Wang JP, Dai Z, Wang JF, Zhang ZY, Lu SH, Huang XW, Wang Z, Qiu SJ, Wang XY, Yang GH, Sun HC, Tang ZY, Wu Y, Zhu HG, Fan J. Plasma MicroRNA Panel to Diagnose Hepatitis B Virus-Related Hepatocellular Carcinoma. *J Clin Oncol.* 2011, 29(33):38.2697. (IF18.970) 使用服务种类: Agilent Human miRNA 芯片服务
- [15] The China Consortium for the Genetics of Autoimmune Thyroid Disease. A genome-wide association study identifies two new risk loci for Graves' disease. *Nature Genetics.* 2011, 43(9):897-901. (IF36.377) 使用服务种类: Illumina Human660-Quad BeadChips 服务
- [16] Jiang J, Ding G, Lin J, Zhang M, Shi L, Lv W, Yang H, Xiao H, Pei G, Li Y, Wu J, Li J. Different developmental potential of pluripotent stem cells generated by different reprogramming strategies. *J Mol Cell Biol.* 2011, 3(3):197-9. (IF13.400) 使用服务种类: Affymetrix Mouse Genome 430 2.0 Array 服务
- [17] Huang PY, He ZY, Ji SY, Sun HW, Xiang D, Liu CC, Hu YP, Wang X, Hui LJ. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature.* 2011, 475(7356):386-9. (IF36.101) 使用服务种类: Agilent 小鼠表达谱芯片服务
- [18] Li KC, Zhang FX, Li CL, Wang F, Yu MY, Zhong YQ, Zhang KH, Lu YJ, Wang Q, Ma XL, Yao JR, Wang JY, Lin LB, Han M, Zhang YQ, Kuner R, Xiao HS, Bao L, Gao X, Zhang X. follistatin-like 1 Suppresses Sensory Afferent Transmission by Activating Na(+),K(+)-ATPase. *Neuron.* 2011, 69(5):974-87. (IF14.027) 使用服务种类: SBC 定制芯片服务
- [19] Fang H, Yang Y, Li C, Fu S, Yang Z, Jin G, Wang K, Zhang J, Jin Y. Transcriptome analysis of early organogenesis in human embryos. *Dev Cell.* 2010, 19(1):174-84. (IF13.363) 使用服务种类: Affymetrix GeneChip Human Genome U133A Array 服务
- [20] Wang K, Liu J, Yan ZL, Li J, Shi LH, Cong WM, Xia Y, Zou QF, Xi T, Shen F, Wang HY, Wu MC. Overexpression of aspartyl-(asparaginyl)-beta-hydroxylase in hepatocellular carcinoma is associated with worse surgical outcome. *Hepatology.* 2010, 52(1):164-73. (IF10.840) 使用服务种类: SBC Human cDNA 表达谱芯片服务
- [21] Wang K, Wang P, Shi J, Zhu X, He M, Jia X, Yang X, Qiu F, Jin W, Qian M, Fang H, Mi J, Yang X, Xiao H, Minden M, Du Y, Chen Z, Zhang J. PML/RARalpha Targets Promoter Regions Containing PU.1 Consensus and RARE Half Sites in Acute Promyelocytic Leukemia. *Cancer Cell.* 2010, 17(2):186-197. (IF25.288) 使用服务种类: Affymetrix Human Promoter 1.0R Arrays 服务 / Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 服务



- [22] Huang J, Zheng DL, Qin FS, Cheng N, Chen H, Wan BB, Wang YP, Xiao HS, Han ZG. Genetic and epigenetic silencing of SCARA5 may contribute to human hepatocellular carcinoma by activating FAK signaling. *J Clin Invest.* 2010, 120(1):223-41. (IF15.387) 使用服务种类: Agilent Whole Human Genome Oligo Microarray 服务
- [23] Fei T, Xia K, Li Z, Zhou B, Zhu S, Chen H, Zhang J, Chen Z, Xiao H, Han JD, Chen YG. Genome-wide mapping of SMAD target genes reveals the role of BMP signaling in embryonic stem cell fate determination. *Genome Res.* 2010, 20(1):36-44. (IF11.342) 使用服务种类: Agilent Promoter Array & Affymetrix Mouse 430 2.0 Array 服务
- [24] Chen C, Chen H, Lin Y-S, Shen J-B, Shan J-X, Qi P, Shi M, Zhu MZ, Huang XH, Feng Q, Han B, Jiang LW, Gao JP, Lin HX. A two-locus interaction causes interspecific hybrid weakness in rice. *Nat Commun.* 2014;5. (IF10.015) 使用服务种类: Affymetrix 水稻表达谱芯片服务
- [25] [24] Zheng JJ, Li SJ, Zhang XD, Miao WY, Zhang D, Yao HS, Y X. Oxytocin mediates early experience-dependent cross-modal plasticity in the sensory cortices. *Nature neuroscience.* 2014. (IF 15.251) 使用服务种类: Oxytocin mRNA microarray 服务
- [26] Huang PY, Zhang LD, Gao YM, He ZY, Yao D, Wu ZT, Cen J, Chen XT, Liu CC, Hu YP, Lai DM, Hu ZL, Chen L, Zhang Y, Cheng X, Ma XJ, Pan GY, Wang X, Hui LJ. Direct Reprogramming of Human Fibroblasts to Functional and Expandable Hepatocytes. *Cell Stem Cell.* 2014. (IF 25.315) 使用服务种类: Agilent 人表达谱芯片服务
- [27] Wang YF, Lei R, Zhuang XQ, Zhang N, Pan H, Li G, Hu J, Pan XQ, Tao Q, Fu D, Xiao JR, Chin YE, Kang B, Yang QF, HuGH. DLC1-dependent parathyroid hormone-like hormone inhibition suppresses breast cancer bone metastasis. *The Journal of Clinical Investigation.* 2014. (IF 12.812) 使用服务种类: Affymetrix GeneChip Human Genome U133 Plus 2.0 Array
- [28] Li L, Tang J, Zhang BH, Yang W, Gao MYL, Wang RY, Tan YX, Fan JL, Chang YX, Fu J, Jiang F, Chen CY, Yang YC, Gu J, Wu DM, Guo LN, Cao D, Li HY, Cao GW, Wu MC, Zhang MQ, Chen L, Wang HY. Epigenetic modification of MiR-429 promotes liver tumour-initiating cell properties by targeting Rb binding protein 4. *Gut.* 2014, doi:10.1136/gutjnl-2013-305715. (IF 10.732)
使用服务种类: Affymetrix GeneChip Human Genome U133 Plus 2.0 Array
- [29] Schaid DJ, Spragg CF, McDonnell SK, Parham LR, Cox CJ, Ejlertsen B, M. Finkelstein D, Rappold E, Curran J, R. Cardon L, E. Goss P. Prospective Validation of HLA-DRB1*07:01 Allele Carriage As a Predictive Risk Factor for Lapatinib-Induced Liver Injury. *Journal of clinical oncology.* 2014. 10.1200/JCO.2013.52.9867. (IF17.879) 使用服务种类: Illumina Human1M-Duo beadchip array
- [30] Gou LT, Dai P, Yang JH, Xue Y, Hu YP, Zhou Y, et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell research.* 2014. (IF10.526) 使用服务种类: Affymetrix GeneChip Mouse Genome 430 2.0 Array 服务
- [31] Zhang L, Chen LH, Wan H, Yang R, Wang Z, Feng J, Yang S, Jones S, Wang S, Zhou W, Zhu H, Killela PJ, Zhang J, Wu Z, Li G, Hao S, Wang Y, Webb JB, Friedman HS, Friedman AH, McLendon RE, He Y, Reitman ZJ, Bigner DD, Yan H. Exome sequencing identifies somatic gain-of-function PPM1D mutations in brainstem gliomas. *Nat Genet.* 2014 Jul;46(7):726-30. doi: 10.1038/ng.2995. Epub 2014 Jun 1. (IF 35.209) 使用服务种类: Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 服务
- [32] Xu C, Ju X, Song D, Huang F, Tang D, Zou Z, et al. An association analysis between psychophysical characteristics and genome-wide gene expression changes in human



adaptation to the extreme climate at the Antarctic Dome Argus. Molecular psychiatry.

2014. (IF15.147) 使用服务种类: Agilent Human Gene Expression 4 × 44K v2 Microarray

[33] Yu B, He ZY, You P, Han QW, Xiang D, Chen F, Wang MJ, Liu CC, Lin XW, Borjigin U, Zi XY, Li JX, Zhu HY, Li WL, Han CS, Wangensteen KJ, Shi Y, Hui LJ, Wang X, Hu YP. Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors. *Cell Stem Cell*.

2013,13(3):328-40. (IF25.315) 使用服务种类: Agilent 人表达谱芯片服务

[34] Lu X, Li X, He Q, Gao J, Gao Y, Liu B, Liu F. miR-142-3p regulates the formation and differentiation of hematopoietic stem cells in vertebrates. *Cell Res*. 2013 Oct 29. (IF10.526)
使用服务种类: Agilent 斑马鱼表达谱服务

[35] Yang H, Liu Z, Ma Y, Zhong C, Yin Q, Zhou C, Shi L, Cai Y, Zhao H, Wang H, Tang F, Wang Y, Zhang C, Liu XY, Lai D, Jin Y, Sun Q, Li J. Generation of haploid embryonic stem cells from *Macaca fascicularis* monkey parthenotes. *Cell Res*. 2013 Jul 16. (IF10.526)

使用服务种类: Agilent SurePrint G3 Rhesus Macaque Genome CGH Microarray 4 × 180K 服务

第四章 测序服务

上海伯豪测序服务一览

基因组测序 (DNA-Seq)	转录组测序 (RNA-Seq)	表观基因组测序
<i>de novo</i> 测序	转录组测序 (无参考基因组)	染色质免疫沉淀测序 (ChIP-Seq)
全基因组重测序	转录组测序 (有参考基因组)	甲基化 DNA 免疫共沉淀测序 (MeDIP-Seq)
DNA 序列捕获重测序	small RNA 测序	全基因组 Bisulfite 甲基化测序
RAD-Seq	微量样本转录组测序	
单细胞全基因组测序	定向转录本测序	
单细胞外显子捕获测序	宏转录组测序	
16S rDNA 分子生态测序	RIP-Seq	
宏基因组测序	环状 RNA 测序	

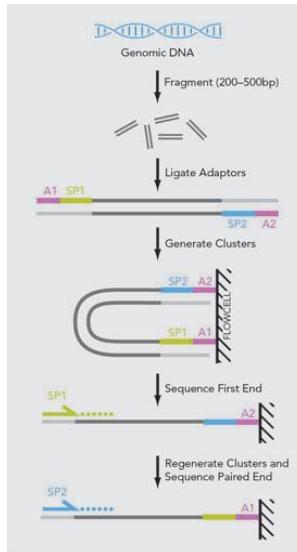
DNA 水平研究

(一) 全基因组重测序服务

全基因组重测序是对已有参考序列 (Reference Sequence) 的物种的不同个体进行基因组测序，并以此为基础进行个体或群体水平的差异性分析。通过全基因组重测序，研究者可以找到大量的单核苷酸多态性位点 (SNP)、拷贝数变异 (Copy Number Variation, CNV)、插入缺失 (Insertion/Deletion,)、结构变异 (Structure Variation, SV) 等变异位点。这在人类疾病及动植物育种研究等方面具有重大的指导意义。

上海伯豪可以帮助您利用有 Illumina HiSeq2500 等平台进行全基因组重测序，并通过生物信息学的手段，分析不同个体基因组间的差异，同时完成注释，从而帮助您在全基因组水平上扫描并检测特定人群的重要遗传性状相关位点。

技术路线



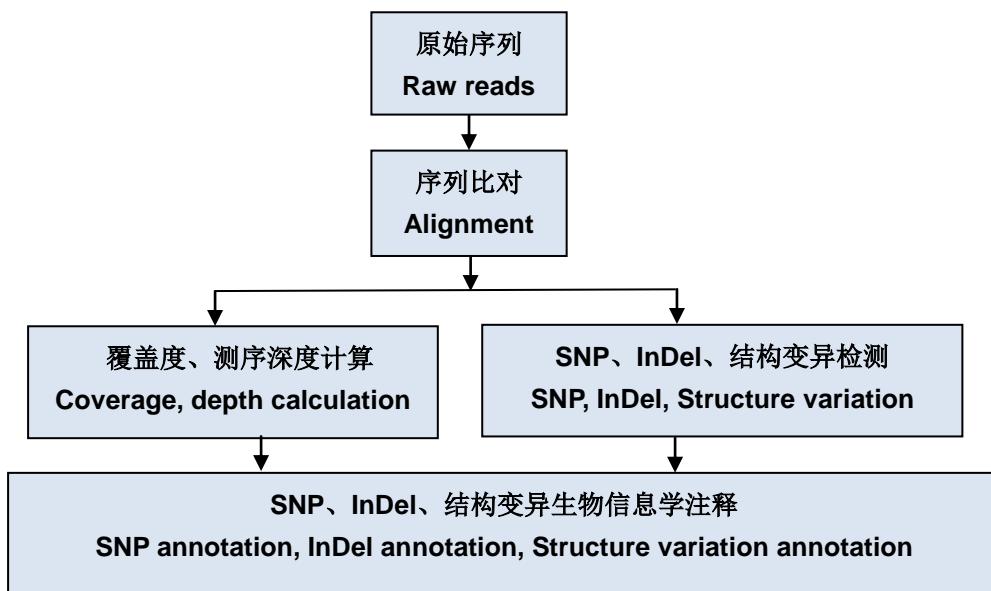
Paired-end Sequencing

服务相关

- 样品纯度：OD 260/280 值应在 1.7~2.0 之间；RNA 应该去除干净。

-
2. 样品浓度：最低浓度不低于 50ng/ μ l。
 3. 样品总量：每个样品总量不少于 4 μ g。
 4. 样品溶剂：要溶解在 H₂O 或 Low TE (pH 8.0) 中。
 5. 样品运输：DNA 低温运输 (-20°C)；且在运输过程中请用 parafilm 将管口密封好，以防出现污染；收到样品后，甲方需要对样品进行检测，最终样品的量和纯度，以甲方的检测结果为准。

数据分析流程和内容



数据基础分析内容：

1. 序列质量评估：应用测序质量 Q 值进行评估。
2. 数据预处理：去除总体质量低于标准的 reads、接头序列去除 reads 中由于测序产生的含糊的 N 碱基。
3. 基因组 mapping 统计，包括 mapped reads ratio, genome coverage, genome depth 等信息的统计。
4. 基因组 SNV 检测结果：碱基突变情况；对应的氨基酸的改变，是否同义突变；对应染色体的物理位置定位信息；关联的 dbSNP 库信息（若是新发现 SNV 则无此信息）；该 SNV 位点相关基因信息（若是新发现 SNV 则无此信息）。
5. 基因组 SV 检测结果：基因组发生 SV 的类型，SV 区段的起始/终止位置，SV 对氨基酸，蛋白质功能的影响。
6. 基因组 Small InDel 检测结果：基因组发生 Small InDel 区段的大小，位置，(仅限于有参考序列，reference sequence, 的区域)；关联的基因信息。

数据高级分析总览

复杂疾病/性状关联研究	肿瘤基因组学研究	群体进化研究	遗传图谱构建
群体 SNP 检测、分型	成对样本的 somatic SNV/InDel/CNV 检测，注释、统计	群体 SNP 检测、分型	群体 SNP 检测、分型
群体 SNP 位点质控	SNP/InDel 数据库分析 (dbSNP/1000G/UCSC)	基于参考单体型集合的缺失基因型推断、	群体 SNP 位点质控

		填充	
SNP 注释与统计 (OMIM, ENCODE)	SNV/InDel COSMIC 数据库注释、筛选	样本质控 (亲缘关系检测、基于 IBD 的样本污染检测、PCA 群体分层检测)	样本质控(系谱矫正，样品混杂检测)
样本质控 (亲缘关系检测、基于 IBD 的样本污染检测、PCA 群体分层检测)	SNV 保守性预测(SIFT、Polyphen-2)	连锁不平衡分析	遗传图谱构建
基于单个位点的关联分析		系统进化树分析	QTL 定位
候选区域 LD 分析		群体多态性分析	
GO、PATHWAY 富集分析		选择分析、驯化基因分析	

应用案例

全基因组重测序揭示小细胞肺癌基因组的变化

吸烟会引发癌症，但是人们对这一过程的分子机制并不清楚，一般认为吸烟过程中释放的大于 60 种致癌物质可与 DNA 链上的鸟嘌呤和腺嘌呤进行化学修饰从而产生大的复合物，该复合物改变了 DNA 双螺旋的结构，如果这些复合物没有被及时纠正，那么在 DNA 复制的过程中就会产生突变，从而引发癌症，Erin D.Pleasance 等利用第二代测序技术(ABI SOLID)对一个小细胞肺癌(Small-cell lung cancer,SCLC) 细胞系 NCI-H209 进行全基因组重测序，探讨了烟气中的致癌物质引发该细胞系基因组中哪些碱基及其周围序列产生突变及细胞损伤修复路径。

原文出处：A small-cell lung cancer genome with complex signatures of tobacco exposure.

Nature,2010,463:184-190.

(二) 全外显子组测序服务

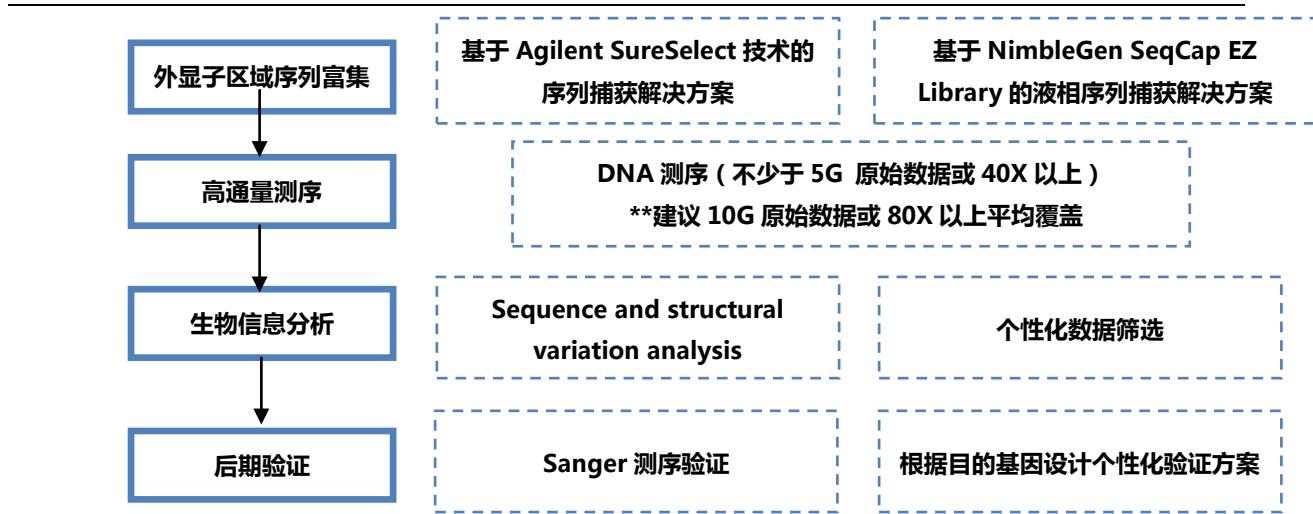
全外显子组测序是利用序列捕获技术将全基因组外显子区域 DNA 捕捉并富集后进行高通量测序的基因组分析方法。由于它具有对常见和罕见变异高灵敏度，因此仅需对 1% 的基因组进行测序就能发现外显子区域的大部分疾病相关变异。全外显子组测序技术具有针对性强，覆盖度深，数据准确性高，简便、经济、高效等特点。目前，已经广泛应用于寻找复杂疾病（如：癌症、糖尿病、肥胖症等）的致病基因和易感基因等的研究。

技术路线

1. 实验设计：

Case-Control 对照组设计

2. 技术流程：



基础数据分析

- 1). 覆盖度、测序深度计算
- 2). SNV 检测及多数据库注释
- 3). InDel 检测及多数据库注释

个性化数据分析总览

复杂疾病/性状关联研究	肿瘤基因组学研究	孟德尔遗传研究
群体 SNP 检测、分型	成对样本的 somatic SNV/InDel/CNV 检测、注释、统计	群体 SNV/InDel 检测、分型、注释
群体 SNP 位点质控	SNP/InDel 数据库分析 (dbSNP/1000G/UCSC)	样本质控 (亲缘关系检测、基于 IBD 的样本污染检测)
SNP 注释与统计 (OMIM, ENCODE)	SNV/InDel COSMIC 数据库注释、筛选	基于家系进行致病区段筛选
样本质控(亲缘关系检测、基于 IBD 的样本污染检测、PCA 群体分层检测)	SNV 保守性预测 (SIFT、Polyphen-2)	
基于单个位点的关联分析		
候选区域 LD 分析		
GO、PATHWAY 富集分析		

服务相关

1. 样品纯度: OD 260/280 值应在 1.7~2.0 之间; RNA 应该去除干净。
2. 样品浓度: 最低浓度不低于 50ng/μl。
3. 样品总量: 每个样品总量不少于 6μg。
4. 样品溶剂: 要溶解在 H₂O 或 Low TE (pH 8.0) 中
5. 样品运输: DNA 低温运输 (-20°C); 且在运输过程中请用 parafilm 将管口密封好, 以防出现污染; 收到样品后, 甲方需要对样品进行检测, 最终样品的量和纯度, 以甲方的检测结果为准。

外显子区域序列富集技术策略

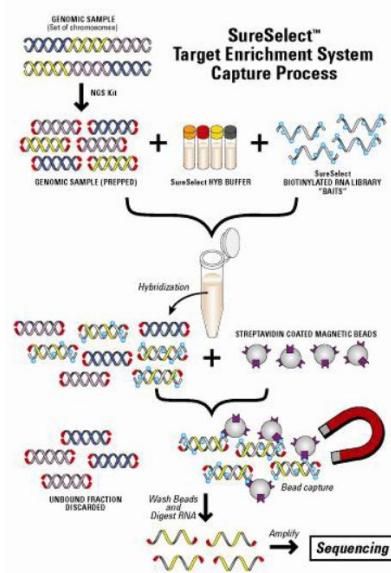
1. 基于 Agilent SureSelect 技术的外显子捕获解决方案

1) 基于 Agilent SureSelect 技术的外显子捕获技术:

Agilent SureSelect 系统自 2009 年初面世以来，已经被超过 100 篇发表的权威学术论文引用。目前，该技术作为癌症研究的一项热点技术，已被广泛应用于发现与癌症等复杂疾病有关的遗传变异，并已被国际癌症协会（ICGC）广泛采用并获得了鉴定认可。

Agilent SureSelect 靶向序列捕获技术是一种溶液型杂交捕获技术，该技术由 Agilent 公司与麻省理工学院-哈佛大学博德研究所共同开发，研究成果随后刊登在《自然生物技术》上。该技术使研究人员能够将测序范围缩小到感兴趣的基因组区域，从而简化新一代测序。120-mer RNA 寡核苷酸能够可靠地捕获更多大小变异，包括 SNP、CNV 和 InDel，具有无与伦比的等位基因平衡。

参考文献：Andreas Gnirke et al., Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol. 2009 Feb;27(2):182-9.

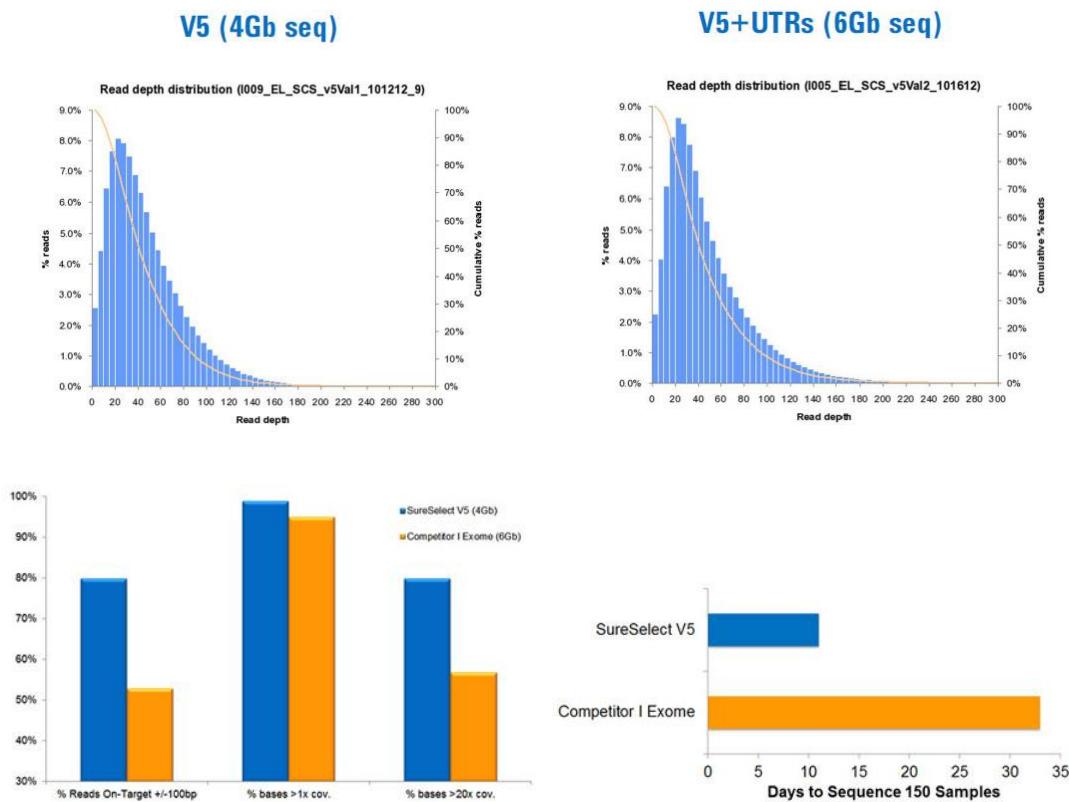


2) 最新解决方案推介：Agilent SureSelect Human All Exon V5 和 V5+UTR 序列捕获方案 HOT!

A. 数据库更新至最新，包括 CCDS, GENCODE, RefSeq 和 TCGA。

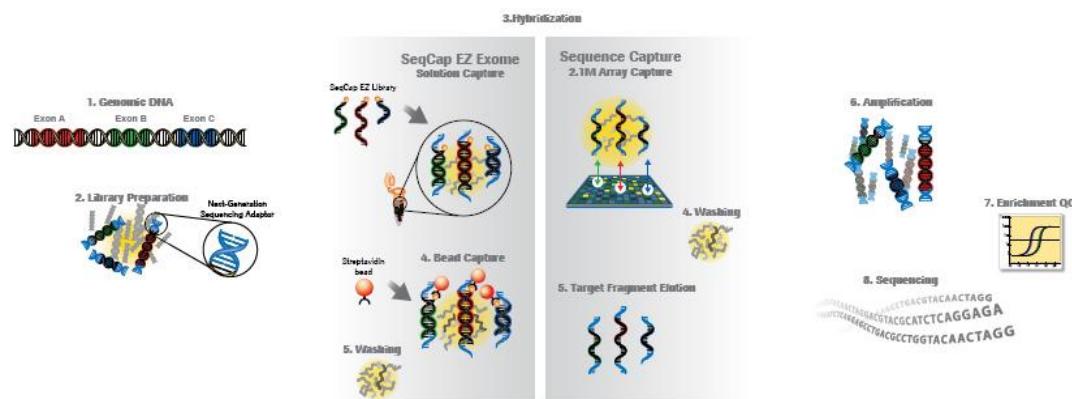
B. 为获得更好的数据质量，推荐测序深度 100X。

C. 以下是部分技术数据：



2. 基于 NimbleGen SeqCap EZ Library 的液相外显子捕获解决方案

1) 基于 NimbleGen SeqCap EZ Library 的液相外显子捕获技术:



序列捕获实验流程图

- a) 基因组 DNA;
- b) 文库制备: 建立基因组 DNA 测序文库, 片段化的基因组序列两端包含测序接头;
- c) 杂交: 基因组 DNA 文库与 SeqCap EZ library 液相芯片进行杂交;
- d) 磁珠捕获: 被链霉素亲和素标记的磁珠将吸附, 被生物素标记的寡核苷酸探针与基因组 DNA 片段杂交成复合物;
- e) 洗脱: 未与液相芯片探针结合的基因组片段被洗脱;
- f) 扩增: 利用 PCR 扩增捕获片段;
- g) 富集实验质量控制: 通过内参基因位点利用 qPCR 测定含量, 判断富集效率;
- h) 高通量测序。

2) 最新解决方案推介:

A. NimbleGen SeqCap EZ Exome Library V3.0 序列捕获方案

可捕获 64M 的基因组序列, 包括所有外显子以及 miRNA, 覆盖了 RefSeq, RefGene, CDS, CCDS 以及 miRBase v14 数据库, 信息更为全面。

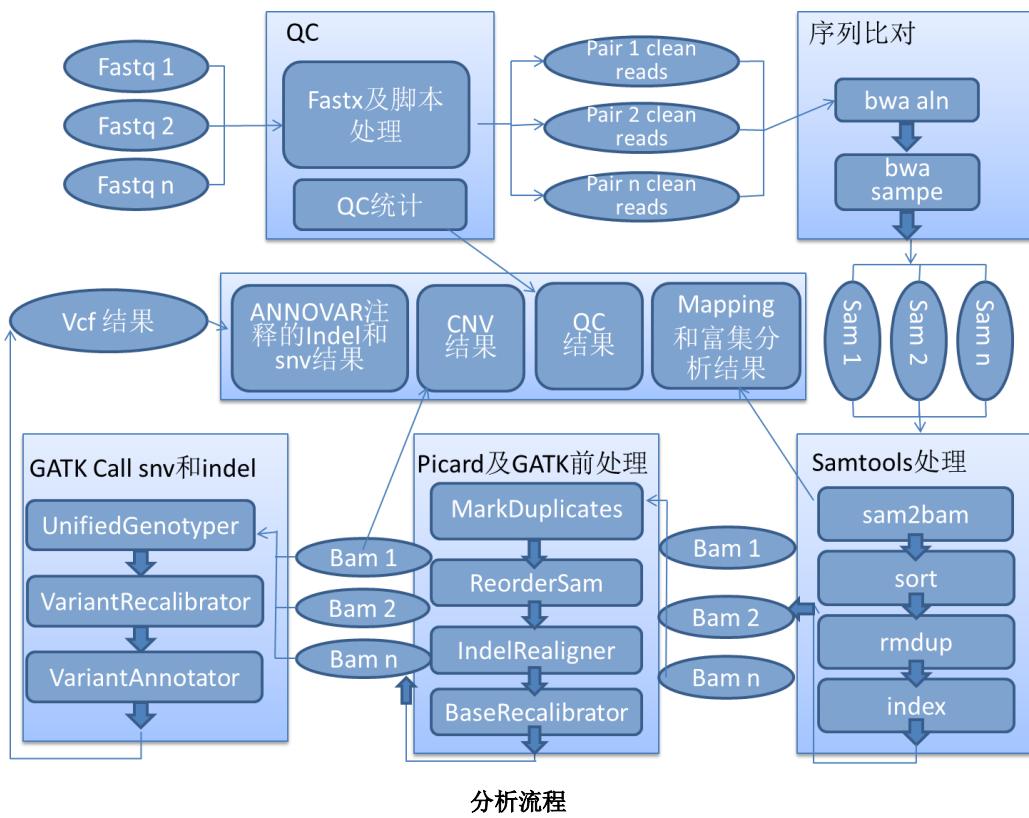
B. NimbleGen SeqCap EZ Exome +UTR Library 序列捕获方案

覆盖了外显子组和 50Mb 的 UTR 区域, 总捕获目标区域高达 96Mb。

C. NimbleGen SeqCap EZ Exome Plus Library 序列捕获方案

在原有 64Mb 的外显子组捕获基础上, 可以由用户自行选择另外的 50Mb 区域进行定制设计。

数据分析流程和内容



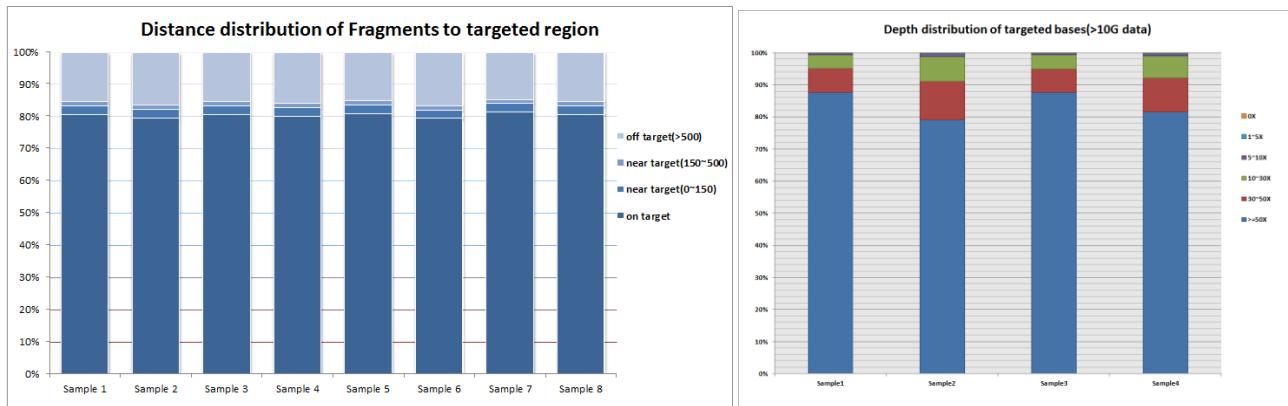
分析流程

1. **Mapping 统计:** 统计总 reads 数, mapped reads 及 unique mapped reads 数目及百分比。

2. 捕获效率统计:

统计来自捕获区域的 Fragment 比例:

统计 target 区域所有的碱基覆盖次数分布:



对每个 target 区域的覆盖和深度统计:

如果客户对某些基因特别感兴趣,想要看看来自这些基因的外显子区域的覆盖情况,可以提供每个 target 或者特定 target 区域的覆盖情况和测序深度统计。

3. SNV 和 InDel 关联数据库:

SNV 和 InDel 结果按照突变的位点是否在捕获的区域之内分成两部分:

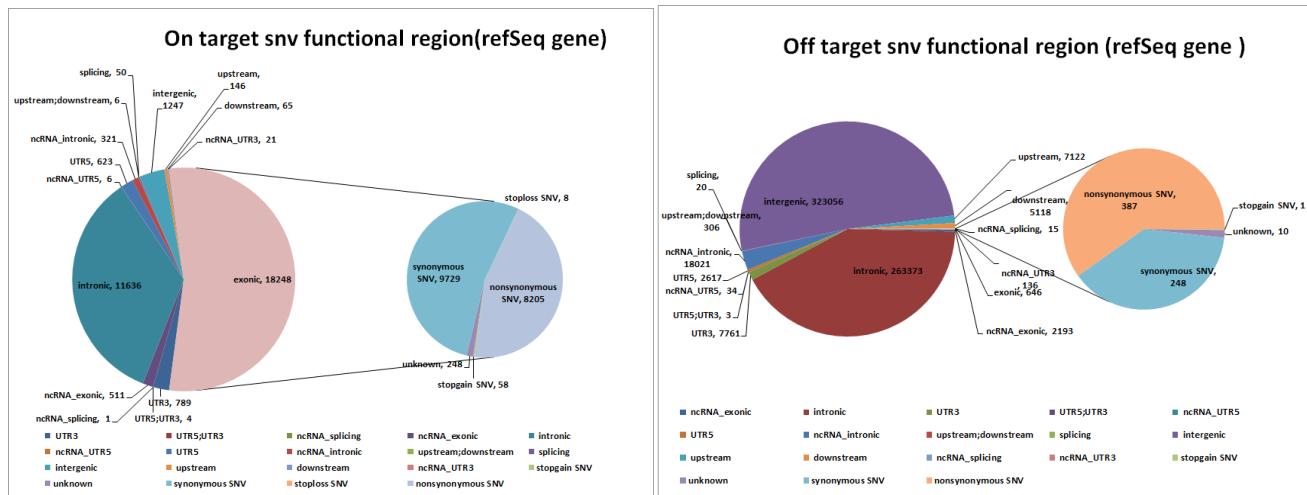
*_target.snv: 突变处于捕获的靶区域内。

*_off_target.snv 或者 *_target.indel: 突变在捕获的靶区域之外。

SNV 和 InDel 结果通过 ANNOVAR 软件与多个数据库关联, 为突变的筛选提供大量的信息。

1) 基因注释:

通过基因注释可以达到以下的目的: 突变的功能定位(在外显子, 内含子, 剪接位点还是基因间区); 突变所在的基因名称或者临近的基因; 突变如果在编码区域, 是否引起氨基酸的改变(同义突变, 非同义突变等); 如果引起氨基酸的改变, 按照 HGVS 命名规则表示--改变的基因 ID, 转录本 ID, 外显子编号, 以及氨基酸改变, 如 NOD2:NM_022162:exon8:c.G2722C:p.G908R。默认使用 refSeq 完成基因注释, 如果有特殊的要求, 可以使用 UCSC known gene, Ensembl, GENCODE, CCDS 等基因注释系统。



2) 1000G 注释:

检测突变位点是否在 1000 Genomes Projects (2012 release) 数据库中检测到, 如果检测到, 显示等位基因频率 (allele frequency)。默认是使用所有人群的数据库, 如果有特定要求, 可以按照要求展示不同人群 (比如 AMR, AFR, ASN, EUR, 中国人, 日本人) 等位基因频率。

3) dbSNP 注释:

检测突变是否在 dbSNP 数据库中, 如果在, 显示 rsID。
默认使用 db SNP135 数据库, 如果有特定的要求, 可以使用 dbSNP129, dbSNP130, dbSNP131, dbSNP132 数据库。

4) SIFT 和 PolyPhen 2 注释:

SIFT 是一款很受欢迎的预测非同义突变位点对蛋白质功能影响的软件, 对应非同义突变位点, 会给出一个打分, 若打分低于 0.05, 则表明突变很可能会影响到蛋白质的功能。PolyPhen 2 和 SIFT 一样, 也是预测非同义突变位点对蛋白质功能影响的软件, ANNOVAR 可以将非同义位点与两者的打分关联起来, 为突变对蛋白质功能的影响提供重要依据。

5) 与 UCSC 的数据库的关联:

[ftp://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/<table-name>.txt.gz](http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/<table-name>.txt.gz), 提供了大量的基因组注释信息, 目前关联的数据库有:

tfbsConsSites: 在人/小鼠/大鼠中保守的转录因子结合位点, 以 transfac Matrix Database (v7.0) 为基础。

wgRNA: snoRNA and miRNA 注释。

targetScanS: TargetScan 预测的 miRNA 靶区域。

gwasCatalog: 已经发表的各种疾病的 GWAS 结果。

genomicSuperDups: 基因组中的重复片段。

phastConsElements46way: 通过 phastCons 对脊椎动物的全基因组比对生成的保守区域，根据用于比对的物种数目，分为 17way, 28way, 30way, 44way 等。默认使用 46way。

如果客户需要关联 UCSC 中其它的数据库，也可以定制。

6) cosmic65:

已观察到的癌症相关突变，显示在 COSMIC 中的 ID (identifiers)，观察到的次数，以及观察到的癌组织。COSMIC 的版本号随数据库的更新而更新。

4. CNV:

通过发表于权威期刊的 exome CNV 检测工具，检测可能的 CNV。推荐成对的样品间(比如：癌 VS 癌旁)的比较。

5. 其它：

分析结果中在保证灵敏度的前提下，通过一系列的筛选条件，提高准确性。

可以通过 GT (genotype) 直接比较样品间的差异(GT 简介：0 表示与 Ref 相同，1 表示与 ALTS 第 1 个碱基相同，2 表示与 ALTS 第 2 个碱基相同)。

#CHROM	Pos	Ref	ALTS	ALT	GT_Sample 1	GT_Sample 2	GT_Sample 3	GT_Sample 4
chr1	865664	C	T	T	0/1	0/0	0/1	0/1
chr1	876499	A	G	G	1/1	1/1	1/1	1/1
chr1	877831	T	C	C	1/1	1/1	1/1	1/1
chr1	881627	G	A	A	1/1	1/1	1/1	1/1
chr1	32136168	G	C, T	C	1/1	1/2	0/1	0/1
chr1	111783982	C	A, T	A	0/0	1/2	0/1	0/1

通过和多个数据库提供的关联精细筛选条件：

Var locat	Gene	Var reslu	Exon	Conservec	Duplication	1000G	dbSNP135	avSIFT	GWAS	TFBS	miRNA or	miRNAtarg	Cosmic
exonic	RNF207	nonsynony	RNF207:NM_	-	-	0.17	rs846111	0.62	Name=QT	i	-	-	-
exonic	RERE	synonymou	RERE:NM_0	-	-	0.78	rs11121172	-	-	Score=846	-	-	ID=COSM6
UTR3	ATXN2L	-	-	486;Name=	-	-	-	0.77	-	-	-	Score=7;N	-
exonic	CPSF3L	synonymou	CPSF3L:NM_0625;Name=	-	-	0.75	rs10907179	0.02	-	-	-	-	-
intronic	NOC2L	-	-	-	0.91	0.65	rs13303010	-	-	-	-	-	-

伯豪案例

HBV 相关的肝细胞癌外显子组研究

利用上海伯豪生物技术有限公司的 Illumina *GAIix* 和 ABI 的 SOLiD 新一代测序服务平台，进行全外显子组测序分析。比较分析原位癌和转移灶的测序结果，发现肿瘤细胞存在 347 个突变基因，平均每个肿瘤样本有 30-40 个基因突变，其中大多数基因突变是第一次在肝癌样本中发现，并且在原发灶和转移灶中同时出现。说明肝癌的发病和转移与基因突变密切相关，并且转移相关基因在原发癌中就已经发生了突变。具体分析基因突变的类型，发现有 5 位患者存在明显的 C:G>A:T 突变，有 4 位患者存在明显的 T:A>A:T 突变。

原文出处： Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. Nat Genet. 2012 Aug 26. (IF35.532) 使用服务类型：外显子组测序。

(三) 目标区域捕获测序服务

目标区域捕获测序是利用特制的探针对客户感兴趣的基因组区域进行富集后再进行测序的研究方法。通过该方法能够获得特定区域的遗传信息，极大地提高了基因组中特定区域的研究效率，显著降低了研究成本。从而帮助广大科研工作者筛选变异位点、识别和研究与疾病、种群进化相关的编码区域内的结构变异。您只需提供样本和候选区域，上海伯豪即可为您提供从探针定制、杂交、测序到生物信息分析的全面解决方案。

基于 AmpliSeq™ 技术的序列捕获解决方案

序列捕获测序能对基因组上某些特定的区域进行分析，但是传统的序列捕获测序技术操作复杂，实验时间长，且对 DNA 样品的数量要求较高。基于 AmpliSeq™ 技术，科学家们能够快速完成某些疾病相关基因的突变筛查。同时，本系统可还可以为有需要的科学家定制目标序列的多重 PCR 体系，快速完成检测流程。

产品介绍：

Ion AmpliSeq™ Ready-to-Use Panels

⊕ Ion AmpliSeq™ Cancer Hotspot Panel v2

针对50个肿瘤相关基因的207个常见突变区域设计了引物，在同一个EP管中进行多重PCR扩增。随后纯化扩增产物，利用Ion Torrent进行测序分析。

技术参数：

Ion AmpliSeq™ Cancer Hotspot Panel v2		
靶向序列	50 个原癌基因或抑癌基因，约 2,800 个 COSMIC 突变	
扩增子长度	111-187bp (平均 154bp)	
引物池大小	1 管内 207 对引物	
起始 DNA 需求量	每个 DNA 样本 10ng	
	规格	实际观测数据 (Ion 314™ chip)
覆盖均一性	≥95%	≥98%
命中目标的读数	≥90%	≥96%
平均覆盖深度	NA	>2,000X
SNP 检测灵敏度	NA	对于平均测序覆盖度为 1,000X 至 4,000X 的位点，5% 变异频率的检测率为 98%

⊕ Ion AmpliSeq™ Comprehensive Cancer Panel

经过精心设计获得Ion AmpliSeq™ Comprehensive Cancer Panel，可同时检测多个基因家族中的编码DNA序列（CDS）和剪接变异数体。这种基于通路的基因选择描绘出了癌症驱动基因的突变谱和药物靶点，以及信号级联、凋亡基因、DNA修复基因、转录调节因子、炎症反应基因和生长因子基因等。此外，它还包含了Ion AmpliSeq™ Cancer Hotspot Panel v2聚焦的全部50个基因。

技术参数：

Ion AmpliSeq™ Comprehensive Cancer Panel		
靶向序列	>400 个原癌基因以及抑癌基因的外显子区域	
扩增子长度	125-175bp (平均 155bp)	
引物池大小	4 管内 16,000 对引物	
起始 DNA 需求量	每个 DNA 样本 40ng, 每个引物池 10ng	
	规格	实际观测数据 (Ion 318™ chip)
覆盖均一性	>90%	94%
命中目标的读数	>95%	97%
平均覆盖深度	NA	350X

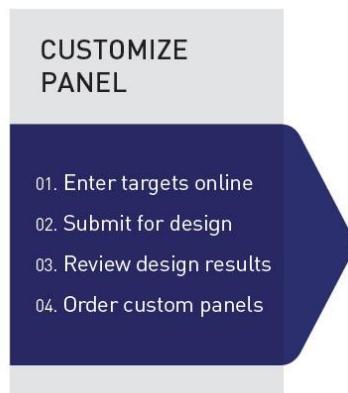
◆ Ion AmpliSeq™ Inherited Disease Panel

技术参数：

Ion AmpliSeq™ Inherited Disease Panel		
靶向序列	超过 300 个基因的外显子，与 700 多种遗传病相关，包括：神经肌肉类、代谢类、心血管类、生长发育类疾病	
扩增子长度	125-225bp (平均 197bp)	
引物池大小	3 管内有超过 10,000 对引物	
起始 DNA 需求量	每个 DNA 样本 30ng, 每个引物池 10ng	
	规格	实际观测数据 (Ion 316™ chip)
覆盖均一性	>85%	90%
命中目标的碱基数	>95%	98%
>20X 覆盖度的目标碱基数	>90%	94%
平均测序深度	NA	275X

Ion AmpliSeq™ Custom DNA Panels

该产品是专为科学家设计的定制产品，目标序列可以为人类或者是小鼠的DNA。上海伯豪可协助您通过Ion AmpliSeq™ Designer创建或者订购您的panel。



Ion AmpliSeq™ custom 工作流程 (Ion AmpliSeq™ Designer 设计 Panel)

Ion AmpliSeq™ Community Panels

⊕ Ion AmpliSeq™ Colon and Lung Cancer Panel

技术参数:

Ion AmpliSeq™ Colon and Lung Cancer Panel	
靶向序列	热点区域，与结肠癌以及肺癌相关的 22 个基因的靶向区域
基因	<i>KRAS, EGFR, BRAF, PIK3CA, AKT1, ERBB2, PTEN, NRAS, STK11, MAP2K1, ALK, DDR2, CTNNB1, MET, TP53, SMAD4, FBXW7, FGFR3, NOTCH1, ERBB4, FGFR1, FGFR2</i>
扩增子长度	90 个扩增子，平均长度 162bp
引物池大小	一个引物池 90 对引物
起始 DNA 需求量	10ng

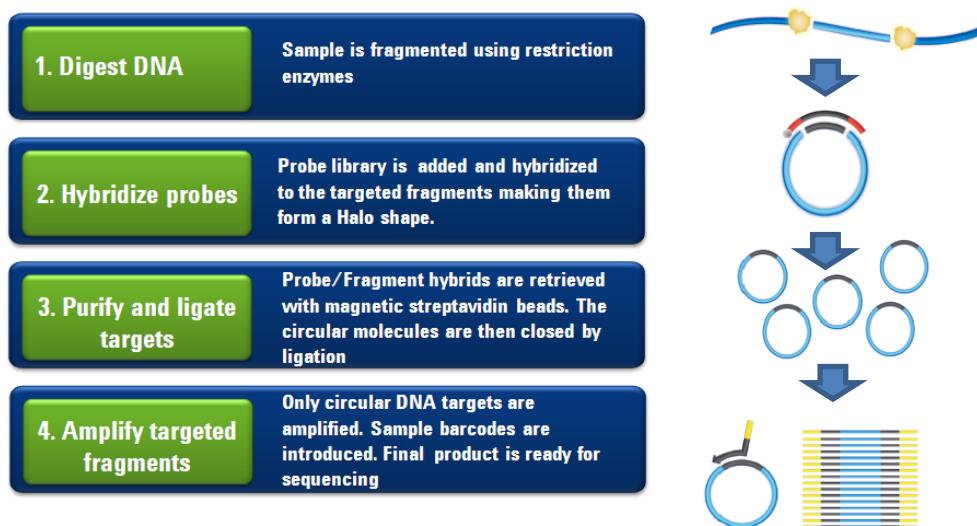
⊕ Ion AmpliSeq™ BRCA1 and BRCA2 Panel

技术参数:

Ion AmpliSeq™ BRCA1 and BRCA2 Panel	
靶向序列	BRCA1 、 BRCA2 基因的编码区域
扩增子长度	平均长度 200bp
引物池大小	3 个引物池共 167 对引物
起始 DNA 需求量	30ng

基于 Agilent HaloPlex 技术的序列捕获解决方案

HaloPlex 技术完美融合了聚合酶链反应系统的速度和特异性以及基于液相杂交体系的可扩展性和捕获片段大小的灵活性（捕获区域大小为 1Kb-5Mb，推荐<500Kb 区段设计）。该实验方案无需文库构建，利用单管操作，从而解决了多重 PCR 由于存在交叉杂交问题而使靶向序列数量受限的问题。



Agilent HaloPlex 技术流程

除了客户自己定制序列，HaloPlex 还针对特定的疾病提供设计好的解决方案，目前有以下几种，如果研究者需要研究其他疾病，可以通过 SureDesign 来查看。

HaloPlex Panels:

目录研究试剂盒:

- ⊕ **HaloPlex Cancer Research Panel**
- ⊕ **HaloPlex Cardiomyopathy panel**

已定制研究试剂盒:

- ⊕ **HaloPlex Arrhythmia**
- ⊕ **HaloPlex Noonan Syndrome**
- ⊕ **HaloPlex ICCG**
- ⊕ **HaloPlex Connective Tissue Disorder**
- ⊕ **HaloPlex X Chromosome**

基于 Agilent SureSelect 技术的序列捕获解决方案

Agilent SureSelect 定制捕获可针对基因组的一小部分区段进行探针设计，从而大大降低测序及分析流程的复杂性。灵活的设计方式以及与各种高通量测序平台良好的对接能力，使得 Agilent SureSelect 定制捕获能够有效地满足科学家的各种定制要求。

Agilent SureSelect DNA panels

- ⊕ **SureSelect Human Kinome panel:** 靶序列为一系列激酶及与激酶相关的基因，靶目标超过 500 种激酶，612 个基因。
- ⊕ **SureSelect X-Chromosome panel:** 靶向 X 染色体的基因。

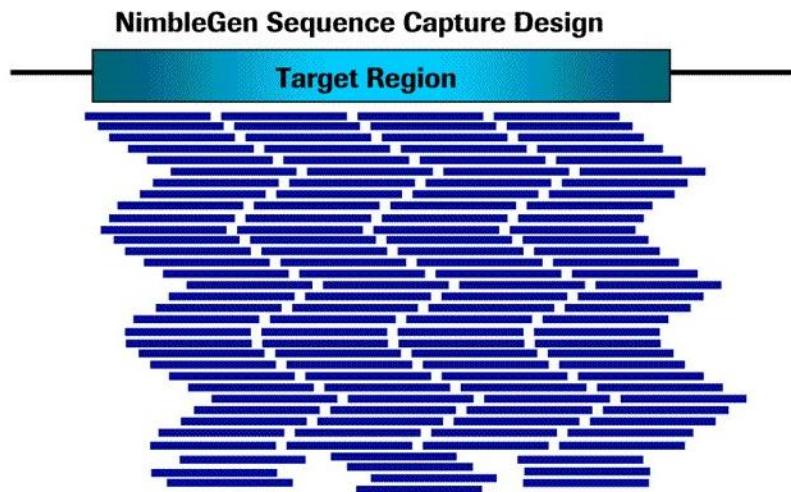
Agilent SureSelect Custom DNA

技术特点:

- 1) 灵活的靶向序列: 1Kb-24Mb。
- 2) 便捷的操作
 研究者可以通过 Sure Design software 或者 Agilent's Design Service 轻松设计程序。
- 3) 有效的工作流程
 整个流程包括 post-capture (XT) 以及 pre-capture (XT2) pooling。

基于 NimbleGen 技术的序列捕获解决方案

NimbleGen SeqCap 捕获系统采用 advanced repeat masking 方法设计探针，并在此基础上，利用实际测序数据再次进行平衡筛选，保证了很好的均一性(uniformity)和覆盖率(coverage)，可以有效捕捉到每个靶向序列，从而对目标区域进行精确细致的分析，有效监测 SNP 突变。



NimbleGen SeqCap 系统可使用多至 2.1M 数量的探针进行捕获设计，产生可靠数据，降低成本，节约时间。

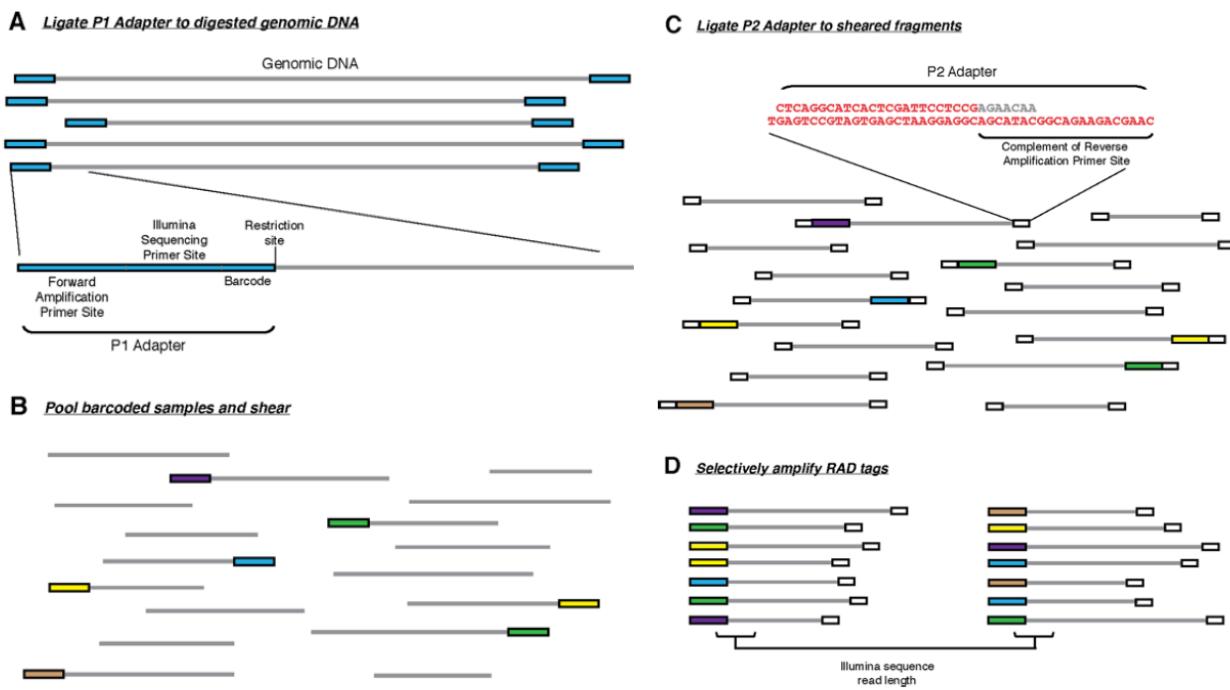
(四) RAD-Seq 服务

基于限制位点相关 DNA (Restriction-site Associated DNA, RAD) 的测序技术，即 RAD-Seq，是一种对酶切产生的基因组标签序列 (Tags) 进行高通量测序的新技术，该技术能够大幅降低基因组的复杂度，可快速在全基因组范围内鉴定出高密度的 SNP 位点。

对没有参考基因组的物种而言，RAD-Seq 技术不受已知基因组序列的限制，即可大规模筛查 SNP 位点；基因组的复杂度被大幅降低，从而降低了测序成本，因而特别适合在群体水平进行研究；以往的技术在获得 SNP 位点信息之后，需要通过设计相应引物，再利用基因分型技术对每个样本进行分型，然而 RAD-Seq 技术获得 SNP 位点信息的同时，也就获得了每个样本的基因型。对有参考基因组的物种而言，数据分析更加简便，并且能够有效开发新的 SNP 位点。另外，我们可以在测序前，通过对限制性内切酶进行选择，灵活调整 Tags 的复杂度，适应不同的研究需要。

获得的大量 SNP 分型信息，主要可用于遗传连锁图谱的构建，QTL 定位和群体遗传分析。此外，还可以在全基因组测序项目中辅助序列的组装。

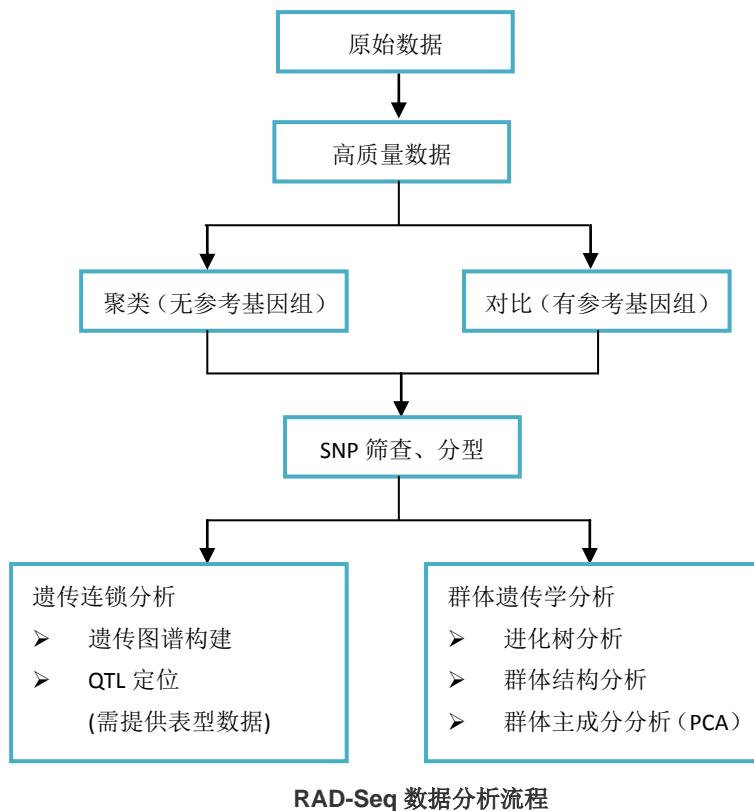
文库构建技术流程



RAD-Seq 文库构建流程 (Baird 等, PLoS ONE, 2008)

- 对基因组进行酶切后，连接 P1 接头（含有 barcode）；
- 随机打断；
- 连接 P2 接头；
- 通过 PCR 筛选同时带有 P1 和 P2 接头的序列；片段大小选择；测序。

数据分析流程



数据分析内容

1. 遗传连锁分析

1) 测序策略

- ⊕ PE100 测序；
- ⊕ 亲本：分别 10X 以上；
- ⊕ 子代：F1、F2 等临时性群体，推荐每个个体平均测序深度为 0.8-1X；RIL、DH 等永久性群体，推荐每个个体平均测序深度为 0.6X。

2) 基础信息分析：

- ⊕ 对于无参考基因组的物种，进行个体 RAD tag 聚类与群体 RAD tag 聚类；对于有参考基因组的物种，与参考基因组进行比对；
- ⊕ SNP 检测。

3) 高级信息分析：

- ⊕ 遗传图谱构建；
- ⊕ QTL 定位（需提供表型数据）。

2. 群体遗传学分析

1) 测序策略

- ⊕ PE100 测序；
- ⊕ 推荐每个个体平均测序深度为基因组大小 1.5 X。

4) 基础信息分析：

- ⊕ 对于无参考基因组的物种，进行个体 RAD tag 聚类与群体 RAD tag 聚类；对于有参考基因组的物种，与参考基因组进行比对；

-
- ⊕ SNP 检测。

5) 高级信息分析:

- ⊕ 进化树分析 (Phylogeny tree);
- ⊕ 群体结构分析 (Structure);
- ⊕ 群体主要成分分析 (PCA)。

RNA 水平研究

(一) 转录组测序服务

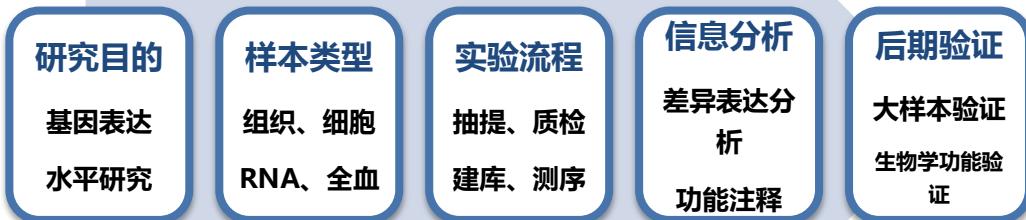
转录组是特定物种、组织或细胞类型转录的所有 RNA (转录本) 的集合，包括 mRNA 和非编码 RNA(Non-coding RNA)，非编码 RNA 又包括：tRNA, rRNA, snoRNA, microRNA, piRNA, lncRNA 等。通过比较转录组或基因表达谱的研究以揭示生物学现象或疾病发生的分子机制是高通量组学研究的一个常用策略。利用高通量测序技术研究转录组在全面快速得到基因表达谱变化的同时，还可以通过测定的序列信息精确地分析转录本的 cSNP (编码序列单核苷酸多态性)、可变剪接等序列及结构变异，另外对于检测低丰度转录本和发现新转录本具有其独特的优势。

技术优势

- ⊕ 直接得到核酸序列信息，除了得到基因表达量的差异，更可以检测 RNA 的结构和结构变异。
- ⊕ 开放性的转录组分析：无需参考基因组信息，无需设计探针，不但能检测已知基因还能够发现新的转录本。
- ⊕ 在测序覆盖率足够大时能够检测到细胞中的低丰度转录本。
- ⊕ 随着测序深度的增加可以获得更广的动态检测范围，能够同时鉴定和定量高丰度转录本和低丰度转录本。

研究内容

- ⊕ **基因表达水平研究：**基因表达是将基因中蕴含的遗传信息通过转录、剪接、翻译等转变成功能产物的所有加工过程，是生物生命活动的基础和关键。基因从 DNA 转录成 RNA 是基因表达的首要步骤。基因转录形成的 RNA 丰度不同很大程度地影响到基因最终功能产物的分子浓度。因此，检测基因在 RNA 水平的表达量即基因转录得到的 RNA 的丰度成为人们研究基因表达的关键方法。



研究流程示例

- 转录本的结构及结构变异研究：利用单碱基分辨率的 RNA-Seq 技术可极大地丰富基因注释的很多方面，包括 5' /3' 边界鉴定、UTRs 区域鉴定以及新的转录区域鉴定等。在发现序列差异(如融合基因鉴定、编码序列多态性研究)方面, RNA-Seq 也展示了其很大的潜力。



研究流程示例

技术路线

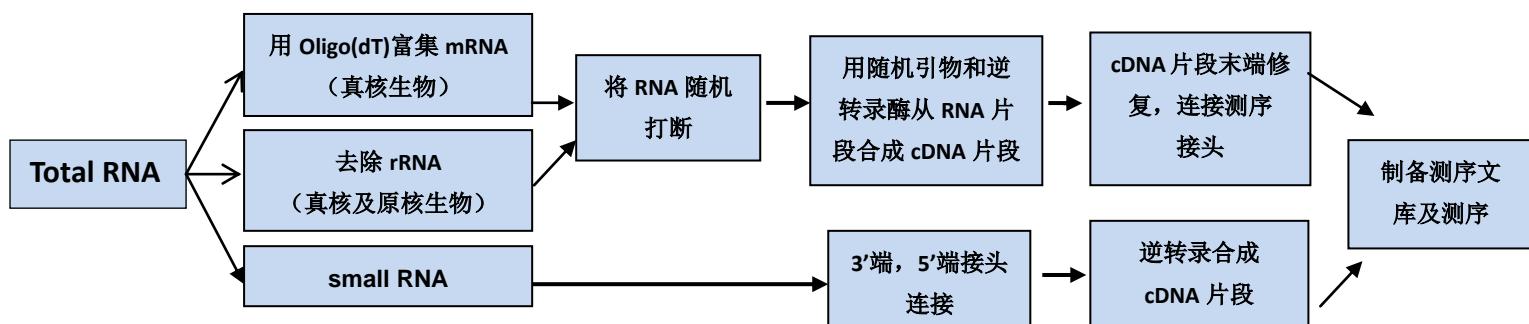
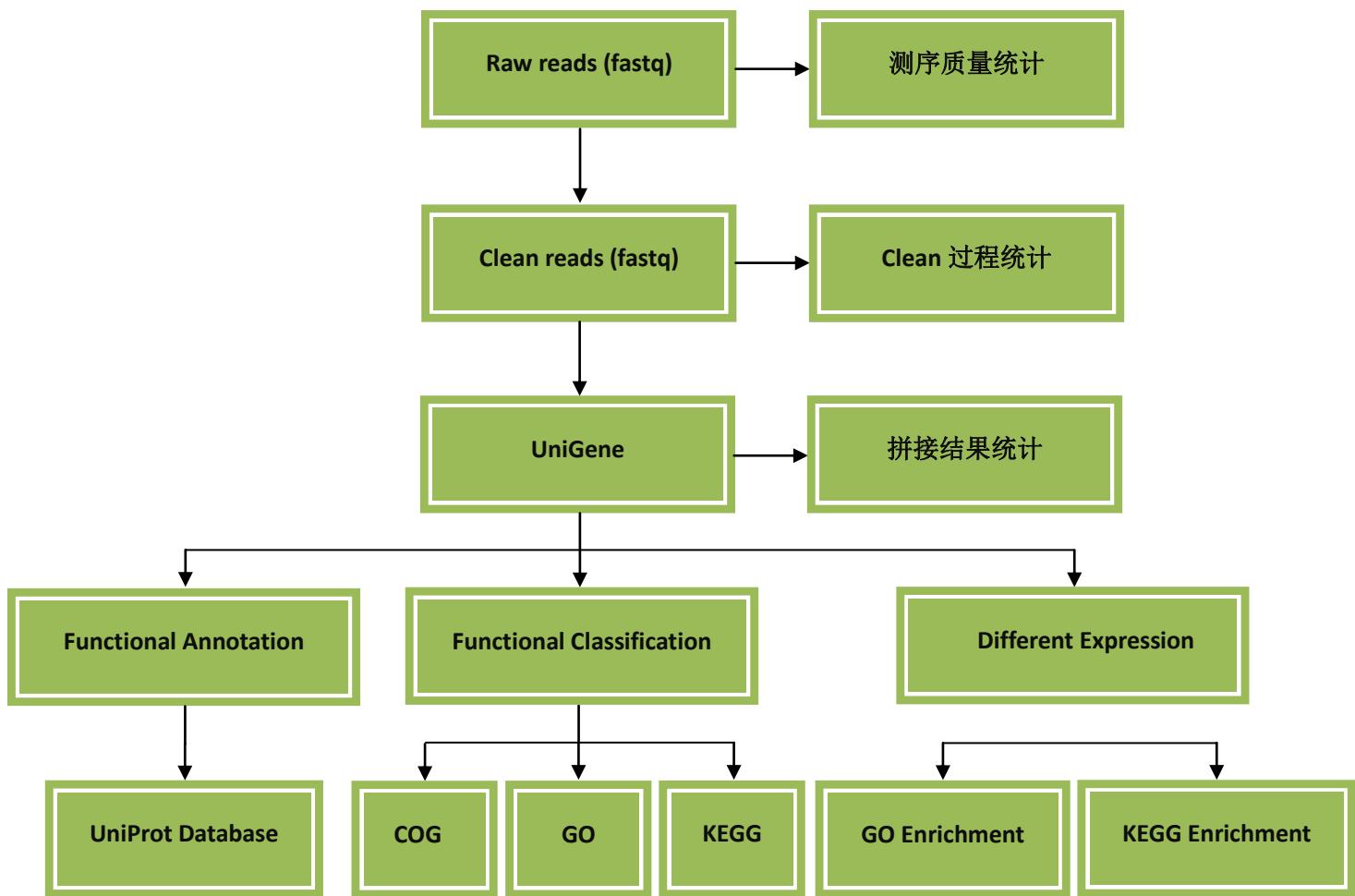


图 RNA-Seq 实验流程

样品要求 (RNA)

- 样品纯度：OD 260/280 值应在 1.9~2.2 之间, RNA 28S:18S \geq 1.5, 推荐 RIN \geq 7; DNA 应该去除干净。
- 样品浓度：最低浓度不低于 100ng/ μ l。
- 样品总量：每个样品总量不少于 7 μ g。
- 样品溶剂：要溶解在 H₂O 或 TE (pH 8.0) 中。
- 样品运输：RNA 用冻存管保存，并用干冰或液氮运输。

数据分析技术路线及数据分析内容（无参考基因组）



转录组测序数据分析（无参考基因组）

1. 数据预处理

目的: 对原始测序数据进行一定程度的过滤。

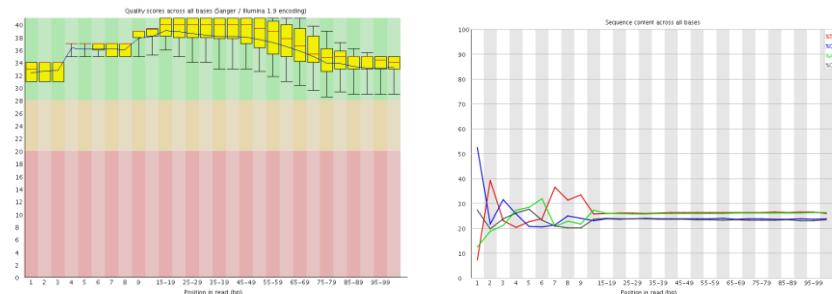
原理: 根据测序接头以及测序质量对原始的测序数据进行预处理，其中，测序质量 Q 与测序错误 E 之间的关系如下：

$$Q = -10 \log_2 E$$

质量与错误率对照表

测序错误率(E)	测序质量值(Q)
5%	13
1%	20
0.1%	30

结果: 对预处理后质量以及碱基分布统计进行统计

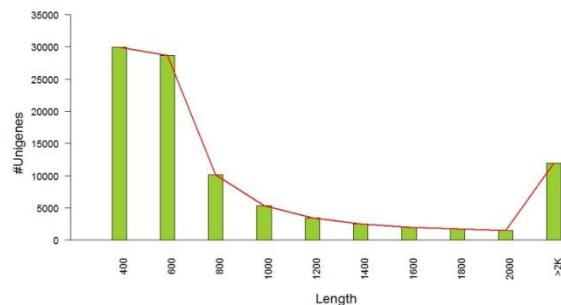


2. UniGene 拼接

目的: 将预处理后 reads 进行拼接, 得到拼接结果。

原理: 应用 de Bruijn graph path 算法对 reads 进行 *de novo* 拼接; 对上一步的拼接结果, 再用 Hamilton Path 算法拼接。

结果: UniGene 序列, UniGene 统计信息, 序列长度分布图

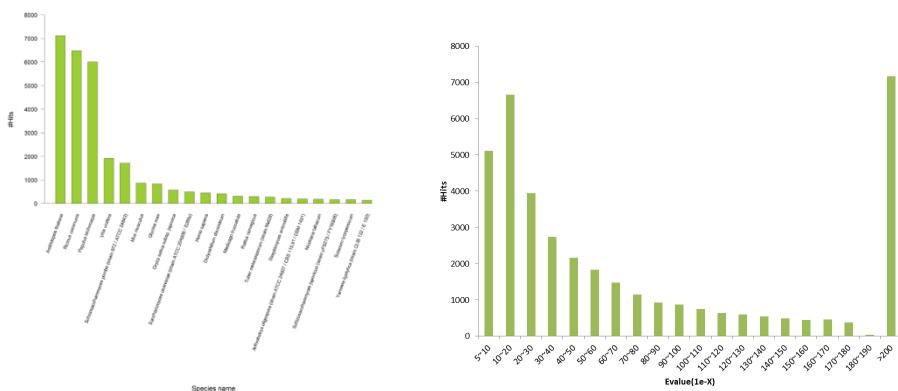


3. 数据库注释

目的: 对拼接得到的 UniGene 进行功能注释

原理: 通过 blast+ 算法将拼接得到的 UniGene 序列与数据库进行比对

结果: 比对结果表格, 物种分布统计和 Evalue 分布统计



4. UniGene 表达分析

目的: UniGene 定量分析。

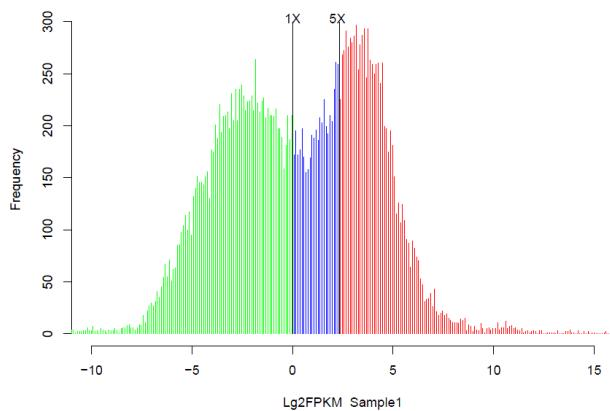
原理: 以 UniGene 为 reference, 分别将每个样本的 reads 进行 reference mapping, 从而得到每个样本在每个 UniGenes 中的一个 reads 覆盖度, 然后应用 RPKM/FPKM 标准化公式对富集片段的数量进行归一化。

RPKM: Reads Per Kilobase of exon model per Million mapped reads, 公式下:

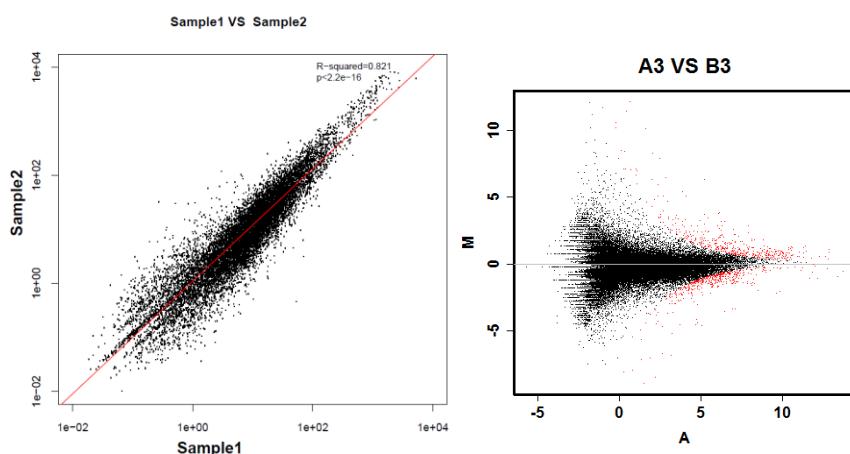
$$\text{RPKM} = \frac{\text{total exon Reads}}{\text{mapped reads}(Millions) \times \text{exon length(KB)}}$$

FPKM: Fragments Per Kilobase of exon model per Million mapped reads, 公式下:

$$\text{FPKM} = \frac{\text{total exon Fragments}}{\text{mapped reads}(Millions) \times \text{exon length(KB)}}$$



UniGene 表达分布图，1X, 5X 分别为 FPKM=1, FPKM=5 分界点，可以大体观察到低表达，中表达以及高表达的比例关系



UniGene 样本间表达相关性散点图 样本间表达差异程度的 MA 图，可以体现差异表达总体偏差

5. UniGene 表达差异分析

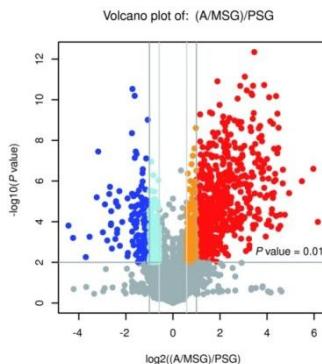
目的：对定量结果进行统计检验分析，找出差异表达 UniGene

原理：双层过滤筛选差异基因

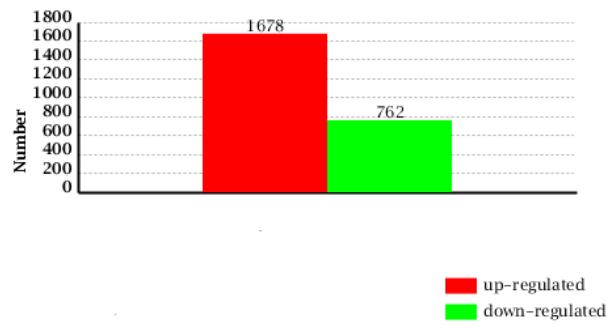
FC 值筛选：采用 Fold-change(FC)，表达差异倍数进行第一层此的差异基因筛选

FDR 检验：一般采用卡方检验中的 fisher 精确检验进行 p 值检验，采用 Benjamini FDR(False discovery ratio)校验方法对 p 值进行假阳性检验，即，通过 FDR 显著性参数进行第二层次的差异基因筛选。

结果展示：



差异基因火山图，可以观察到差异基因总体分布



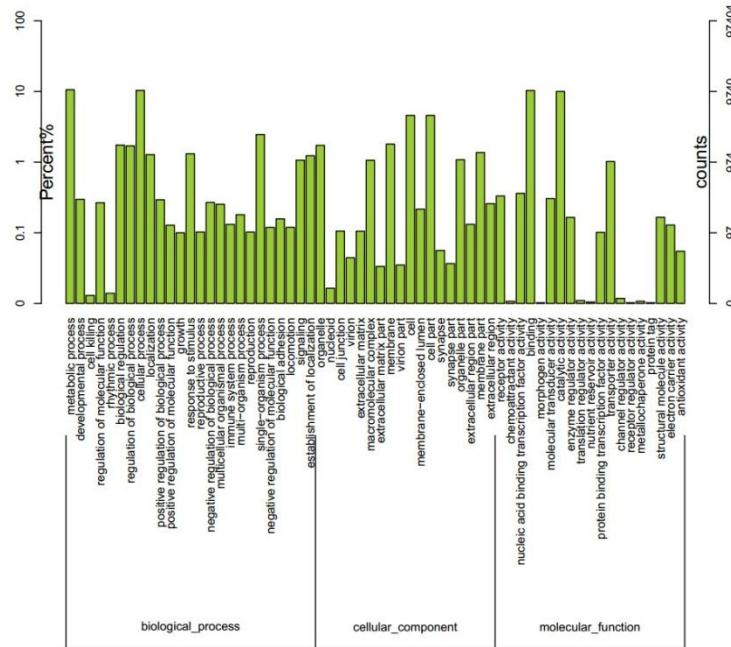
组间差异基因上调与下调个数统计, 可以通过
此图观察上调与下调的一个总体趋势

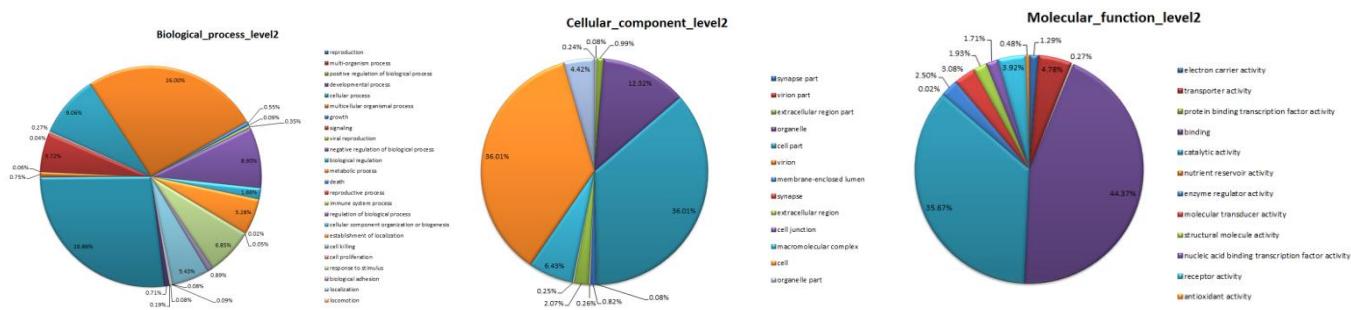
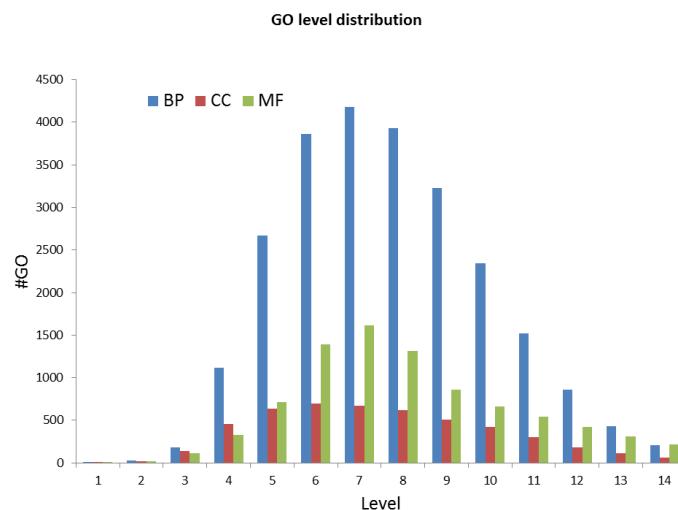
6. GO 功能分类

目的: 利用数据库注释信息将 UniGene 进行 GO 功能分类。

原理: 利用数据库的注释结果, 应用 blast2GO 算法进行 GO 功能分类, 得到所有序列在 Gene Ontology 的三大类: molecular function, cellular component, biological process 的各个层次所占数目, 一般取到 14 层。

结果: MF, BP, CC 三大分类结果文件以及 UniGene2GO 关系列表, 三大类别中第二层次上的柱状分布图和饼图, GO 功能的层次分布图。



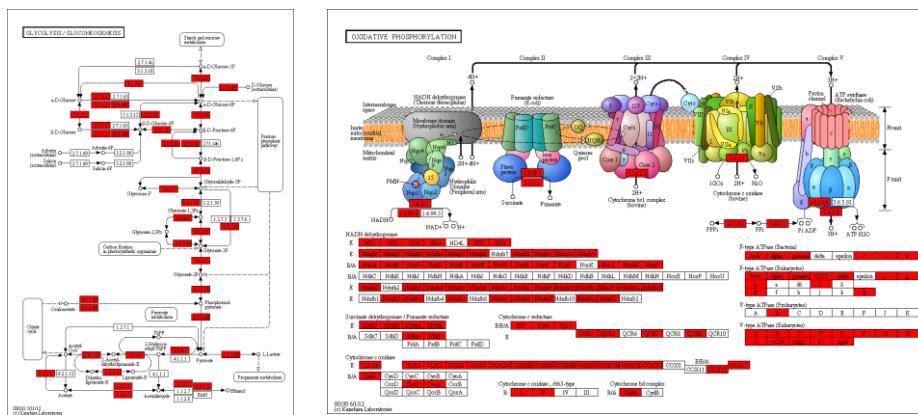


7. KEGG 代谢通路分析

目的：对拼接得到 UniGene 进行 KEGG pathway 映射。

原理：应用 KEGG KAAS 在线 pathway 比对分析工具对拼接得到的 UniGene 进行 KEGG 映射分析。

结果：标记的 Pathway 通路图。



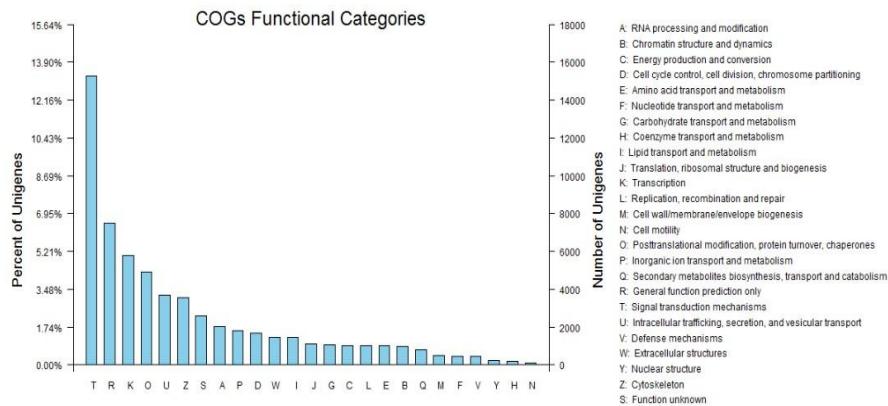
8. COG 注释

目的：对拼接得到 UniGene 进行 COG 功能分类。

原理: 利用 blast+ 算法将拼接得到的 UniGene 与 CDD 库中的 COG/KOG 库进行比对, 进行 COG

功能分类预测，将其映射到 COG 分类中。

结果： COG 分类分布情况图。



9. SSR 重复序列注释

目的：对拼接得到 UniGene 进行 SSR 简单重复序列的查找。

原理：筛选标准：单核苷酸重复的次数在 10 次或 10 次以上，二核苷酸重复的次数在 6 次或 6 次以上，三至六核苷酸重复的次数在 5 次或 5 次以上。同时，也筛选中间被少数碱基（间隔小于 100 或等于 100）打断的不完全重复的 SSR。

结果：重复序列的信息文件以及统计文件。

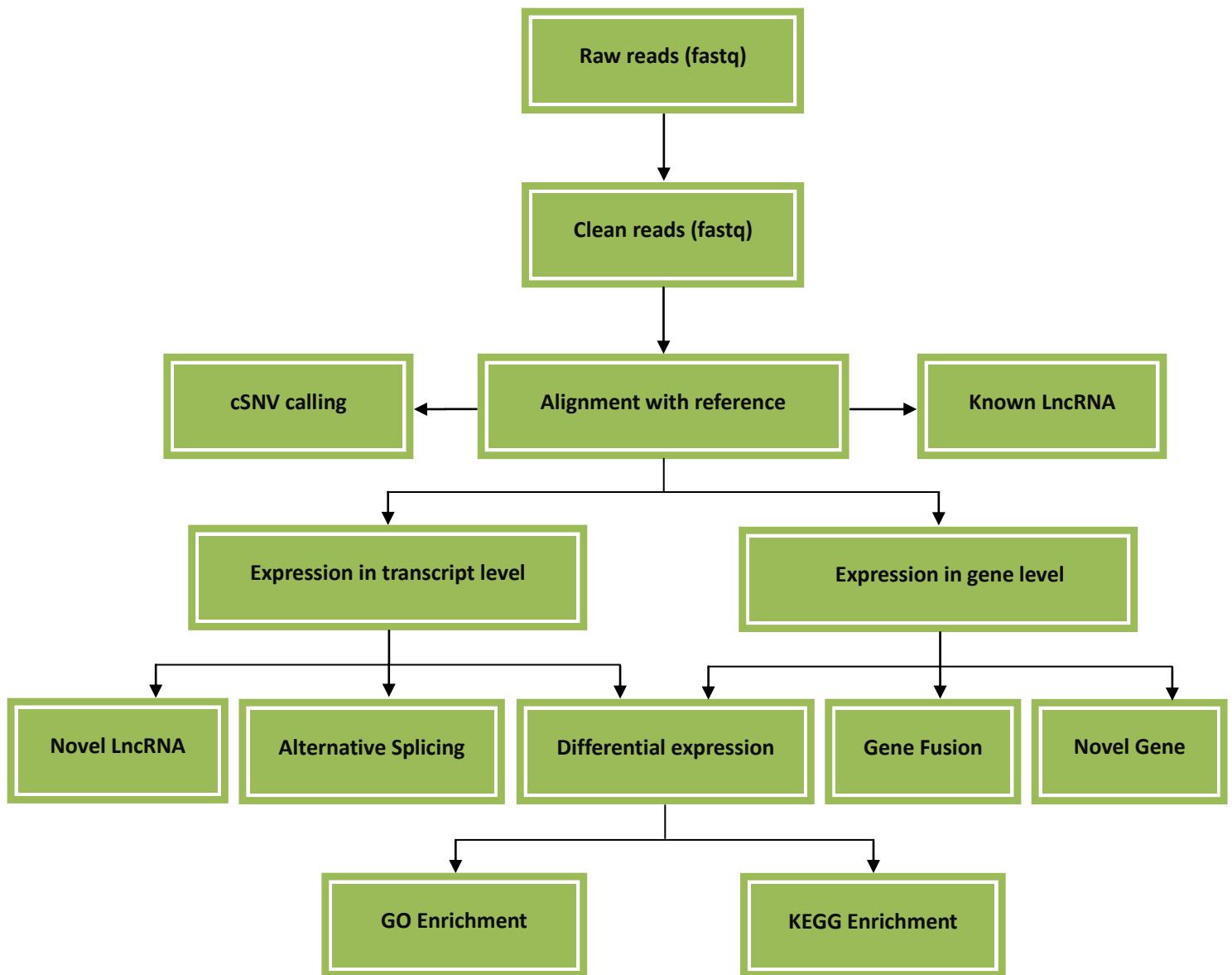
10. LncRNA 预测

目的：对拼接得到的 UniGene 进行 LncRNA(Long noncoding RNA)预测。

原理：通过以下过程对 UniGene 进行过滤，最终得到候选 LncRNA 序列。

- 1) Unigene length > 200bp;
- 2) Unigene ORF(Open Reading Frame) length < 300;
- 3) 将满足长度条件的 UniGene 与多个近源物种进行进化分析，得到序列的保守性和进化特性；
- 4) 根据上述的特性和已知数据库中 coding、noncoding 区域的特性建立编码筛选模型；
- 5) 将符合 noncoding 模型的 UniGene 与 Pfam 等蛋白域数据库进行同源性比对，进一步去除可能的编码特性，最终得出 LncRNA 预测结果。

数据分析技术路线及数据分析内容（有参考基因组）



转录组测序数据分析（有参考基因组）

1. 数据预处理

目的：对原始测序数据进行一定程度的过滤。

原理：根据测序接头以及测序质量对原始的测序数据进行预处理，其中，测序质量 Q 与测序错误 E 之间的关系如下：

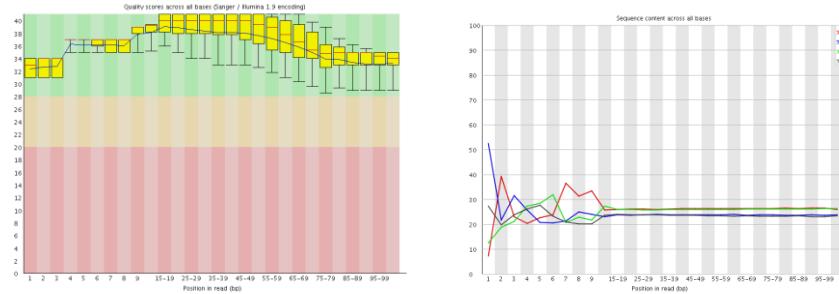
$$Q = -10 \log_2 E$$

质量与错误率对照表

测序错误率(E)	测序质量值(Q)
5%	13

1%	20
0.1%	30

结果展示：对预处理后质量以及碱基分布统计进行统计。



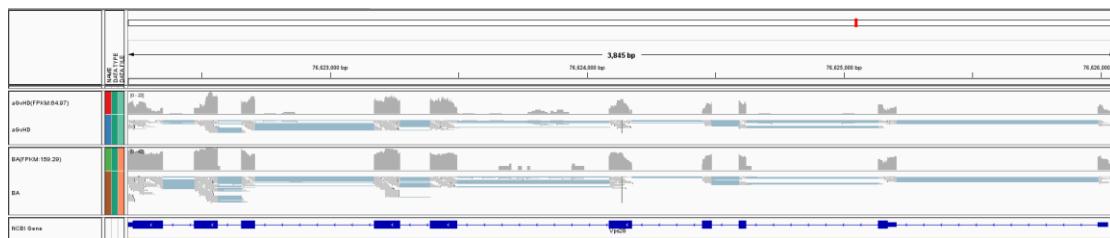
2. 比对基因组

目的：将经过预处理的测序数据与参考基因组进行相似性比对。

原理：Burrower-Wheeler 转换算法与 splicing 比对算法。

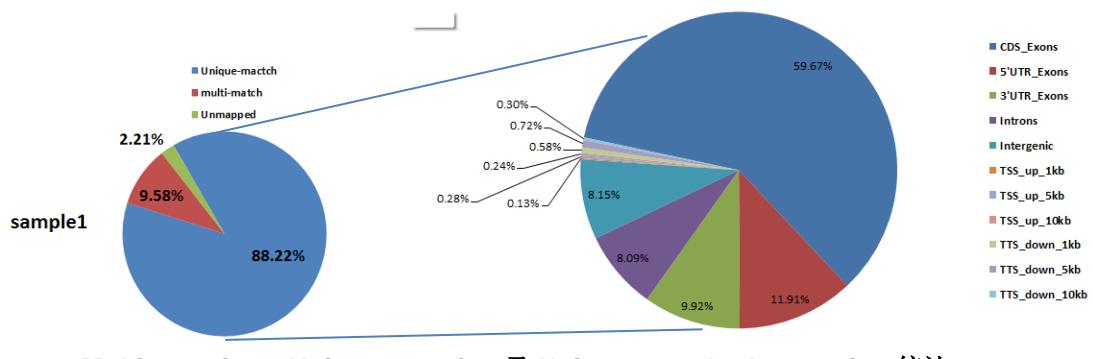
1) **Burrower-Wheeler 转换算法：**由于测序数据量非常大，与整条基因组比对所需资源与时间是较为巨大的。目前，我们采用 Burrower-Wheeler(BWT)算法对基因进行建立索引、碱基压缩等过程，这样可以很大程度上加快比对速度，减少比对过程中所需资源。

2) **splicing 比对算法：**即分段比对算法，当某条测序序列位于转录本剪切位点时，也就是这条序列同时属于两个外显子，如果将它与参考基因组进行比对，由于基因组两个外显子之间含有 intron 区，那么它将无法找到它合适的位置；但是应用分段比对算法就可以将这条测序序列分割变成多段子序列，然后应用这些段子序列与基因组进行比对，这样就可以找到它们真正的位置。

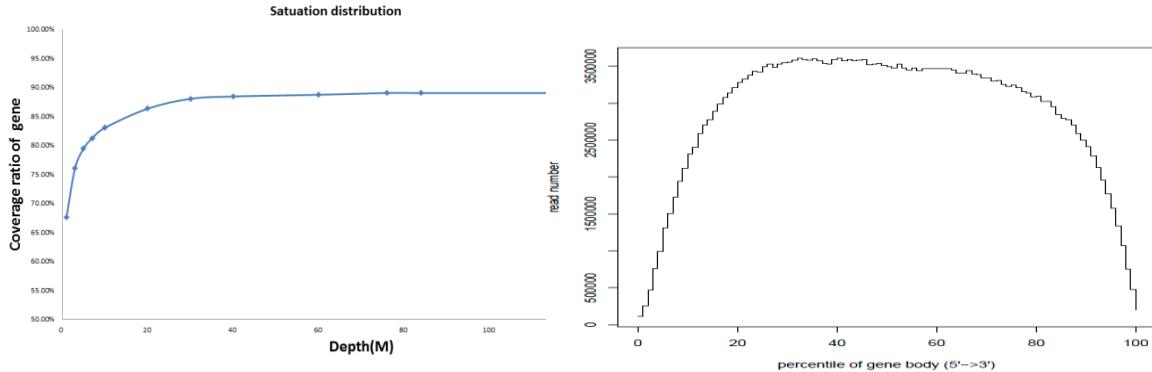


Vps28 基因的一个分段比对的结果，蓝线连接的两端即为被分割的子序列，可见此种算法非常的适用于转录组测序。

结果展示：应用比对结果进行一些相关 mapping 统计，测序饱和度及测序 5'，3' bias 统计。



Multi mapping, Unique mapping 及 Unique gene-body mapping 统计。



饱和度分析，当 reads 达到一定测序量后，基因覆盖率基本达到饱和。

测序 3',5' 偏好性统计，测序主要集中于基因 body 区，两端偏向性较轻。

3. 基因表达水平研究

目的：应用基因组比对结果进行基因定量。

原理：从指定物种基因模型(基因结构)中得到 gene、exon、intron 以及 UTR 等位置信息，通过基因组比对结果计算出在不用区域富集片段数目，然后应用 RPKM/FPKM 标准化公式对富集片段的数量进行归一化。

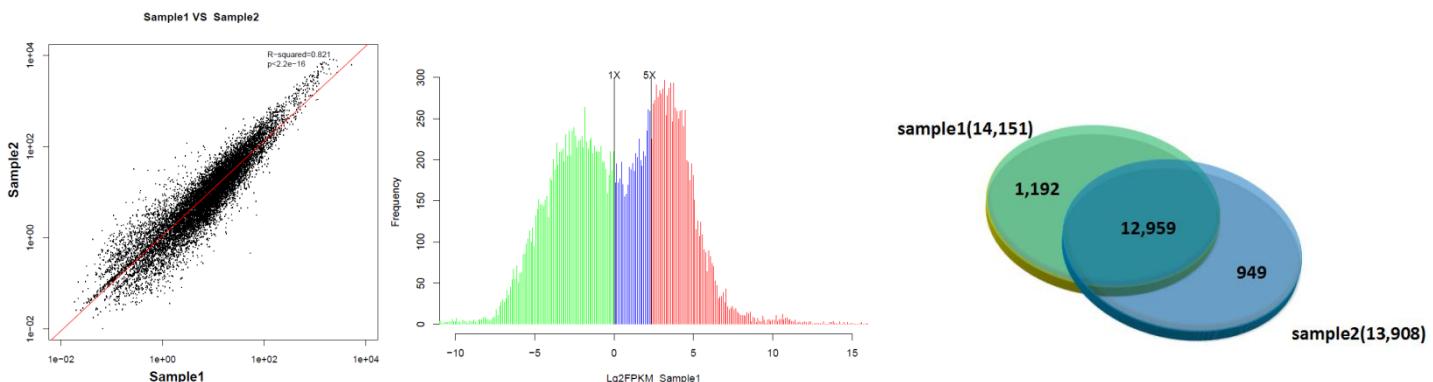
RPKM: Reads Per Kilobase of exon model per Million mapped reads, 公式如下:

$$\text{RPKM} = \frac{\text{total exon Reads}}{\text{mapped reads}(Millions) \times \text{exon length(KB)}}$$

FPKM: Fragments Per Kilobase of exon model per Million mapped reads, 公式如下:

$$\text{FPKM} = \frac{\text{total exon Fragments}}{\text{mapped reads}(Millions) \times \text{exon length(KB)}}$$

结果展示：表达量相关性图以及表达量分布图。



样本表达相关性，通过相关系数 R^2 与显著性 p 值进行相关性检验。

基因表达分布图，1X, 5X 分别为 FPKM=1, FPKM=5 分界点，可以大体观察到低表达，中表达以及高表达的比例关系。

样本间表达基因关系饼图，可以看出共同表达以及独有表达的概况。

4. 差异表达分析

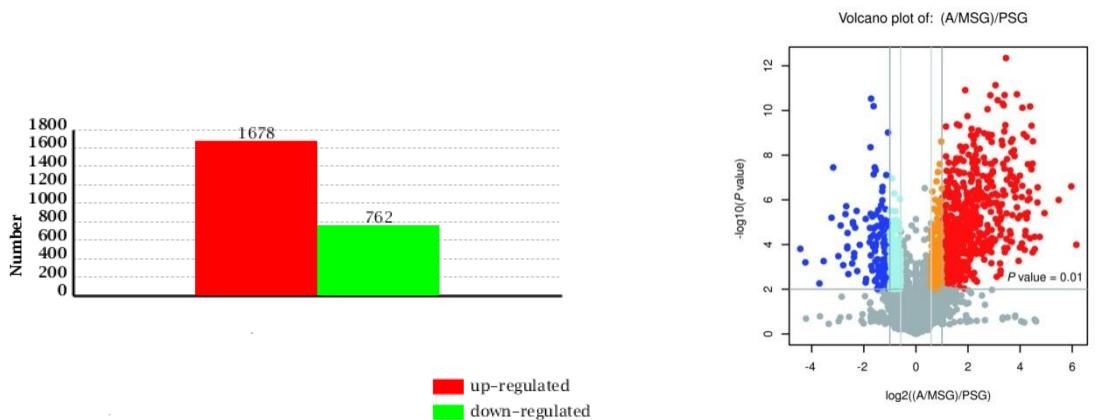
目的：应用统计学方法对基因在样本间的表达差异进行分析。

原理：双层过滤筛选差异基因。

FC 值筛选：采用 Fold-change(FC)，表达差异倍数进行第一层此的差异基因筛选。

FDR 检验：一般采用卡方检验中的 fisher 精确检验进行 p 值检验，采用 Benjamini FDR(False discovery ratio)校验方法对 p 值进行假阳性检验，即，通过 FDR 显著性参数进行第二层次的差异基因筛选。

结果展示：



组间差异基因上调与下调个数统计，可以通过此图观察上调与下调的一个总体趋势

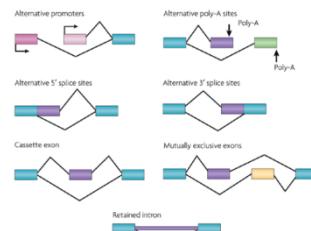
差异基因火山图，可以观察到差异基因总体分布

5. 转录本结构分析

目的：侦测不同类型的可变剪切事件。

原理：通过测序序列的 splicing 事件来侦测可能发生剪切连接的候选 exon，通过已有可变剪切方式进行验证，最终得出真实的可变剪切事件。

结果展示：对常见的可变剪切方式进行统计分析。



6. 新转录本预测

目的：预测 antisense transcript 以及 intron transcript。

原理：通过测序序列在基因组上富集的方向性进行反义转录本预测，如果有富集区域方向与基因转录本方向相反且达到一定的富集阈值，即可认为其为 antisense transcript。将完全位于 intron 区的一段富集片段作为 intron transcript。

7. 新基因预测

目的：预测 intergenic 区可能存在的新基因并对新基因进行功能注释。

原理：首先，得到在基因间区有测序序列富集的一些段区域；然后，排除那些已经有注释的那些段区域作为候选的新基因。

结果展示：

novel_gene_ID	chr_ID	start	end	strand	S1_RPKM	S2_RPKM
XLOC_009269	chr3L	23712607	23712829	+	141.962095	24.358539
XLOC_006159	chr2RHET	2773549	2774009	+	123.868536	4.645177
XLOC_006062	chr2R	15728925	15729265	+	113.252178	10.787401
XLOC_016026	chrX	8235861	8235980	-	110.055983	129.838445
XLOC_002947	chr2L	22888366	22888609	+	108.881655	27.201717
XLOC_006112	chr2R	20626356	20626726	-	105.763197	8.669353
XLOC_002884	chr2L	15107516	15107665	-	103.177186	165.740292
XLOC_009103	chr3L	9060330	9061299	+	98.347576	163.966484
XLOC_009160	chr3L	12701929	12702611	+	93.382837	103.650211
XLOC_009231	chr3L	20756764	20757020	+	91.713319	10.731791

新基因预测结果，给出了新基因起始、终止位置以及在样本间的表达值。

8. 基因融合分析

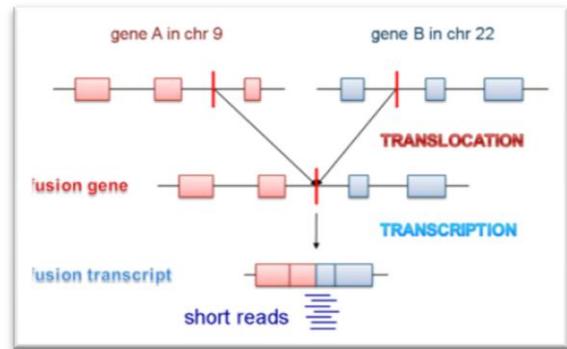
目的：寻找可能发生融合功能的基因

原理：通过测序片段的 splicing 事件以及 pair-end 测序的

距离信息进行基因融合位点的定位，

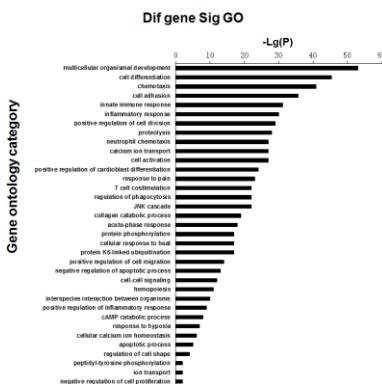
如果一个测序片段的一个子片段与 geneA 匹配，另一子片段与 geneB 匹配，那么 geneA 与 geneB 有可能为一个融合基因，而当 pair-end 双向测序时，一对测序片段中一个与 geneA 匹配，另一个与 geneB 匹配，那么 geneA 与 geneB 有可能为一个融合基因。如果同时满足两个条件，那么融合发生的可能性就较大。

结果展示见右图。



9. GO 富集分析

目的：对差异基因相关 GO 功能进行富集分析。

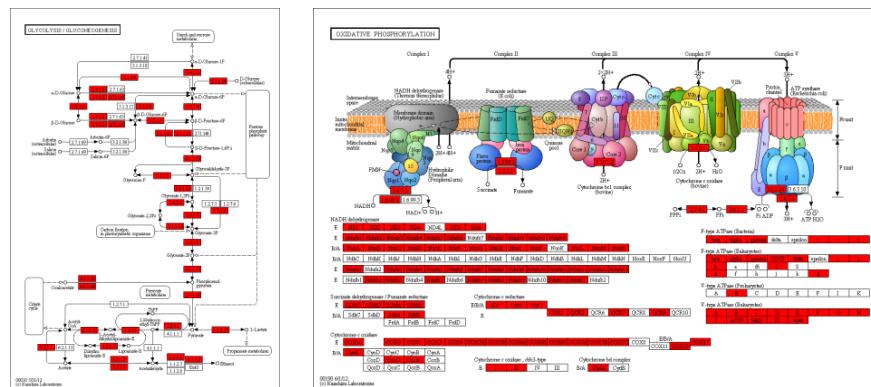


显著富集 GO 功能图形统计

10. KEGG 富集分析

目的：对差异基因进行 KEGG 通路富集分析。

原理：应用物种自己的 KEGG pathway 进行富集分析，富集结果更加贴近物种现实功能实现的通路，尤其对目前功能注释尚不完全的物种，如，大豆、玉米、葡萄、杨树、白菜、牛、羊等物种的 KEGG 通路分析。



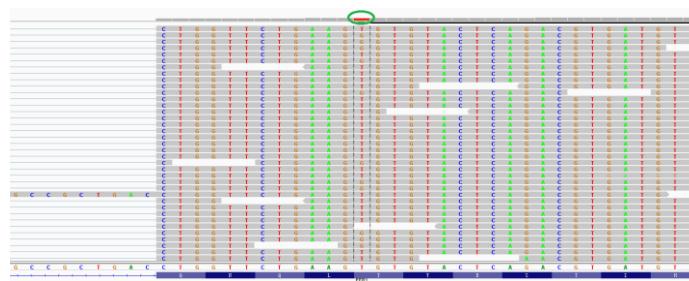
有显著富集功能 KEGG 通路图，其中，红色标记为差异基因

11. cSNV 查找

目的：在转录水平找出变异位点或者片段。

原理：通过测序数据得到基因组每个位点的碱基富集情况；然后，统计每个点富集富集的碱基种类，得出可能存在的变异(即，与参考基因组碱基不同且富集程度较高的碱基类别)。

结果展示：



12. LncRNA 预测

目的：对新转录本进行 LncRNA(**Long noncoding RNA**) 预测。

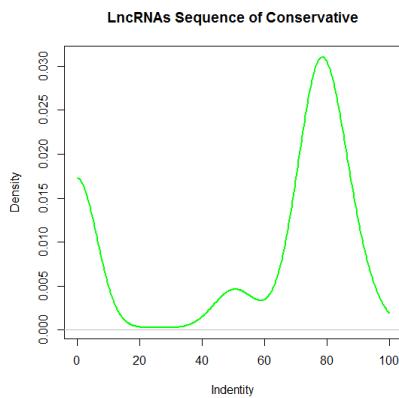
原理：通过以下过程对新转录本进行过滤，最终得到候选 LncRNA 序列：

- 1) 通过基因组比对得到 4 类新转录本：Intergenic transcript、Full intron transcript、Antisense transcript、Overlapped with known transcript，将这些新转录本用于 LncRNA 预测；
- 2) New Transcript length > 200bp；
- 3) New Transcript ORF(Open Reading Frame) length < 300；
- 4) 将满足长度条件的 New Transcript 与多个近源物种进行进化分析，得到序列的保守性和进化关系；
- 5) 根据上述的特性以及已知数据库中 coding、noncoding 区域的特性建立编码筛选模型；
- 6) 将符合 noncoding 模型的 New Transcript 与 Pfam 等蛋白域数据库进行同源性比对，进一步去除可能的编码特性，最终得出 LncRNA 预测结果。

13. LncRNA 保守度分析

- 1) lncRNA 序列保守性分析 lncRNA 序列保守性分析

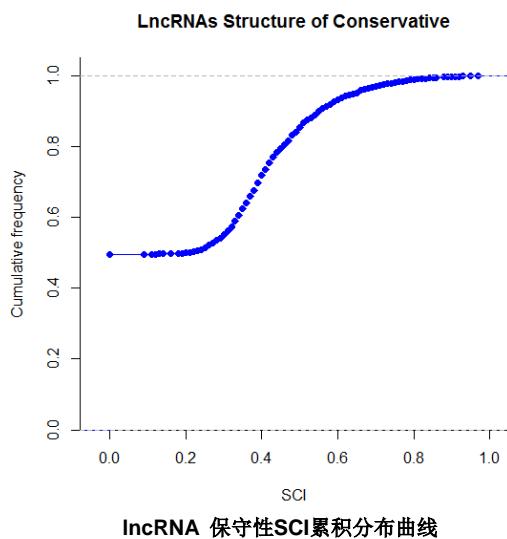
基于multiple genome alignment进行序列保守性分析。lncRNA的序列保守性用其序列比对的identity来表示。



Identity score的密度分布图

2) lncRNA 结构保守性分析

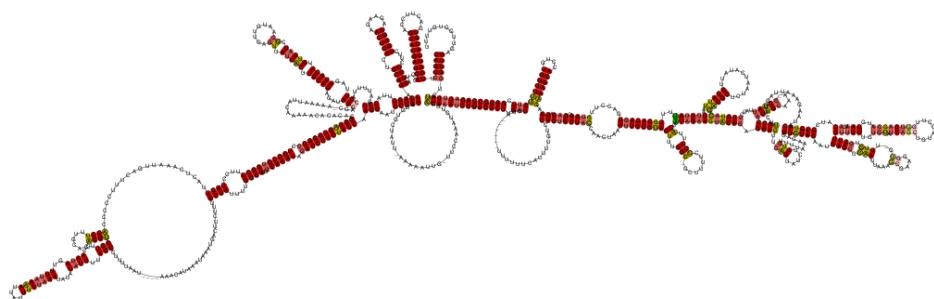
通过使用 RNAz 软件，推断基于多序列比对的一致结构，比较二级结构自由能与一致结构的自由能，得到每一段 aligned region 的 SCI 值，并以此作为二级结构的表征，SCI 越大则结构越保守。



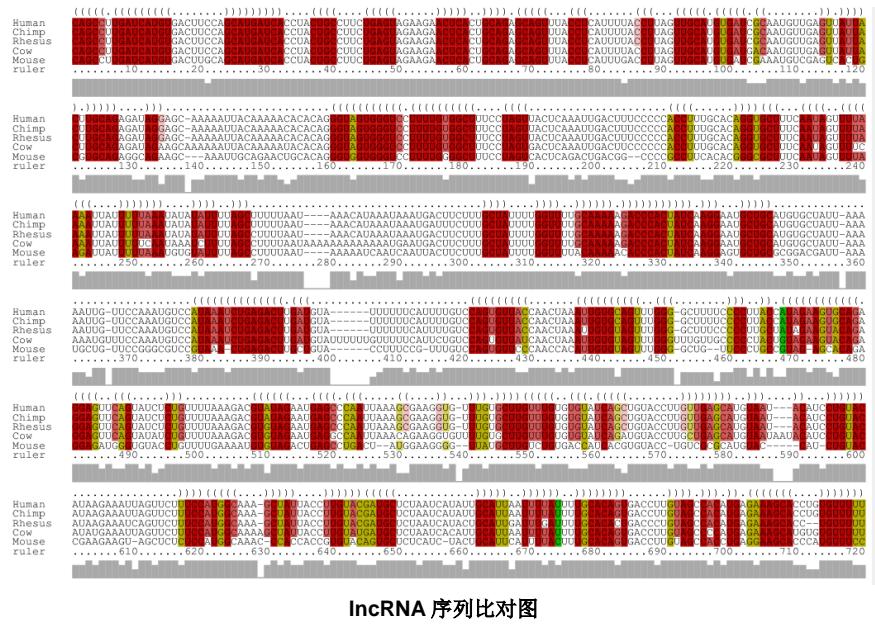
lncRNA 保守性SCI累积分布曲线

3) LncRNA 二级结构预测

根据 lncRNA 的序列保守性，使用 RNAalifold 软件，可以同时看到其序列比对结果和保守的二级结构。



lncRNA 保守二级结构图



IncRNA 序列比对图

伯豪案例

玉米 Proline responding 1(pro1)突变在蛋白合成和细胞周期调控中起到关键作用

本研究证实玉米 Pro1 基因(Zm P5CS2)的突变造成了突变体细胞中脯氨酸(proline)合成受阻，从而导致 proline 积累的减少。Pro1 突变基因的克隆和功能分析解析了脯氨酸在植物生长发育过程中对蛋白合成和细胞周期 G1 到 S 期转换的调控机制。

原文出处： Proline responding1 Plays a Critical Role in Regulating General Protein Synthesis and the Cell Cycle in MaizeC. The Plant Cell.2014. (IF9.251)

芥菜型油菜种皮转录组 *de novo* 拼接及类黄酮生物合成基因的鉴定

本研究是湖南农业大学油料作物研究所刘忠松教授课题组完成的。该研究首次实现芥菜型油菜种皮转录组测序，研究中所获得的基因不仅有利于阐明芥菜型油菜种皮着色的分子机制，且为该物种今后的基因组学研究提供了基础。

原文出处： De Novo Transcriptome of Brassica juncea Seed Coat and Identification of Genes for the Biosynthesis of Flavonoids. PLoS One. 2013, 8(8):e71110. (IF3.730)

(二) miRNA 测序服务

上海伯豪生物技术有限公司应用 Illumina 第二代高通量测序技术对 microRNA 及 Small RNA 文库进行测序，可以一次获得数百万条 microRNA 序列，能够快速鉴定出不同物种、不同组织、不同发育阶段、不同疾病状态下已知和未知的 microRNA 及其表达差异。同时，新一代测序技术在 microRNA 互补链、microRNA 编辑、microRNA 异构体检测、新 microRNA 预测及 microRNA 靶基因分析方面也具有重要应用意义，为研究 microRNA 对细胞进程的作用及其生物学影响提供了有力工具。

技术特点

高灵敏度：理论上可以检测单个细胞中一个拷贝的 microRNA；

高精度：可以检测 microRNA 单个碱基的差异；

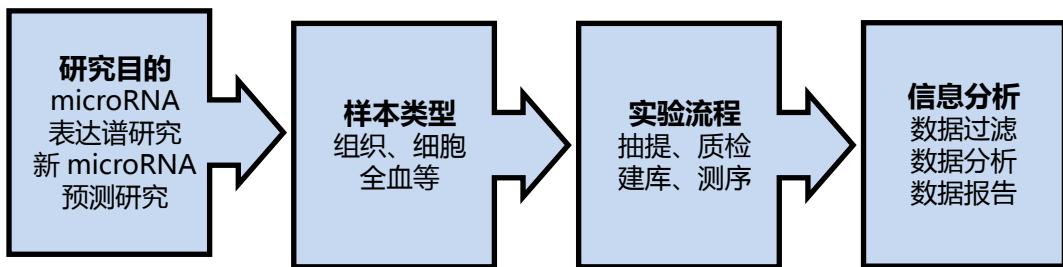
不受先验信息的干扰，既能鉴定已知 microRNA，又有能力发现新的 microRNA 和其它 Small RNA；

保留定向信息，用于链特异性的表达分析；
利用Barcode在单次运行中经济地分析多个样品。

服务内容

对客户提供的total RNA或microRNA样品进行定量检测，检测合格后进行microRNA文库构建、簇生成及上机测序。上海伯豪生物microRNA sequencing服务，为您提供从样本抽提到数据分析的全面综合服务，服务内容包括：样品抽提质检，文库建立，测序，生物信息学分析以及数据报告。

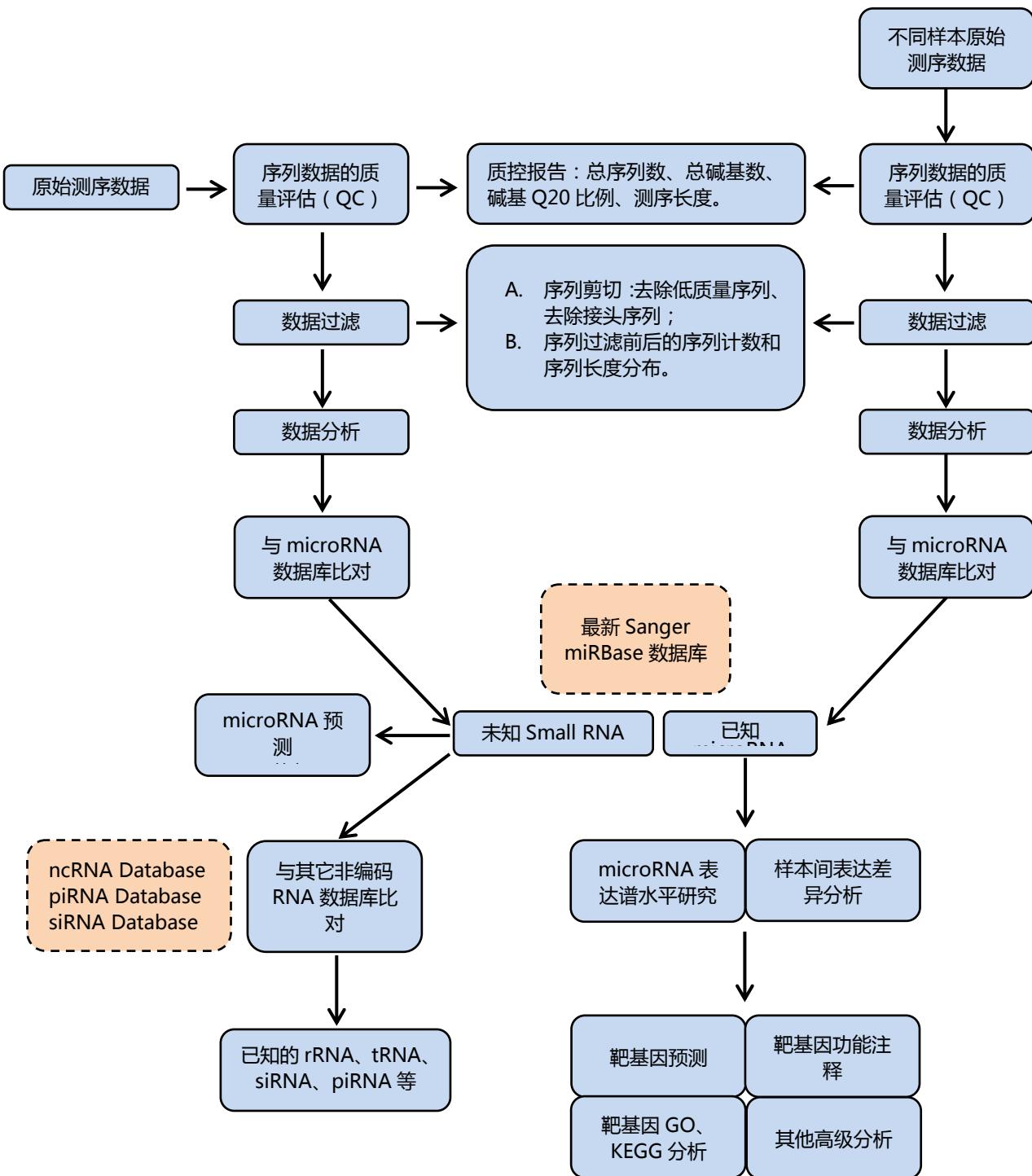
服务流程



文库建立流程



数据分析流程



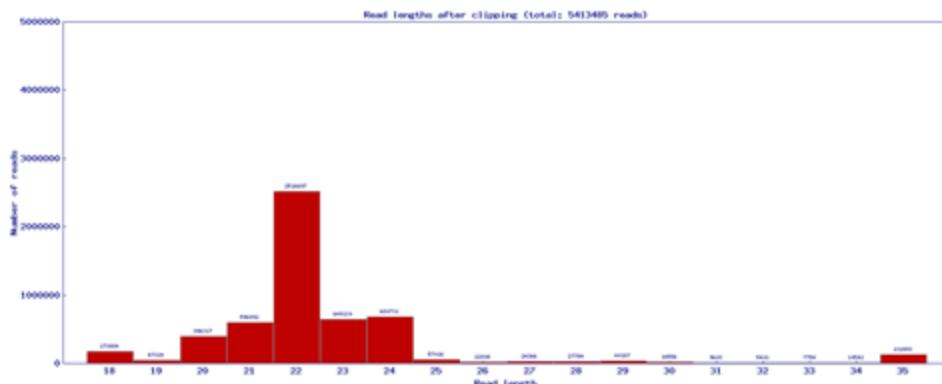
microRNA(Small RNA) sequencing 数据分析流程

数据分析内容

基础分析	高级分析
原始数据处理，数据产出统计，对原始测序数据去接头污染，去低质量reads	差异的已知miRNA的靶基因GO注释和KEGG
18~30 nt 小RNA 测序结果的长度分布统计 小RNA的分类注释及统计 ✓ 小RNA与miRBase 的已知的 miRNA的比对 ✓ 小RNA与rRNA、tRNA、snRNA、snoRNA的比对信息	novel miRNA的靶基因GO注释和KEGG Novel miRNA 差异分析（2个或2个以上样品）和聚类分析
已知miRNA的表达谱构建 ✓ 单个样本中miRNA表达量分析（TPM） ✓ 组间差异miRNA的筛选	miRTarBase（已有实验证据的miRNA靶基因数据库） 关联分析
miRNA表达模式聚类分析	差异miRNA及靶基因的网络图
差异的已知miRNA的靶基因预测	miRNA测序结果与mRNA表达谱数据的关联分析（需要有mRNA表达谱数据）
预测新的miRNA，绘制新的miRNA 的二级结构图	差异miRNA的primary-miRNA与转录因子的关联分析
novel miRNA的靶基因预测	

1. MicroRNA长度分布统计以验证试验可靠性

应用fastx(fastx_toolkit-0.0.13.2)对测序原始reads进行预处理，去除接头序列以及低质量序列（包括模糊碱基N，碱基质量小于10以及长度小于18nt的序列），最终给出处理结果统计表以及长度分布图。



图为经过长度统计，各长度上序列的分布情况

2. MicroRNA比对注释统计

将测序得到的序列与最新Sanger miRBase数据库，已知rRNA、tRNA、重复区、RefSeq数据库，以及其他非编码数据库如ncRNA, piRNA, Rfam数据库比对，对已知microRNA进行注释。

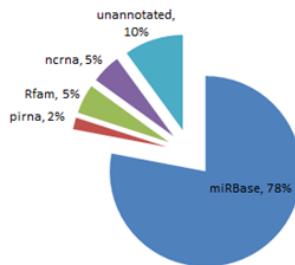
下表为经过注释的结果，其中分别列出和miRBase数据库，piRNA数据库，Rfam数据库以及ncRNA数据库的比对情况。

Resource	Sequences in resource	Sequences found	Percentage found
miRBase (Homo sapiens)	1,600	705	44.1%
piRNA	171,551	655	0.4%
Rfam v10	444,417	25,940	5.8%
Homo_sapiens.GRCh37.68.ncrna	20,677	2,409	11.7%

下图为针对miRBase种Sus scrofa物种进行的比对注释统计：

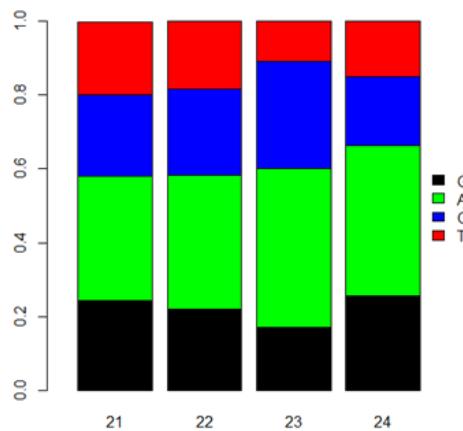
Organism	Total	Mature 5' total	Mature 5' exact matches	Mature 5' length variants	Mature 5' mutant variants	Non-mature total	Mature 3'	Precursor
Sus scrofa	9,323,383	764,675	394,087	308,794	61,794	8,558,708	8,409,720	148,988

由之前所得的注释结果，可以作图来更进一步展示其结果：

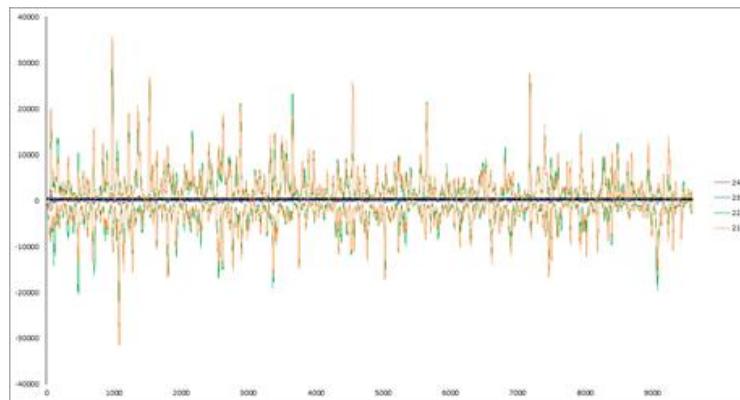


对整体的注释结果，还可以采取进一步的分析，例如：

- (1) 统计碱基偏好性，下图就是测序所得序列分别在21, 22, 23, 24长度上的5' 碱基分布情况。

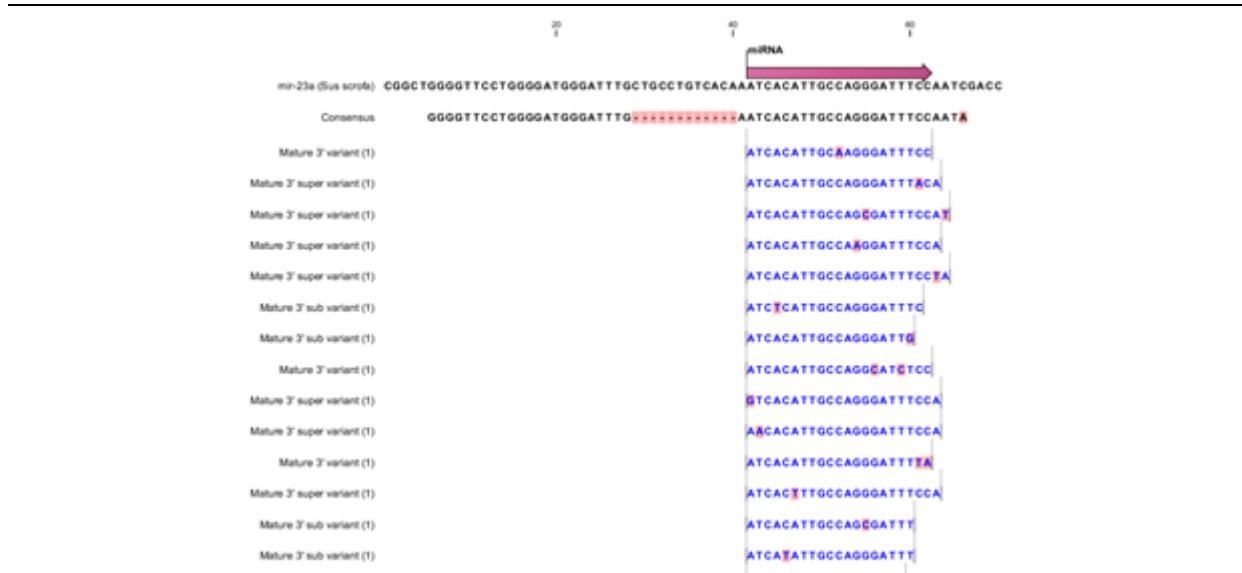


- (2) 对于测序所得序列，可以统计出其正负链分布情况，以找寻生物学上的特征。

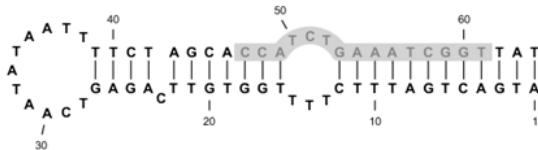


针对某单一microRNA，也可以对其进行更深度的分析。

例如，对其序列的匹配情况进行分别统计：



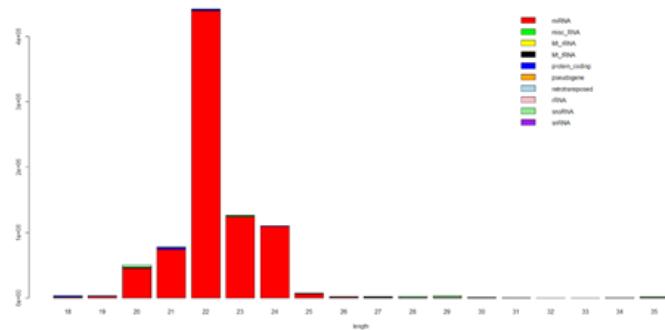
还可以对其对应的microRNA前体二级结构进行观察。



3. 其它Small RNA的分类注释

将测序得到的序列与物种所对应的基因组数据库比对，对有注释的reads的来源进行分类统计，鉴定并统计出已知的microRNA及其它各种不同种类的small RNA分子。

如下图，经过与数据库进行分别比对，可以鉴定并统计出包括tRNA、rRNA、snoRNA、snRNA的数量及分布。



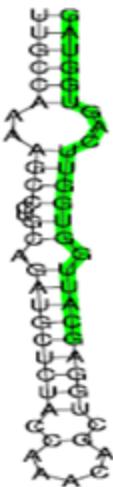
4. 新microRNA预测

MicroRNA前体的标志性发卡(hairpin)结构能够用来预测新的microRNA。对于测序结果中未比对注释上的序列，我们将其与该物种的全基因组序列进行比对分析，通过折叠模型分析，若有序列位于茎环结构上，则初步判定该序列为一个候选的新microRNA。

对于预测出的新microRNA，会统计并列出其所位于的染色体，起始位置，终止位置，正负链，以及数目，长度，GC含量，最小自由能等数值。

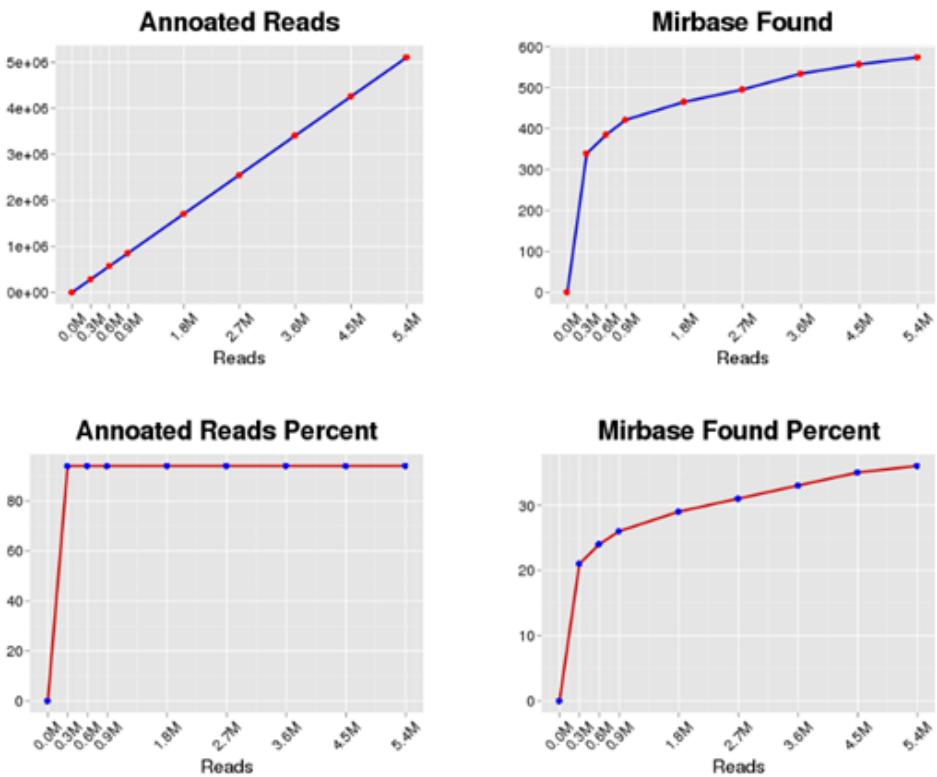
Chromosome	Start	End	Orientalis	Abundance	Sequence	nDNA length	# Genomic Hits	Escherichia Length	Escherichia % G/C content	Minimum Free Energy	Adjusted Rand Index
1	16414479	16414499+		5	TGTTTGCTTGTGTTGTA	23	6	83	43.27349398	-41	-49.3976 0.00990
1	64725663	64725684	9	AACATCTTGGATAGTC		22	1	70	47.14285714	-29.4	-42 0.00990
1	1.29E+08	1.29E+08	6	CCTGGCTGCGAAGCTG		22	1	156	47.43589744	-65.4	-29.7436 0.00960
1	1.88E+08	1.88E+08	5	TTTGGCGGTTGATATG		22	1	84	42.85714286	-32.9	-39.1667 0.00990
1	1.81E+08	1.81E+08	5	GTTTGGGAGGGAGGAC		20	1	87	64.36781609	-37.1	-42.6437 0.0990
1	2.02E+08	2.02E+08	21	ACGGCGGGGGGGGGGG		23	1	55	70.90909091	-33.2	-46.3836 0.01980
1	2.46E+08	2.46E+08	8	TGTTGGCTTGTGTTGTA		22	1	95	53.68421053	-47.6	-56.1053 0.00990
1	3.03E+08	3.03E+08	23	CAGGGGGGGGGGGGGG		23	2	50	68	-27.3	-54.6 0.00990
10	26508853	26508872	5	ATGGTGGGGAGGGAGC		20	1	100	44	-31.7	-31.7 0.05040
10	26718209	26718230	23	TGTTGGCTGGGGGCTA		22	16	120	55	-64.1	-53.4167 0.00990
10	27752746	27752766	13	TAGGGGGGGGGGGGAA		23	2	81	35.80246914	-53.8	-66.4198 0.00990
10	37874204	37874224	6	TGGCTGGGGCTGGGAA		23	1	51	70.58823529	-23.1	-45.2941 0.05940
11	10414941	10414999	13	CTGGGGGGGGGGCTG		22	1	76	68.4210563	-49	-52.6316 0.02970
11	4056392	4056412	33	ATGCTACGCTGCTG		21	1	127	52.75590581	-61.5	-48.4282 0.00990

对于新microRNA，我们还会计算并绘制出其前体的二级结构，以及其与成熟microRNA之间的位置关系。



5. 饱和度分析图表

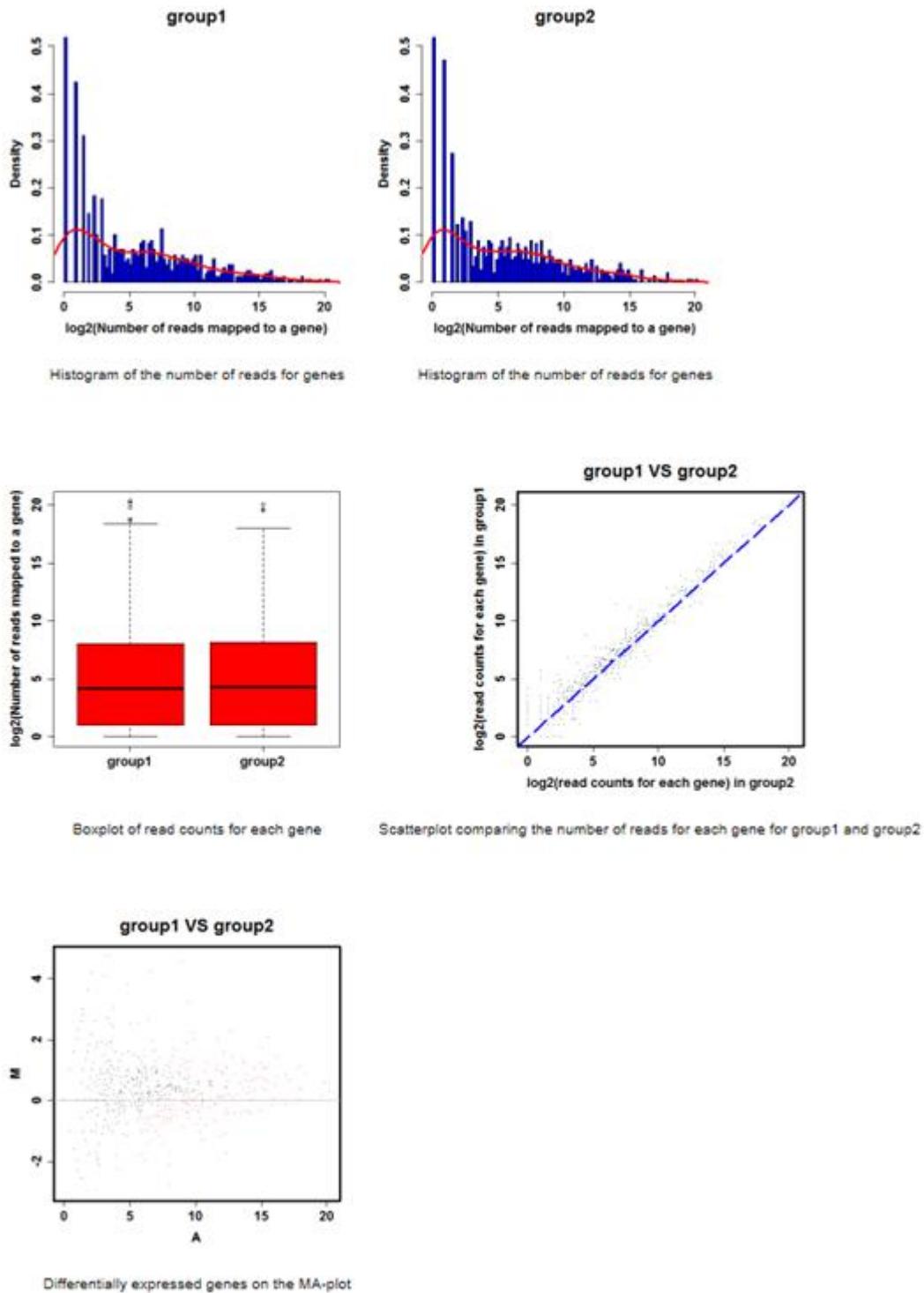
饱和度分析：随着测序量的增加，检测到的miRBase上的microRNA是否随之上升。我们会将注释结果按比例划分作图，以观察样品测序结果注释的趋势，发现其在生物学上的合理性。



6. 样本间microRNA表达差异的筛选

采用DEGseq R语言包结合perl脚本将样品按照客户的分组情况（如：对照组与实验组），进行microRNA表达量的比较分析。在差异分析中，采用TPM（Transcripts per million，公式为：单一microRNA reads数 $\times 10^6$ /总reads数）作为标准化数据。

结果展示如下：



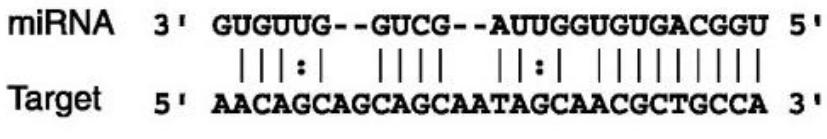
差异microRNA表达模式聚类分析

可以对组间差异表达的

7. 靶基因预测

采用miranda软件，对microRNA序列以及对应物种的基因组cDNA序列进行可能的靶位点预测。

Miranda软件比对结果示意图如下：



Match Score: 115

或采用TargetScan数据库关联microRNA的靶基因，结果形式包含microRNA和靶基因的相关信息。

8. GO功能显著性富集分析（靶基因）

9. Pathway显著性富集分析（靶基因）

伯豪案例

应用案例一：雌雄同株异花的毛白杨 microRNA 的性能分析以及对靶基因表达的影响

本研究首次报道了与花发育相关的 miRNAs 以及它们的靶基因，根据基因注释，将所有的靶基因分为四类。第一类，花发育相关：Pto-F6, Pto-F11, Pto-F14, Pto-F19, Pto-25, Pto-36, Pto-F51, Pto-F54 以及 Pto-F66。第二类，Ca²⁺转运相关：Pto-F28 和 Pto-F45(雌花特异性 miRNA), Pto-F16(雄花特异性 miRNA)。第三类，激素调节相关：10 个 miRNAs (除 Pto-F6 以外均为雌花特异性)。第四类，与 DNA 甲基化相关：3 个 miRNAs。

原文出处：Sexual dimorphism floral microRNA profiling and target gene expression in andromonoecious poplar (*Populus tomentosa*). PLoS One. 2013, 8(5):e62681. (IF3.730)

应用案例二：MeCP2 调控 DGCR8/Drosha 复合物抑制 miRNA 加工和树突发育。

MeCP2 是一种甲基化 DNA 结合蛋白，负责招募转录抑制复合物并且关闭基因表达，重要转录调控因子。

人类 MeCP2 基因的突变或者拷贝数增多均会导致孤独症 (Autism) 等发育性神经系统疾病。中科院上海生科院神经科学研究所仇子龙研究组利用高通量测序技术对 MeCP2 基因敲除小鼠海马区的成熟小 RNA 进行定量分析，发现 MeCP2 抑制了大量小 RNA 的产生。该研究揭示了孤独症相关蛋白 MeCP2 参与小 RNA 剪切加工的新功能，并提示此功能很是 MeCP2 基因突变导致发育性神经系统疾病的相关致病机理。

原文出处：MeCP2 Suppresses Nuclear MicroRNA Processing and Dendritic Growth by Regulating the DGCR8/Drosha Complex p547. Developmental Cell. 2014, 28, 547 - 60. (IF: 10.366)

(三) CircRNA-Seq

环形 RNA 是 mRNA 在剪接的过程中，上游 exon 的 5' 端与下游 exon 的 3' 端剪接到一起，从而形成的首尾相接的环状 RNA 分子。近来的研究表明，高等动物中环状 RNA 的种类和含量远远超过预期。最近的研究发现 circular RNA 可以作为“microRNA 海绵”，竞争结合 microRNA 从而解除这些 microRNA 对其他靶标的调控；同时保守型分析发现环状 RNA 上潜在具有多种 RNA 结合蛋白的结合位点，暗示 circular RNA 可以调控 RNA 结合蛋白的功能。更为重要的是，circular RNA 的表达呈现出很强的组织特异性，暗示这些 circular RNA 具有重要的生物学功能。但目前对 circular RNA 的研究刚刚起步，这类 RNA 的生物学功能还有待挖掘，因此是一个值得深入的研究方向。

环状 RNA 特点

1. 大多存在于细胞质中，且序列高度保守；
2. 一般细胞或组织中，环状 RNA 较稳定，半衰期在 48h，其他 mRNA（线性）半衰期约为 10h；血清中的外显子环状 RNA 不稳定，半衰期<15s。
3. 不同细胞中表达的环状 RNA 可能不同；
4. 同一个基因可能即产生线性 mRNA，又可以转录成环状 RNA；
5. 总量估计，外显子环状 RNA (exonic circRNA) 的量约为 polyA mRNA (线性 mRNA) 的 1%；占所有非核糖体 RNA 的 0.8%；
6. 大部分为 ncRNA。

样本要求

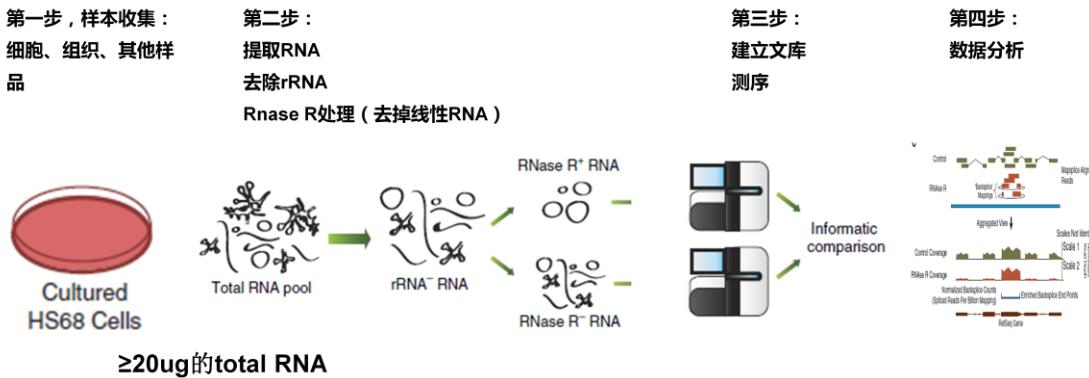
样本类型：细胞、新鲜组织或足量 total RNA

样品量：≥20ug 的 total RNA (需保证提供的细胞及组织样本可以提取到足量 RNA)，最低浓度不低于 100ng/μl

样本质量：OD 260/280 值应在 1.9~2.2 之间；DNA 应该去除干净，RNA 无降解；RIN 值大于等于 7

样品保存运输：RNA 用离心管保存，封口膜封口后用干冰运输，组织样品则用冻存管保存，干冰或液氮运输

CircRNA-Seq 实验流程



数据分析内容

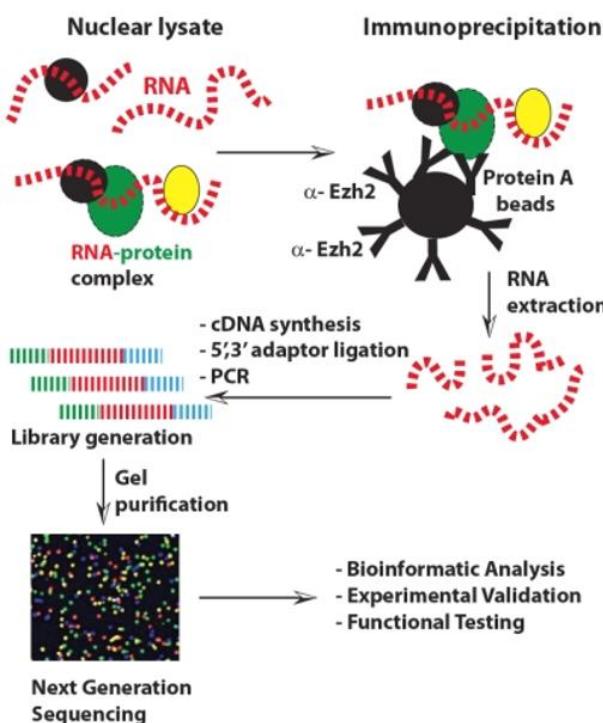
基础分析	高级分析
CircRNA-Seq 原始数据	与 circBase 数据库进行比对 (circBase 现在仅收录有人、小鼠和线虫的环状 RNA 数据)
原始数据质量控制 ✓ reads 质量评估统计 ✓ reads 碱基分布统计 ✓ readsGC 含量评估	功能基因关联：通过 junction 位置与基因进行关联，然后关联基因的功能注释
数据预处理，过滤，Mapping	miRNA 互作预测：与线性的 RNA 关联原理相同，但发生的作用可能不同，可参考 ceRNA 作用机制
基于 Backsplice 的 circRNA 预测	与常规转录组数据进行差异关联
circRNA 表达量统计	
长度分布统计	
差异 circRNA 分析：可对多组样本进行共同 circRNA 查找并得出差异 circRNA	

(四) RIP-Seq 服务

RNA Immunoprecipitation (RIP) 是研究细胞内 RNA 与蛋白结合情况的技术，是了解转录后调控网络动态过程的有力工具，能更有效地发现 miRNA 的调节靶点。RIP 技术针对目标蛋白的抗体把相应的 RNA-蛋白复合物沉淀下来。之后，经过分离纯化就可以对结合在复合物上的 RNA 进行测序、芯片等高通量分析。

上海伯豪将 RIP 技术结合新一代测序技术，推出了 RIP-Seq 服务，将有助于广大科研工作者更高通量地了解癌症以及其它疾病整体水平的 RNA 变化。

实验原理



- 用抗体或表位标记物捕获细胞核内或细胞质中内源性的 RNA 结合蛋白。
- 防止非特异性的 RNA 的结合。
- 免疫沉淀把 RNA 结合蛋白及其结合的 RNA 一起分离出来。
- 结合的 RNA 序列通过高通量测序 (RIP-Seq) 方法来鉴定。

数据分析服务内容

- 原始数据的 QC、数据预处理
- RIP-Seq 分析
 - 将 reads 比对到基因组：将预处理 reads 与 reference genome 进行 mapping，最后得到 mapping 的 sam 结果文件，给出 mapping 结果统计。
 - peak detect：应用 sam 文件进行 peak 富集区查找。
 - motif detect：查找结合位点区结构特性，寻找转录因子结合区域的 motif 结构。
 - 基因关联：应用结合位点区位置，确定其周围所涉及的基因。
 - 关联基因 GO 差异分析：以参考基因组为背景集对结合位点关联基因进行 GO 富集分析。

伯豪案例

被 TGF- β 活化的 lncRNA-ATB 促进肝癌的侵袭-转移级联反应

该课题组研究人员首次发现了一个介导转化生长因子- β (TGF- β) 促转移作用的 lncRNA，并将这个新 lncRNA 命名为“被 TGF- β 活化的长链非编码 RNA (lncRNA activated by TGF- β , lncRNA-ATB)”。研究发现 lncRNA-ATB 在肝癌组织中高表达，在肝癌转移灶中进一步上调表达，并与肝癌的侵袭特征成正相关，指示着肝癌病人更差的预后。因此 lncRNA-ATB 显著促进肝癌的侵袭-转移级联反应。这些研究结果表明，lncRNA-ATB 作为 TGF- β 信号通路促癌作用的介导因子，让肝癌病人易发生转移，可作为抗转移治疗的潜在靶点。

原文出处： A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. Cancer Cell 2014;25(5):666-81. (IF24.755)

表观遗传学水平

(一) 全基因组甲基化测序服务

全基因组甲基化测序结合了亚硫酸氢盐转化 (bisulfite conversion) 方法与新一代高通量测序技术，可在单碱基分辨率水平上高效地检测全基因组 DNA 甲基化状态。亚硫酸氢盐处理可以使 DNA 中未发生甲基化的胞嘧啶脱氨基转变成尿嘧啶，而甲基化的胞嘧啶保持不变，PCR 扩增所需片段，则尿嘧啶全部转化成胸腺嘧啶。对 PCR 产物进行高通量测序，与参考序列比对，即可判断 CpG/CHG/CHH 位点是否发生甲基化。

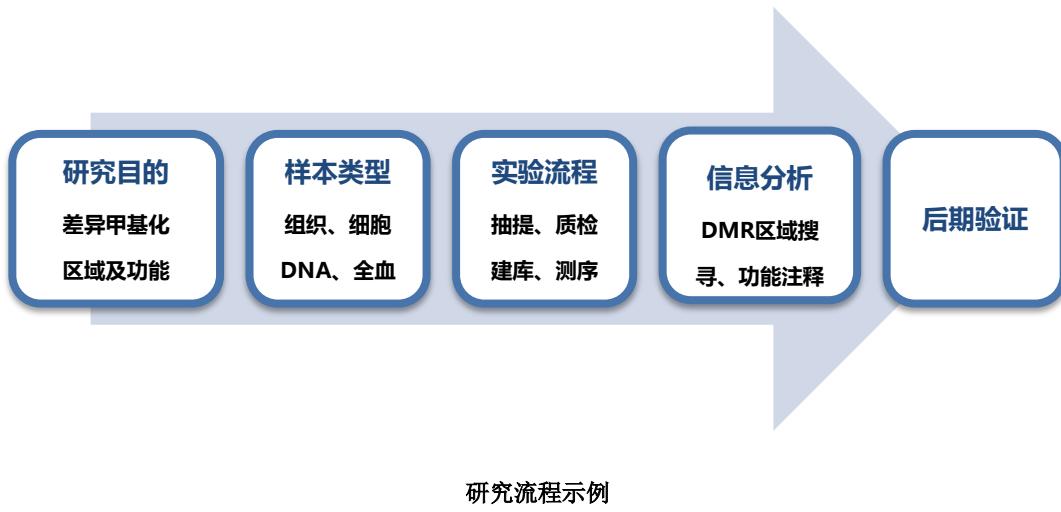
研究内容

- 全基因组甲基化图谱 (DNA methylation profiling) 及单样品高低甲基化位点在基因组不同区域的分布：全基因组甲基化测序可全面、精确地检测全基因组 DNA 甲基化状态，为更深入的表观遗传调控分析奠定基础。在全基因组甲基化图谱的基础上，对基因组不同染色体、不同基因区域、不同序列环境、不同基因功能元件如启动子、外显子、内含子、UTR 区的胞嘧啶甲基化状态进行分析，有助于从多个角度研究基因的功能及调控。



- 多样品间的差异性甲基化区域 (DMR) 分析：全基因组甲基化测序不但可使研究者在基因组规模洞察不同生物之间在DNA 甲基化水平和模式上的差异，也可用于研究同一物种多个样品间的差异性甲基化区域 (differentially methylated regions DMRs)。
- 基因组的甲基化状态具有一定的时空特异性。对某一物种，不同的环境、不同的发育阶段、不同类型

的样品（细胞、组织、个体）所反映的甲基化状态可能是不同的，存在着差异性甲基化区域。其中有些是与基因调控有关的功能区域。鉴定、分析这些区域及其所在的基因或CpG岛，对发育和肿瘤等生命过程研究有一定意义。



研究流程示例

技术路线

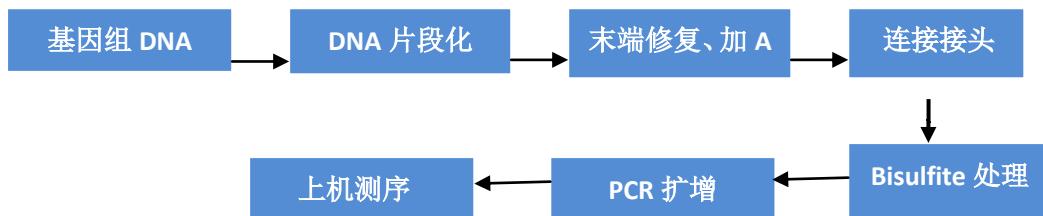


图 全基因组甲基化测序实验流程

样品要求（基因组DNA）

- 样品纯度：OD 260/280值应在1.8~2.0 之间； RNA 应去除干净；
- 样品浓度：最低浓度不低于50ng/ μ l；
- 样品总量：每个样品总量不少于15 μ g；
- 样品溶剂：应溶解在H₂O或TE (pH 8.0)中；
- 样品运输：DNA低温运输 (-20°C)； 在运输过程中用parafilm将管口密封好，以防出现污染。

数据分析技术路线

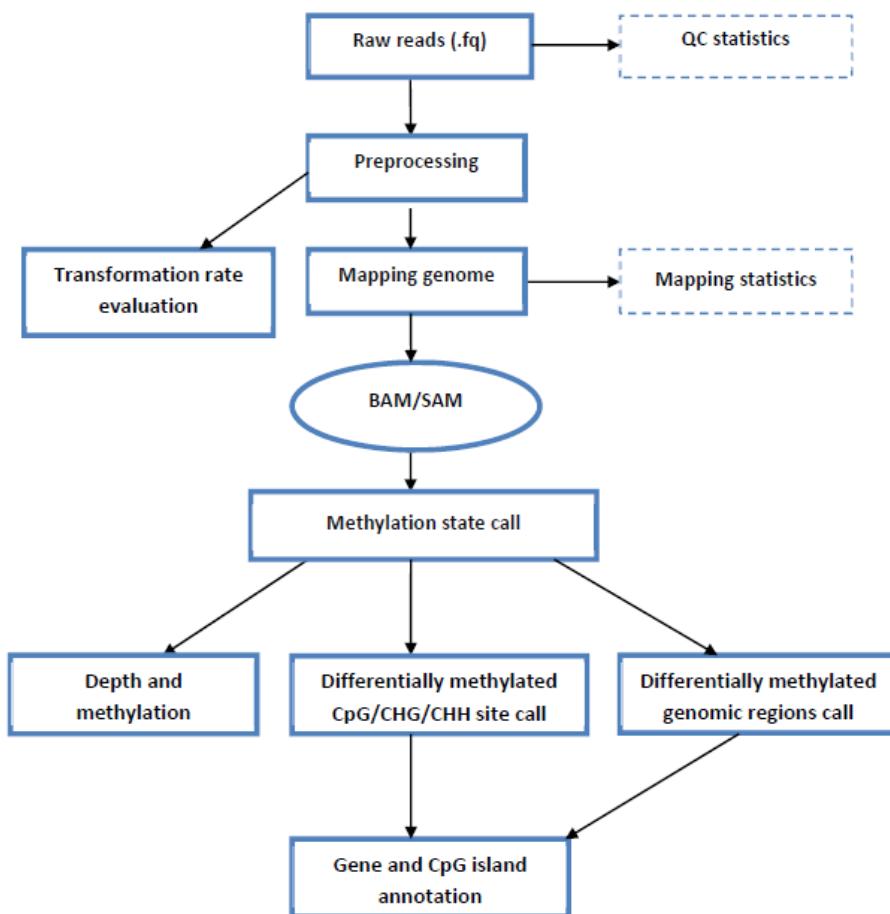


图 全基因组甲基化测序数据分析

1. 测序数据质量控制

包括：总 reads 数、总 base 数，Q20 比例、Q20 位点分布图、碱基分布图。

2. 测序数据预处理

包括：去除总体质量偏低的 reads，将质量大于 20 碱基所占比例小于 50% 的 reads 去除；去除 3' 端质量 Q 低于 20 的碱基；去除 reads 中所含有的接头序列；去除长度小于 20 的测序片段(reads)，去除含有低质量 N 碱基的 reads。

3. 转化效率评估

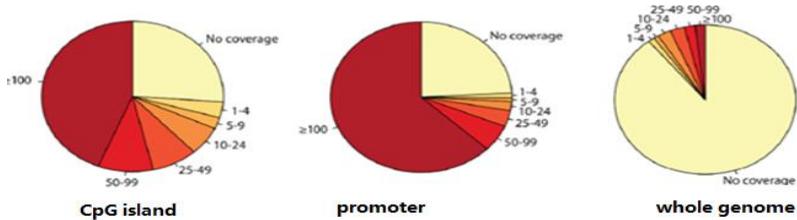
使用 lamda 噬菌体基因组作为内参，评价 BS 转化效率。

	测得 lamda DNA reads 数	总共检测 C 数量	C->T 转化数量	BS 转化效率
Sample 1	2988205 (1.1%)	143433482	143382275	99.96%
Sample 2				

4. 全基因组比对和 mapping 率统计

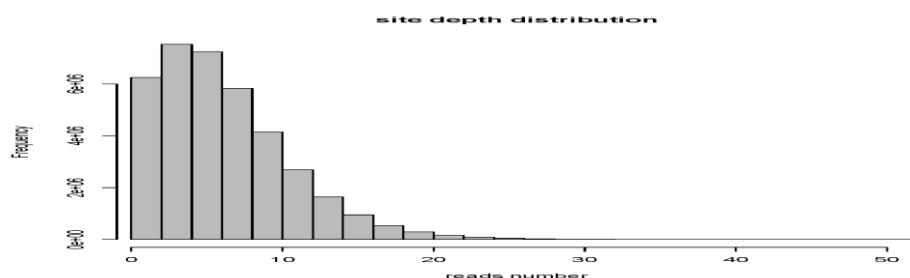
使用 bismark 软件将预处理后的 reads 比对到基因组，统计 mapping 率、样本整体甲基化水平，每个 CpG\CHG\CHH 位点的测序深度和甲基化比例。

5. CpG 岛区域、启动子区域及整个基因组的覆盖及不同深度关系

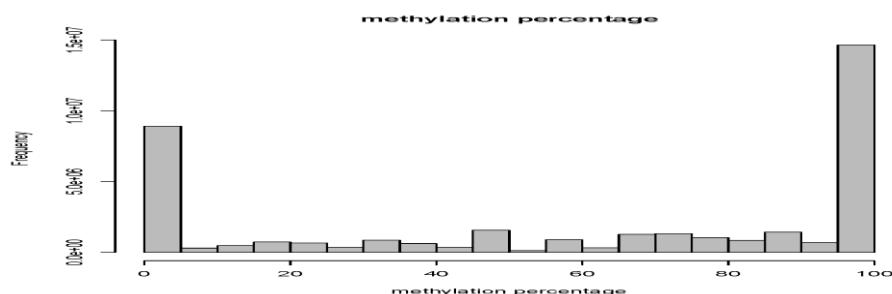


6. CpG/CHG/CHH 位点覆盖率、测序深度、甲基化比例和染色体分布

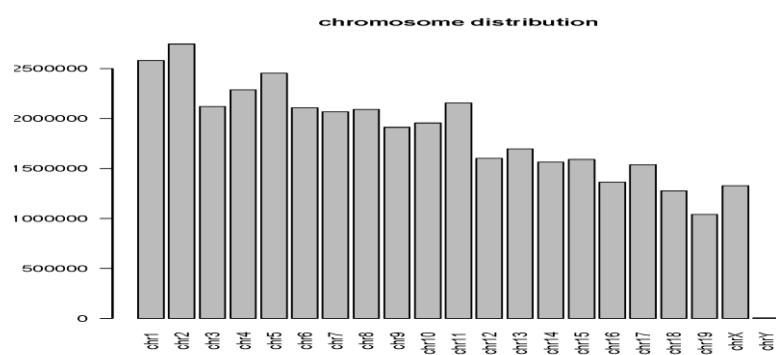
1) CpG/CHG/CHH 位点测序深度



2) CpG/CHG/CHH 位点甲基化比例

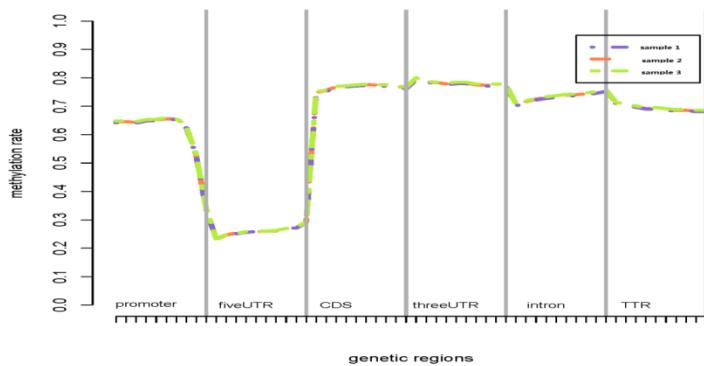


3) CpG/CHG/CHH 位点染色体分布

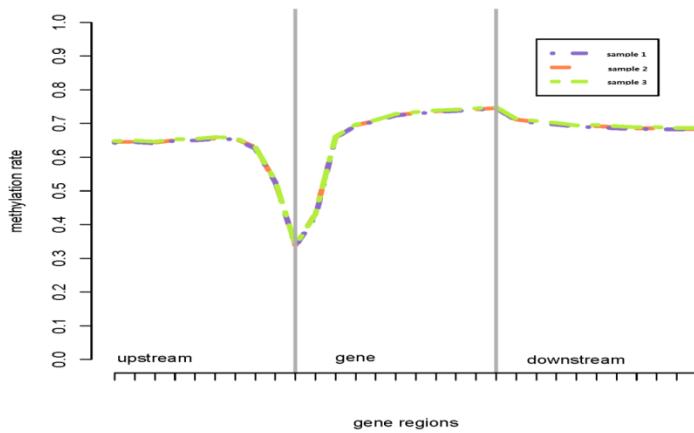


4) 高甲基化位点（甲基化比例>70%）和低甲基化位点（甲基化比例<30%）的基因功能区分布。

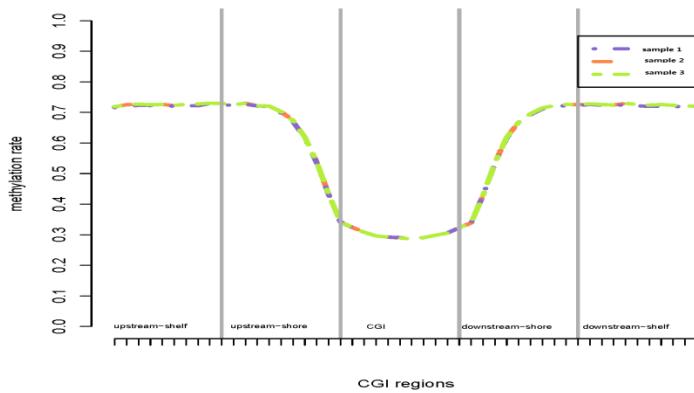
5) 甲基化程度 (%) 对于基因的不同功能区域的分布。



6) 甲基化程度 (%) 对于基因本体和上下游区域的分布。



7) 甲基化程度 (%) 对于 CpG island, shore 和 shelf 区域的分布。



7. 差异甲基化位点和甲基化区段筛选

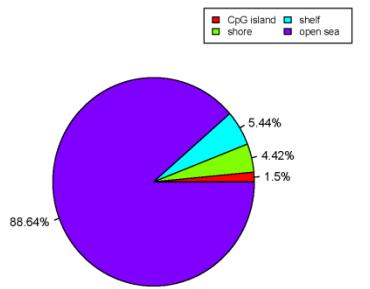
使用 fisher 检验筛选差异甲基化 CpG/CHG/CHH 位点和甲基化区段。

8. 基因 promoter、TSS、gene body 区域甲基化水平计算和筛选

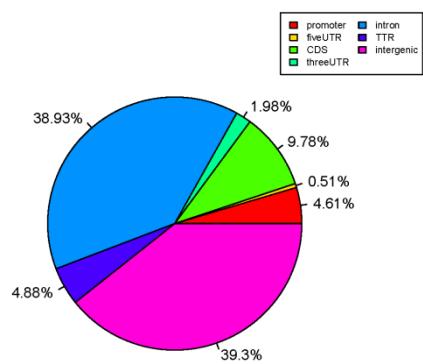
根据已有基因注释，计算基因的 promoter、TSS、gene body 等区域的甲基化情况，并据此筛选差异甲基化的基因。

9. 差异甲基化位点或区段的 CpG 岛注释和基因注释

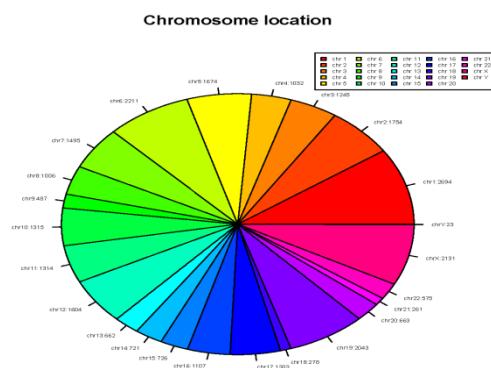
1) 差异甲基化位点或区段 CpG 岛分布



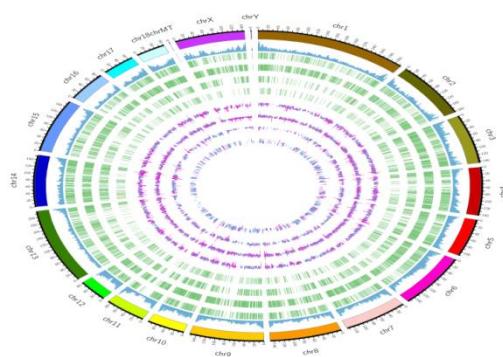
2) 差异甲基化位点或区段不同基因功能元件分布



3) 差异甲基化位点或区段染色体分布



10. 全基因组甲基化图谱





11. 甲基化和表达谱数据的联合分析

将甲基化数据和表达谱数据结合分析，将基因的差异甲基化和差异表达进行关联，可进行差异基因的甲基化情况聚类热图、不同基因区域甲基化和表达谱相关性计算和作图。

12. GO 和 KEGG 等功能分析和作图

13. 客户定制分析

可根据客户的研究需求，协商定制个性化分析内容。

应用案例

全基因组甲基化测序揭示蚕基因组稀疏的甲基化

昆虫的表观遗传学调控可以反映其多样的生物反应过程，本研究采用 Illumina 高通量亚硫酸盐测序，检测了模式生物蚕在单碱基水平的甲基化情况。保守估计，基因组中 0.11% 的胞嘧啶被甲基化，且可能全部存在于 CG 二核苷酸中。基因区域 CG 甲基化是非常丰富的，这可能与基因表达水平相关，同时也暗示了甲基化在基因转录中扮演了一个正面的角色。研究人员发现转录元素、启动子及核糖体 DNAs 低甲基化，相反，基因区域与基因座匹配的小 RNAs 被充分甲基化。该研究有助于我们对于昆虫表观遗传的理解。

原文：Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. Nature Biotechnology.2010,28:216-20.

（二）甲基化 DNA 免疫共沉淀测序服务

甲基化 DNA 免疫共沉淀测序 (Methylated DNA Immunoprecipitation Sequencing, MeDIP-Seq) 先用 5'-甲基胞嘧啶抗体富集胞嘧啶甲基化的基因组片断，然后对富集的片段进行高通量测序。该方法检测范围覆盖全基因组，所需测序数据量较少，但不能在单碱基分辨率水平上检测 DNA 胞嘧啶甲基化状态。

研究内容

- ⊕ 全基因组不同水平甲基化区域的分布：分析 MeDIP-Seq 读序在全基因组、不同功能元件区域 (CpG 岛、重复序列、启动子、非翻译区、外显子、内含子和基因间等) 的分布情况。识别序列富集峰 (peak)。结果可显示不同样本间 DNA 甲基化修饰模式的差异。
- ⊕ 基于富集峰的基因功能及其调控机制的预测：对于单个样本，分析与 peak 相关的基因；对于多样本，进行基于 peak 的全基因组甲基化差异分析。通过分析差异甲基化区段 (DMR)，对差异基因进行 GO 功能富集分析及 pathway 功能分析；也可结合表达谱数据，预测受甲基化调控的基因。



研究流程示例

技术路线

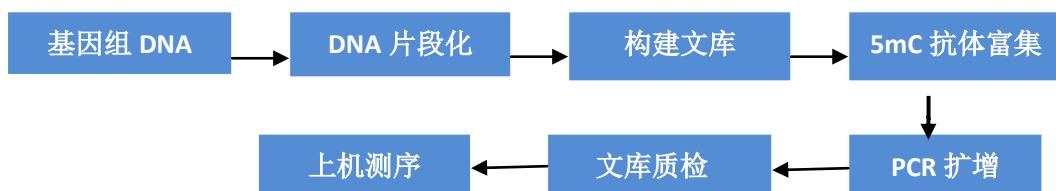
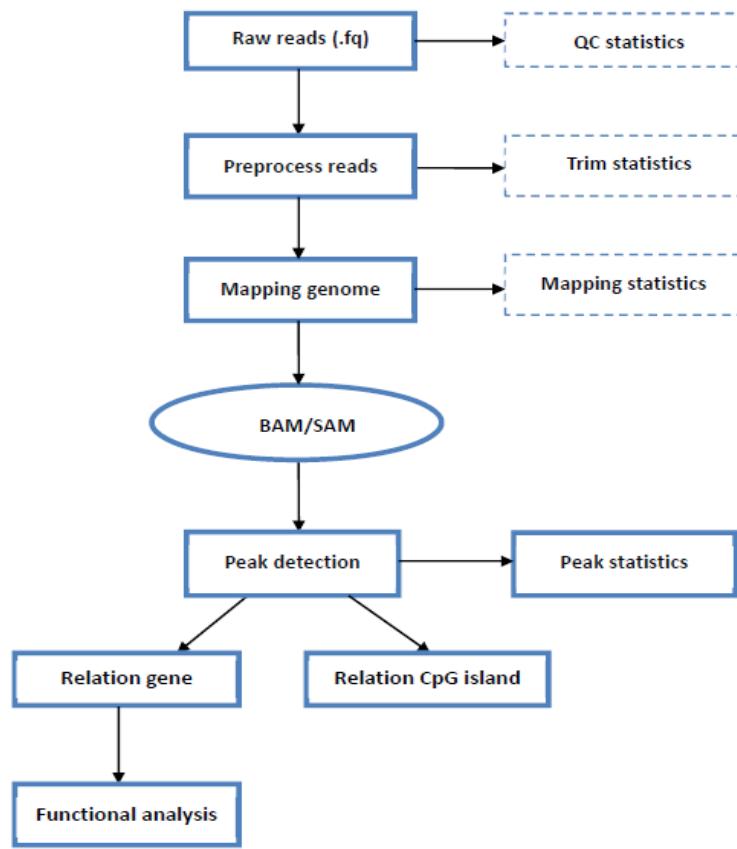


图 MeDIP-Seq 实验流程

样品要求 (基因组 DNA)

- ◆ 样品纯度：OD 260/280值应在1.8~2.0 之间；RNA 应去除干净。
- ◆ 样品浓度：最低浓度不低于50ng/ul。
- ◆ 样品总量：每个样品总量不少于10ug。
- ◆ 样品溶剂：溶解在H₂O或TE (pH 8.0)中。
- ◆ 样品运输：DNA低温运输 (-20℃)；且在运输过程中请用parafilm将管口密封好，以防出现污染。

数据分析技术路线



数据分析内容

1. 测序数据质量控制

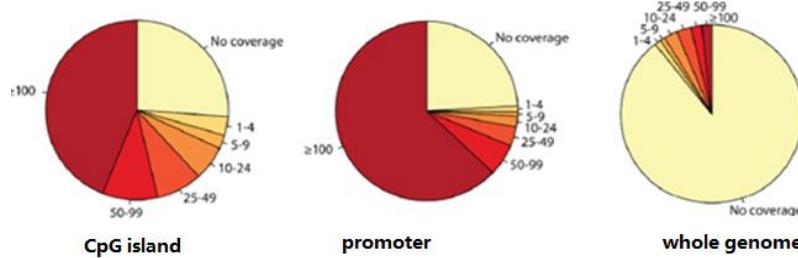
包括：总 reads 数、总 base 数，Q20 比例、Q20 位点分布图。

2. 测序数据预处理

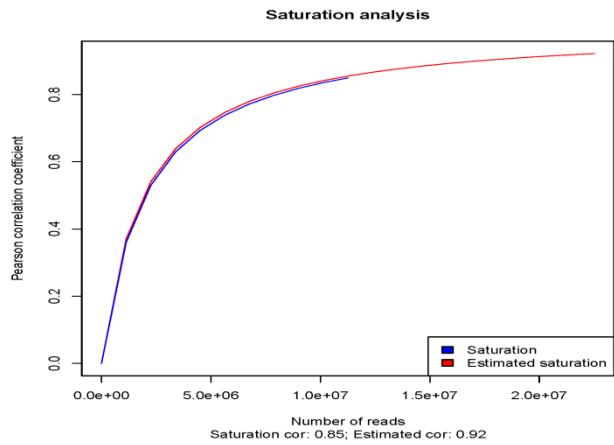
包括：去除总体质量偏低的 reads，将质量大于 20 碱基所占比例小于 50% 的 reads 去除；去除 3' 端质量 Q 低于 20 的碱基；去除 reads 中所含有的接头序列；去除长度小于 20 的测序片段(reads)。

3. 将预处理后的 reads 比对到参考基因组并统计富集效率

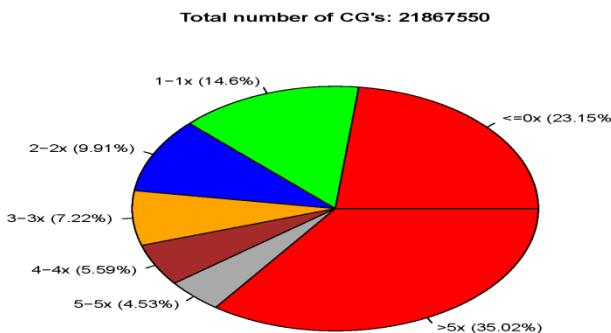
包括：Reads 唯一比对率，在 CpG 岛区域、启动子区域及整个基因组的覆盖及不同深度关系。



4. 测序深度是否足够质量控制



5. 甲基化富集效果质量控制

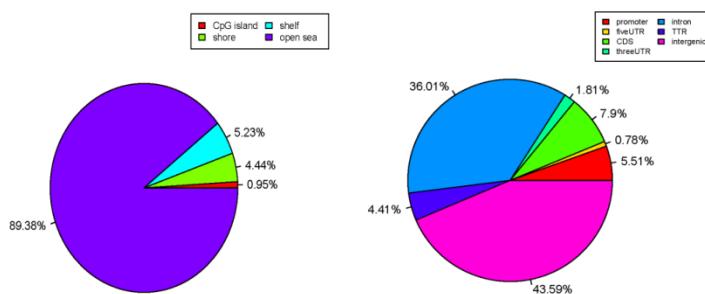


943120 of 26044572 reads (3.62%) do not cover a pattern

6. Peak 查找和统计

应用 MACS 软件对基因组 mapping 得到的 bam 文件进行 peak 富集区查找，并统计长度和染色体分布。

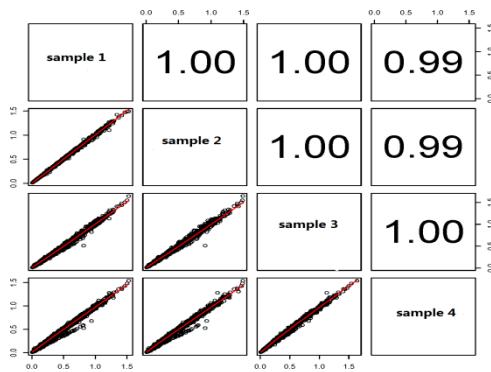
7. Peak 在 CpG 岛、不同基因功能元件中的分布



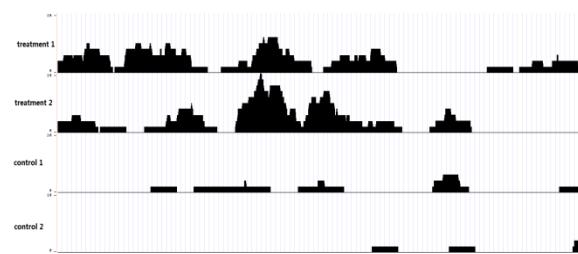
8. 基于 peak 的样本间差异分析

将各样本有重叠的 peak 合并，统计每个样本在合并后的 peak 中是否富集，统计组间差异的个数并基于 RPKM (Reads Per Kilobase per Million mapped reads)计算 peak 在组间差异是否显著。

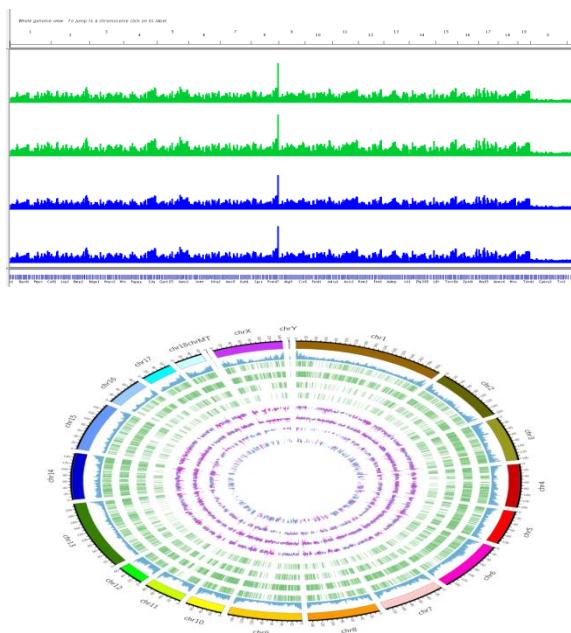
9. 样本间富集程度相关性分析



10. 样本间差异 peak 区域测序深度作图



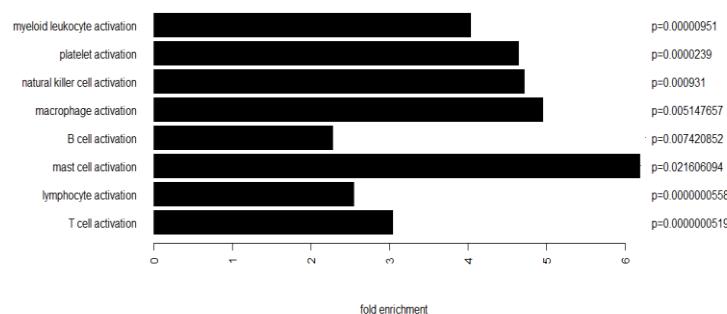
11. 全基因组甲基化图谱



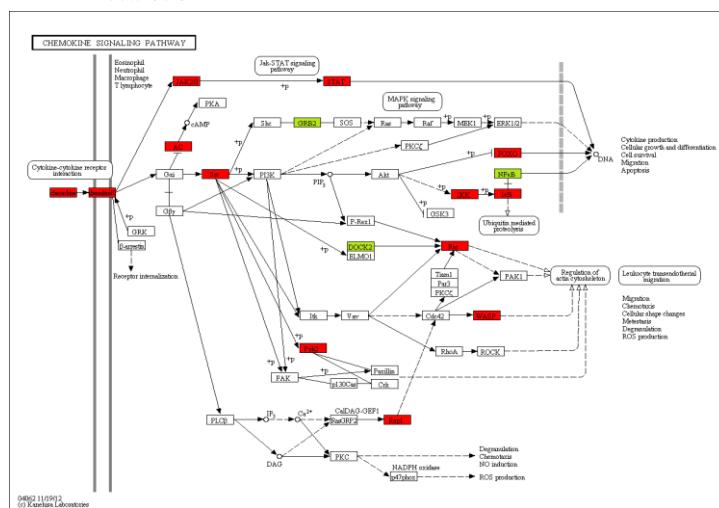


12. 对差异基因进行 GO 功能富集分析及 pathway 功能分析

GO 功能分析：



KEGG 功能分析：



13. MeDIP-seq 数据和表达谱数据结合分析

将 MeDIP-seq 数据和表达谱数据结合分析，对富集 peak 区域和差异表达基因进行关联，可进行差异基因的甲基化情况聚类热图、不同基因区域甲基化和表达谱相关性计算和作图。

14. 客户定制分析

可根据客户的研究需求，协商定制个性化分析内容。

伯豪案例

DNMT3A 突变驱动了慢性粒单核细胞白血病

在 20% 的急性髓性白血病案例中会发生 DNMT3A 的突变，突变热点为 Arg882 (R882)。本研究通过运用逆转病毒转导及骨髓移植 (BMT) 的方法来研究 DNMT3A-Arg882His (R882H) 突变体的转化能力。研究人员发现突变的基因能够引起造血干细胞/祖细胞的异常增殖。BMT 1 年后，所有小鼠都出现了血小板增多，且患上 CMML。RNA 芯片分析结果显示一些造血作用相关基因异常表达，DNA 甲基化实验证实基因区域甲基化类型发生相应的变化。此外，DNMT3A-R882H 增加了 CDK1 蛋白水平，加强了细胞周期活力，这些都导致白血病的生成。上海交通大学医学院附属瑞金医院的陈竺院士、陈赛娟院士以及王月英博士是这篇论文的共同通讯作者。

原文出处：DNMT3A Arg882 mutation drives chronic myelomonocytic leukemia through disturbing gene expression/DNA methylation in hematopoietic cells. Proc Natl Acad Sci U S A. 2014 Feb 4. 使用服务类型：MeDIP-SEQ

（三）染色质免疫共沉淀测序服务

染色质免疫共沉淀技术 (Chromatin Immunoprecipitation, ChIP) 是研究体内蛋白质与 DNA 相互作用的有力工具，通常用于转录因子结合位点或组蛋白特异性修饰位点的研究。将 ChIP 与新一代测序技术相结合的 ChIP-Seq 技术，能够高效地在全基因组范围内检测与组蛋白、转录因子等互作的 DNA 区段。

ChIP-Seq 首先通过染色质免疫共沉淀技术 (ChIP) 特异性地富集目的蛋白结合的 DNA 片段，并对其进行纯化与文库构建，然后对富集得到的 DNA 片段进行高通量测序。研究人员通过将获得的数百万条序列标签精确定位到基因组上，从而获得全基因组范围内与组蛋白、转录因子等互作的 DNA 区段信息。

技术优势

- ◆ 检测覆盖范围广：全基因组范围内扫描目标区域；
- ◆ 检测分辨率高：更利于精确定位目标区域；
- ◆ 样本需要量低：需要的免疫沉淀后的DNA量可低至5-10ng；
- ◆ 可靠性好：避免了非特异性杂交，背景低；
- ◆ 性价比高：花费较少即可检测全基因组，获取准确丰富的信息。

研究内容

- ◆ 通过富集peak检测，研究转录因子、RNA 聚合酶或其它蛋白在基因组上结合位点及其相关基因。提取peak区域的序列，寻找结合位点的motif，并对peak 进行基因、CpG 岛、GO和KEGG 等注释和分析。对于多个样品，可进行差异分析；
- ◆ 用于组蛋白修饰的表观遗传学研究。通过不同的特异性抗体检测DNA链的某一位置会出现何种组蛋白修饰，以及组蛋白修饰与基因表达的关系。



研究流程示例

技术路线

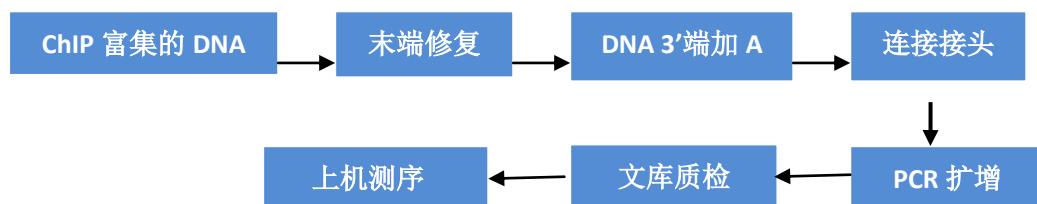
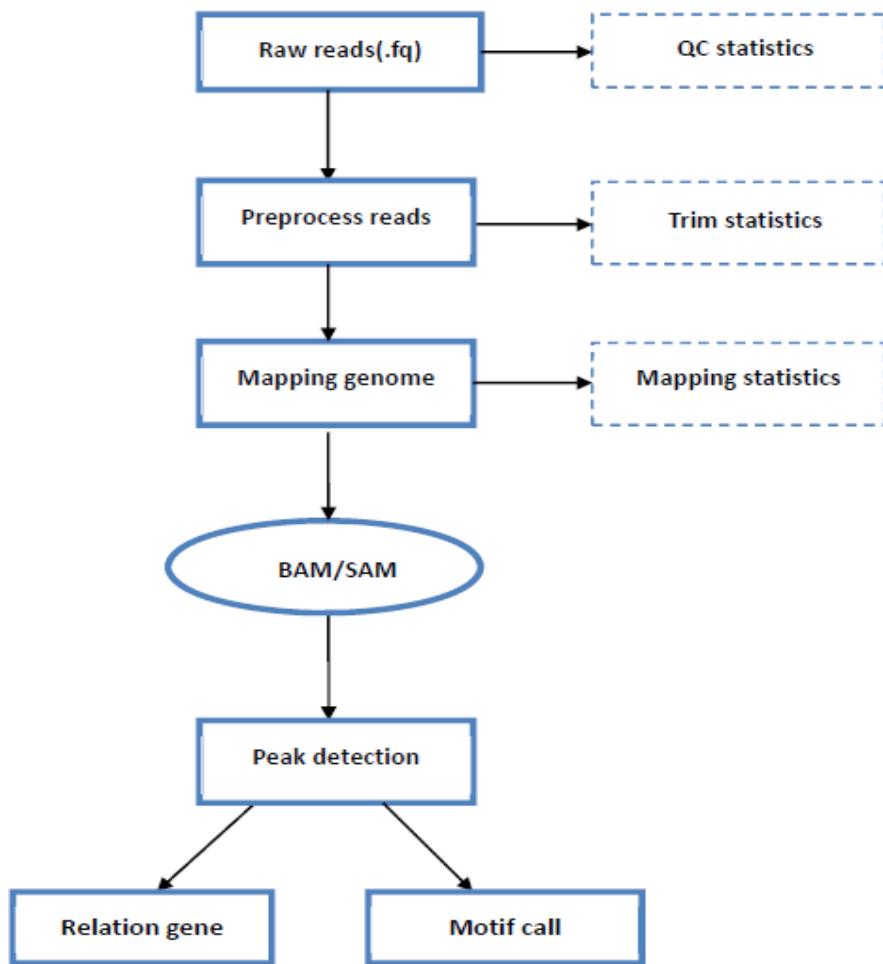


图 ChIP-Seq 实验流程

样品要求（ChIP 富集的 DNA）

- ◆ 样品纯度：OD 260/280 值应在 1.8~2.0 之间。
- ◆ 样品浓度：最低浓度不低于 20ng/μl。
- ◆ 样品总量：每个样品总量不少于 30ng。
- ◆ DNA 片段大小：长度分布小于 500bp，峰在 300bp 左右。
- ◆ 样品溶剂：溶解在 H₂O 或 low TE (pH 8.0) 中。
- ◆ 样品运输：低温运输 (-20°C)；建议用 LoBind 管，parafilm 将管口密封好，以免污染。

数据分析技术路线



数据分析内容

对 ChIP-seq 测序数据进行分析，包括数据预处理、基因组比对、比对结果相关统计、对于基因组的覆盖和深度、富集 peak 检测、富集 peak 在启动子和不同基因功能区的分布统计、富集 peak 区段 Fasta 序列提取、motif 查找、相关注释和富集分析（基因、CpG 岛、GO 和 KEGG 等）。

1. 测序数据质量控制

包括：总 reads 数、总 base 数，Q20 比例、Q20 位点分布图。

2. 测序数据预处理

包括：去除总体质量偏低的 reads；去除 3' 端质量 Q 值低于 20 的碱基；去除 reads 中所含有的接头序列；去除长度小于 20 的测序片段(reads)。

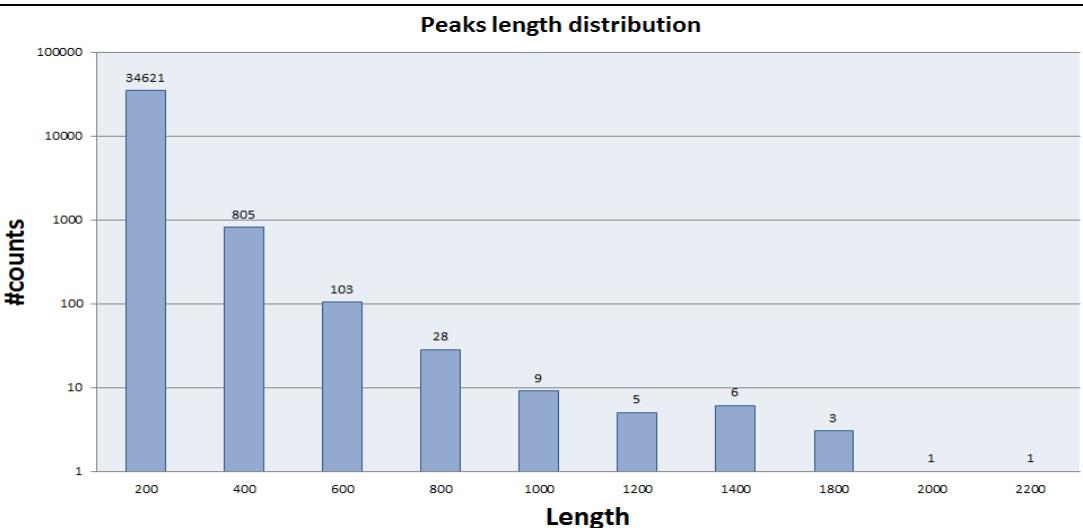
3. 全基因组 mapping 率

4. Peak 查找

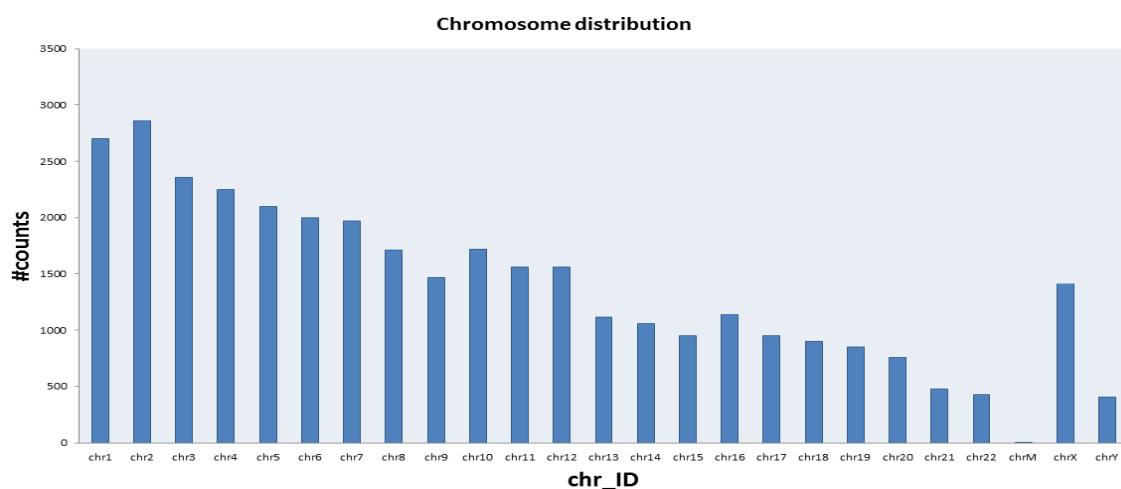
应用 MACS 软件对基因组 mapping 得到的 bam 文件进行 peak 富集区查找。

5. Peak 长度和染色体分布

Peak 长度分布：



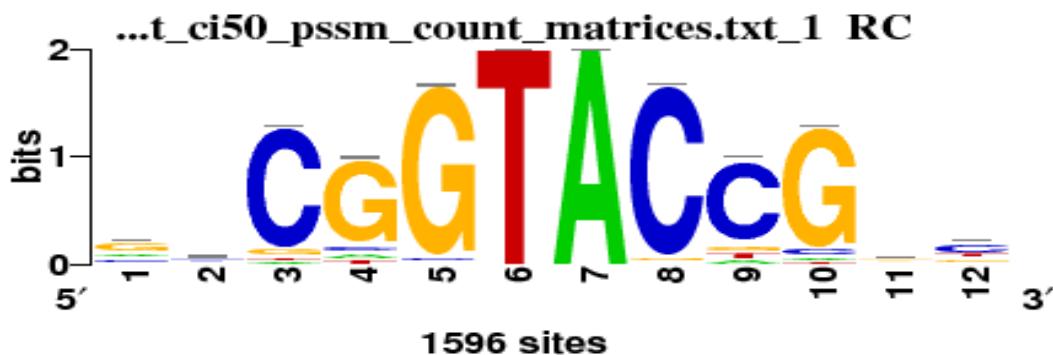
Peak 染色体分布：



6. 对 Peak 进行基因和 GO 功能注释

7. 提取 peaks 区域的 fasta 序列并寻找结合 motif

motif 功能区示意图：



8. 客户定制分析

可根据客户的研究需求，协商定制个性化分析内容。

应用案例

在哺乳动物中，ZBED6 是一个重要的转录因子，影响着发育，细胞增殖和生长。最近科学家们通过 Chip-Seq 技术，发现家猪 IGF2 基因（Insulin-like growth factor 2）第三内含子单碱基 G-A 的替换能导致抑制因子 ZBED6 在该基因上结合位点丧失，从而使 IGF2 在骨骼肌中表达上调（表达量是正常情况的 3 倍），这一突变能刺激肌肉生长，心脏增大和脂肪沉积。

原文出处：ZBED6, a Novel Transcription Factor Derived from a Domesticated DNA Transposon Regulates IGF2 Expression and Muscle Growth. Plos Biology, 2009, 7(2):e1000256.

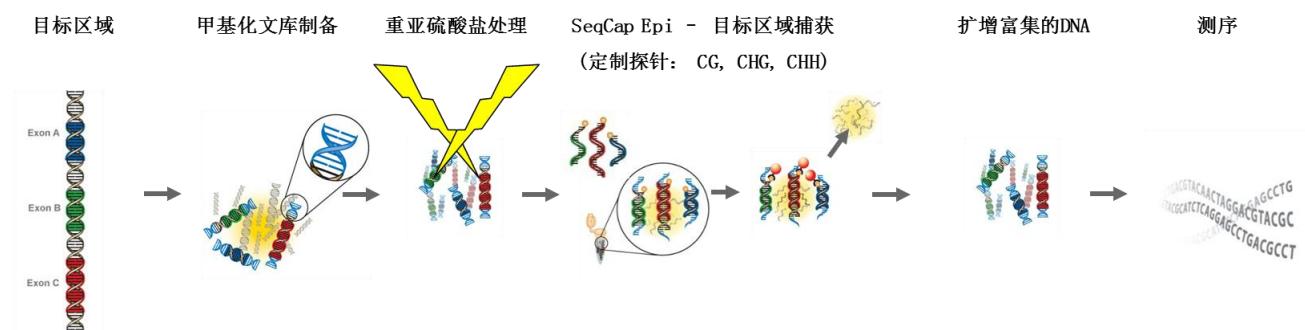
（四）NimbleGen SeqCap Epi 甲基化序列富集方案 ——通过“超级”探针设计实现单碱基分辨率的定向甲基化测序

上海伯豪全新推出 NimbleGen SeqCap Epi 基因组 DNA 甲基化序列富集服务，伯豪公司利用该产品并结合二代测序，可进行单碱基分辨率的样品甲基化分析。上海伯豪该服务系列包括：固定内容的、人类全基因组水平的可甲基化位点的 DNA 片段富集方案；以及，适合于各种不同应用的由客户自选区域的目标序列富集服务。SeqCap Epi 可以为表观遗传研究人员提供全新的研究工具，相比现有其他技术，可以在更大 DNA 序列广度，更高测序深度和更大样本量水平下进行 DNA 甲基化研究。

服务特色

- ✓ 研究具有生物学意义的甲基化改变
- ✓ 样品间具有良好的可重复性
- ✓ 探针具有捕获各种不同甲基化状态，不同甲基化程度的能力

实验流程（先重亚硫酸处理后捕获）



服务介绍

- ✓ 目录产品（包装数：4, 48, 384 Rxn）

SeqCap Epi CpGiant Enrichment Kit

- 全基因组水平的甲基化位点富集，包括了 Illumina's Infinium Human Methylation 450 BeadChip 的所有位点及其附近的区域
- 覆盖人类基因组上>5.5 million 个 CpG 位点

- ✓ SBC designed Content

SeqCap Epi MouseCpGi Enrichment Kit

-
- 小鼠全基因组水平的甲基化位点富集，设计覆盖 27465 个 CpG 岛区域及 35507 个基因的启动子区域
 - 覆盖小鼠基因组上>1.8 million CpG 位点

SeqCap Epi RatCpGi Enrichment Kit

- 大鼠全基因组水平的甲基化位点富集，设计覆盖 25133 个 CpG 岛区域及 26519 个基因的启动子区域
- 覆盖大鼠基因组上> 1.6 million 个 CpG 位点

✓ **Custom Content** (包装数: 12, 48, 384 Rxn)

SeqCap Epi Choice Enrichment Kits

- 人类基因组特定区域定制设计
- 小区域 <30Mb, 中等区域 30-60Mb, 以及大区域 60-90Mb 均可设计

SeqCap Epi Developer Enrichment Kits

- 非人类基因组特定区域定制以及超大区域的人类基因组
- 小区域 <30Mb, 中等区域 30-60Mb, 大区域 60-90Mb, 以及超大区域 90-210Mb 均可设计

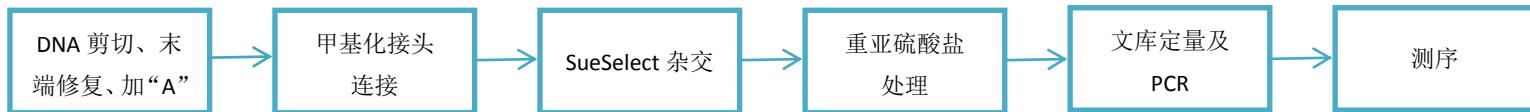
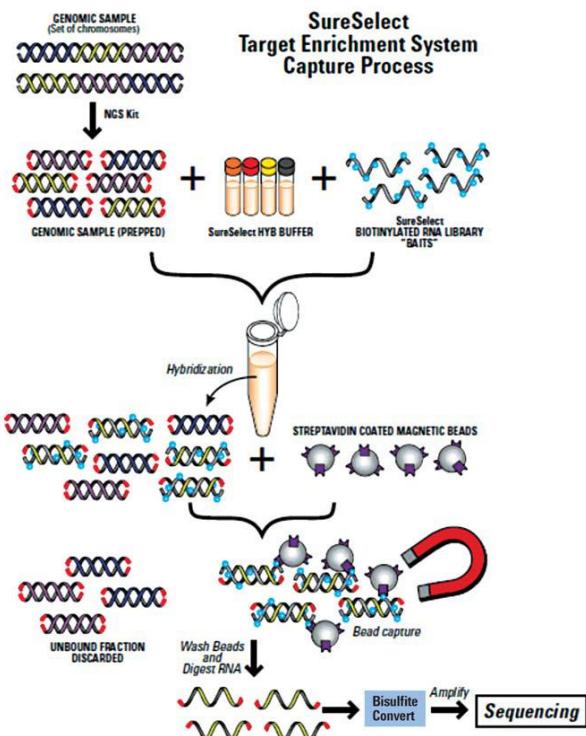
数据分析内容

测序数据质量控制	捕获区域每 100bp 的测序深度和甲基化比例
测序数据预处理	捕获区域单个位点和 100bp 区段的甲基化水平比较和差异筛选
将预处理后的 reads 比对到参考基因组	差异甲基化位点和区段的基因、CpG 岛和功能注释
统计落在捕获区域的 reads 比例	对差异甲基化基因进行 GO 功能富集分析及 pathway 功能分析
捕获区域单个 CpG/CHG/CHH 位点的测序深度和甲基化比例	

(五) Agilent sureselect 甲基化序列富集方案

上海伯豪与Agilent公司携手，为国内外客户提供Agilent SureSelect^{XT} 甲基化测序服务，该测序服务系统基于液相杂交，可以分析人类基因组中低甲基化与过度甲基化的胞嘧啶位点。该系统将靶向序列捕获平台SureSelect与亚硫酸盐测序结合在一起，挑选了与表观遗传学研究最相关的基因组序列，包含了与多种疾病（例如，癌症、基因组印记疾病、行为和精神障碍等等）相关的区域，实现了前所未有的序列覆盖范围。作为“Agilent亚太地区最佳技术服务供应商”，上海伯豪将为您提供全面、准确的Agilent SureSelect^{XT} 甲基化测序服务。

实验流程



服务介绍

上海伯豪可以为您提供人及小鼠 SureSelect^{XT} Methyl-Seq 服务，服务具体介绍如下：

✓ SureSelect^{XT} Human Methyl-Seq

该服务基于 Illumina 测序平台，捕获靶区域为 84Mb。

设计详情	参数	Methyl-Seq
	CpG 岛	19.6Mb/1.7M CpG
	癌症和组织特异性 DMRs	9.8 Mb/293K CpG
	Gencode 启动子	37Mb/1.3M CpG
	CpG 岛 shores/shelves±4kb	48Mb/2.1M CpG
	DNase I 高敏感位点	
	增强子	
	Ensembl 调控特征	

✓ SureSelect^{XT} Mouse Methyl-Seq

Site 分类	靶标数目	总碱基覆盖
CpG 岛	16,027	10,512,276bp
组织特异性 DMRs	33,456	10,452,692bp
Ensembl 调控特征	171,796	91,799,015bp
◊ CpG shores 及 shelves ($\pm 4\text{kb}$)		
◊ DNase I 高敏感位点		
◊ 组蛋白修饰		
◊ TFBS		
◊ 聚合酶		
Open Regulatory Annotation (ORegAnno)	14,951	9,983,957bp
◊ 启动子		
◊ 增强子		
◊ TFBS		
◊ 调控多态性		

数据分析内容

测序数据质量控制	捕获区域 100bp 区段的测序深度和甲基化比例
测序数据预处理	捕获区域单个位点和 100bp 区段的甲基化水平比较和差异筛选
将预处理后的 reads 比对到参考基因组	差异甲基化位点和区段的基因、CpG 岛和功能注释
统计落在捕获区域的 reads 比例	对差异甲基化基因进行 GO 功能富集分析及 pathway 功能分析
捕获区域单个 CpG/CHG/CHH 位点的测序深度和甲基化比例	

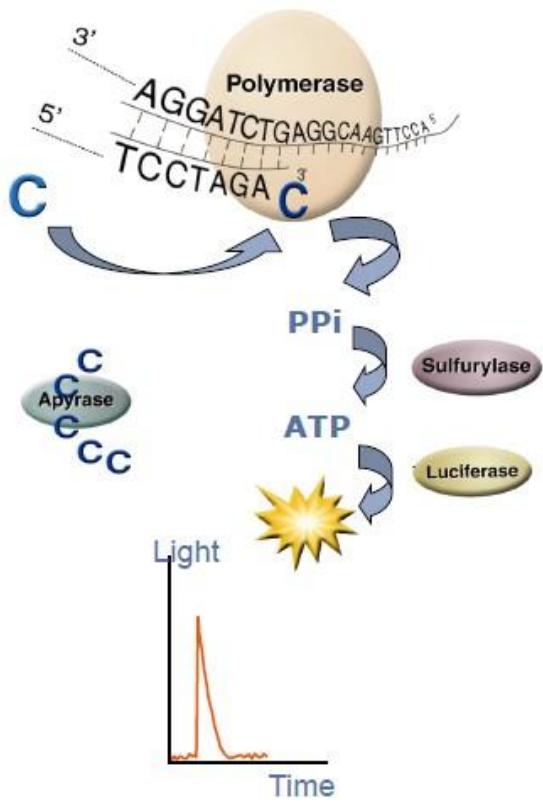
(六) 焦磷酸测序法检测甲基化

焦磷酸测序技术 (Pyrosequencing) 作为一种新的序列分析技术，能够快速地检测甲基化的频率，对样品中的甲基化位点进行定性及定量检测，为甲基化研究提供了新的途径。上海伯豪为您提供基于QIAGEN公司焦磷酸测序平台的服务。

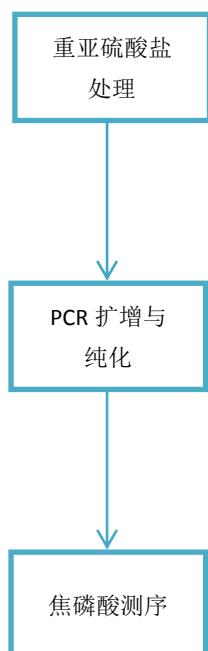
技术原理

焦磷酸测序技术是由4种酶 (DNA 聚合酶、ATP硫酸化酶、荧光素酶和三磷酸腺苷双磷酸酶) 催化的同一反应体系中的酶级联化学发光反应。具体过程如下：

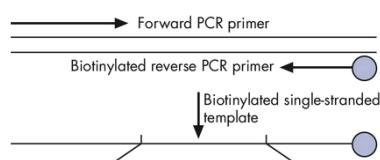
- ✓ 仪器向待检测样品中每次只加入一个核苷酸碱基(dNTP)
- ✓ 互补的碱基结合在模板上时，会释放出焦磷酸基团Pyrophosphate (PPi)
- ✓ PPi在ATP硫酸化酶(ATP Sulfurylase)作用下，合成ATP，ATP作为能量分子，驱动荧光素在荧光素酶(Luciferase)的催化下转化为氧化荧光素，并发出与ATP能量成正比的荧光信号
- ✓ 信号被CCD数码相机捕捉，形成峰形Pyrogram™
- ✓ 未与模板结合的碱基将被三磷酸腺苷双磷酸酶(Apyrase)降解
- ✓ 加入下一种dNTP



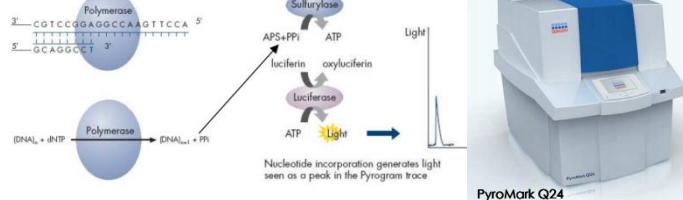
实验流程



- 待分析序列: **AGTTACGAC**
- AGTTAC^mGAC** and **AGTTACGAC**
- 焦磷酸盐处理后序列: **AGTTACGAT** and **AGTTATGAT**



- PCR扩增出包含待测序列的区域，其中一条引物带生物素标记；
- 扩增后，利用链霉素珠子纯化

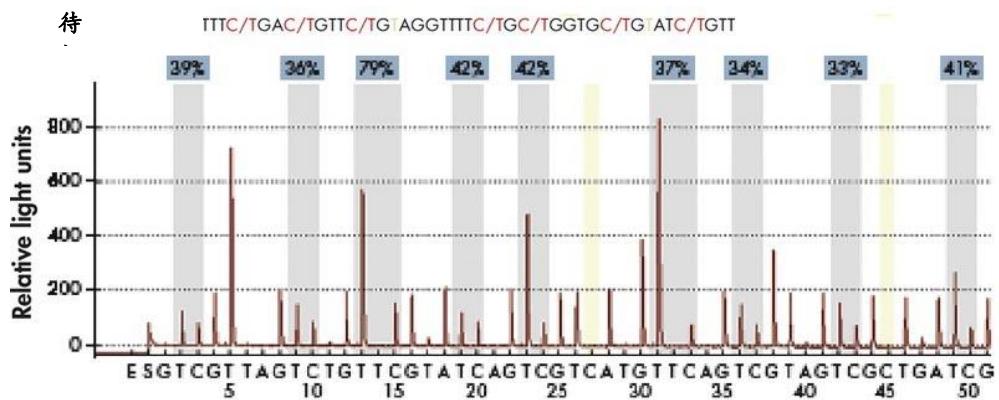


- 类似高通量测序，边合成边测序

焦磷酸测序在DNA甲基化中的应用

- ✓ 基因组整体水平的甲基化检测
- ✓ 特异位点的甲基化检测
- ✓ 寻找新的甲基化位点

结果展示



应用案例

焦磷酸测序检测宫颈癌中UTF1启动子区域甲基化水平

研究者利用 Illumina HumanMethylation27 芯片对9例样本（4例正常，5例宫颈癌）进行全基因组启动子区甲基化扫描，结果显示宫颈癌患者UTF1启动子区域呈现出高的甲基化状态。为了验证芯片筛选的结果，采用焦磷酸测序方法首先对这些样本的UTF1启动子区域进行深入分析，结果表明正常宫颈中UTF1启动子区域呈现低甲基化状态，而相同的位点在SCC中呈现高甲基化状态。为了进一步验证，研究者扩大样本量，对47例不同类型的石蜡包埋组织样本及10例白细胞进行焦磷酸测序检测，与之前的结论保持一致。结合 RNA 水平和蛋白水平的实验，研究者最后提出UTF1启动子甲基化状态可以作为宫颈癌诊断的一个重要的分子标志物。

原文出处：Aberrant Promoter Methylation and Expression of UTF1 during Cervical Carcinogenesis. Plos One. 2012, 7(8):e42704.

(七) PacBio RSII 测序法检测甲基化

国家人类基因组南方研究中心拥有亚洲首台 PacBio RS II 系统，是美国 Pacific Bioscience 公司认证示范实验室。是中国首家成功构建 20kb 文库实验室。2014 年度起，携手上海伯豪生物技术有限公司，对国内外用户提供 PacBio RSII 测序服务。



PacBio可以直接读出多种碱基修饰，包括A的甲基化、C的甲基化、C的羟甲基化等，所以PacBio在碱基修饰研究中，有着独特的优势。该平台进行过芽孢杆菌等的全基因组甲基化测序。

PacBio测序进行DNA甲基化研究

SMRT技术采用的是对DNA聚合酶的工作状态进行实时监测的方法，聚合酶合成每一个碱基，都有一个时间段，而当模板碱基带有修饰时，聚合酶会慢下来，使带有修饰的碱基两个相邻的脉冲峰之间的距离和参考序列的距离之间的比值如果大于1，由此就可以推断这个位置有修饰。甲基化研究中关于5-mC和5-hmC（5 mC的羟基化形式）是甲基化研究中的热点。但现有的测序方法无法区分5-mC和5-hmC。美国芝加哥大学利用SMRT测序技术和5-hmC的选择性化学标记方法来高通量检测5-hmC。通过聚合酶动力学提供的信息，可直接检测到DNA甲基化，包括N6 甲基腺嘌呤、5-mC和5-hmC，为表观遗传学研究打开了一条通路。

PacBio RS系统检测修饰碱基的优势

1. 分辨率可达到1个碱基，可适用于多种碱基修饰的检测；
2. 读数较长；
3. 无需双链DNA扩增，起始模板量要求低（500bp文库的起始模板量最低只要250ng），可检测到DNA链上的特定修饰，比如半甲基化；
4. 通过测序动力学效应可测定未知的修饰。

Pacbio RS II 测序平台的特点

测序快速：48小时内完成文库构建和测序。

序列超长：平均读长可达7.9 kb，最大读长可达32 kb（20kb文库,P5+C3）。

精度极高：>99.999%。

代表性好：无明显GC或AT偏好。

方式灵活：1-16个SMRT CELL测序可灵活安排。

多重样本：已建立BarCode建库体系，多个样本同时测序。

单分子测序：文库构建过程中无PCR扩增，可检测到罕见突变。

实时测序：测序过程中可直接获得碱基修饰信息。

Gap 数少：基因组<3M的样品，gap数0-10个。

应用广泛：真核和原核基因组、PCR产物、全长cDNA测序，全基因组甲基化等测序。

测序服务应用文章精选（2010-2014）

-
- [1] Song Y, Ma K, Ci D, Zhang Z, Zhang D. Sexual dimorphism floral microRNA profiling and target gene expression in andromonoecious poplar (*Populus tomentosa*). *PLoS One.* 2013, 8(5):e62681. (IF4.092) 使用服务种类: **miRNA 测序服务**
- [2] Zou W, Chen D, Xiong M, Zhu J, Lin X, Wang L, Zhang J, Chen L, Zhang H, Chen H, Chen M, Jin M. Insights into the increasing virulence of the swine-origin pandemic H1N1/2009 influenza virus. *Sci Rep.* 2013, 3:1601. 使用服务种类: **RNA 测序服务**
- [3] Wang Q, Du HB, Li M, LI Y, Liu SN, Gao P, Zhang XL, Cheng J. MAPK Signal Transduction Pathway Regulation: A Novel Mechanism of Rat HSC-T6 Cell Apoptosis Induced by FUZHENGHUAYU Tablet. *Evid Based Complement Alternat Med.* 2013. (IF4.774) 使用服务种类: **miRNA 测序服务**
- [4] Wang T, Cui Y, Jin J, Guo J, Wang G, Yin X, He QY, Zhang G. Translating mRNAs strongly correlate to proteins in a multivariate manner and their translation ratios are phenotype specific. *Nucleic Acids Res.* 2013 Mar 21. (IF8.026) 使用服务种类: **RNA 测序服务**
- [5] Yang J, Duan S, Zhong R, Yin J, Pu J, Ke J, Lu X, Zou L, Zhang H, **Zhu Z**, Wang D, **Xiao H**, Guo A, Xia J, Miao X, Tang S, Wang G. Exome Sequencing Identified NRG3 as a Novel Susceptible Gene of Hirschsprung's Disease in a Chinese Population. *Mol Neurobiol.* 2013 Jan 12. (IF5.735) 使用服务种类: **外显子捕获测序数据分析服务**
- [6] Li J, Li X, Chen Y, Yang Z, Guo S. Solexa sequencing based transcriptome analysis of *Helicoverpa armigera* larvae. *Mol Biol Rep.* 2012, 39(12):11051-9. (IF2.929) 使用服务种类: **mRNA 测序服务**
- [7] Ma L, Yang S, Zhao W, Tang Z, Zhang T, Li K. Identification and analysis of pig chimeric mRNAs using RNA sequencing data. *BMC Genomics.* 2012, 13(1):429. (IF4.073) 使用服务种类: **mRNA 测序服务**
- [8] Ye L, Su X, Wu Z, Zheng X, Wang J, Zi C, Zhu G, Wu S, Bao W. Analysis of Differential miRNA Expression in the Duodenum of *Escherichia coli* F18-Sensitive and -Resistant Weaned Piglets. *PLoS One.* 2012, 7(8):e43741. (IF4.092) 使用服务种类: **miRNA 测序服务**
- [9] Huang J, Deng Q, Wang Q, Li KY, Dai JH, Li N, **Zhu ZD**, Zhou B, Liu XY, Liu RF, Fei QL, Chen H, Cai B, Zhou B, **Xiao HS**, Qin LX, Han ZG. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. *Nat Genet.* 2012 Aug 26. (IF35.532) 使用服务种类: **全外显子组测序服务**
- [10] Wu H, Qu S, Lu C, Zheng H, Zhou X, Bai L, Deng Z. Genomic and transcriptomic insights into the thermo-regulated biosynthesis of validamycin in *Streptomyces hygroscopicus* 5008. *BMC Genomics.* 2012, 13:337. (IF4.073) 使用服务种类: **Roche 454 测序服务/Agilent earray 定制芯片服务**
- [11] Tao XY, Xue XY, Huang YP, Chen XY, Mao YB. Gossypol-enhanced P450 gene pool contributes to cotton bollworm tolerance to a pyrethroid insecticide. *Mol Ecol.* 2012 Apr 20. (IF6.457) 使用服务种类: **Roche 454 测序/NimbleGen 芯片服务**
- [12] Gai S, Zhang Y, Mu P, Liu C, Liu S, Dong L, Zheng G. Transcriptome analysis of tree peony during chilling requirement fulfillment: Assembling, annotation and markers discovering. *Gene.* 2011 Dec 16. (IF2.266) 使用服务种类: **Roche 454 测序服务**
- [13] Wei X, Ju X, Yi X, Zhu Q, Qu N, Liu T, Chen Y, Jiang H, Yang G, Zhen R, Lan Z, Qi M, Wang J, Yang Y, Chu Y, Li X, Guang Y, **Huang J**. Identification of sequence variants in genetic disease-causing genes using targeted next-generation sequencing. *PLoS One.* 2011;6(12):e29500. (IF4.411) 使用服务种类: **序列捕获测序服务**
- [14] Wu DQ, Ye J, Ou HY, Wei X, Huang X, He YW, Xu Y. Genomic analysis and temperature-dependent transcriptome profiles of the rhizosphere originating strain *Pseudomonas aeruginosa* M18. *BMC Genomics.* 2011, 12:438. (IF4.206) (客户发表文章) 使用服务种类: **Roche 454 测序服务 / Agilent earray 定制芯片 (假单胞菌 M18) 服务 / SBC Analysis System 分析服务**

-
- [15] Zhong J, Cui Y, Guo J, Chen Z, Yang L, He QY, Zhang G, Wang T. Resolving Chromosome-Centric Human Proteome with Translating mRNA Analysis: A Strategic Demonstration. *J Proteome Res.* 2013.(IF 5.056) 使用服务种类: **RNA 测序服务**
- [16] Qin L, Chen Y, Liu X, Ye S, Yu K, Huang Z, Yu J, Zhou X, Chen H, Mo D. Integrative Analysis of Porcine microRNAome during Skeletal Muscle Development. *PLoS One.* 2013, 8(9):e72418.(IF 3.73) 使用服务种类: **miRNA 测序服务**
- [17] Zhang Y, Lu D, Ju T, Wang L, Lin S, Zhao Y, Wang C, He H, Du Y. Biodegradation of Phenol Using *Bacillus cereus* WJ1 and Evaluation of Degradation Efficiency Based on a Graphene-Modified Electrode. *Int. J Electrochem. Sci.* 2013, 8: 504-519.(IF 1.956) 使用服务种类: **Ion Torrent 测序服务**
- [18] Sun X, Ruan R, Lin L, Zhu C, Zhang T, Wang M, Li H, Yu D. Genomewide investigation into DNA elements and ABC transporters involved in imazalil resistance in *Penicillium digitatum*. *FEMS Microbiol Lett.* 2013 (IF 2.049) 使用服务种类: **全基因组重测序服务**
- [19] Liu X, Lu Y, Yuan Y, Liu S, Guan C, Chen S, Liu Z. De Novo Transcriptome of *Brassica juncea* Seed Coat and Identification of Genes for the Biosynthesis of Flavonoids. *PLoS One.* 2013, 8(8):e71110. (IF 3.73) 使用服务种类: **RNA 测序服务**
- [20] Chen G, Zhang C, Jiang F, Wang Y, Xu Z, Wang C. Bioinformatics analysis of hemocyte miRNAs of scallop *Chlamys farreri* against acute viral necrobiotic virus (AVNV). *Fish Shellfish Immunol.* 2014, (IF 2.964) 使用服务种类: **RNA 测序服务**
- [21] Jiang JH, Liu YF, Ke AW, Gu FM, Yu Y, Dai Z, Gao Q, Shi GM, Liao BY, Xie YH, Fan J, Huang XW, Zhou J. Clinical significance of the ubiquitin ligase E3C in hepatocellular carcinoma revealed by exome sequencing. *Hepatology*, 2014, (IF 12.003) 使用服务种类: **外显子组测序服务**
- [22] Xu J, Wang YY, Dai YJ, Zhang W, Zhang WN, Xiong SM, Gu ZH, Wang KK, Zeng R, Chen Z, Chen SJ. DNMT3A Arg882 mutation drives chronic myelomonocytic leukemia through disturbing gene expression/DNA methylation in hematopoietic cells. *PNAS.* 2014,(IF 9.737) 使用服务种类: **MeDIP-Seq 服务**
- [23] Cheng TL, Wang ZZ, Liao QM, Zhu Y, Zhou WH, Xu WQ, Qiu ZL. MeCP2 Suppresses Nuclear MicroRNA Processing and Dendritic Growth by Regulating the DGCR8/Drosha Complex. *Developmental Cell.* 2014,(IF 12.861) 使用服务种类: **miRNA 测序服务**
- [24] Wang J, Yang Y, Liu X, Huang J, Wang Q, Gu J, Lu YM. Transcriptome profiling of the cold response and signaling pathways in *Lilium lancifolium*. *BMC genomics.* 2014, 15:203 (IF 4.397) 使用服务种类: **转录组测序服务**
- [25] Gu Y, Lu K, Yang G, Cen Z, Yu L, Lin L, Hao J, Yang Z, Peng J, Cui S, Huang J. Mutation Spectrum of Six Genes in Chinese Phenylketonuria Patients Obtained through Next-Generation Sequencing. *PLoS One.* 2014,9(4):e94100.(IF 3.73) 使用服务种类: **ION PGM 测序服务**
- [26] Xiao CL, Mai ZB, Lian XL, Zhong JY, Jin JJ, He QY, Zhang G. FANSe2: A Robust and Cost-Efficient Alignment Tool for Quantitative Next-Generation Sequencing Applications. *PloS one.* 2014;9(4):e94250 (IF 3.73) 使用服务种类: **RNA 测序服务**
- [27] Schaid DJ, Spraggs CF, McDonnell SK, Parham LR, Cox CJ, Ejlertsen B, M. Finkelstein D, Rappold E, Curran J, R. Cardon L, E. Goss P. Prospective Validation of HLA-DRB1*07:01 Allele Carriage As a Predictive Risk Factor for Lapatinib-Induced Liver Injury. *Journal of clinical oncology.* 2014,(IF 17.879) 使用服务种类: **测序服务 (GWAS)**
- [28] Chen R, Zhang F, Song L, Shu YN, Lin YL, Dong LH, Nie X, Zhang DD, Chen C, Han M. Transcriptome profiling reveals that the SM22alpha-regulated molecular pathways contribute to vascular pathology. *J Mol Cell Cardiol.* 2014,72:263-72.(IF 5.148) 使用服务种类: **RNA 测序服务**

-
- [29] Yuan JH, Yang F, Wang F, Ma JZ, Guo YJ, Tao QF, Liu F, Pan W, Wang TT, Zhou CC, Wang SB, Wang YZ, Yang Y, Yang N, Zhou YP, Yang GS, Sun SH. A Long Noncoding RNA Activated by TGF- β Promotes the Invasion-Metastasis Cascade in Hepatocellular Carcinoma. *Cancer Cell.* 2014, 12;25(5):666-81.(IF 24.755) 使用服务种类: **RIP-SEQ 服务**
- [30] Xia ZH, Peng J, Li YQ, Chen L, Li S, Zhou T, Fan ZF. Characterization of Small Interfering RNAs Derived from Sugarcane Mosaic Virus in Infected Maize Plants by Deep Sequencing. *PLoS one.* 2014, 9;(5)e97013.(IF 3.73) 使用服务种类: **small RNA 测序服务**
- [31] Dong ZG, Ge JC, Xu ZQ, Dong XH, Cao SS, Pan JL, Zhao QS. Generation of Myostatin b Knockout Yellow Catfish (*Tachysurus Fulvidraco*) Using Transcription Activator-Like Effector Nucleases. *Zebrafish.* 2014, (IF 2.883) 使用服务种类: **RNA 测序服务**
- [32] Li X, Zhang BT, Wu Q, Ci XP, Zhao RR, Zhang ZQ, Xia SY, Su D, Chen J, Ma G, Fu LY, Dong JT. Interruption of KLF5 acetylation converts its function from tumor suppressor to tumor promoter in prostate cancer cells. *Int J Cancer.* 2014,(IF 6.198) 使用服务种类: **RNA 测序服务**
- [33] Wang G, Zhang JS, Wang GF, Fan XY, Sun X, Qin HL, Xu N, Zhong MY, Qiao ZY, Tang YP, Songa R., Proline responding1 Plays a Critical Role in Regulating General Protein Synthesis and the Cell Cycle in MaizeC. *Plant Cell.*2014,(IF 9.251) 使用服务种类: **RNA 测序服务**
- [34] Gong WZ, Qi PF, Du JB, Sun X, Wu XL, Song C, Liu WG, Wu YS, Yu XB, Yong TW, Wang XC, Yang F, Yan YH, Yang WY. Transcriptome analysis of shade-induced inhibition on leaf size in relay intercropped soybean. *PLoS One.* 2014 ,9 (6): e98465.(IF 3.73) 使用服务种类: **RNA 测序服务**
- [35] Xiong YY, Huang JF, Li XX, Zhou LL, Dong F, Ye H, Gan L. Deep sequencing of the tilapia (*Oreochromis niloticus*) liver transcriptome response to dietary protein to starch ratio. *Aquaculture.* 2014, 433,20:299-306.(IF 2.009) 使用服务种类: **RNA 测序服务**
- [36] Wang G, Zhang J, Wang G, Fan X, Sun X, Qin H, Xu N, Zhong M, Qiao Z, Tang Y, Song R. Proline responding1 Plays a Critical Role in Regulating General Protein Synthesis and the Cell Cycle in Maize. *Plant Cell.* 2014, 26(6): 2582-600.(IF 9.575) 使用服务种类: **RNA 测序服务**
- [37] Liu X, Wang Q, Gu J, Lü Y. Vernalization of Oriental hybrid lily ‘Sorbonne’: changes in physiology metabolic activity and molecular mechanism. *Mol Biol Rep.* 2014,(IF 1.958) 使用服务种类: **RNA 测序服务**
- [38] Wang H, Tong W, Feng L, Jiao Q, Long L, Fang RJ, Zhao WG. De novo transcriptome analysis of mulberry (*Morus L.*) under drought stress using RNA-Seq technology. *Russ J Bioorg Chem.* 2014, 40(4):423-32.(IF 0.623) 使用服务种类: **RNA 测序服务**
- [39] Huang WL, Wang WL, Shi WS, Lei ZF, Zhang ZY, Chen RZ, Zhou BB. Use low direct current electric field to augment nitrification and structural stability of aerobic granular sludge when treating low COD/NH4-N wastewater. *Bioresource Technol.* 2014,(IF 5.039) 使用服务种类: **Ion Torrent 测序服务**
- [40] Wu XH, Xu L, Gu W, Xu Q, He QY, Sun XS, Zhang G. Iterative Genome Correction Largely Improves Proteomic Analysis of Nonmodel Organisms. *J Proteome Res.*2014,(IF 5.001) 使用服务种类: **基因组测序**
- [41] Wang J, Wang Q, Yang Y, Liu X, Gu J, Li W, Ma S, Lu Y.. De novo assembly and characterization of stress transcriptome and regulatory networks under temperature, salt and hormone stresses in *Lilium lancifolium*. *Mol Biol Rep.* 2014,(IF 1.958) 使用服务种类: **RNA 测序服务**
- [42] Luo H, Wu Q, Ye X, Xiong Y, Zhu J, Xu J, Diao Y, Zhang D, Wang M, Qiu J, Miao J, Zhang W, Wan J. Genome-Wide Analysis of miRNA Signature in the APPswe/PS1 Δ E9 Mouse Model of Alzheimer’s Disease. *Plos One.* 2014,22;9(8):e101725.(IF 3.534) 使用服务种类: **RNA 测序服务**
- [43] Chen J, Quan M, Zhang D. Genome-wide identification of novel long non-coding RNAs in *Populus tomentosa* tension wood, opposite wood and normal wood xylem by RNA-seq. *Planta.* 2014,(IF 3.376) 使

用服务种类: RNA 测序服务

[44] Zhang Y, Xie S, Zhou Y, Xie Y, Liu P, Sun M, Xiao H, Jin Y, Sun X, Chen Z, Huang Q, Chen S. H3K36 Histone Methyltransferase Setd2 Is Required for Murine Embryonic Stem Cell Differentiation toward Endoderm. 2014,25;8(6):1989-2002 (IF 7.207) 使用服务种类: **RNA 测序服务**

[45] Liu XH, Huang J, Wang JM, Lu YM. RNA Seq Analysis Reveals Genetic Bases of the Flowering Process in Oriental Hybrid Lily cv. Sorbonne. RUSS J PLANT PHYSIOL+. 2014,(IF 0.759) 使用服务种类: **RNA 测序服务**

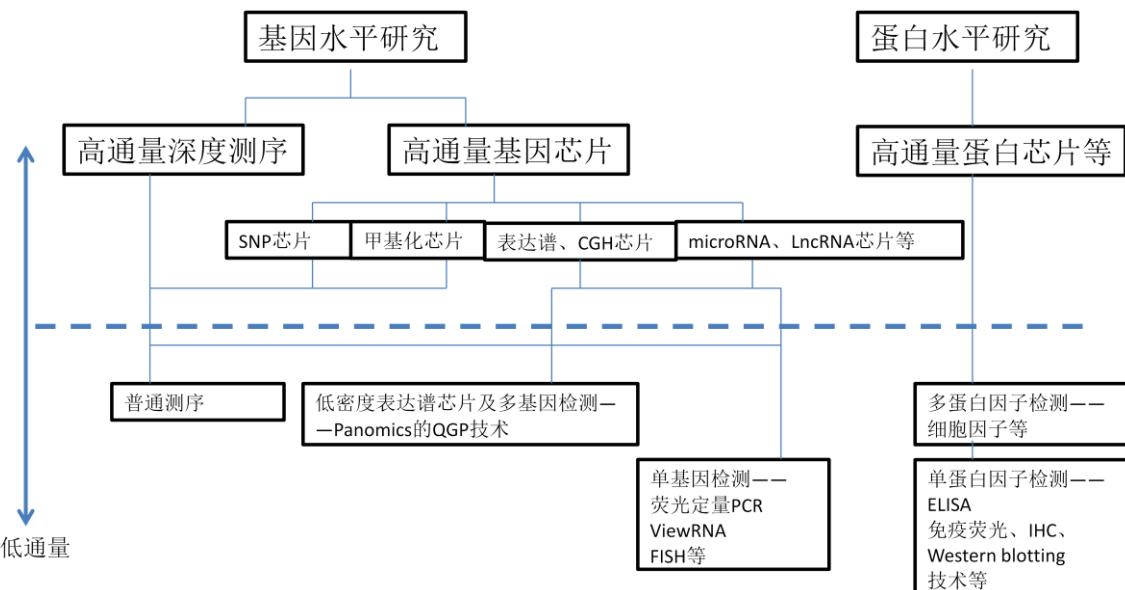
[46] Wang YJ , Zhao YH,Sun C, Hu W,Zhao J,Li GC, Zhang LZ, Liu M,Liu Y,Ding F,Yang YM, Gu XS. Chitosan Degradation Products Promote Nerve Regeneration by Stimulating Schwann Cell Proliferation via miR-27a/FOXO1 Axis. Mol Neurobiol. 2014,(IF 5.286) 使用服务种类: **miRNA 测序**

[47] Xia J, Yuan J, Xin L, Zhang Y, Kong S, Chen Y, Yang S, Li K. Transcriptome analysis on the inflammatory cell infiltration of nonalcoholic steatohepatitis in bama minipigs induced by a long-term high-fat, high-sucrose diet. PLoS One. 2014,21;9(11):e113724.(IF 3.534) 使用服务种类: **RNA 测序服务**

第五章 生物标志物验 证及分子检测服务

上海伯豪生物技术有限公司生物标志物验证及分子检测平台负责核酸和蛋白标志物的检测以及有关技术的研发和服务工作，也针对本公司微阵列芯片平台和新一代测序平台的结果提供相应的验证服务。该平台采用多种技术，利用目标区域测序、定量PCR、QGP等技术方法检测或验证靶标；同时也采用高灵敏度、准确和快速的方法对蛋白标记物进行验证；该平台还提供基因组编辑技术(CRISPR/Cas9)服务。

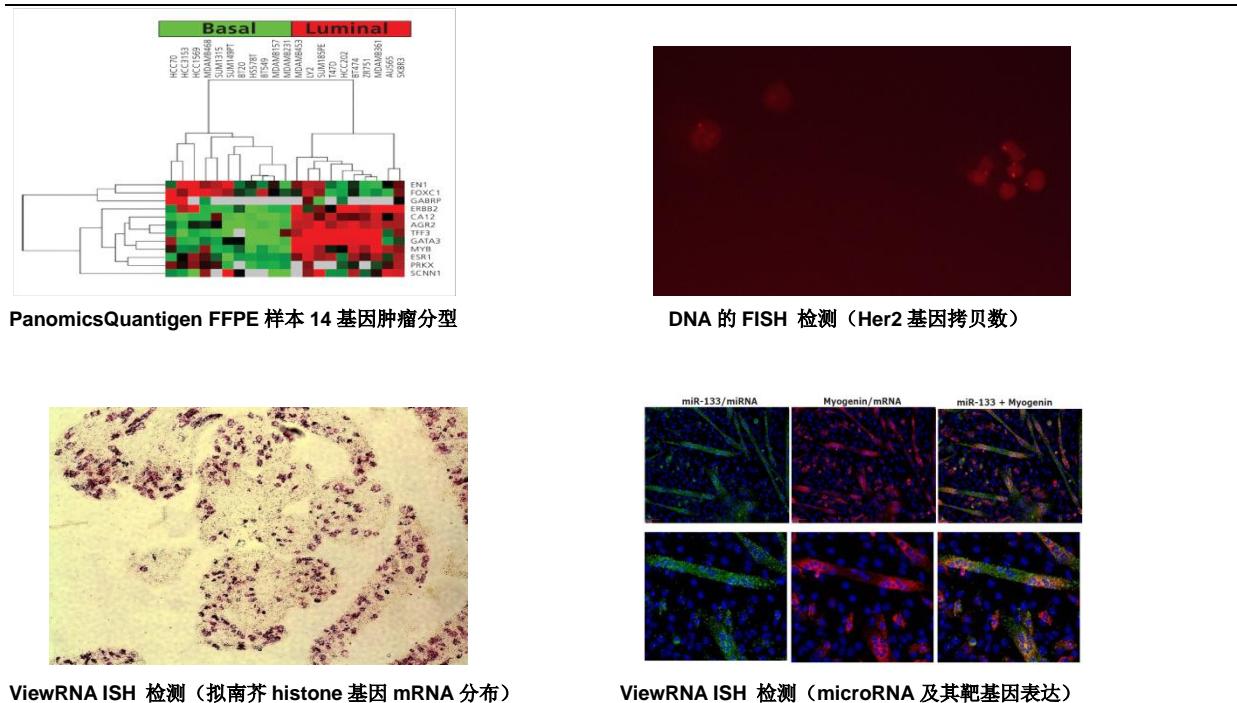
技术路线



生物标志物验证及分子检测平台拥有以下几类技术：核酸检测技术、蛋白免疫检测技术平台、siRNA筛选及细胞模型建立技术、基因克隆及蛋白表达技术、基因组编辑CRISPR/Cas9技术。

服务一览

中低通量核酸 检测技术	蛋白免疫检测技术	标志物功能筛选 及验证工具
QuantiGene 服务	蛋白芯片	siRNA 筛选
QGP 多基因检测	Panomics 等多细胞因子检测	细胞模型
荧光定量 PCR	ELISA	基因组编辑技术（CRISPR/Cas9）
PCR Array	免疫荧光、IHC/Western Blotting	基因克隆蛋白表达纯化
Sanger 测序		
焦磷酸测序		
FISH 及 ViewRNA		
MassArray 检测		
以及其它相关技术		



(一) 中低通量核酸检测

核酸检测主要包括：定量PCR的Taqman法、sybgreen法等检测；branch DNA技术的mRNA,miRNA,LncRNA或基因拷贝数检测；PCR产物测序或克隆测序；焦磷酸测序；MassArray检测；PCR Array检测等内容。

技术及应用

✓ 低通量多基因验证技术

- 采用Panomics的branch DNA-QGP技术对mRNA, miRNA等进行检测
- 拥有Qiagen的PCR-Aarray低密度芯片技术，进行基因Pathway水平检测
- 拥有ABI的TLDA低密度芯片技术，进行低通量核酸水平检测

✓ 常规测序技术

- 进行甲基化、SNP、突变验证检测，PCR产物测序或克隆测序

✓ 焦磷酸测序技术

- 适合进行DNA甲基化检测，也可进行突变检测等遗传学分析

✓ 荧光定量PCR技术

- Taqman探针进行SNP位点检测，Taqman探针及Sybgreen探针进行多个或单个基因或MicroRNA表达水平等验证、拷贝数等变化等检测

✓ 荧光核酸检测技术

——ViewRNA及FISH技术

✓ **MassArray 检测**

——可用于SNP分型、甲基化分析、CNV分析等。本平台承接用该技术检测SNP业务

基于 bDNA 技术的检测服务

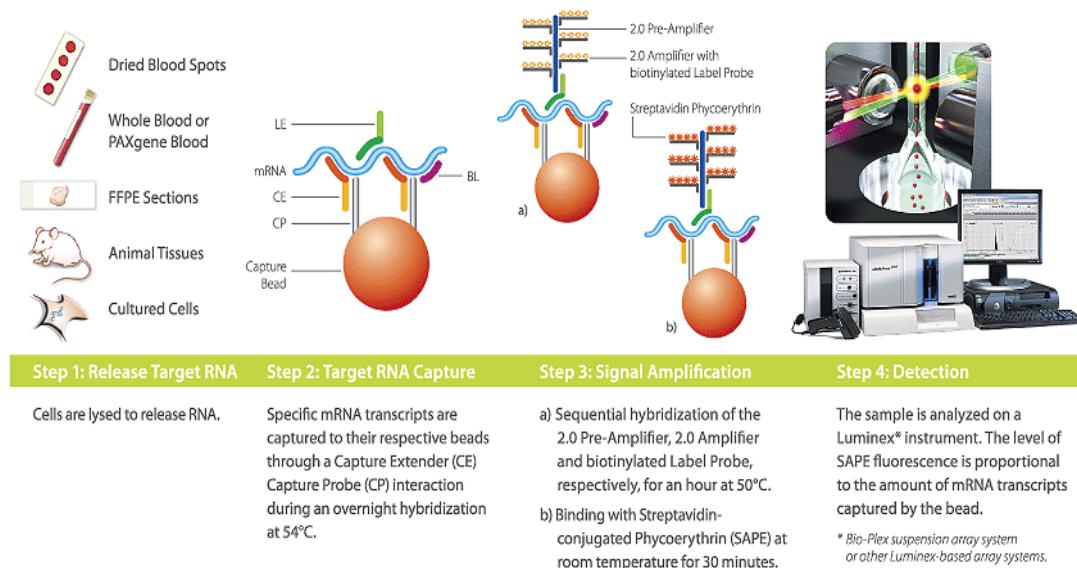
Affymetrix公司旗下的Panomics开发出的Branch-DNA (bDNA) 技术是一种三明治结构的核酸杂交方法，该方法采用 bDNA 分子放大捕获的目标 RNA 信号。bDNA 测定的优势是直接从样本中获得RNA进行分析，无需对 RNA 进行提纯或酶处理，从而避免了这些过程中可能引起的分析效率下降以及RNA降解。

该技术目前已应用到基础研究和临床前期研究等各领域中。SBC还将 bDNA技术与Luminex / BioPlex 平台相结合，为您提供蛋白分析、核酸分析和RNA原位分析等服务。

服务内容：核酸分析

核酸分析原理

Affy-Panomics 技术所有技术基于 Branched DNA (bDNA) 技术原理：是一种三明治结构的核酸杂交方法，该方法采用 bDNA 分子放大捕获的目标 RNA 信号。



技术特点

- ✓ 同时定量测量多个 RNA 或 DNA 目标，准确性与精确度无与伦比
- ✓ RNA 与 DNA 定量直接来源于培养细胞、全血或组织（新鲜、冰冻或福尔马林固定、石蜡包埋 (FFPE)）
- ✓ 无需抽提纯化 RNA 或 DNA
- ✓ RNA 无需反转录
- ✓ 无目标放大
- ✓ 简单的测定流程

技术组成

✓ **定量检测技术：**

单基因：Quantigene

——实现单个基因的准确定量检测，具有很高的灵敏度，可检测低至 200–500 个拷贝到核酸分子；
 多基因定量检测技术：Quantigene Plex
 ——同一反应中可同时检测 3–36 个基因的表达水平，同时实现这些核酸分子的准确定量检测；
 ——特别适用于血液样本和保存，以及多年后 mRNA 高度降解的 FFPE 样本；
 ——同时适用于基因拷贝数的检测，可以分辨 0.5–4 个拷贝，以及 MicroRNA、LncRNA 等等检测；操作简便，优于 Q-PCR 技术。

✓ 荧光基因原位杂交检测技术：

ViewRNA 技术：

基于 b-DNA 技术原理，进行原位杂交实验，结果为图像，可以定性和半定量分析目标基因的表达及分布情况。
 ——单细胞水平检测低至一个拷贝的 mRNA，在定量与结构生物学水平原位检测多基因的 mRNA，在抗体无法获得情况下，对免疫细胞化学研究是一种极好的补充。

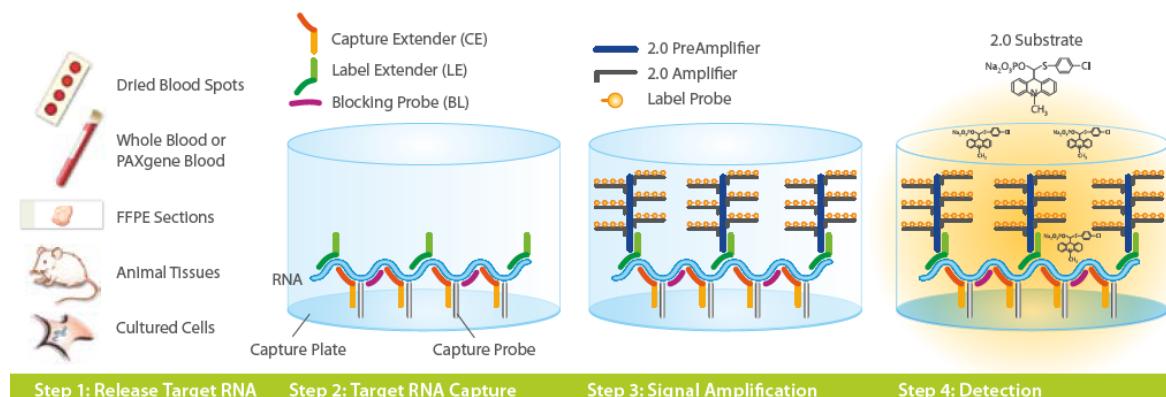
1. 单基因表达精确定量服务

产品推荐：QuantiGene® 2.0 系列产品

服务特点：

- ✓ 无需抽提纯化 RNA、无需反转录、无需 PCR 扩增；
- ✓ 可分析各种样本（组织、细胞、血液、福尔马林浸泡、石蜡包埋等），只需用特定裂解液裂解即可，对样品量下限无要求（只需 mRNA 拷贝数在检测范围内）；
- ✓ 该技术也可用于转基因动植物的 DNA 拷贝数检测、DNA 拷贝数变化等分析用途。

实验流程：



样品用特定裂解液裂解后加入孔中 过夜杂交，探针组 (CE, LE, BL) 特异性识别目标 mRNA 信号经逐层放大。LE 的数目决定了检测的灵敏度 加入底物产生化学发光，光强度与待测 mRNA 成比例

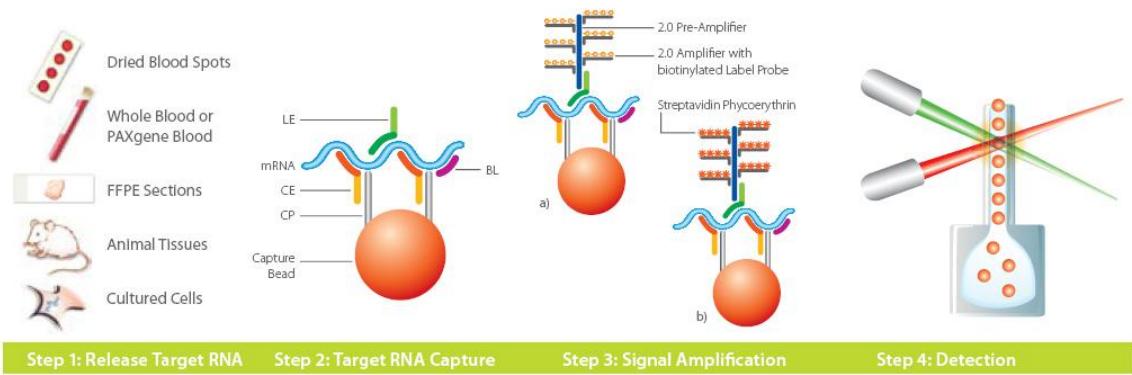
2. 多基因表达精确定量服务

产品推荐: QuantiGene® Plex 2.0 系列产品

服务特点:

- ✓ 同一反应中可同时对 3-36 个基因;
- ✓ 在定量与结果细胞生物学水平原位检测多基因的核酸分子, 特别适用于血液样本和保存
- ✓ 多年后 mRNA 高度降解的 FFPE 样本;
- ✓ 在抗体无法得到的情况下, 对免疫细胞化学研究是一种极好的补充。

实验流程:



样品用特定裂解液 过夜杂交, 探针组 (CE、LE、
裂解 BL) 特异地识别目标 mRNA,
并将其连结早特定的微珠上 信号经逐层放大。LE 上连接的是 SAPE 荧
光基团 用 Luminex/Bio-Plex 检测，
特定微珠上的 SAPE 荧光强
度与特定 mRNA 量成比例

Quantigene 定量检测技术的应用:

- ✓ 表达谱芯片数据的验证
- ✓ DNA 拷贝数的检测
- ✓ 临床 FFPE 样本和血液样本的多基因表达检测
- ✓ 一信号传导通路检测
- ✓ 一药物筛选及 siRNA 筛选样本特定目标基因的检测
- ✓ 标志物基因与药物靶点验证, 等

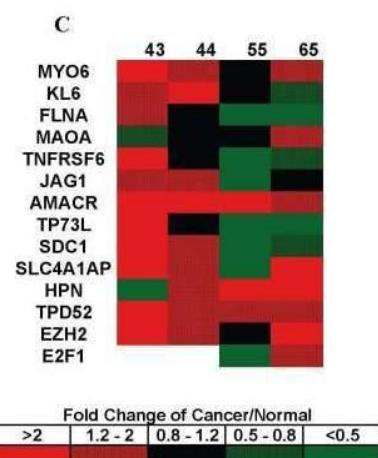
应用举例

一、多基因的表达水平变化

在前列腺肿瘤石蜡样本中进行多基因检测:

Knudsen et al J Mol Diagnostics 2008

该文提到在石蜡样本中成功进行 14 个前列腺癌相关的标志基因检测。



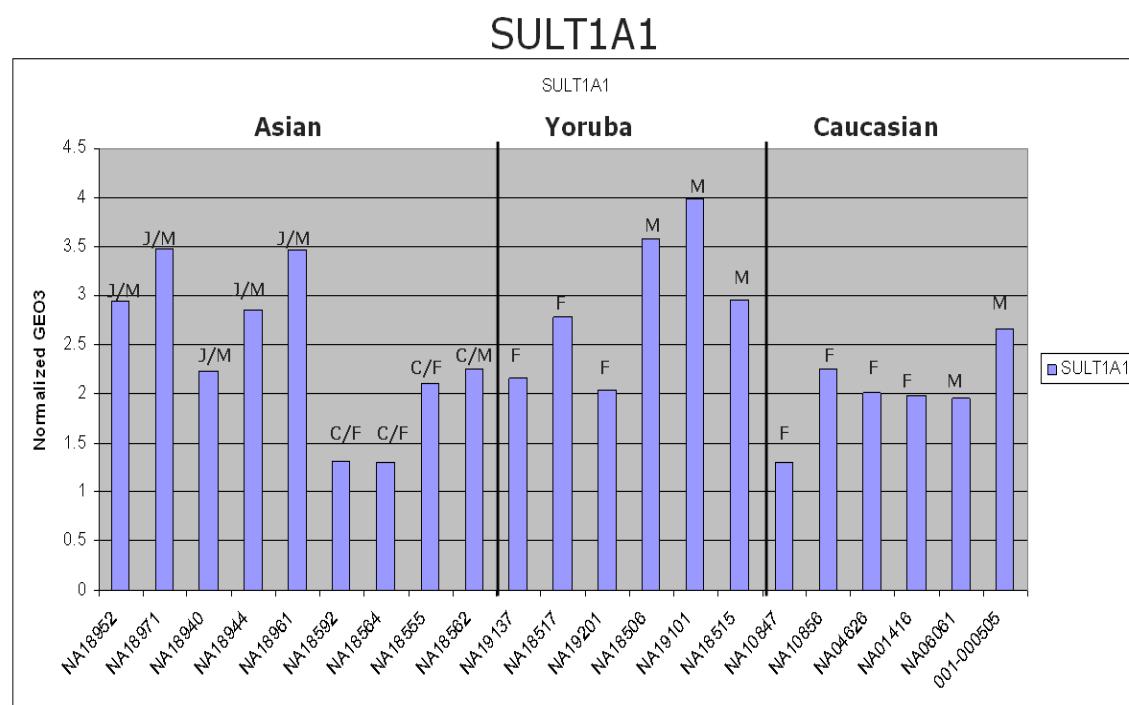
Macro dissected FFPE samples from matched tumor and normal tissues homogenized in Tissue Homogenization Solution (THS). 14 prostate cancer gene panel measured in tumor and adjacent normal tissues. Gene expression normalized to housekeeping gene RPL19. The ratio between tumor and normal tissues is calculated and shown in a 5-tiered categorical scale. “In macrodissected tissues from 9 – 13 year old blocks with poor RNA quality, the QuantiGene® Reagent System correctly identified the over-expression of known cancer genes (arrows).”

“In conclusion, the QuantiGene® Reagent System appears to be well suited for clinical analysis of FFPE tissues with diagnostic or prognostic gene expression biomarker panels.”

二、基因拷贝数的验证

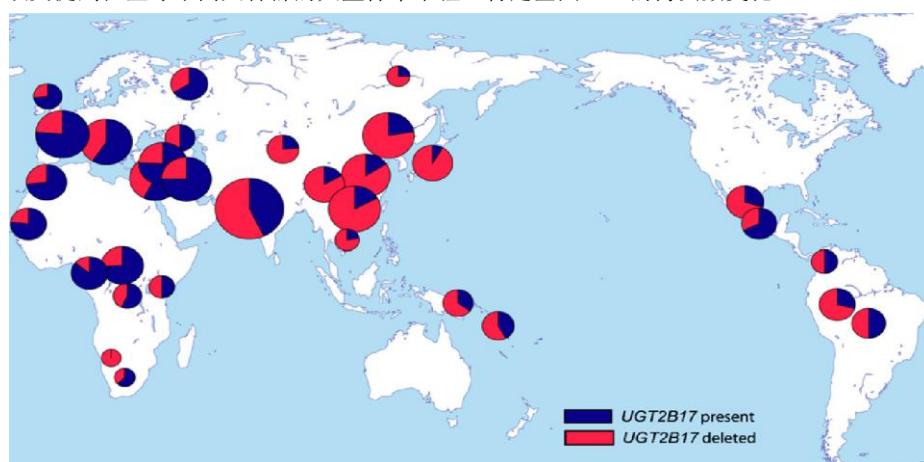
SULT1A1 DNA Copy Number Ethnic Groups, Human Molecular Genetics 2007

该文提到在不同美国人种群的大量样本中验证特定基因 SULT1A1 的拷贝数变化。



Yali Xue et al The American Journal of Human Genetics 83, 337 – 346, September 12, 2008

该文提到在全球不同人种群的大量样本中验证特定基因 UGT 的拷贝数变化。



服务内容：核酸原位分析

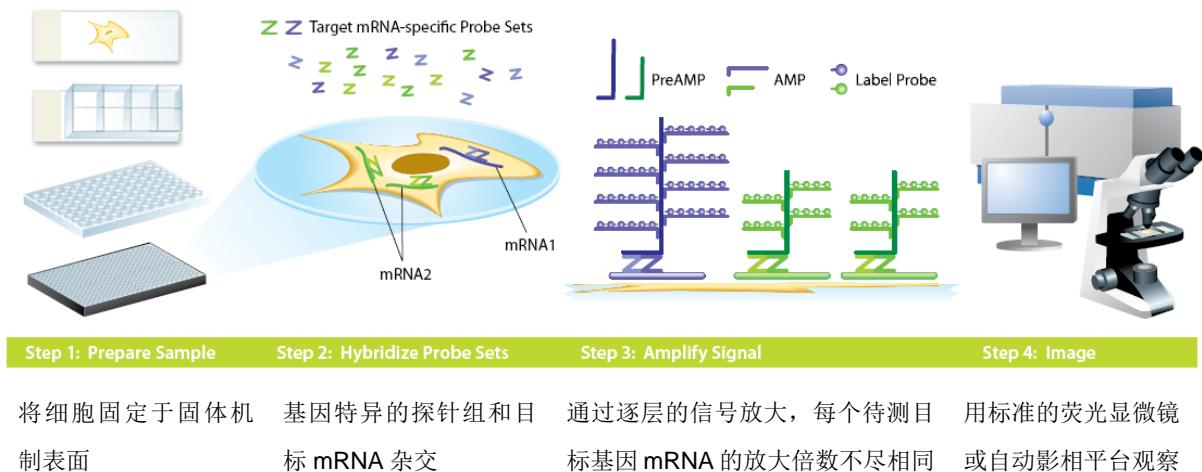
1 单细胞多基因单拷贝核酸分子原位检测服务

产品推荐：QuantiGene® ViewRNA 系列产品

服务特点：

- ✓ 单细胞水平检测单拷贝核酸分子；
- ✓ 在定量与结构细胞生物学水平原位可检测 2 个以上的核酸分子；
- ✓ 在抗体无法得到的情况下，对免疫细胞化学研究是一种极好的补充。

实验流程：



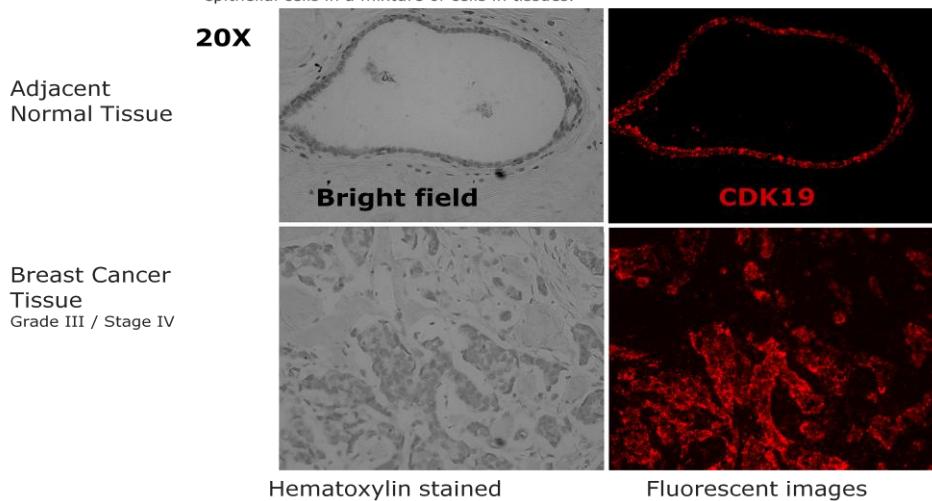
单基因荧光原位杂交 ViewRNA 检测技术应用：

- ✓ mRNA 原位表达水平及分布检测
- ✓ microRNA 原位表达水平及分布检测
- ✓ lncRNA 原位表达水平及分布检测
- ✓ 基因拷贝数原位表达水平及分布检测

应用举例

一、应用 ViewRNA 技术检测 CK19 在乳腺癌样本中的表达和分布

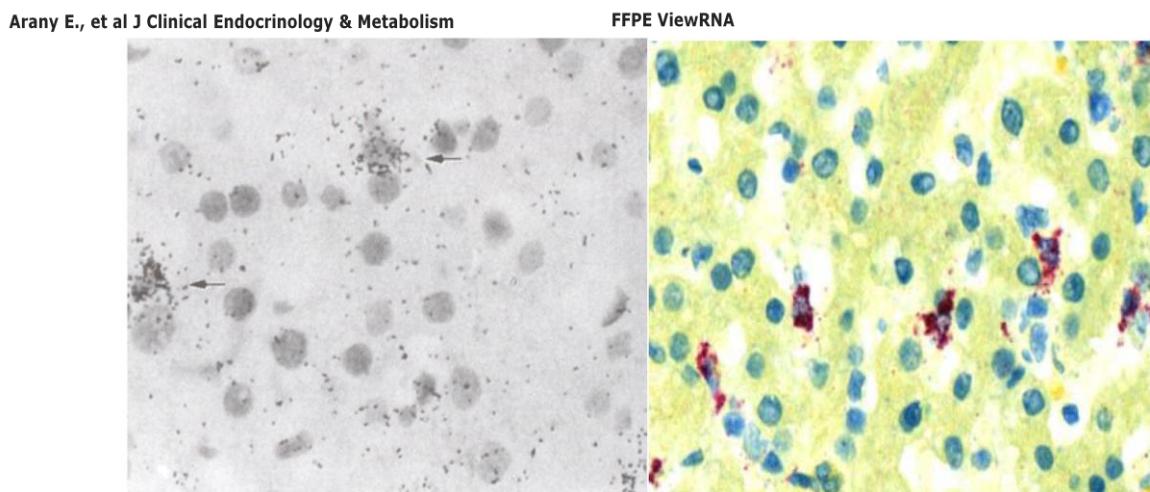
CDK19 (Cytokeratin 19) is an epithelial cell specific marker that can be used to identify epithelial cells in a mixture of cells in tissues.



二、ViewRNA 技术与放射性原位杂交技术比较

Arany E., et al J Clinical Endocrinology & Metabolism

该文比较了 ViewRNA 技术与放射性原位杂交技术，ViewRNA 技术更方便，快捷，安全。



基于 PCR-array 技术的检测服务

RT2 Profiler PCR Arrays 芯片

基于 real time RT PCR 技术，以阵列方式高通量集成检测，研究通路或疾病相关基因的表达。阵列选择覆盖癌症、凋亡、信号转导、细胞因子、炎症反应、细胞毒理、药物代谢、干细胞及发育生物学、神经科学等方面。操作简单，无需优化。

技术优势

✓ 结果高度敏感性，精确性，可复制性：

1) 科学和完整的参照系统：

对每个样本都设立了多个 housekeeping gene (HKG)、genomic DNA contamination 参照、逆转录效率参

照 (RTC)、阳性 PCR 反应条件参照 (positive PCR control)，确保了每一步实验的精确性符合发表 q-PCR 数据的国际标准 M I Q E。

2) 严格的质控过程：

PCR Array 三个步骤涉及的试剂，引物实验条件均经实验验证，确保均一和高度扩增 效率 (Aplification Efficiency 100%)，特异性(单峰分离曲线)，敏感性 (1ng/96-well plate) 和宽广的线性范围(wide dynamic range 108)

✓ 是操作简单，节约时间的即用型反应系统：

在几个小时内一次性检测 84 or 370 个基因的表达水平，所有实验条件均已设定，经过验证，无需优化条件。利用在线分析工具可以在 15 分钟内处理数千个 Ct 值，用户只需经过简单培训

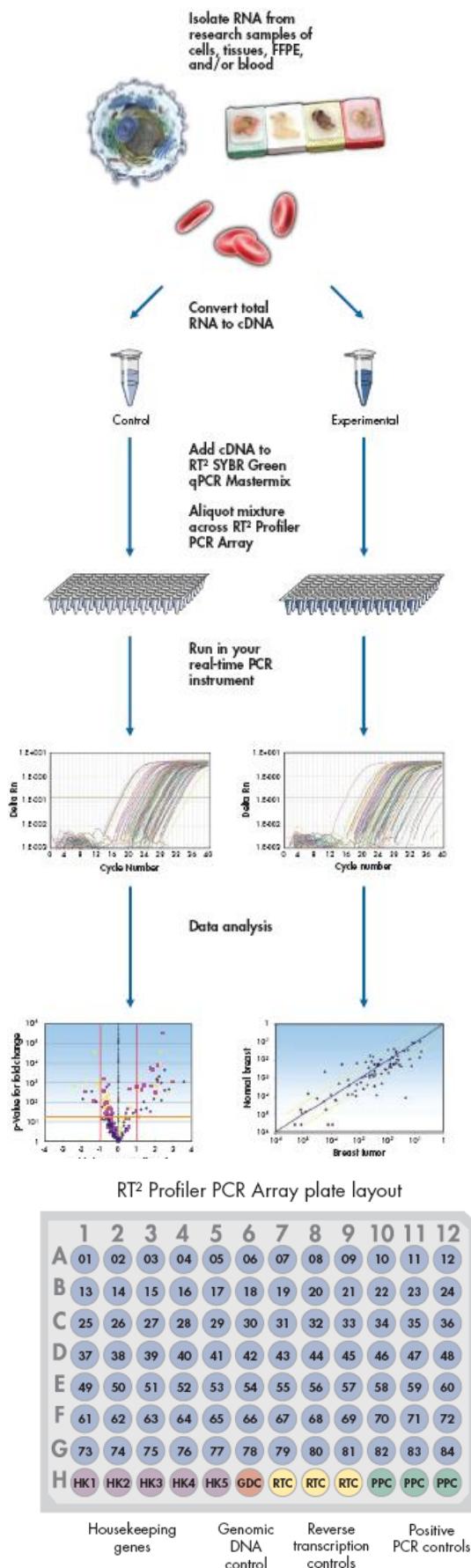
✓ 基因选择具有科学性和时效性，目标在于解决生物学问题：

PCR Arrays 上基因的选择过程。这些被选中的基因并不仅仅是简单的包括某个信号通路中的所有基因，所选基因必须对于相应的信号转导途径相当重要，并且是在 mRNA 水平上调控的。研发人员及生物信息学科学家通过阅读已发表文献，从疾病发病，进展，治疗，耐药性，等多方面入手。研发部门做了大量实验来证明所选基因的相关性。另外，我们与学术界的权威密切合作，所选的所有基因都通过了他们的认可。

✓ 产品应用广泛：

有大于 1000 多篇文献引用了 PCR Array 数据，其中包括在 Science, PNAS 等杂志上发表的文献。全球有 1000 多个实验室（包括 300 多个大学）在使用该产品，中国各地机构有北京/上海/武汉中科院，中山肿瘤医院等。

✓ 与多种 PCR 仪兼容



Popular RT ² Profiler PCR Arrays for research	Cat. no.* SAP # / PCR Array #
<u>Angiogenesis</u>	330231 / PAXX-024Y
<u>Apoptosis</u>	330231 / PAXX-012Y
<u>Autophagy</u>	330231 / PAXX-084Y
<u>Breast Cancer and Estrogen Receptor Signaling</u>	330231 / PAXX-005Y
<u>Cancer PathwayFinder</u>	330231 / PAXX-033Y
<u>Cell Cycle</u>	330231 / PAXX-020Y
<u>Chemokines and Receptors</u>	330231 / PAXX-022Y
<u>Diabetes</u>	330231 / PAXX-023Y
<u>DNA Damage Signaling Pathway</u>	330231 / PAXX-029Y
<u>Drug Metabolism</u>	330231 / PAXX-002Y
<u>EGF/PDGFR Signaling Pathway</u>	330231 / PAXX-040Y
<u>Embryonic Stem Cells</u>	330231 / PAXX-081Y
<u>Endothelial Cell Biology</u>	330231 / PAXX-015Y
<u>Epigenetic Chromatin Modification Enzymes</u>	330231 / PAXX-085Y
<u>Epithelial to Mesenchymal Transition</u>	330231 / PAXX-090Y
<u>Extracellular Matrix and Adhesion Molecules</u>	330231 / PAXX-013Y
<u>GPCR Signaling Pathway</u>	330231 / PAXX-071Y
<u>Growth Factors</u>	330231 / PAXX-041Y
<u>Heat Shock Proteins</u>	330231 / PAXX-076Y
<u>Hedgehog Signaling Pathway</u>	330231 / PAXX-078Y
<u>Hematopoietic Stem Cells & Hematopoiesis</u>	330231 / PAXX-054Y
<u>Hepatotoxicity</u>	330231 / PAXX-093Y
<u>Hypoxia Signaling Pathway</u>	330231 / PAXX-032Y
<u>Inflammatory Cytokines and Receptors</u>	330231 / PAXX-011Y
<u>Innate and Adaptive Immune Response</u>	330231 / PAXX-052Y
<u>Interferon α/β Response</u>	330231 / PAXX-016Y
<u>JAK/STAT Signaling Pathway</u>	330231 / PAXX-039Y
<u>MAP Kinase Signaling Pathway</u>	330231 / PAXX-061Y
<u>Mitochondrial Energy Metabolism</u>	330231 / PAXX-008Y
<u>Nephrotoxicity</u>	330231 / PAXX-094Y
<u>NF-κB Signaling Pathway</u>	330231 / PAXX-025Y
<u>Oxidative Stress and Antioxidant Defense</u>	330231 / PAXX-065Y
<u>p53 Signaling Pathway</u>	330231 / PAXX-027Y
<u>Wnt Signaling Pathway</u>	330231 / PAXX-043Y

(二) 蛋白标志物的免疫检测技术

检测基于免疫反应，主要用于检测蛋白水平的表达及定量定性分析。

检测技术

✓ 蛋白芯片

——采用Fullmoon蛋白芯片。所有抗体都通过专利的3-D聚合材料固定在高质量的玻片表面，从而确保抗体芯片产品高度的有效性和特异性。为了增大数据的可信度，每个抗体都进行了六次重复。芯片主要有广谱磷酸化蛋白芯片，广谱及功能特异的信号传导通路蛋白芯片，肿瘤标志物蛋白芯片等的通量化检测。上海伯豪固相蛋白芯片产品明细前已述及。

✓ 低通量多蛋白因子定量检测技术

——Panomics的Cytokine检测、Biorad的Cytokine、磷酸化检测试剂盒等

✓ 单因子定量检测技术

——免疫酶联反应的ELISA技术

✓ 单因子定性检测技术

——免疫荧光、Western Blotting、免疫组化 IHC 等检测技术

基于 xMAP 技术的检测服务

Luminex 的代表产品 Luminex® 100/200 以及新推出的 FLEXMAP 3DTM 和 BIO-RAD 公司的 BioPlex system 都是基于 xMAP 技术原理，整合了荧光编码微球、激光检测、应用流体学、最新的高速数字信号和计算机运算法则等多项技术的液相悬浮芯片，真正实现了“高通量”检测，并荣获 2005 年度国际临床诊断技术革新奖。是唯一得到美国 FDA 批准的，也是唯一被纳入美国临床实验室质控网络的高通量诊断技术。被国际业界专家评价为临床诊断的趋势性技术之一。

xMAP技术原理



xMAP技术优势

- ✓ **高通量，高速度：** 每个微球作为单独的检测体，可同时进行大量的生物检测，只需要 10~20 μl 的样本量就可以一次检测多达 100 个指标（FLEXMAP 3DTM 可多达 500 个指标），最快可达 10000 次测试/小时，真正实现了“高通量”与“高速度”。
- ✓ **多功能性：** xMAP 技术可以运用到多种生物检测中，包括免疫分析、基因分型、基因表达、酶分析等。既能检测蛋白，又能检测核酸。除了用于临床外，也能用于科研、CDC、血站、农牧业、生物及制药专业实验室等。
- ✓ **灵活性高：** 微球上可连接特异性的探针、抗原或抗体等来满足不同客户的需求。
- ✓ **灵敏度高：** 检测低限可达 0.01 pg/ml。
- ✓ **重复性好：** 类均相反应模式，每个指标有 1000-5000 个反应单元，分析 100 次取中位均值。
- ✓ **准确性高：** 检测范围达 3.5-6 个数量级，与 ELISA 和质谱分析具有很强的一致性。
- ✓ **成本低：** 流式荧光技术联检的试剂用量少，能有效降低临床应用的成本。

服务内容：蛋白分析

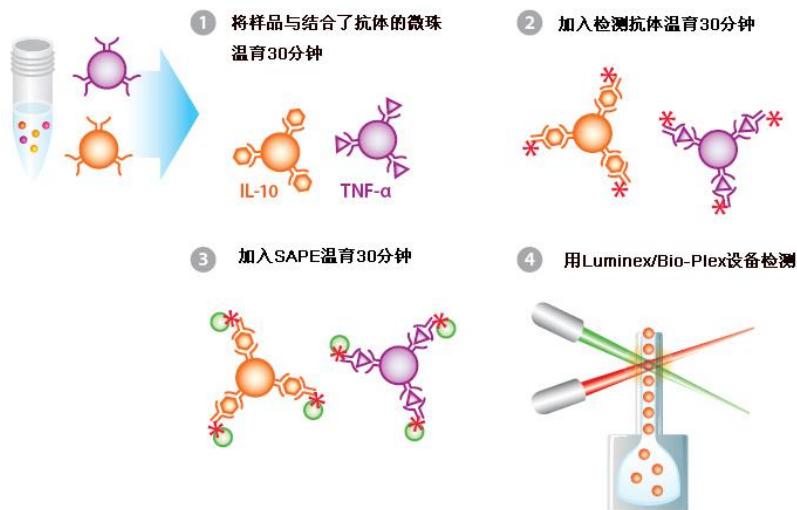
1 细胞因子/趋化因子高通量检测服务

产品推荐：Procarta Cytokine Plex Profiling Kit

服务内容：

细胞因子是机体分泌的蛋白质，在先天免疫、凋亡、血管生成、细胞生长和分化过程中扮演着重要角色。该检测服务可在同一反应中同时检测最多 50 种蛋白，目前已有 77 中 Human, 34 种 Mouse, 28 种 Rat, 43 种 non-human primate 和 6 种 porcine 细胞因子/趋化因子可供选择。并可根据您的需要随意组合。

实验流程：



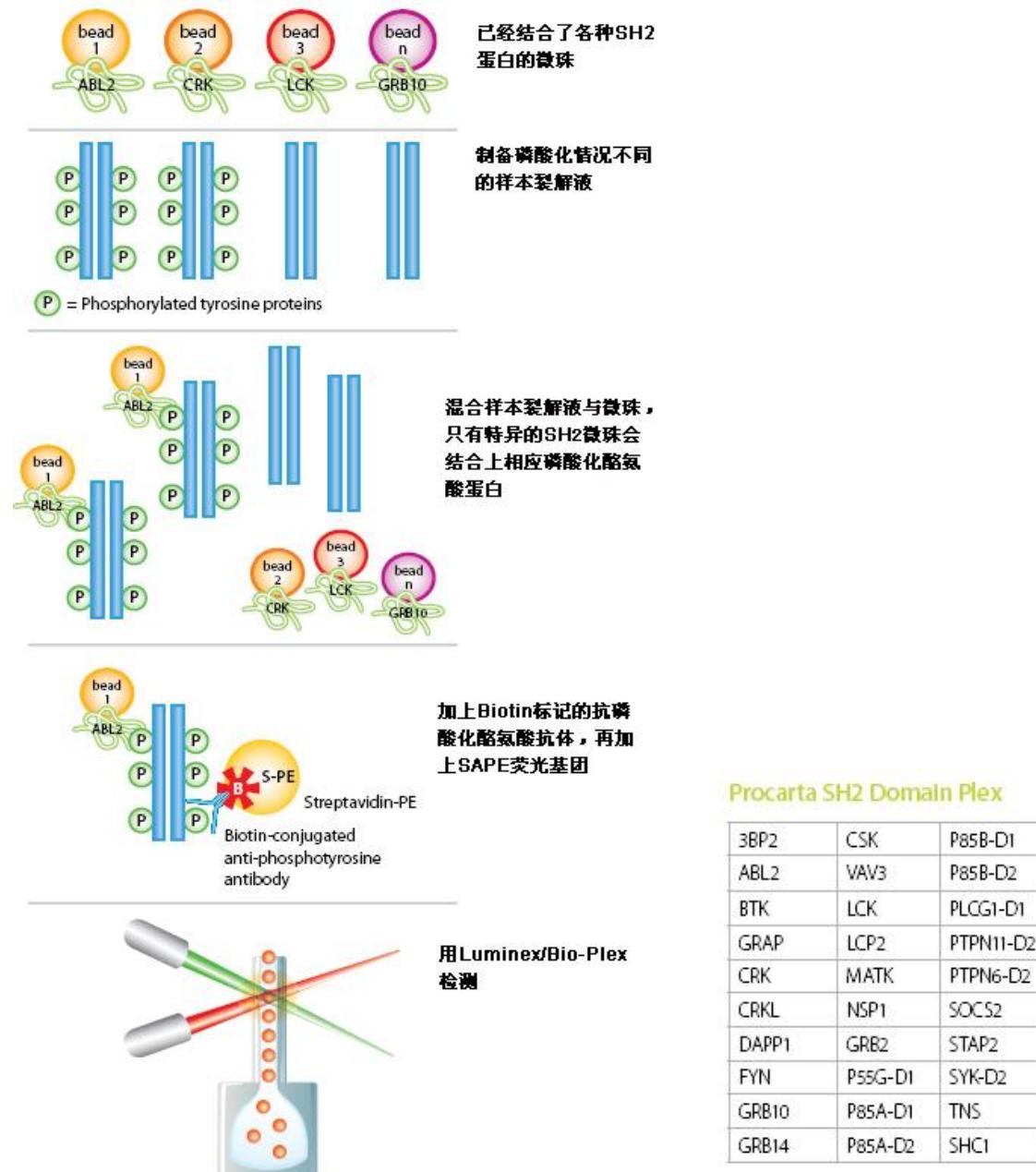
2、蛋白磷酸化检测服务

产品推荐：Procarta SH2 Domain Plex Kit

服务内容：

SH2 蛋白可特异结合磷酸化后的受体，因此可以利用不同的 SH2 蛋白来了解受体磷酸化的变化。该检测服务可以自由选择不同的 SH2 蛋白对样本的磷酸化情况做定量定性分析。

实验流程及检测因子：



服务内容：其他用于基础研究的液相悬浮芯片技术服务

细胞信号传导分析	36-40 phosphorylation protein
细胞因子、趋化因子、生长因子分析	Cytokine/Chemokine/Growth factor Cytochrome p450
粘连因子	Human Adhesion panel
金属基质酶	Human MMP panel
转录因子研究	Transcription factor panel
启动子甲基化	Promoter Methylation Array
抗体分型	IgG、IgM 分型 (Biorad)

服务内容：肿瘤标志物及疾病检测服务

1. 肿瘤标志物及相关因子检测

肿瘤标志物分析	Multiple Tumor marker assay kit 4plex cancer antigen Panomics 16plex human cancer biomarker
血管生成	Angiogenesis 9-plex

2. 心血管疾病、肥胖、糖尿病检测

肥胖疾病的检测	Diabetes / obesity-related hormone test 9-12 adipose factor test
代谢综合征检测	Metabolic Disease
激素的检测	Diabetes / obesity-related hormone test Endocrine Gut Hormone
心脏病标志物分析	Cardiovascular Disease
其它常见病	HLA Testing(One Lambda) ANA Test System 自身免疫病 QUANTA PLEX ENA PROFILE Immunoassay Kit 自身免疫性疾病抗可提取性核抗原 CF test 囊性纤维病变 Allergy Testing 过敏原 FIDIS celiac disease 腹腔疾病 Alzheimer's Disease Panel 阿尔茨海默氏症

3. 病原体和传染病

病原体基因检测	Pneumococcal Immunity 肺炎球菌 ResPlex I Panel 呼吸道感染疾病 Infectious Disease 54 种常见传染病同时检测
感染性疾病检测等	HPV Genotyping kit HPV 分型 EBV IgG System EB 病毒检测

4 其他

器官移植配型	Donor Specific Antibody Testing
新生儿筛查	Newborn Screening test
产前检查	Torch test system
疫苗接种效果检测	Vaccine testing

神经生物学	Human neurogenerative Disease panel
皮肤、骨代谢	Human Skin Bone Panel
药物毒理毒性研究	Toxicity
凋亡因子	Human Sepsis/Apoptosis
神经生物学	Human neurogenerative Disease panel

5. SBC 开发试剂盒

胃炎	Gastritis 3plex
冠心病	Coronary heart disease (CHD) 3plex
肝纤维化	Hepatic fibrosis 5plex
糖尿病分型	Classification of diabetes mellitus 4plex
过敏原	Allergen 8plex
唐氏综合症	Down syndrome 4plex

项目介绍:

- ✓ 国家科技部 863 项目——课题编号：(2006AA02A252)《抗体功能化制备新技术》
- ✓ 上海市科委启明星项目——(07QB14026)《糖尿病诊断液态芯片关键技术的攻关研究》
- ✓ 为大型药企 Lxxx，交大附属医院、复旦肿瘤医院等国内多家大型临床单位及科研单位进行服务。

(三) 标志物功能筛选及验证

siRNA筛选及细胞模型

该技术平台建立于2006-2007年，拥有Tecan高通量液体分装设备，高灵敏度酶标仪，P2级细胞培养实验室。可以进行100-7000个siRNA基因文库的筛选工作，以及各类细胞模型的构建及表型检测工作。

平台组成

- ✓ 通量化及多基因siRNA文库筛选——10~7000基因同时Knock down筛选基因的功能与表型。
平台主要运用siRNA技术实现规模化基因沉默，可以进行Thermo, Qiagen等公司的商业大型siRNA文库或功能分类文库，也可以进行客户指定文库的筛选实验。
 1. siRNA 高通量干扰技术进行肿瘤新基因功能研究
 2. siRNA 方法进行肿瘤功能基因的信号转导通路研究
 3. 基于 siRNA 技术规模化筛选新的肿瘤药物靶点

4. 基于 siRNA 技术规模化筛选肿瘤药物的增敏基因或耐受基因

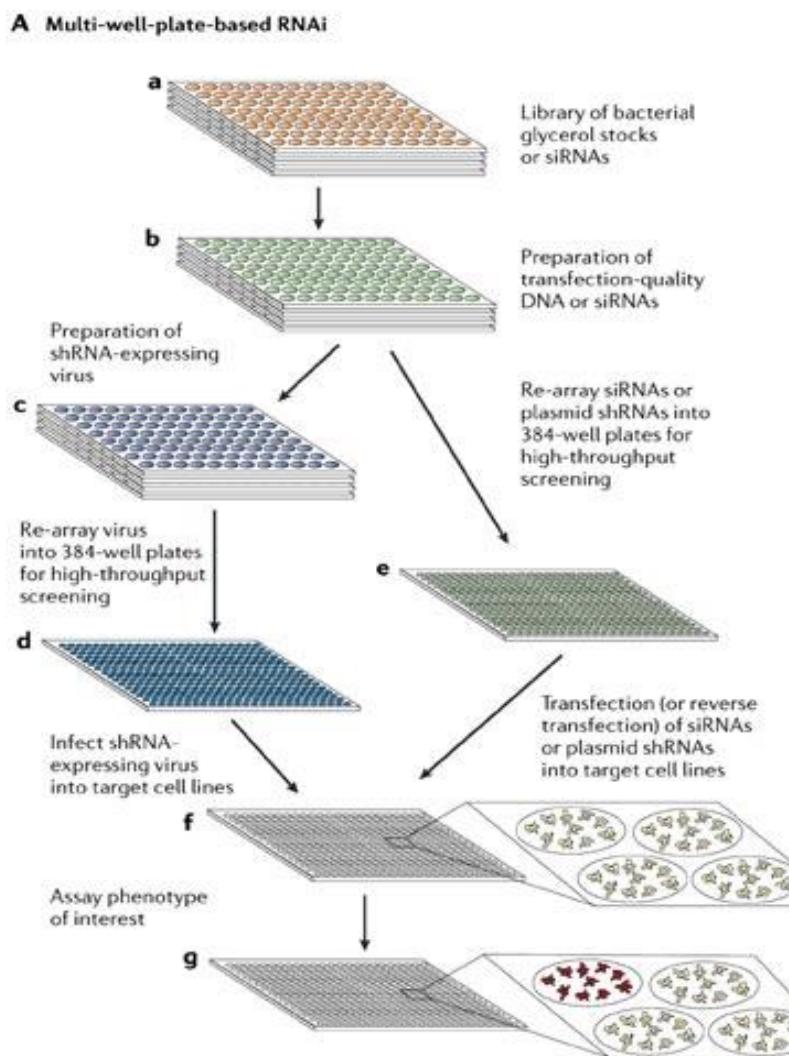
5. 筛选和开发新的 siRNA 肿瘤基因治疗药物

✓ 细胞模型构建及表型检测——特定的单个基因（包括miRNA）高表达或低表达细胞模型构建，以及后续的药物活性；细胞增殖，分化，迁移，厌氧培养，细胞周期等表型检测

✓ 细胞库资源——各类细胞株：人、鼠、肿瘤、血管内皮、原代等300多种细胞株

siRNA筛选流程

可以将高通量测序及高通量表观遗传学在全基因组得到的差异化的具有潜在功能标志物基因群，进行规模化的siRNA的knock down的功能筛选研究：



参考文献：

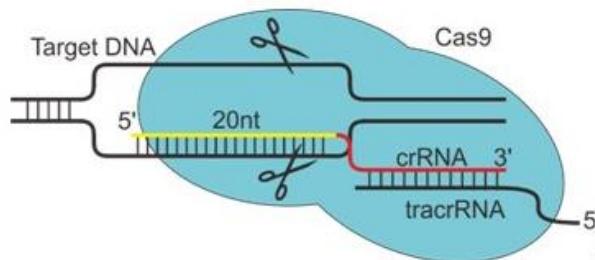
徐晓晶, 殷弘为 肝癌中高通量RNA干扰技术及其应用[J];中华肿瘤防治杂志2010 -12-14

项目介绍:

- ✓ 平台已完成上海市科委创新行动计划基础研究重点项目——(08JC1416500)《siRNA HTS数据分析方法在靶位点筛选中的应用》
- ✓ 已与 LXXXX, AXXXX 等大型药物研发机构进行服务

(四) 基因组编辑技术-CRISPR/Cas9技术

CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats) 是最新出现的一种由 RNA 指导的 Cas9 核酸酶对靶向基因进行编辑的技术。它基于细菌和古细菌为应对病毒和质粒不断攻击而演化来的一种获得性免疫防御机制。在这一系统中, crRNA (CRISPR-derived RNA) 通过碱基配对与 tracrRNA (trans-activating RNA) 结合形成双链 RNA, 此 tracrRNA/crRNA 二元复合体指导 Cas9 蛋白在 crRNA 引导序列靶定位点切断双链 DNA (见下图)。在当前进行的基因组编辑过程中, tracrRNA 和 crRNA 通常融合成为 1 条 RNA (sgRNA) 表达, 同样可以起到靶向剪切的作用。CRISPR/Cas9 技术对基因组的编辑效率高。需要对某一个靶位点编辑的时候, 只需要表达相应的 sgRNA 即可, 不需要对 Cas9 核酸酶进行改造。它可对任何物种的基因组进行高效率的定向编辑。为了克服脱靶效应, 本公司采用了一个突变的 Cas9 酶 Cas9n 对 DNA 单链进行切割。



上海伯豪已完成:

- ✓ 胃癌、肺癌、HeLa、293T、肝癌等细胞的建系;
- ✓ 200 多个基因的 sgRNA 构建;
- ✓ 20 多种肿瘤细胞系 CRISPR/Cas9 测试;
- ✓ 15 个 Gene 的敲除和 2 个 miRNA 的删除;
- ✓ 正在试验 293T 细胞 SNP、胰腺癌细胞定点插入等。

(五) 基因克隆蛋白表达纯化平台

蛋白表达平台主要负责表达载体构建, 原核/真核表达蛋白, 蛋白的纯化等, 2005年在国家科技部863项目的支持下, 瞄准重大疾病, 把握国际人类基因组发展趋势, 利用我国丰富的疾病遗传资源, 对我国发生率高、危害性大的重要疾病进行了研究, 获得了一批与疾病的发生和发展有关的基因及其蛋白质, 建立了规模化蛋白表达、大肠杆菌、酵母、哺乳动物细胞等多种表达系统, 以及蛋白纯化平台, 同时还建立了特定功能细胞色素p450酶的研究平台。

平台组成

特殊功能基因研究平台：p450表达平台。

基因克隆：从头克隆及亚克隆

——自2005年至今完成肿瘤与疾病相关基因、产品开发、活性功能基因的克隆约500多种

蛋白表达：

——大肠杆菌、酵母、哺乳动物细胞等表达系统实现300多种蛋白的表达

蛋白纯化：

——建立His6及GST等标签纯化与天然蛋白纯化技术，制备了超过100个纯化的原核、真核表达的重组蛋白

细胞色素p450酶平台：

——细胞色素p450酶在植物生理、药物天然合成、药物代谢、个体化用药等方面有广泛的用途，已建立具有活性的细胞色素p450酶的体外表达平台

项目介绍：

✓ 参加多项973、863项目：

国家科技部863项目——课题编号：（2002AA2Z2002）《基因、蛋白和组织芯片的研究与开发》等

✓ 为国外IXXX、国内中国科学院植物生理所，复旦大学等单位进行服务，其中协助复旦表达的蛋白的部分工作，文章发表在国际著名期刊《Science》上。

上海伯豪生物技术有限公司生物标志物验证及分子检测平台技术服务内容

基于 B-DNA 技术核酸标志物相关验证与检测服务

Panomics 多基因（多 mRNA 检测、多 DNA 的 copynumber 检测、多 microRNA 检测、多 lncRNA 检测等）

Quntigene 单基因检测（mRNA、DNA 的 copynumber、microRNA、lncRNA 等）

ViewRNA 原位荧光检测等（mRNA、DNA 的 copynumber、microRNA、lncRNA 等）

核酸检测服务其他技术

Qiagen 的 PCR-array 技术用于 Pathway 研究

TDLA 低密度芯片检测

Sanger 法普通测序—ABI3730

Pyrosequencing 测序（DNA 变异检测和甲基化检测）

荧光定量 PCR 检测（mRNA、DNA 的 copynumber、microRNA、lncRNA 等）

MassArray 检测（SNP 分型）

DNA FISH （检测 DNA 拷贝数及融合基因等）

Bio-Plex/Luminex 等基于免疫反应的蛋白因子相关验证与检测服务

Luminex 多细胞因子检测；多磷酸化蛋白检测

ELISA 单蛋白因子检测

目标蛋白因子免疫荧光检测

目标蛋白因子 Western Blotting 检测

siRNA 高通量服务

siRNA 高通量干扰技术进行肿瘤新基因功能研究
siRNA 方法进行肿瘤功能基因的信号转导通路研究
基于 siRNA 技术规模化筛选新的肿瘤药物靶点
基于 siRNA 技术规模化筛选肿瘤药物的药物作用靶点、增敏基因或耐受基因
筛选和开发新的 siRNA 肿瘤基因治疗药物
siRNA、转录编码 shRNA 的 DNAs 的制备、转录编码 shRNA 的质粒载体构建
siRNA 靶位的设计和实验选择分析

细胞模型的建立及相关服务

细胞模型的筛选和构建：
可培养多种细胞并能应用表达载体或慢病毒系统在细胞内稳定或瞬时转染特定的药物靶点受体基因、
microRNA 高表达细胞模型
可培养多种细胞并能应用表达载体或慢病毒系统在细胞内稳定或瞬时转染特定的药物靶点受体基因、
microRNA 低表达细胞模型
基因组编辑技术服务(CRISPR/Cas9 技术服务)

细胞培养和与药物作用下的细胞检测实验（Cell-based assay）服务

常规细胞操作：细胞复苏，实验细胞培养，细胞扩增
细胞行为与形态学实验：细胞增殖实验，细胞迁移实验，细胞形态学观测，细胞衰老实验，凋亡测定，细胞分化相关的实验，细胞侵袭转移相关的实验
细胞的药物生长抑制 IC50 实验

蛋白及抗体制备相关服务

工程细胞/菌株的筛选与培养：构建表达重要工程蛋白药物（如疫苗、抗体等）和酶类工程菌株，包括构建、筛选、表达、工程菌株细胞的培养和中试发酵
基因从头克隆、亚克隆、载体构建、突变基因构建
细胞和工程菌蛋白/抗体提取纯化
细胞色素 p450 酶活性蛋白表达

生物标志物验证及分子检测平台技术服务应用文章精选 (2010-2014)

- [1] Wang HP, Wang YY, Pan J, Cen R, Cai YK. Evaluation of specific fecal protein biochips for the diagnosis of colorectal cancer. World J Gastroenterol. 2014 Feb 7;20(5):1332-9. (IF 2.547) 使用服务种类：蛋白芯片
- [2] Wang L, Zhao C, Peng Q, Shi J, Gu G. Expression levels of CD28, CTLA-4, PD-1 and Tim-3 as novel indicators of T-cell immune function in patients with chronic hepatitis B virus infection. Biomed Rep. 2014 Mar;2(2):270-274. 使用服务种类：免疫检测
- [3] Chen J, Ruan X, Wang S, Zhang B, Liu B, Sun Z, Liu Q. Antitumor effect and biological pathways of a recombinant adeno-associated virus as a human renal cell carcinoma suppressor. Tumour Biol. 2014 Nov;35(11):10993-1003. (IF 2.840) 使用服务种类：一代测序
- [4] Dai Z, Zhou SL, Zhou ZJ, Bai DS, Xu XY, Fu XT, Chen Q, Zhao YM, Zhu K, Yu L, Yang GH, Wang Z,

-
- Wu WZ, Zhou J, Fan J. Capn4 contributes to tumour growth and metastasis of hepatocellular carcinoma by activation of the FAK-Src signaling pathways. *J Pathol.* 2014 Nov;234(3):316-28. (IF7.330) 使用服务种类: 蛋白芯片
- [5] Wei T, Xiong F, Wang S, et al. Flavonoid ingredients of Ginkgo biloba leaf extract regulate lipid metabolism through Sp1-mediated carnitine palmitoyltransferase 1A up-regulation[J]. *Journal of biomedical science*, 2014, 21(1): 87. (IF2.740) 使用服务种类: 一代测序
- [6] Liu Y, Huang X, Bi R, Yang W, Shao Z. Similar prognoses for invasive micropapillary breast carcinoma and pure invasive ductal carcinoma: a retrospectively matched cohort study in China. *PLoS One* 2014 Sep 4:9(9):e160564 (IF3.534) 使用服务种类: Pannomics
- [7] Ouyang LL, Du DH, Yu SY, Li CY, Zhang CW, Gao HJ, Zhou ZG. Expressed sequence tags analysis revealing the taxonomic position and fatty acid biosynthesis in an oleaginous green microalga, *Myrmecia incisa* Reisigl (Trebouxiophyceae, Chlorophyta). *Chinese Science Bulletin*. 2012 使用服务种类: 一代测序
- [8] Zhou SL, Dai Z, Zhou ZJ, Wang XY, Yang GH, Wang Z, Huang XW, Fan J, Zhou J. Overexpression of CXCL5 mediates neutrophil infiltration and indicates poor prognosis for hepatocellular carcinoma. *Hepatology*. 2012 Dec;56(6):2242-54. (IF11.9) 使用服务种类: 蛋白芯片
- [9] Hu Y, Pioli PD, Siegel E, Zhang Q, Nelson J, Chaturbedi A, Mathews MS, Ro DI, Alkafeef S, Hsu N, Hamamura M, Yu L, Hess KR, Tromberg BJ, Linskey ME, Zhou YH. EFEMP1 suppresses malignant glioma growth and exerts its action within the tumor extracellular compartment. *Mol Cancer*. 2011 Sep 28;10:123. (IF5.400) 使用服务种类: 蛋白芯片

第六章 生物信息分析 服务

一、服务总览

生物芯片分析服务			
基因组研究层次	转录组研究层次	表观遗传研究层次	SBC 特色分析服务
SNP 芯片分析服务 CGH 芯片分析服务	表达谱芯片分析服务 miRNA 芯片分析服务 lncRNA 芯片分析服务	DNA 甲基化芯片分析服务 ChIP-on-Chip 分析服务	各类芯片联合分析服务 SBC Analysis System 分析服务 GEO 上传服务
新一代测序分析服务			
基因组研究层次	转录组研究层次	表观遗传研究层次	SBC 特色分析服务
<i>de novo</i> 测序分析服务 DNA 重测序分析服务 外显子组测序分析服务	RNA 测序分析服务 Small RNA 测序分析服务 lncRNA 测序分析服务	全基因组甲基化测序分析服务 MeDIP-SEQ 分析服务 ChIP-SEQ 分析服务 RIP-Seq 分析服务	16S r DNA 测序分析服务 宏基因组测序分析服务 宏转录组测序分析服务 PacBio RSII 测序分析服务

二、生物芯片分析服务

1. 基因组研究层次分析服务

■ SNP 芯片分析服务

基本数据分析	原始数据预处理、SNP 基因型、CNV 信息、CNV 在染色体分布
高级数据分析	连锁分析、关联分析、单倍型分析

■ CGH 芯片分析服务

基本数据分析	原始数据 z-score 归一化、CNV、CNV 在染色体分布、log2ratio 分布图
高级数据分析	探针改变百分比、变异区段作图

2. 转录组研究层次分析服务

■ 表达谱芯片分析服务—3'IVT 芯片

基本数据分析	原始数据归一化处理、Box Plot 图、PCA 图、聚类图、基因水平表达差异筛选、散点图、火山图
高级数据分析	GO 富集分析、KEGG 富集分析、基因共表达网络分析

■ 表达谱芯片分析服务—WT 芯片

基本数据分析	原始数据归一化处理、Box Plot 图、PCA 图、聚类图、基因水平表达差异筛选、外显子水平表达差异筛选、散点图、火山图、
--------	--

	差异基因的染色体分布、lncRNA 靶基因预测、GO 富集分析、KEGG 富集分析
高级数据分析	可变剪切分析、lncRNA 与 mRNA 共表达分析

- miRNA 芯片分析服务

基本数据分析	原始数据归一化处理、Box Plot 图、聚类图、PCA 图、样品相关性系数图、miRNA 表达差异筛选、散点图、火山图、聚类热图、miRNA 靶基因预测
高级数据分析	差异 mRNA 靶基因的 GO 富集分析、差异 miRNA 靶基因的 KEGG 富集分析、miRNA 与 mRNA 共表达分析

- lncRNA 芯片分析服务

基本数据分析	原始数据归一化处理、差异 lncRNA 的筛选、差异 lncRNA 分类、差异 lncRNA 靶基因的预测
高级数据分析	差异 lncRNA 与靶基因共表达网络、差异 lncRNA 与差异 mRNA 的共表达分析、差异 lncRNA 靶基因的 GO analysis、差异 lncRNA 靶基因的 pathway analysis、差异 lncRNA 的转录因子的预测、lncRNA 分子诊断分析服务

3. 表观遗传研究层次分析服务

- DNA 甲基化芯片分析服务

基本数据分析	差异甲基化的筛选、非监督聚类、差异甲基化基因的功能分析、差异甲基化基因的 pathway 分析、转录因子预测分析、甲基化的染色体分布图
高级数据分析	特定基因在不同基因区段的甲基化水平变化分析、甲基化与表达谱的关联分析、网络图的构建、差异甲基化基因中邻近启动子区域 motif 的识别

- ChIP-on-Chip 分析服务

4. SBC 特色分析服务

- 各类芯片联合分析服务
- SBC Analysis System 分析服务

基本数据分析	差异倍数计算(Fold Change), t 检验 (t-Test), SAM 检验 (Significant Analysis of Microarray, SAM), 一维方差分析 (one-way ANOVA), 二维方差分析(two-way ANOVA), 相关性分析 (Correlation analysis) , 聚类分析 (Unsupervised)
--------	---

	hierarchical clustering), 主成份分析(Principal Component Analysis, PCA), GO 富集分析 (GO enrichment analysis), pathway 富集分析(Pathway enrichment analysis), microRNA 检索, 蛋白质相互作用检索(protein interaction)
--	--

- GEO 上传服务

三、新一代测序分析服务

1. 基因组研究层次分析服务

- *de novo* 测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、 <i>de novo</i> 序列拼接、基因注释、SSR 分析、功能注释 (GO, Pathway)
高级数据分析	致病岛预测, 原噬菌体预测, 转座子预测, 共线性分析, 图形展示等

- DNA 重测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、全基因组比对、SNV 检测及注释、InDel 检测及注释、CNV 检测及注释
高级数据分析	SV 精细分析、目标区段 GO/KEGG 富集分析、全基因组关联分析

- 外显子组测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、全基因组比对、单样本 Small InDel 分析、单样本 SNV 分析、样本间 Small InDel 差异分析、样本间 SNV 差异分析
高级数据分析	SV 精细分析、目标区段 GO/KEGG 富集分析、全基因组关联分析、个性化数据库注释、高频变异区段挖掘、CNV 检测及注释

2. 转录组研究层次分析服务

- RNA 测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、基因组比对、全基因组覆盖可视化分析、饱和度分析、基因表达分析、差异基因富集分析
高级数据分析	时间趋势分析(三个时间点以上)、融合基因分析 (数据量要求至少 10G, 推荐肿瘤样本)、新基因分析、转录本表达分析、差异剪接分析 (模式生物)、lncRNA 分析、lncRNA 与 mRNA 共表达网络分析、蛋白互作网络分析 (模式生物)、cSNP 分析

- Small RNA 测序分析服务

基本数据分析	原始数据处理、数据过滤及长度分布统计、比对注释，鉴定已知 miRNA、miRNA 表达水平分析、样品间表达差异分析、Small RNA 分类注释、饱和度分析
高级数据分析	Novel miRNA 预测、miRNA 作用靶基因预测、靶基因功能分析与 KEGG 通路分析、miRNA-mRNA-network 构建

- lncRNA 测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、基因组比对、全基因组覆盖可视化分析、饱和度分析、lncRNA 表达分析、差异表达及富集分析
高级数据分析	差异 lncRNA 分类、Novel lncRNA 预测、lncRNA 靶基因预测、lncRNA 保守性、二级结构分析、lncRNA 与 mRNA 共表达网络分析

3. 表观遗传研究层次分析服务

- 全基因组甲基化测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、BS 转化基因组比对及统计、转化效率评估
高级数据分析	CpG 岛区域、启动子区域及整个基因组的覆盖及不同深度统计分析，CpG 岛、CpG shelf、CpG shore 以及基因的不同区域的甲基化程度分布图、CpG/CHG/CHH 位点覆盖率、测序深度、甲基化比例和染色体分布、差异甲基化位点和区段筛选、差异甲基化位点或区段的 CpG 岛注释和基因注释、全基因组甲基化图谱、差异区段全基因组分布环图图、转录组关联分析

- MeDIP-SEQ 分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、基因组比对、Peak Detect、基因功能关联注释、样本间比较分析
高级数据分析	CG 位点测序深度统计，差异甲基化区段、差异 CpG 岛注释与基因注释、差异区段染色体分布图、全基因组甲基化图谱、差异区段测序深度图谱、表达谱关联分析

- ChIP-SEQ 分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、基因组比对、Peak Detect、Motif Detect、基因组 mapping、样本间分析
--------	--

高级数据分析	表达谱关联分析、转录因子模式预测
--------	------------------

■ RIP-Seq 分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、基因组比对、基因表达定量、差基因分析、功能富集分析、Peak Detect 和注释、Motif 分析
高级数据分析	LncRNA 关联分析、mRNA 关联分析

4. SBC 特色分析服务

■ 16S r DNA 测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、Box Plot 图评估数据质量、序列质量统计、序列长度及分布统计、有效序列统计、OTU 聚类及丰度分析、物种分类及样本间差异比较、群落结构分析
高级数据分析	OTU 及物种丰度聚类热图分析、组间相似及差异性分析、环境因子 RDA/CCA 分析、进化分析

■ 宏基因组测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、基因组组装、物种组成和丰度分析、基因预测、基因功能注释
高级数据分析	多样品比较分析

■ 宏转录组测序分析服务

基本数据分析	原始数据处理、数据过滤及统计、测序序列质量评估、序列拼接、NR 库比对、基因组分分析、基因定量、物种分类及物种基因表达丰度分析、基因功能注释、基因差异表达分析、差异表达基因 KEGG 、GO 富集分析
高级数据分析	多样品相似度比较、物种聚类分析、PCA 分析（5 个以上样本）、PCoA 分析（5 个以上样本）、RDA/CCA 分析（10 个样本以上）、系统发生树构建

■ PacBio RSII 测序分析服务(Genome Assembly)

基本数据分析	测序 Reads 过滤及统计、pcabio 数据质量校正、 <i>de novo</i> 序列拼接、基因注释、SSR 分析、功能注释（GO, Pathway）
高级数据分析	致病岛预测，原噬菌体预测，转座子预测，共线性分析，图形展示等

第七章 试剂耗材服务

(一) RNAsafety 组织 RNA 保存液

RNAsafetyTM 组织 RNA 保存液是一种无毒的可直接使用的样品保存液，其原理是抑制 RNase 活性，保护新鲜组织样本里的 RNA 免受降解。本产品目前已被中科院上海生命科学研究院、国家基因组南方研究中心、上海交通大学医学院、中国医科大学、哈尔滨医科大学、汕头大学医学院、唐都医院、江南大学等单位广泛使用。

产品优势

- ※ 迅速稳定组织样品中的 RNA，保护 RNA 免受降解。
- ※ 室温下操作，安全简便。
- ※ 无需使用液氮或干冰转送和运输。
- ※ 在-20℃/-70℃可用于长期保存组织。
- ※ 与大部分 RNA 提取方法兼容，如 Trizol 法。

RNAsafety 溶液的使用方法

将切割后的组织（尺寸<0.5cm）没入 5–10 倍体积的室温 RNAsafety 溶液中（如 0.5g 的样品需要 2.0-5.0mL 的 RNAsafety 溶液）。溶液能迅速渗入细胞，稳定样品中的 RNA。随后样品可置于室温（25℃）保存 1 周，4℃一个月，或者在-20℃长期保存。进行 RNA 分离时，只需将组织从 RNAsafety 溶液中取出，就如刚采集后一样操作。RNAsafety 溶液能够成功地维持 RNA 的完整性，甚至是在那些已知 RNA 酶含量较高的组织，从这些组织中通常难以提取完整 RNA（例如脾脏）。

RNAsafety 溶液的多种用途

RNAsafety 溶液已被测试用于多种哺乳动物组织，如脂肪、肾脏、膀胱组织、小肠组织、肌肉、软骨、肾脏、心脏、睾丸、皮肤、脾脏、肺脏等，同时进行了细菌、真菌 RNA 保存测试。

RNAsafety 溶液适用于：

- 1.在不能立即进行 RNA 分离的地点收集样品。
- 2.在富含 RNA 酶的组织中保护 RNA 的完整性，例如脾脏或者胰脏在不同时间点收集样品，而不用马上处理来自每个时间点的样本。
- 3.处理后的样品可用于将来的显微分离。
- 4.填充动物腔，或者将器官浸没在 RNAsafety 溶液中，在长时间，冗长的解剖中稳定 RNA。
- 5.组织材料可以从 RNAsafety 溶液中取出，切割使用，重新放回继续保存，而不影响组织 RNA 的质量。
- 6.RNAsafety 溶液不影响组织结构，可在室温下切割和称量组织块。
- 7.如果需要过夜运输，可在常温条件下运输样品。

订购信息

货号	产品名称	规格	价格
REA80001	RNAsafety TM 组织 RNA 保存液	100ml/瓶	688 元/瓶
REA80002	RNAsafety TM 组织 RNA 保存液	2ml/支， 50 支/包	1000 元/包

REA80003	RNAsefree™组织 RNA 保存液	2ml/支, 100 支/包	2000 元/包
REA80004	RNAsefree™组织 RNA 保存液	5ml/支, 50 支/包	1250 元/包
REA80005	RNAsefree™组织 RNA 保存液	5ml/支, 100 支/包	2500 元/包

温馨提示：该产品仅用于科学研究，不可用于临床诊断及药物治疗。



产品实物图

产品咨询邮箱：market@shbiochip.com

订购邮箱：biosales@shbiochip.com

(二) Agilent PLUF 系列试剂耗材产品

Agilent 2100 生物芯片分析系统配套耗材

Agilent 2200 TapeStation 全自动电泳系统配套耗材

Agilent 3100 等电聚焦分析分离仪等分析系统配套耗材

2013年1月，上海伯豪生物技术有限公司成为Agilent PLUF系列耗材全国总代理。

Agilent PLUF系列耗材可以用于Agilent 2100生物芯片分析系统、Agilent 2200 TapeStation全自动电泳系统、Agilent 3100等电聚焦分析分离仪等分析系统，完成DNA、RNA以及蛋白样品的快速而精确的分析和检测。

2013年起，上海伯豪将以更快速的物流、更优质的服务为广大用户提供Agilent PLUF系列耗材。



重点推荐 — Agilent 2100 生物芯片分析系统配套试剂耗材

上海伯豪现货供应产品列表

货号	产品名称	产品用途	试剂盒包装大小	目录价
5067-1504	DNA 1000 Kit	用于测定 25-1000bp DNA 片段的大小和含量，最小分辨率 5bp，定量范围 0.1ng/ul-50ng/μl。	25 张芯片，12 个样品/片，300 个样品/试剂盒。	¥6668
5067-1506	DNA 7500 Kit	用于测定 100-7500bp DNA 片段的大小和含量，100-1000bp 为 5% 分辨率，1000-7500bp 为 15% 分辨率，定量范围 0.1-50ng/μl。	25 张芯片，12 个样品/片，300 个样品/试剂盒。	¥6668
5067-4626	High Sensitivity DNA Kit	对 pg/ul 浓度范围内的极微量 DNA 片段进行准确的定量分析，分离片段分析范围 50 - 7000bp。	10 张芯片，11 个样品/片，110 个样品/试剂盒。	¥4961
5067-1511	RNA 6000 Nano Kit	用于总 RNA 和 mRNA 的分析和定量，定量范围为 25 - 500 ng/μL。	25 张芯片，12 个样品/片，300 个样品/试剂盒。	¥7098
5067-1513	RNA 6000 Pico Kit	用于总 RNA 和 mRNA 的分析和定量，定量范围为 50 pg/μL(总 RNA)、250 pg/μL(mRNA)。	25 张芯片，11 个样品/片，275 个样品/试剂盒。	¥7670

安捷伦 2100 芯片生物分析仪是目前最为成功的基于微流控的商业化平台，提供用于 DNA、RNA、蛋白质和细胞分析的解决方案，30 分钟内即可给出结果，提供自动化的、高品质的数字化数据。在这个多功能平台上，可进行 PCR/多重 PCR/RFLP 产物分析；RNA 样品质量控制，小 RNA 分子，如 miRNA, siRNA 分析；蛋白质质量控制，蛋白质表达/纯化分析；还可以进行综合的细胞分析。



其它 Agilent 2100 生物芯片分析系统配套试剂耗材列表:

货号	产品名称	产品用途	试剂盒包装大小
5067-1508	DNA 12000 kit	用于测定 100-12000bp DNA 片段的大小和含量, 100-1000bp 为 5% 分辨率, 1000-12000bp 为 10% 分辨率, 定量范围 0.1-50ng/ μ l。	25 张芯片, 12 个样品/片, 300 个样品/试剂盒。
5067-1548	Small RNA kit	对 small RNA 进行定量和质控。分离片段分析范围为 6-150nt, 灵敏度高达 50pg/ μ l。	25 张芯片, 11 个样品/片, 275 个样品/试剂盒。
5067-1515	Protein 80 kit	用于蛋白纯化和质控, 5-80kd, 分辨率 10% (起码 500 个道尔顿, 多达 8000)。	25 张芯片, 10 个样品/片, 250 个样品/试剂盒。
5067-1517	Protein 230 kit	用于蛋白纯化和质控, 14-230kd, 分辨率 10% (起码 1400 个道尔顿, 多达 23000)。	25 张芯片, 10 个样品/片, 250 个样品/试剂盒。
5067-1575	High Sensitivity Protein 250 kit	高灵敏度蛋白分析试剂盒, 10-250kd, 分辨率: 10% (最低 1000kd), 准确度 10%CV。	10 张芯片, 10 个样品/片, 100 个样品/试剂盒。
5067-1519	Cell kit	用于细胞分析, 单个细胞的荧光参数检测。	25 张芯片, 6 个样品/片, 150 个样品/试剂盒。

Agilent 2200 TapeStation 全自动电泳系统配套试剂耗材

Agilent 2200 TapeStation 全自动电泳系统能够带来可扩展的通量, 同时大幅缩短得到结果的时间, 是新一代测序 (NGS)、生物芯片分析和 qPCR 工作流程以及蛋白质纯化和抗体生产过程中对生物样品进行质量控制(QC)的理想解决方案。

Agilent 2200 TapeStation 系统使用一次性 Agilent ScreenTape 预制胶条, 可用于 DNA、RNA 和蛋白质应用。只需要将 ScreenTape 预制胶条装载到 2200 TapeStation 上, 然后使用移液器将样品加载到 16 联管或 96 孔微量滴定板上, 1 分钟便可得到检测结果。

Agilent 2200 TapeStation 全自动电泳系统配套试剂耗材列表:

货号	产品名称	货号	产品名称
5067-5367	R6K 检测预制胶	5067-5368	R6K 配套试剂
5067-5369	高灵敏度 R6K 预制胶	5067-5370	高灵敏 R6K 度配套试剂
5067-5371	P200 预制胶	5067-5372	P200 配套试剂
5067-5361	D1K DNA 检测预制胶	5067-5362	D1K DNA 配套试剂
5067-5363	高灵敏度 D1K 检测预制胶	5067-5364	高灵敏度 D1K 配套试剂
5067-5365	基因组 DNA 检测预制胶	5067-5366	基因组 DNA 检测配套试剂

Agilent 3100 等电聚焦分析分离仪配套试剂耗材

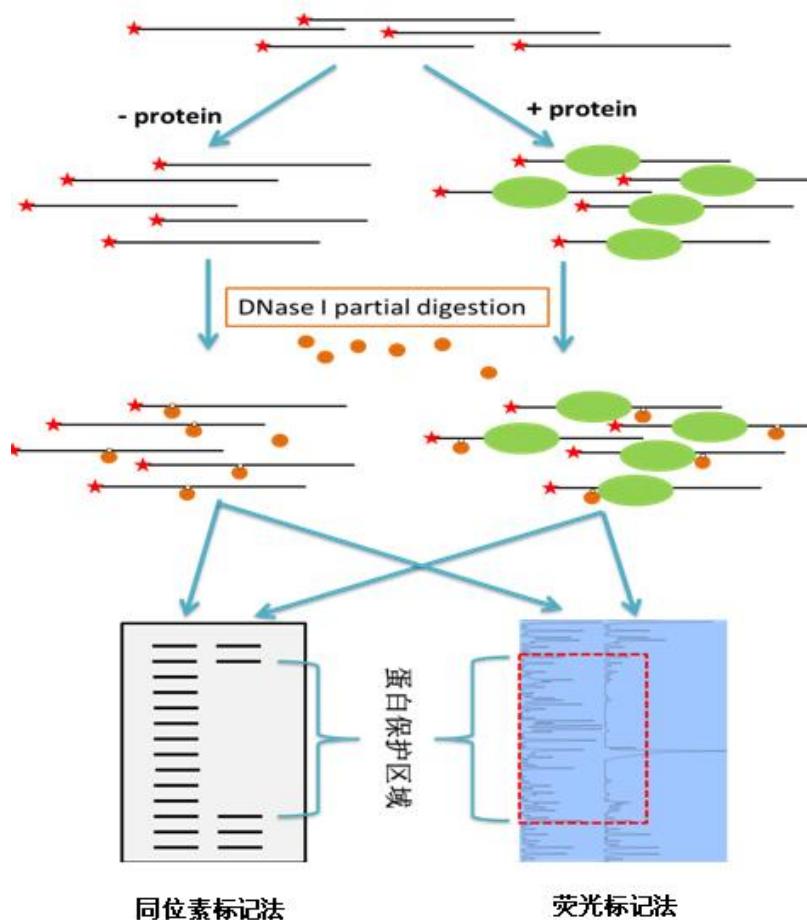
配套试剂耗材请致电（021-51370731）咨询！

（三）DNase I footprinting assay 服务

DNase I footprinting assay (DNase I 足迹分析实验) 可以精确鉴定 DNA 结合蛋白 (如转录调控蛋白) 在 DNA 分子上的结合位点。转录调控蛋白通过与启动子区域结合来调控靶标基因的转录。目前，凝胶阻滞实验 (EMSA) 和 DNase I 足迹分析实验是重要的一套体外研究工具：凝胶阻滞实验可以确定转录因子是否与 DNA 直接结合，而 DNase I 足迹分析实验则可以精确鉴定其结合的 DNA 序列，从而帮助我们研究基因转录调控的机制。经典的 DNase I 足迹分析实验需要使用同位素进行标记，并在反应完成后通过跑原始的 DNA 测序胶来检测；该方法不仅危险，而且费时、对操作者的技术要求高、成功率低。针对这种情况，上海伯豪目前推出利用“荧光标记”结合 ABI 测序仪的方法来进行 DNase I 足迹分析实验。

“荧光标记”的优势：

1. 相对于“同位素标记”，“荧光标记”更安全、无污染；
2. “荧光标记”，实验的速度快、重复性好、灵敏度高。

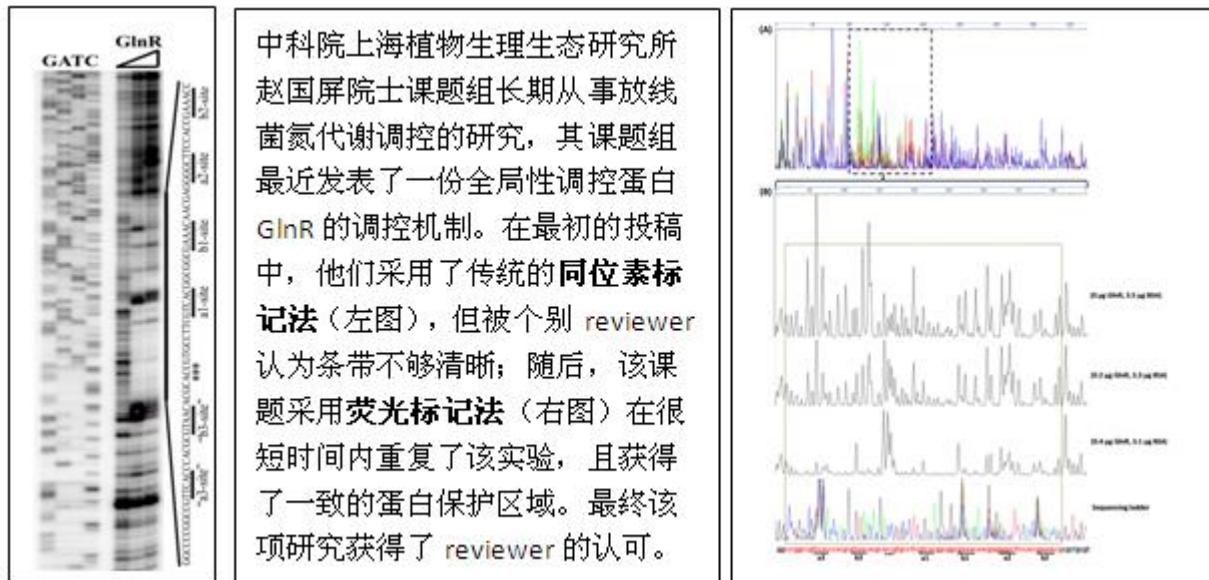


- 荧光标记法，由于可以进行 DNA 的精确定量，并可以在常规的实验台上面操作；同位素标记法只能通过放射性强度来估测，且操作时需要特别的保护装置，操作不便。因此，荧光标记法的可重复性要远远好于同位素操作。
- 原始的 PAGE 测序胶对操作者的技术要求很高，且条带的清晰度受到多种实验因素的影响；ABI 测序仪的操作体系非常成熟，灵敏度高，重复性好。因此，荧光标记法更灵敏度，获得的保护区域更清晰，图片的质量更好。

实验及数据分析流程：



实例分析：



原文：Characterization of a New GlnR Binding Box in the Promoter of amtB in Streptomyces coelicolor Inferred a PhoP/GlnR Competitive Binding Mechanism for Transcriptional Regulation of amtB. Journal of Bacteriology. 2012,194(19):5237-44.

服务说明：

- 该服务适用于：已经通过 EMSA 等实验证明转录调控蛋白可以结合 DNA，但是需要通过 DNase I 足迹分析实验来进一步精确鉴定蛋白结合的 DNA 序列，以便研究其转录调控的机制。客户需要提供纯化的蛋白，DNA 样品及引物序列，同时需要提供 EMSA 结果图及实验参数以便于接下来的 DNase I 足迹分析实验。
- 公司同时提供蛋白纯化或者 EMSA 实验验证，但需另计费用。
- 本服务同样适用于 mRNA 转录起始位点的鉴定，原理与 DNase I 足迹分析实验类似，即在 Primer extension 实验或者 S1 mapping 实验中用荧光标记法替代同位素标记法。客户需要提供高质量的 RNA 样品及目的基因序列。

近年来使用该技术发表的部分论文：

- 1 Structural and functional analysis of the transcriptional regulator Rv3066 of Mycobacterium tuberculosis. Nucleic Acids Res. 2012,40(18):9340-55.
- 2 LVIS553 Transcriptional Regulator Specifically Recognizes Novobiocin as an Effector Molecule. J Biol Chem. 2010,285(22):16921-30
- 3 Detailed analysis of the DNA recognition motifs of the Xanthomonas type III effectors AvrBs3 and AvrBs3Δ rep16. Plant J. 2009,59 (6): 859 – 871.
- 4 A key developmental regulator controls the synthesis of the antibiotic erythromycin in Saccharopolyspora erythraea. Proc Natl Acad Sci U S A. 2008,105(32):11346-51.
- 5 A Pseudomonas aeruginosa transcription factor that senses fatty acid structure. Mol Microbiol. 2007,66(3):622-32.

-
- 6 Anaplasma phagocytophilum p44 mRNA expression is differentially regulated in mammalian and tick host cells: involvement of the DNA binding protein ApxR. *J Bacteriol.* 2007;189 (23): 8651-9.

2012–2014 年度经典论文

1. Zhang L, Chen LH, Wan H, Yang R, Wang Z, Feng J, Yang S, Jones S, Wang S, Zhou W, Zhu H, Killela PJ, Zhang J, Wu Z, Li G, Hao S, Wang Y, Webb JB, Friedman HS, Friedman AH, McLendon RE, He Y, Reitman ZJ, Bigner DD, Yan H. Exome sequencing identifies somatic gain-of-function PPM1D mutations in brainstem gliomas. *Nat Genet.* 2014 Jul;46(7):726-30. doi: 10.1038/ng.2995. Epub 2014 Jun 1. (影响因子 35.209)

使用服务种类: Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 服务

2. Pengyu Huang, Ludi Zhang, Yimeng Gao, Zhiying He, Dan Yao, Zhitao Wu, Jin Cen, Xiaotao Chen, Changcheng Liu, Yiping Hu, Dongmei Lai, Zhenlei Hu, Li Chen, Ying Zhang, Xin Cheng, Xiaojun Ma, Guoyu Pan, Xin Wang, Lijian Hui. Direct Reprogramming of Human Fibroblasts to Functional and Expandable Hepatocytes. *Cell Stem Cell*, 28 February 2014.(影响因子 25.315)

使用服务种类: Agilent 人表达谱芯片服务

3. Schaid DJ, Spraggs CF, McDonnell SK, Parham LR, Cox CJ, Ejlertsen B, M. Finkelstein D, Rappold E, Curran J, R. Cardon L, E. Goss P. Prospective Validation of HLA-DRB1*07:01 Allele Carriage As a Predictive Risk Factor for Lapatinib-Induced Liver Injury. *Journal of clinical oncology.* 2014. 10.1200/JCO.2013.52.9867. (影响因子 18.038)

使用服务种类: Illumina Human1M-Duo beadchip array 及测序 (GWAS)

4. Zheng JJ, Li SJ, Zhang XD, Miao WY, Zhang D, Yao H, et al. Oxytocin mediates early experience-dependent cross-modal plasticity in the sensory cortices. *Nature neuroscience.* 2014.(影响因子 15.251)

使用服务种类: Oxytocin mRNA microarray 服务

5. Xu C, Ju X, Song D, Huang F, Tang D, Zou Z, et al. An association analysis between psychophysical characteristics and genome-wide gene expression changes in human adaptation to the extreme climate at the Antarctic Dome Argus. *Molecular psychiatry.* 2014.(影响因子 15.147)

使用服务种类: Agilent Human Gene Expression 4 × 44K v2 Microarray Kit

6. Tian-Lin Cheng,,Zhizhi Wang, Qiuming Liao, Ying Zhu, Wen-Hao Zhou, Wenqing Xu, and Zilong Qiu. MeCP2 Suppresses Nuclear MicroRNA Processing and Dendritic Growth by Regulating the DGCR8/Drosha Complex. *Developmental Cell.* 2014. (影响因子 12.861)

使用服务种类: miRNA 测序服务

7. Wang Y, Lei R, Zhuang X, Zhang N, Pan H, Li G, et al. DLC1-dependent parathyroid

hormone-like hormone inhibition suppresses breast cancer bone metastasis. The Journal of clinical investigation. 2014. (影响因子 12.812)

使用服务类型: Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 服务

8. Jiang JH, Liu YF, Ke AW, Gu FM, Yu Y, Dai Z, et al. Clinical significance of the ubiquitin ligase E3C in hepatocellular carcinoma revealed by exome sequencing. Hepatology (Baltimore, Md).

2014.(影响因子 12.003)

使用服务类型: 外显子组测序服务

9. Li L, Tang J, Zhang BH, Yang W, Gao MYL, Wang RY, Tan YX, Fan JL, Chang YX, Fu J, Jiang F, Chen CY, Yang YC, Gu J, Wu DM, Guo LN, Cao D, Li HY, Cao GW, Wu MC, Zhang MQ, Chen L, Wang HY. Epigenetic modification of MiR-429 promotes liver tumour-initiating cell properties by targeting Rb binding protein 4. Gut. 2014, doi:10.1136/gutjnl-2013-305715. (影响因子 10.732)

使用服务种类: Affymetrix GeneChip Human Genome U133 Plus 2.0 Array / 测序服务

10. Gou LT, Dai P, Yang JH, Xue Y, Hu YP, Zhou Y, et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell research. 2014. (影响因子 10.526)

使用服务种类: Affymetrix GeneChip Mouse Genome 430 2.0 Array 服务

11. Chen C, Chen H, Lin Y-S, Shen J-B, Shan J-X, Qi P, et al. A two-locus interaction causes interspecific hybrid weakness in rice. Nat Commun. 2014;5. (影响因子 10.015)

使用服务种类: Affymetrix 水稻表达谱芯片服务

12. Yu B, He ZY, You P, Han QW, Xiang D, Chen F, Wang MJ, Liu CC, Lin XW, Borjigin U, Zi XY, Li JX, Zhu HY, Li WL, Han CS, Wangensteen KJ, Shi Y, Hui LJ, Wang X, Hu YP. Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors. Cell Stem Cell. 2013, 13(3):328-40.. (影响因子 25.315)

使用服务种类: Agilent 人表达谱芯片服务

13. Liu X, Sun H, Qi J, Wang L, He S, Liu J, Feng C, Chen C, Li W, Guo Y, Qin D, Pan G, Chen J, Pei D, Zheng H. Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT-MET mechanism for optimal reprogramming. Nat Cell Biol. 2013, 15(7):829-38. (影响因子 20.761)

使用服务种类: Agilent 小鼠表达谱芯片服务

14. Han W, Xin Z, Zhao Z, Bao W, Lin X, Yin B, Zhao J, Yuan J, Qiang B, Peng X. RNA-binding protein PCBP2 modulates glioma growth by regulating FHL3. J Clin Invest. 2013, 123(5):2103-18. (影响因子 13.069)

使用服务种类: Agilent 人表达谱芯片服务

15. Cai J, Guan H, Fang L, Yang Y, Zhu X, Yuan J, Wu J, Li M. MicroRNA-374a activates Wnt/ β -catenin signaling to promote breast cancer metastasis. J Clin Invest. 2013 Jan 16. (影响因子 13.069)

使用服务种类: Agilent Human miRNA 芯片服务

-
16. Yan XL, Jia YL, Chen L, Zeng Q, Zhou JN, Fu CJ, Chen HX, Yuan HF, Li ZW, Shi L, Xu YC, Wang JX, Zhang XM, He LJ, Zhai C, Yue W, Pei XT. Hepatocellular carcinoma-associated mesenchymal stem cells promote hepatocarcinoma progression: Role of the S100A4-miR155-SOCS1-MMP9 axis. *Cell Res.* 2013 Jan 12. (影响因子 11.665)
使用服务种类: Agilent 人表达谱芯片服务
17. Lu X, Li X, He Q, Gao J, Gao Y, Liu B, Liu F. miR-142-3p regulates the formation and differentiation of hematopoietic stem cells in vertebrates. *Cell Res.* 2013 Oct 29. (影响因子 10.526)
使用服务种类: Agilent 斑马鱼表达谱服务
18. Yang H, Liu Z, Ma Y, Zhong C, Yin Q, Zhou C, Shi L, Cai Y, Zhao H, Wang H, Tang F, Wang Y, Zhang C, Liu XY, Lai D, Jin Y, Sun Q, Li J. Generation of haploid embryonic stem cells from Macaca fascicularis monkey parthenotes. *Cell Res.* 2013 Jul 16. (影响因子 10.526)
使用服务种类: Agilent SurePrint G3 Rhesus Macaque Genome CGH Microarray 4×180K 服务
19. Li W, Shuai L, Wan H, Dong M, Wang M, Sang L, Feng C, Luo GZ, Li T, Li X, Wang L, Zheng QY, Sheng C, Wu HJ, Liu Z, Liu L, Wang L, Wang XJ, Zhao XY, Zhou Q. Androgenetic haploid embryonic stem cells produce live transgenic mice. *Nature*. 2012, 490(7420):407-11. (影响因子 36.280)
使用服务种类: Affymetrix GeneChip Mouse Genome 430 2.0 芯片服务
20. Huang J, Deng Q, Wang Q, Li KY, Dai JH, Li N, Zhu ZD, Zhou B, Liu XY, Liu RF, Fei QL, Chen H, Cai B, Zhou B, Xiao HS, Qin LX, Han ZG. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. *Nat Genet*. 2012 Aug 26. (影响因子 35.532)
使用服务种类: 全外显子组测序服务
21. Yang H, Shi LY, Wang BA, Liang D, Zhong CQ, Liu W, Nie YZ, Liu J, Zhao J, Gao X, Li DS, Xu GL, Li JS. Generation of Genetically Modified Mice by Oocyte Injection of Androgenetic Haploid Embryonic Stem Cells. *Cell*. 2012, 149(3):605-617. (影响因子 32.403)
使用服务种类: Affymetrix Mouse Genome 430 2.0 Array 服务
22. Zhu S, Pan W, Song X, Liu Y, Shao X, Tang Y, Liang D, He D, Wang H, Liu W, Shi Y, Harley JB, Shen N, Qian Y. The microRNA miR-23b suppresses IL-17-associated autoimmune inflammation by targeting TAB2, TAB3 and IKK- α . *Nat Med*. 2012 Jun 3. (影响因子 22.464)
使用服务种类: Affymetrix GeneChip Human Genome U133 Plus 2.0 芯片服务
23. Huang Y, Li W, Yao X, Lin QJ, Yin JW, Liang Y, Heiner M, Tian B, Hui J, Wang G. Mediator Complex Regulates Alternative mRNA Processing via the MED23 Subunit. *Mol Cell*. 2012 Jan 18. (影响因子 14.148)
使用服务种类: Affymetrix GeneChip Human Exon 1.0 ST Array 服务
24. Jiang L, Lin C, Song L, Wu J, Chen B, Ying Z, Fang L, Yan X, He M, Li J, Li M. MicroRNA-30e* promotes human glioma cell invasiveness in an orthotopic xenotransplantation model by

disrupting the NF- κ B/I κ B α negative feedback loop. *J Clin Invest.* 2012, 122(1):33-47. (影响因子 13.069)

使用服务种类: Agilent Human miRNA 芯片服务

25. Huang W, Hu J, Yang DW, Fan XT, Jin Y, Hou YY, Wang JP, Yuan YF, Tan YS, Zhu XZ, Bai CX, Wu Y, Zhu HG, Lu SH. Two MicroRNA Panels to Discriminate Three Subtypes of Lung Carcinoma in Bronchial Brushing Specimens. *Am J Respir Crit Care Med.* 2012 Oct 4. (影响因子 11.080)

使用服务种类: Agilent Human miRNA 芯片服务

26. Wang S, Wang L, Bayaxi N, Li J, Verhaegh W, Janevski A, Varadan V, Ren Y, Merkle D, Meng X, Gao X, Wang H, Ren J, Kuo WP, Dimitrova N, Wu Y, Zhu H. A microRNA panel to discriminate carcinomas from high-grade intraepithelial neoplasms in colonoscopy biopsy tissue. *Gut.* 2012 Apr 25. (影响因子 10.111)

使用服务种类: Agilent Human miRNA 芯片服务

27. Huang PY, He ZY, Ji SY, Sun HW, Xiang D, Liu CC, Hu YP, Wang X, Hui LJ. The China Consortium for the Genetics of Autoimmune Thyroid Disease. A genome-wide association study identifies two new risk loci for Graves' disease. *Nature Genetics.* 2011, 43(9):897-901. (影响因子 36.377)

使用服务种类: Illumina Human660-Quad BeadChips 服务

28. Huang PY, He ZY, Ji SY, Sun HW, Xiang D, Liu CC, Hu YP, Wang X, Hui LJ. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature.* 2011, 475(7356):386-9.

(影响因子 36.101) (客户发表文章)

使用服务种类: Agilent 小鼠表达谱芯片服务

29. Zhou J, Yu L, Gao X, Hu J, Wang JP, Dai Z, Wang JF, Zhang ZY, Lu SH, Huang XW, Wang Z, Qiu SJ, Wang XY, Yang GH, Sun HC, Tang ZY, Wu Y, Zhu HG, Fan J. Plasma MicroRNA Panel to Diagnose Hepatitis B Virus-Related Hepatocellular Carcinoma. *J Clin Oncol.* 2011, 29(38):2697. (影响因子 18.970)

使用服务种类: Agilent Human miRNA 芯片服务

30. Li KC, Zhang FX, Li CL, Wang F, Yu MY, Zhong YQ, Zhang KH, Lu YJ, Wang Q, Ma XL, Yao JR, Wang JY, Lin LB, Han M, Zhang YQ, Kuner R, Xiao HS, Bao L, Gao X, Zhang X. Follistatin-like 1 Suppresses Sensory Afferent Transmission by Activating Na⁺,K⁺-ATPase. *Neuron.* 2011, 69(5):974-87. (影响因子 14.027)

使用服务种类: SBC 定制芯片服务

31. Song L, Dai T, Xie Y, Wang C, Lin C, Wu Z, Ying Z, Wu J, Li M, Li J. Upregulation of miR-1245 by c-myc targets BRCA2 and impairs DNA repair. *J Mol Cell Biol.* 2011 Dec 8. (影响因子 13.400)

使用服务种类: Agilent 人 CGH 芯片 (4×180K) 服务

32. Jiang J, Ding G, Lin J, Zhang M, Shi L, Lv W, Yang H, Xiao H, Pei G, Li Y, Wu J, Li J. Different developmental potential of pluripotent stem cells generated by different reprogramming strategies. *J Mol Cell Biol.* 2011, 3(3):197-9. (影响因子 13.400)

使用服务种类: Affymetrix Mouse Genome 430 2.0 Array 服务

33. Wang K, Wang P, Shi J, Zhu X, He M, Jia X, Yang X, Qiu F, Jin W, Qian M, Fang H, Mi J, Yang X, Xiao H, Minden M, Du Y, Chen Z, Zhang J. PML/RARalpha Targets Promoter Regions Containing PU.1 Consensus and RARE Half Sites in Acute Promyelocytic Leukemia. *Cancer Cell*. 2010, 17(2):186-197. (影响因子 25.288)

使用服务种类: Affymetrix Human Promoter 1.0R Arrays 服务 / Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 服务

34. Huang J, Zheng DL, Qin FS, Cheng N, Chen H, Wan BB, Wang YP, Xiao HS, Han ZG. Genetic and epigenetic silencing of SCARA5 may contribute to human hepatocellular carcinoma by activating FAK signaling. *J Clin Invest*. 2010, 120(1):223-41. (影响因子 15.387)

使用服务种类: Agilent Whole Human Genome Oligo Microarray 服务

35. Fang H, Yang Y, Li C, Fu S, Yang Z, Jin G, Wang K, Zhang J, Jin Y. Transcriptome analysis of early organogenesis in human embryos. *Dev Cell*. 2010, 19(1):174-84. (影响因子 13.363)

使用服务种类: Affymetrix GeneChip Human Genome U133A Array 服务

36. Fei T, Xia K, Li Z, Zhou B, Zhu S, Chen H, Zhang J, Chen Z, Xiao H, Han JD, Chen YG. Genome-wide mapping of SMAD target genes reveals the role of BMP signaling in embryonic stem cell fate determination. *Genome Res*. 2010, 20(1):36-44. (影响因子 11.342)

使用服务种类: Agilent Promoter Array & Affymetrix Mouse 430 2.0 Array 服务

37. Wang K, Liu J, Yan ZL, Li J, Shi LH, Cong WM, Xia Y, Zou QF, Xi T, Shen F, Wang HY, Wu MC. Overexpression of aspartyl-(asparaginyl)-beta-hydroxylase in hepatocellular carcinoma is associated with worse surgical outcome. *Hepatology*, 2010, 52(1), 164-73. (影响因子 10.840)

使用服务种类: SBC Human cDNA 表达谱芯片服务

经典服务案例

1. miR-146a/b 通过调控 UHRF1 基因来调节胃癌的入侵及迁移

Zhou L, Zhao X, Han Y, Lu Y, Shang Y, Liu C, Li T, Jin Z, Fan D, Wu K. Regulation of UHRF1 by miR-146a/b modulates gastric cancer invasion and metastasis. *FASEB J*. 2013 Aug 27. (IF5.704)
研究者通过分析 miR-146a/b 以及 UHRF1 基因在人类胃癌组织中的含量发现 miR-146a/b 与 UHRF1 基因的表达呈负相关。该发现描绘了一幅 UHRF1 基因调控的新机制及胃癌组织中异常 DNA 甲基化的蓝图,新发现的 miR-146a/b/UHRF1 路线为胃癌迁移灶发生的研究提供了新思路,同时也为阻止胃癌迁移提供了新的治疗方式。**本研究中 Infinium HumanMethylation450 BeadArray 服务由上海伯豪提供。**

2. miR-142-3p 调控脊椎动物造血干细胞的形成和分化

Lu X, Li X, He Q, Gao J, Gao Y, Liu B, Liu F. miR-142-3p regulates the formation and differentiation of hematopoietic stem cells in vertebrates. *Cell Res*. 2013 Oct 29. (IF10.526) .

本文研究者发现在斑马鱼和小鼠中, miR-142-3p 在造血干细胞 (HSCs) 中特异性表达。敲除斑

马鱼的 miR-142a-3p 基因导致主动脉-性腺-中肾(AGM)区的 HSCs 数量降低，同时胸腺中的 T 细胞数量也降低。从机制上来讲，miR-142a-3p 通过抑制干扰素调节因子 7 (irf7) 介导的炎症信号通路来调节 HSCs 的形成和分化。此外，研究者还证实 miR-142-3p 在小鼠的 AGM 区 HSCs 的形成中也发挥了重要的作用，这也暗示了它在脊椎动物中扮演着一个高度保守的角色。该研究还揭示了 miR-142a-3p 通过抑制 irf7 信号通路在 HSCs 的形成和发生过程中起关键作用。**本研究中 Agilent Zebrafish Oligo Microarrays 服务由上海伯豪提供。**

3. 被 TGF-β 活化的 lncRNA-ATB 促进肝癌的侵袭-转移级联反应

Yuan JH, Yang F, Wang F, Ma JZ, Guo YJ, Tao QF, Liu F, Pan W, Wang TT, Zhou CC, Wang SB, Wang YZ, Yang Y, Yang N, Zhou YP, Yang GS, Sun SH. A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* 2014;25(5):666-81.
(IF24.755)

该课题组研究人员首次发现了一个介导转化生长因子-β (TGF-β) 促转移作用的 lncRNA，并将这个新 lncRNA 命名为“被 TGF-β 活化的长链非编码 RNA (lncRNA activated by TGF-β, lncRNA-ATB)”。研究发现 lncRNA-ATB 在肝癌组织中高表达，在肝癌转移灶中进一步上调表达，并与肝癌的侵袭特征成正相关，指示着肝癌病人更差的预后。因此 lncRNA-ATB 显著促进肝癌的侵袭-转移级联反应。这些研究结果表明，lncRNA-ATB 作为 TGF-β 信号通路促癌作用的介导因子，让肝癌病人易发生转移，可作为抗转移治疗的潜在靶点。**课题相关 RIP-SEQ 服务由上海伯豪生物技术有限公司提供。**

4. 多囊卵巢综合征病人基因组 DNA 甲基化免疫沉淀反应分析

Wang X, Ao H, Zhai L, Bai L, He W, Yu Y, Wang C. Genome-wide effects of DNA methyltransferase inhibitor on gene expression in double-stranded RNA transfected porcine PK15 cells. *Genomics*. 2013 Oct 31. (IF3.010)

本研究主要致力于探究 IR 在 PCOS 发生过程中的分子机制，运用基因组 DNA 甲基化免疫沉淀技术 (methylated DNA immunoprecipitation, MeDIP)，分析了正常人与 PCOS 病人基因组中有差异的甲基化基因，此外，还鉴别了胰岛素不耐受的 PCOS 病人 (PCOS-non-insulin resistance, PCOS-NIR) 与胰岛素耐受的 PCOS 病人 (PCOS insulin resistance, PCOS-IR) 差异的甲基化基因。由上海仁济医院的狄文教授与苏州大学神经科学研究所的曹聪教授共同合作完成的，其团队主要采用 MeDIP 技术完成了对 PCOS 疾病的分子机制的研究，**该服务由上海生物芯片公司提供。**

5. DNMT3A Arg882 的突变通过扰乱造血细胞基因表达/DNA 甲基化诱导慢性粒单核细胞白血病

Xu J, Wang YY, Dai YJ, Zhang W, Zhang WN, Xiong SM, et al. DNMT3A Arg882 mutation drives chronic myelomonocytic leukemia through disturbing gene expression/DNA methylation in hematopoietic cells. *Proc Natl Acad Sci U S A*. 2014 Feb 4. (IF9.737)

上海交通大学医学院附属瑞金医院的研究人员证实，在造血细胞中 DNA 甲基转移酶 3A(DNMT3A) Arg882 突变，通过破坏基因表达/DNA 甲基化驱动了慢性粒单核细胞白血病 (CMML)。这一研究成果发表在 2 月 4 日的《美国科学院院刊》(PNAS) 杂志上。上海交通大学医学院附属瑞金医院的陈竺 (Zhu Chen) 院士、陈赛娟 (Sai-Juan Chen) 院士以及王月英 (Yue-Ying Wang) 博士是这篇论文的共同通讯作者。**课题相关 MeDIP-SEQ 技术服务由上海伯豪生物技术有限公司提供。**

6. 短期的和持续的 ABA 区别地与多种调控植物生存和生长的激素相互作用

Yang C, Liu J, Dong X, Cai Z, Tian W, Wang X. Short-term and continuing ABA differentially

interplay with multiple hormones to regulate plant survival and growth. Molecular Plant.2014,Feb 4. (IF6.126)

本文研究人员采用 1 小时及 9 天的 ABA 处理植物，基因组范围内分析发现，两种条件下基因调节有很大的区别。采用一些激素相关的突变，研究人员发现了相应激素信号通路中的候选者（包括：受体或转录调控者），是对 ABA 作出反应所必须的。揭示了依赖 ABA 的压力信号与不同水平的激素相互作用，它们与植物在短暂压力下的生存及适应持续的环境压力相关。**本研究中 Agilent Arabidopsis Expression gene chips 服务由上海伯豪提供。**

7. CXCL5(epithelial neutrophil-activating peptide-78)促进中性粒细胞浸润并可作为肝癌预后的判断指标

Zhou SL, Dai Z, Zhou ZJ, Wang XY, Yang GH, Wang Z, Huang XW, Fan J, Zhou J Overexpression of CXCL5 Mediates Neutrophil Infiltration and Indicates Poor Prognosis for Hepatocellular Carcinoma.Hepatology.2012, 56(6) :2242-54. (IF10.8)

复旦大学中山医院肝癌研究所癌变与侵袭教育部重点实验室，复旦大学生物医学研究院的研究人员证实趋化因子 CXCL5 在肝癌的肿瘤生长、侵袭及预后中发挥重要的作用。单独的 CXCL5 的过表达或其与瘤内中性粒细胞的存在状态相结合，都可以作为一个新的预后衡量指标，同时 CXCL5 也可以作为肝癌一个新的治疗靶点。**课题相关蛋白芯片服务由上海伯豪生物技术有限公司提供。**

8. 芥菜型油菜种皮转录组 *de Novo* 拼接及类黄酮生物合成基因的鉴定

Liu X, Lu Y, Yuan Y, Liu S, Guan C, Chen S, Liu Z. De Novo Transcriptome of Brassica juncea Seed Coat and Identification of Genes for the Biosynthesis of Flavonoids. PLoS One. 2013, 8(8):e71110. (IF3.730)

本研究是湖南农业大学油料作物研究所刘忠松教授课题组完成的。该研究首次实现芥菜型油菜种皮转录组测序，研究中所获得的基因不仅有利于阐明芥菜型油菜种皮着色的分子机制，且为该物种今后的基因组学研究提供了基础。**研究中所涉及的 Illumina HiSeq 2000 测序服务以及数据分析服务由上海伯豪生物技术有限公司提供。**

9. 雌雄同株异花的毛白杨 microRNA 的性能分析以及对靶基因表达的影响

Song Y, Ma K, Ci D, Zhang Z, Zhang D. Sexual dimorphism floral microRNA profiling and target gene expression in andromonoecious poplar (*Populus tomentosa*). PLoS One. 2013, 8(5):e62681. (IF3.730)

本研究首次报道了与花发育相关的 miRNAs 以及它们的靶基因，根据基因注释，将所有的靶基因分为四类。第一类，花发育相关：Pto-F6, Pto-F11, Pto-F14, Pto-F19, Pto-25, Pto-36, Pto-F51, Pto-F54 以及 Pto-F66。第二类，Ca²⁺转运相关：Pto-F28 和 Pto-F45（雌花特异性 miRNA），Pto-F16（雄花特异性 miRNA）。第三类，激素调节相关：10 个 miRNAs（除 Pto-F6 以外均为雌花特异性）。第四类，与 DNA 甲基化相关：3 个 miRNAs。**课题相关小 RNAs 测序服务 Illumina HiSeq 2000 由上海伯豪生物技术有限公司提供。**

10. 有效提高非模式生物蛋白质组鉴定的新策略——逐次修正基因组法

Wu XH, Xu L, Gu W, Xu Q, He QY, Sun XS, Zhang g. Genome Correction Largely Improves Proteomic Analysis of Nonmodel Organisms. J Proteome res. 2014. (IF5.001)

暨南大学翻译组学实验室提出利用迭代修正的方法不断矫正已知近缘物种的基因组序列，以获得研究物种相对精确的基因组。这种修正方法是基于该实验室自行开发的 mapping 算法：FANSe。

本课题相关测序服务由上海伯豪生物技术有限公司提供。

11. 玉米 Proline responding 1(pro1)突变在蛋白合成和细胞周期调控中起到关键作用

Wang G, Zhang JS, Wang GF, Fan XY, Sun X, Qin HL, Xu N, Zhong MY, Qiao ZY, Tang YP, Songa R. Proline responding1 Plays a Critical Role in Regulating General Protein Synthesis and the Cell Cycle in MaizeC. *The Plant Cell*. 2014. (IF9.251)

本研究证实玉米 Pro1 基因(Zm P5CS2)的突变造成了突变体细胞中脯氨酸(proline)合成受阻，从而导致 proline 积累的减少。Pro1 突变基因的克隆和功能分析解析了脯氨酸在植物生长发育过程中对蛋白合成和细胞周期 G1 到 S 期转换的调控机制。**课题相关 RNA-seq 技术服务由上海伯豪生物技术有限公司提供。**

12. 肿瘤特异性染色体错误分离通过保持肿瘤异质性来控制癌症的可塑性

Zhang X, Jia Z, Jadus M, LImoli C, Linskey M, Xing J, Zhou Y. Tumor-Specific Chromosome Mis-Segregation Controls Cancer Plasticity by Maintaining Tumor Heterogeneity. *PLoS One*. 2013 Nov. (IF3.730)

非整倍体染色体的不稳定性是癌症的一大特点。研究者研究了由原代细胞培养出的神经胶质瘤细胞的 7 号染色体 (Chr7) 拷贝数变异 (CNV) 情况，发现普遍出现在胶质瘤中的 Chr7-CNV 的肿瘤异质性细胞，在高级别的胶质瘤（携带超过 2 个拷贝数的 Chr7）中出现的百分比较低级别的胶质瘤更高。本研究首次发现染色体的不稳定性与保持肿瘤异质性相关，这也为癌症的生长、维持、耐药性奠定了基础。**研究中所使用的 Agilent Human Genome CGH 244 k Microarray 服务由上海伯豪提供。**

13. Agilent 表达谱芯片揭示体细胞重编程分子机制

Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT-MET mechanism for optimal reprogramming. *Nat Cell Biol*. 2013, 15(7):829-38. (IF20.761)

中科院广州生物医药与健康研究院的裴端卿教授和郑辉研究员，通过优化转化因子的导入顺序，开发了一个体细胞重编程新方案，证实相比于传统的方法其可使重编程效率大大提高，并由此揭示了重编程起始阶段按顺序经历的一个先发生 EMT(上皮细胞向间充质转化)，再发生 MET(间充质向上皮转化)的分子过程。该研究或将有助于推动诱导多能干细胞研究的理论和实践突破。**该课题所用 Agilent 小鼠表达谱芯片服务由上海伯豪提供。**

14. Affymetrix GeneChip Mouse Genome 430 2.0 芯片助力单倍体胚胎干细胞 (haESCs) 研究

Androgenetic haploid embryonic stem cells produce live transgenic mice. *Nature*. 2012, 490(7420):407-11. (IF36.280)

中科院动物研究所、中国科学院大学与东北农业大学的研究人员获得了胚胎干细胞研究的突破性成果，成功建立了来自孤雄囊胚单倍体胚胎细胞系，并进一步验证将这些细胞注入到卵母细胞后产生了健康的小鼠。该研究证实了实验室培育雄性单倍体多能干细胞能够在功能上像配子一样形成受精卵，并可生成活体后代，或许有可能为辅助生殖提供一个宝贵的资源。**文章中 Affymetrix GeneChip Mouse Genome 430 2.0 Array 服务由上海伯豪完成。**

15. Affymetrix GeneChip Human Genome U133 Plus 2.0 芯片揭示 miR-23b 相关的自身免疫病

The microRNA miR-23b suppresses IL-17-associated autoimmune inflammation by targeting TAB2, TAB3 and IKK- α . *Nat Med*. 2012, 18(7):1077-86. (IF22.464)

中科院健康所钱友存研究员带领的研究组与上海交通大学医学院附属仁济医院沈南教授带领的研究组共同

合作，通过高通量 microRNA 芯片筛选发现了 miR-23b 是自身免疫病病人（类风湿性关节炎、红斑狼疮）和自身免疫病小鼠模型（CIA、MRL/lpr、EAE）的炎症病理组织中共同下调的 microRNA。miR-23b 通过靶向炎症性细胞因子 IL-17、TNFa、IL-1b 的下游信号通路中的关键信号分子（TAB2、TAB3、IKKa）来抑制自身免疫病。该工作首次阐述了非免疫细胞来源的 miRNA 参与免疫性疾病的机制，miR-23b 可以成为治疗自身免疫病的一个新靶点，有可能被开发成有效缓解甚至治愈病症的新药。研究课题中 **Affymetrix GeneChip Human Genome U133 Plus 2.0 芯片服务由上海伯豪提供**。

16. GWAS 发现 Graves 病两个新易感基因位点

A genome-wide association study identifies two new risk loci for Graves' disease. Nature Genetics. 2011, 43(9):897-901. (IF36.377)

上海交通大学医学院附属瑞金医院、国家人类基因组南方研究中心、徐州医学院附属医院、山东省临沂人民医院、蚌埠医学院第一附属医院等十几家医疗机构的研究人员，利用先进的全基因组关联分析技术（GWAS）发现了 Graves 病（格雷夫斯氏病，简称甲亢，GD）的两个新易感基因位点。这项新研究不仅找到了与甲亢发病相关的两处新的染色体变异区域，首次揭示了致病基因差异是导致甲亢药物治疗效果因人而异的原因，初步解开了甲亢发病机制的谜团。并为甲亢治疗方法的选择和预后的评估提供了重要的生物标志。文章中 **Illumina Human660-Quad BeadChips 服务由上海伯豪完成**。

17. 再生医学新进展—科学家首次获得转化型肝脏细胞

Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature. 2011, 475(7356):386-9. (IF36.101)

中国科学院上海生命科学研究院生物化学与细胞生物学研究所惠利健研究小组成功将小鼠尾部成纤维细胞诱导转化为肝脏细胞，并且在小鼠动物实验中成功利用转化型肝脏细胞重建受体小鼠肝脏。这一发现在国际上尚属首次，为将来在肝移植术中，利用病人自身体细胞诱导获得肝脏细胞进行移植的应用奠定了基础。

相关的 Agilent 小鼠表达谱芯片服务由上海伯豪提供。

18 Affymetrix SNP 6.0 芯片揭示先天多毛症遗传缺失

Copy-number mutations on chromosome 17q24.2-q24.3 in congenital generalized hypertrichosis terminalis with or without gingival hyperplasia. Am J Hum Genet. 2009, 84(6):807-13. (IF12.303)

先天性全身多毛症（congenital generalized hypertrichosis，CGHT）是一种极其罕见的体毛生长异常。中国医学科学院-北京协和医学院张学教授领导的联合团队首先收集了 3 个常染色体显性遗传的中国人 CGHT 家系和 1 名患重度 CGHT 伴牙龈增生的散发个体；继之通过全基因组扫描将 CGHT 的遗传位点定位到染色体 17q24.2-q24.3 区域；进一步应用全基因组拷贝数检测、荧光定量 PCR 和荧光原位杂交（FISH）等多项先进技术，确定染色体 17q24.2-q24.3 拷贝数突变对无论有无牙龈增生的 CGHT 都起到决定性的作用。该课题相关 **Affymetrix Genome-Wide Human SNP Array 6.0 服务在生物芯片上海国家工程研究中心完成**。

19. Affymetrix GeneChip Human Mapping 100K 芯片揭示亚洲人群迁徙规律

Mapping Human Genetic Diversity in Asia. Science. 2009, 326(5959):1541-5. (IF29.747)

由人类基因组机构（Human Genome Organization, HUGO）泛亚太地区 SNP 联盟（Pan-Asian SNP Consortium）发起的一项大型研究首次通过 SNP 分析获得了亚洲人群 SNP 遗传多元化图谱，并以此分析了亚洲人群的历史移民路线。生物芯片上海国家工程研究中心作为作者单位之一承担了中国样品的 SNP 实验。研究表明，亚洲大陆的人口主要是通过某单一的迁徙活动而形成的。**该课题相关 Affymetrix GeneChip Human Mapping100KArray 服务由上海伯豪提供。**

20. Illumina SNP 芯片发现汉族人群红斑狼疮易感基因

Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nature Genetics*. 2009, 41(11):1234-7. (IF34.284)

安徽医科大学第一附属医院张学军教授领衔的研究团队通过对 12000 多例中国汉族红斑狼疮患者和健康对照样本进行研究，发现了 5 个与汉族人群发病密切相关的易感基因 ETS1、IKZF1、RASGRP3、SLC15A4 和 TNIP1，并确定了 4 个新的易感位点；研究同时验证了在欧洲人中发现的 7 个易感基因在汉族人中同样存在。**该课题相关 Illumina Human 610-Quad BeadChips 服务由上海伯豪提供。**

21. 利用表达谱芯片筛选肝癌预后标志物

Overexpression of aspartyl-(asparaginyl)-beta-hydroxylase in hepatocellular carcinoma is associated with worse surgical outcome. *Hepatology*, 2010, 52(1), 164-73. (IF10.840)

这项成果由第二军医大学东方肝胆外科医院沈锋教授带领的课题组历时 5 年研究完成。他们首先对 8 名肝癌术后患者的样品进行了表达谱 cDNA 芯片（使用了 SBC 人 cDNA 表达谱芯片）杂交实验，发现肝癌 AAH 高表达预示肝切除术预后更差，且极易引起复发。这一研究成果对肝癌 AAH 高表达的手术患者具有重要意义，医生可根据患者手术后 AAH 的表达水平预测肝癌复发的危险度，进而筛选出高危的复发人群。**该课题相关 SBC Human cDNA 表达谱芯片服务由上海伯豪提供。**

22. 表达谱芯片筛选发现肝癌扩散转移相关基因

Genetic and epigenetic silencing of SCARA5 may contribute to human hepatocellular carcinoma by activating FAK signaling. *J Clin Invest*. 2010, 120(1):223-41. (IF15.387)

该研究使用 Agilent 人全基因组表达谱芯片对 15 个经过 DAC+TSA 处理和没有经过处理的肝癌细胞系进行了表达谱分析，发现 SCARA5 基因沉默可导致肝癌癌细胞扩散转移。*J Clin Invest* 评审专家认为，这不仅是发现了一个与癌症相关的重要基因，为肝癌提供了潜在新型诊断和治疗方法，还采用了新的研究方法，拓展了研究思路。该课题相关的芯片技术服务在生物芯片上海国家工程研究中心完成。本文第一作者黄建博士这样评价我们的工作：“全面的解决方案；先进的技术平台；优质的服务水平。”**相关 Agilent Whole Human Genome Oligo Microarray 服务由上海伯豪提供。**

23. 基因芯片筛选发现大肠癌分子标志物

A microRNA panel to discriminate carcinomas from high-grade intraepithelial neoplasms in colonoscopy biopsy tissue. *Gut*. 2013, 62(2): 280-289. (IF10.111)

该研究课题首先收集了 133 例大肠癌样品，利用激光共聚焦捕获技术(laser capture microdissection, LCM)切取组织样品中的上皮细胞，随后提取总 RNA，利用 Agilent microRNA 芯片筛选，RT-PCR 验证等方法找到了 3 个 microRNA 的组合，最后，利用这 3 个 microRNA 的组合来对 58 个 FFPE 样本进行早期诊断，发现准确率达到 89% (AUC=0.918)。这一研究成果对于大肠癌的及时发现和治疗有重大意义。其中 **microRNA 基因芯片技术服务由上海伯豪完成。**

24. 基因芯片筛选发现乙型肝癌血清血浆 microRNA 生物标志物

Plasma MicroRNA Panel to Diagnose Hepatitis B Virus-Related Hepatocellular Carcinoma. *J Clin Oncol*. 2011, 29(36):2697. (IF18.970)

复旦大学附属中山医院、复旦大学肝癌研究所樊嘉教授领衔的课题组利用 microRNA 芯片对比肝癌患者 VS 正常人群，肝癌患者 VS 慢性乙型肝炎 (chronic hepatitis B, CHB) 患者，肝癌患者 VS 肝硬化 (cirrhosis) 患者血清血浆 microRNA 的表达谱，筛选出 15 个明显差异表达的 microRNA。随后，通过 RT-PCR 验证等方法寻找到 7 个 microRNA 的组合。最后，科学家们又利用 390 份样品对建立好的诊断模型进行验证，该模型区分肝癌患者和正常人群的正确率能达到 94.1%。这项研究成果对于肝癌的早期诊断和治疗有重大

意义, microRNA 生物芯片技术服务由上海伯豪提供。

25. 运用 ChIP-chip 技术寻找癌蛋白靶基因

PML/RARalpha Targets Promoter Regions Containing PU.1 Consensus and RARE Half Sites in Acute Promyelocytic Leukemia. *Cancer Cell.* 2010, 17(2):186-197. (IF25.288)

PML/RAR α 是一个癌基因, 在急性早幼粒性白血病 (APL) 中发挥着重要作用。本文作者利用 ChIP-chip 技术筛选得到与癌蛋白 PML/RAR α 特异结合的启动子序列 (共 2979 个), 并用生物信息学方法对这 2979 个结合序列进行保守性分析, 发现大部分的结合序列均含有转录因子结合位点 PU.1 和 RAREh (即 PU.1-RAREh)。在后续的验证实验中, 作者发现癌蛋白 PML/RAR α 致癌的机制正是与转录因子 PU.1 结合, 从而导致下游基因转录的变化。该研究相关的芯片技术服务在生物芯片上海国家工程研究中心完成。

26. HBV 相关的肝细胞癌外显子组研究

Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. *Nat Genet.* 2012 Aug 26. (IF35.532)

利用上海伯豪生物技术有限公司的 Illumina GAIIx 和 ABI 的 SOLiD 新一代测序服务平台, 进行全外显子组测序分析。比较分析原位癌和转移灶的测序结果, 发现肿瘤细胞存在 347 个突变基因, 平均每个肿瘤样本有 30-40 个基因突变, 其中大多数基因突变是第一次在肝癌样本中发现, 并且在原发灶和转移灶中同时出现。说明肝癌的发病和转移与基因突变密切相关, 并且转移相关基因在原发癌中就已经发生了突变。具体分析基因突变的类型, 发现有 5 位患者存在明显的 C:G>A:T 突变, 有 4 位患者存在明显的 T:A>A:T 突变。其中外显子组测序服务由上海伯豪提供。

27. DNMT3A 突变驱动了慢性粒单核细胞白血病

DNMT3A Arg882 mutation drives chronic myelomonocytic leukemia through disturbing gene expression/DNA methylation in hematopoietic cells. *Proc Natl Acad Sci U S A.* 2014 Feb

在 20% 的急性髓性白血病案例中会发生 DNMT3A 的突变, 突变热点为 Arg882 (R882)。本研究通过运用逆转病毒转导及骨髓移植 (BMT) 的方法来研究 DNMT3A-Arg882His (R882H) 突变体的转化能力。研究人员发现突变的基因能够引起造血干细胞/祖细胞的异常增殖。BMT 1 年后, 所有小鼠都出现了血小板增多, 且患上 CMML。RNA 芯片分析结果显示一些造血作用相关基因异常表达, DNA 甲基化实验证实基因区域甲基化类型发生相应的变化此外, DNMT3A-R882H 增加了 CDK1 蛋白水平, 加强了细胞周期活力, 这些都导致白血病的生成。上海交通大学医学院附属瑞金医院的陈竺院士、陈赛娟院士以及王月英博士是这篇论文的共同通讯作者。课题相关 MeDIP-SEQ 技术服务由上海伯豪生物技术有限公司提供。

28 肿瘤样品之间甲基化谱的差异分析

Regulation of UHRF1 by miR-146a/b modulates gastric cancer invasion and metastasis. *FASEB J., August, 2013, 27.*

本研究中, 科学家们利用 DNA 甲基化芯片分析了正常组织、腺瘤、原发性恶性肿瘤、转移性恶性肿瘤样品之间甲基化谱的差异。结果发现, 四种样品之间存在明显的甲基化谱差异, 恶性肿瘤组织存在广泛的低甲基化现象, 而且不同信号通路的甲基化水平不同。发现描绘了一幅 UHRF1 基因调控的新机制及胃癌组织中异常 DNA 甲基化的蓝图, 新发现的 miR-146a/b/UHRF1 路线为胃癌迁移灶发生的研究提供了新思路, 同时也为阻止胃癌迁移提供了新的治疗方式。本研究由西京医院及第四军医大学共同合作完成, 文章的通讯作者为樊代明教授和吴开春教授。本研究中 Infinium HumanMethylation450 BeadArray 服务由上海伯豪提供。