

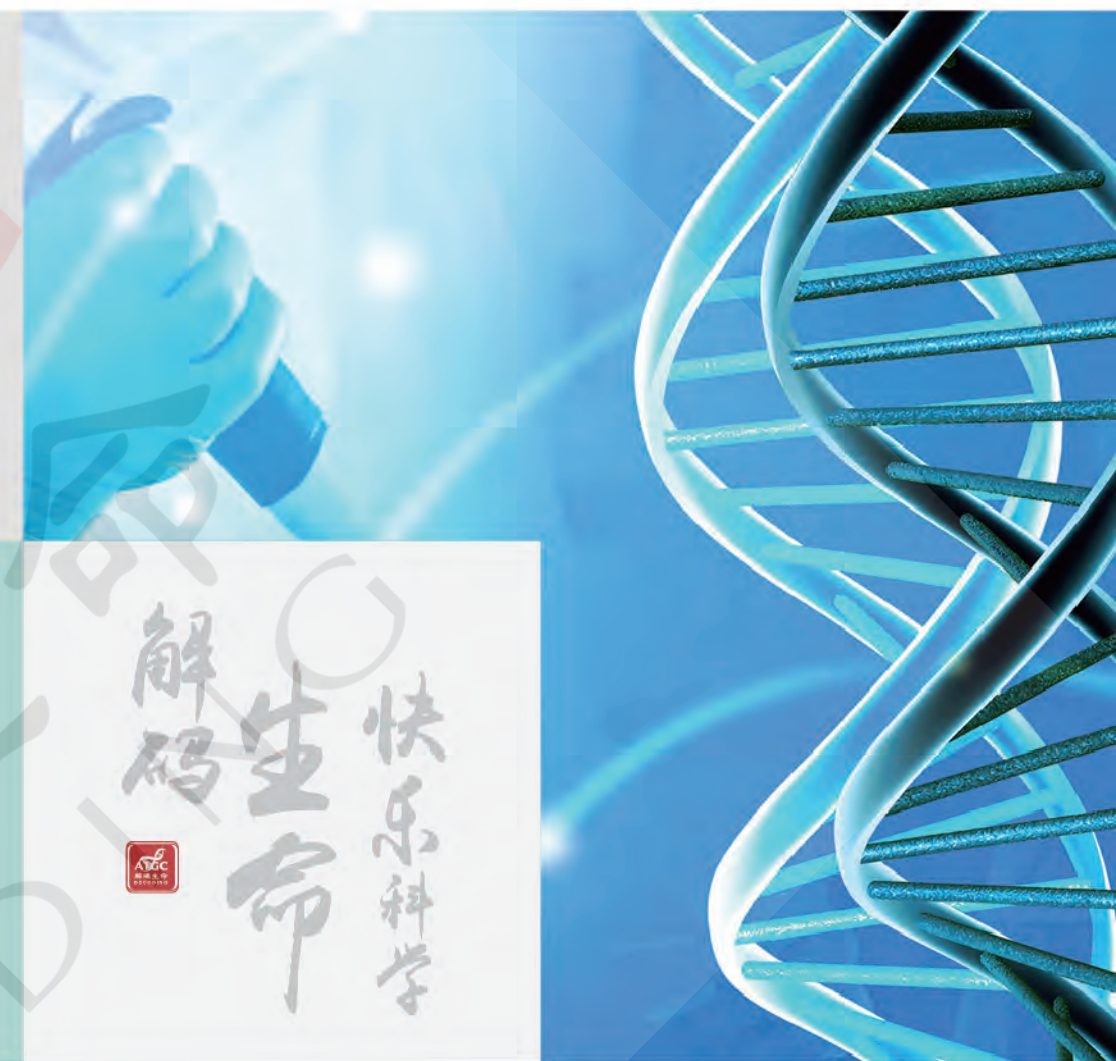
技术手册



扫一扫 关注更多

上海市徐汇区华山路2018号1902室
021-64152178 021-54072208

decodingbio@163.com
www.decodinglife.cn/



研究院简介

INTRODUCTION

当人类的脚步踏入21世纪的第二个十年时，科学家们对人类健康与疾病的认识，已经逐渐由组织、细胞和蛋白质层面，跨入决定蛋白质和细胞结构与功能的基因及基因组层面。对于已经明确病因的单基因疾病，单从技术角度而言，进行早期诊断已不在话下。但对于复杂形状的疾病，即对人类危害最大的多基因疾病或环境与遗传相互作用所致的疾病，人们还只是初窥门径。对于尚未发病，但具有发病遗传学基础的个体，及早进行基因或全基因组的精确检测，分析并捕捉与疾病相关的遗传变异，及早进行预防或干预，将极大地提升人类的健康水平、生活质量，乃至全民的人口整体素质！

正是在先进技术方法不断涌现、崭新健康理念日渐深入人心的关键时代，“上海浦东解码生命科学研究院”（简称浦东解码生命科学研究院）应运而生。浦东解码生命科学研究院的缔造者贺林院士秉承“解码生命，造福人类”的宗旨，团结并凝聚了一大批中外在生命科学和医学研究及实践中颇有造诣的科学家和医学家，力图用一种新的模式，对人类健康与疾病领域中的关键问题，进行协作攻关。使志同道合的同仁们，能够在一个宽松和高效的平台上，愉快地将基础和临床的研究成果，迅速地转化为能够为世人所用的技术。让那些饱受健康问题困扰和疾病折磨的人们，能够尽快从科学的发展中受益，从而享受到生命的幸福和尊严！

浦东解码生命科学研究院具备完善的一代、二代和三代基因测序平台，全面的芯片检测平台，强大的生物信息学分析平台。在这些平台上，可以高效地完成基因组学领域的各类研究。

浦东解码生命科学研究院已经在基因身份证、个体营养状况评估、个体化用药、单基因遗传病基因检测和染色体异常筛选等方面获得重要进展。经与有关医疗机构联合初步应用，已经显示出巨大的市场前景。

真诚欢迎生命科学和医学领域的有识之士与浦东解码生命科学研究院一起在这个温暖的大家庭里，让我们共同实现为人类健康造福的伟大理想！

解码生命
快乐科学

目 录

CONTENTS

第一章 基因组学.....	01
一. 全基因组测序.....	01
二. 全外显子组测序.....	07
三. 目标区域测序.....	12
四. 比较基因组芯片.....	17
五. Chip-on-chip芯片.....	19
第二章 表观遗传组学.....	21
一. 全基因组甲基化测序.....	21
二. DNA甲基化芯片.....	26
三. DNA甲基化研究策略.....	28
第三章 转录组学.....	31
一. 转录组测序.....	31
二. lncRNA测序.....	33
三. miRNA测序.....	34
四. circRNA测序.....	35
五. 全转录组测序.....	36
六. 表达谱芯片检测.....	41
七. microRNA芯片检测.....	45
第四章 微生物组学.....	47
一. 16S/18S/ITS测序芯片检测.....	47
二. 宏基因组测序芯片检测.....	54
第五章 蛋白质组学.....	59
一. 蛋白质组定性检测.....	59
（一）蛋白质全谱分析.....	60
（二）蛋白质胶条/混合液鉴定.....	61
（三）蛋白质胶点鉴定.....	62
二. 蛋白质组定量检测.....	64
（一）iTRAQ.....	65
（二）TMT.....	66
（三）Label-free.....	67
（四）SILAC.....	68
（五）2D双向凝胶电泳.....	68
（六）DIGE荧光差异双向凝胶电泳.....	69
三. 蛋白质化学修饰组学检测.....	71
四. 蛋白质组学检测平台.....	73
（一）AB SciexTriple TOF 5600.....	73
（二）Thermo Scientific Q-Exactive.....	73
（三）Thermo Scientific LTQ XL线性离子阱质谱仪.....	74

（四）Orbitrap FusionTM TribridTM 质谱仪.....	74
第六章 代谢组学.....	75
一. 非靶向代谢组学.....	76
二. 定量代谢组学.....	77
（一）脂质检测.....	78
（二）胆汁酸检测.....	79
（三）游离脂肪酸检测.....	80
（四）肠道菌群和宿主代谢物检测.....	80
三. 代谢组学检测平台.....	85
（一）GC-TOF/MS.....	85
（二）LC-MS/MS.....	85
第七章 测序平台.....	87
一. ABI 3730XL毛细管电泳测序系统.....	87
二. PGM/Ion Torrent测序系统.....	89
三. Illumina HiSeq 3000/4000测序系统.....	91
四. 焦磷酸测序(Pyrosequencing)系统.....	94
五. Agena MassARRAY系统.....	97
第八章 芯片平台.....	99
一. Illumina微阵列扫描仪系统.....	99
二. 安捷伦DNA微阵列扫描仪系统.....	100
三. 基因芯片扫描仪系统.....	101
四. 罗氏MS200扫描仪.....	102
第九章 生物信息学平台.....	103
一. 平台简介.....	103
二. 测序分析流程.....	103
三. 结果报告.....	104
四. 生信平台优势.....	110
第十章 分子生物学平台.....	111
一. 基因编辑(CRISPR-Cas9).....	111
二. MLPA.....	119
三. RT-qPCR.....	121
四. TaqMan检测.....	122
五. 基因定点突变.....	125
六. 载体构建.....	126
第十一章 蛋白质和细胞研究平台.....	129
一. Western Blot(WB).....	129
二. ELISA.....	130
三. 免疫荧光.....	131
四. 免疫组织化学.....	133
第十二章 外泌体研究.....	134

第一章 基因组学

一.全基因组测序

(一) 背景介绍

全基因组测序 (Whole Genome Sequencing, WGS) 是利用高通量测序平台对人类不同个体或群体进行全基因组测序, 并在个体或者群体水平上进行生物信息分析的技术手段。全基因组测序可以全面挖掘DNA水平的遗传变异, 包括较大的结构性变异, 为筛选疾病的致病及易感基因, 研究发病及遗传机制, 以及推断种群迁移和进化等提供重要信息。全基因组测序可以分为de novo测序和重测序。全基因组de novo测序不参考任何已有的DNA序列信息, 直接测序某物种的基因组, 因此被称为从头测序。从头测序技术首先获取某物种基因组DNA样品, 检测样品质量, 评估基因组特征, 构建梯度插入片段文库上机测序, 采用生物信息学分析手段将测序产生的海量DNA序列片段 (reads) 进行排序、拼接, 从而组装出该物种完整的基因组DNA序列图谱。对已有参考基因组序列的物种进行个体或群体全基因组测序称为全基因组重测序。全基因组重测序结果与已有参考基因组序列进行比对, 检测出全基因组范围内的单核苷酸多态性位点 (SNP)、插入缺失突变 (InDel)、拷贝数变异 (CNV) 和基因组重排导致的结构变异 (SV) 等变异信息, 获得个体或群体分子遗传特征, 广泛应用于遗传变异检测、性状基因定位、遗传图谱构建和遗传进化研究等。

(二) 测序原理和技术流程

WGS测序属于常规二代测序, 其原理为边合成边测序 (Sequencing by synthesis, SBS)。

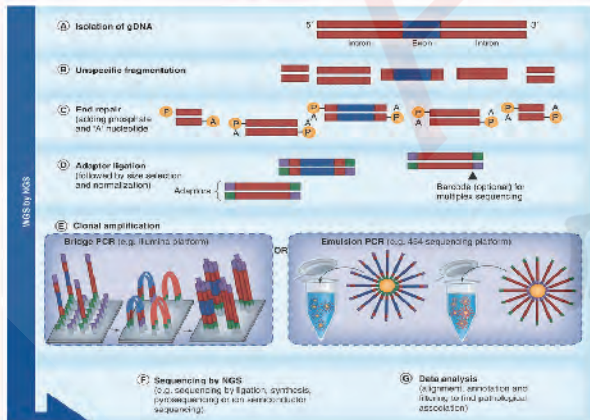


图1. WGS测序原理和技术流程

(三) 生物信息学分析

疾病基因组学

基本分析	◆ 数据质控：去除接头污染和低质量数据 ◆ 与参考序列比对、统计测序深度及覆盖度 ◆ SNP/InDel/CNV/SV 检测，注释及统计	
单基因病高级分析	◆ 突变位点筛选 ◆ 显/隐性遗传模式分析 - 显性遗传模式分析 - 隐性遗传模式分析 ◆ CNV/SV检测与注释	◆ 候选基因功能注释 ◆ 基因功能及通路分析 ◆ 家系连锁分析 ◆ 纯合子区域（ROH）分析 ◆ 共有突变筛选
复杂疾病高级分析	◆ 突变位点筛选 ◆ 显/隐性遗传模式分析 - 显性遗传模式分析 - 隐性遗传模式分析 ◆ CNV/SV检测与注释	◆ 新生突变筛选 ◆ 共有突变筛选 ◆ 候选基因功能注释 ◆ 候选基因功能富集 ◆ 基因显著性分析（推荐30对case/control or trios样本） ◆ 蛋白互作网络分析（PPI）
个性化分析	◆ 药物效应多态性的遗传学机理研究 ◆ 生存分析 ◆ 疾病显著性关联位点分析（建议基于150对以上的case/control）	◆ 疾病显著性关联基因分析（建议基于100对以上case/control） ◆ 基于HLA型别的关联分析

癌症基因组学（肿瘤成对样本）

标准分析	◆ 数据质控：去除接头污染和低质量数据 ◆ 与参考序列比对、统计测序深度及覆盖度 ◆ Somatic mutation/InDel/SV/CNV检测、注释及统计（成对样本）	
高级分析	◆ 易感基因筛查 ◆ 高频突变基因统计及通路富集分析（>2对样本） ◆ NMF突变特征及突变频谱分析 ◆ NoyoDriver已知驱动基因筛选 ◆ MRT高频突变基因相关性分析 ◆ OncodriveCLUST驱动基因检测 ◆ 突变位点分布情况分析 ◆ 高频 CNV 分析	◆ ABSOLUTE肿瘤纯度及倍性分析 ◆ 杂合性缺失（LOH）分析 ◆ 瘤内异质性及克隆结构分析 - 单样本克隆结构分析（Pyclone、Expands） - 体细胞突变CCF计算 ◆ NovoDrug高频突变基因靶向用药预测 ◆ NovoDR耐药性突变筛选
个性化分析	◆ 肿瘤进化树分析 ◆ 病毒整合分析 ◆ 多样本克隆结构分析	◆ 临床数据整合 ◆ 肿瘤分型分析 ◆ 突变频谱3D展示图

(四) 应用范围

全基因组测序适用于癌症基因组学、遗传病研究、药物基因组学、遗传图谱构建、功能基因挖掘和群体进化等研究领域。

(1)在癌症基因组学研究中，每个癌症样品呈现在研究人员眼前的是一个发生了改变的基因组，其中包含独特且不可预测的多个点突变、插入缺失、易位、融合及其他畸变。由于许多变异之前从未观察到，也不一定存在于基因组的编码区域中，故全基因组测序逐渐被视为唯一严谨的、可以发现癌症基因组中所有变异的方法。

(2)在遗传病研究中，全基因组测序可在全基因组范围内搜寻疾病相关候选位点或区域变异，特别适用于有较少分子研究基础的疾病类型或者研究导致疾病的基因组水平大的结构变异。

(3)在人类进化的研究中，可利用全基因组测序获得的遗传信息进行种群进化、种群特异性基因及区域等的筛选。

(五) 浦东解码生命科学研究院优势

拥有ABI公司和Illumina公司的全系列二代测序平台，包括Nextseq500、Hiseq 3000/4000、Miseq、Ion Torrent、Ion Proton等，能够准确迅速地对客户提供的样本进行各种个性化、针对性的全基因高通量测序服务。

拥有的深度测序平台多年使用经验，在基因变异的检测中有独特的创新之处。

浦东解码生命科学研究院具有强大的生物信息学分析能力，能够高效率地帮助用户完成海量数据分析，捕捉有用的DNA信息。同时，可根据项目的不同为客户提供量身定制的个性化数据分析服务。

根据样品的不同采用最优的实验流程。具备多年的分子生物学服务经验，可对gDNA、细胞、血液、新鲜组织、冷冻组织、石蜡切片（FFPE）、制备好的文库等多种类型的样品进行最佳方案设计及检测流程规划。

(六) 案例分析

案例1.WGS揭示多种转移性实体瘤的全基因组特征

癌细胞转移是癌症相关死亡的主要原因之一，并与不良的治疗效果有关。癌细胞转移通常会导致癌细胞和肿瘤微环境发生急剧变化，并对当前的治疗方法产生抗性。因此，更好地了解转移性肿瘤的特征有助于找到合适的个性化治疗方法，减少过度治疗并改善预后。

作者运用全基因组测序方法阐明22种转移性实体瘤样本的基因组变异。首先，作者对来自2399位癌症患者的2520对肿瘤和非癌血液样本进行了全基因组测序（WGS）和特征分析，共鉴定出7000多万个体细胞变异，包括59,472,629个单核苷酸变异，839,126个多核苷酸变异，9,598,205个DNA片段插入和缺失，以及653,452个结构变体，获得了每个转移肿瘤样本的遗传突变目录。

作者发现，转移的肿瘤组织中，普遍存在全基因组加倍（WGD）事件，在一些癌种中甚至高达80%，而在原发灶中只有30%。这可能与转移灶对化疗的耐受性有关。

除此之外，作者还发现，在单一的转移灶中，肿瘤细胞的异质性非常低，高达96%的基因突变是克隆性的。还有高达80%的抑癌基因通过不同的方式双等位失活。最后，作者发现，通过全基因组测序，能帮助62%的患者找到已经获批或者正在开展研究的抗癌药物。

作者所用研究方法总结如下：

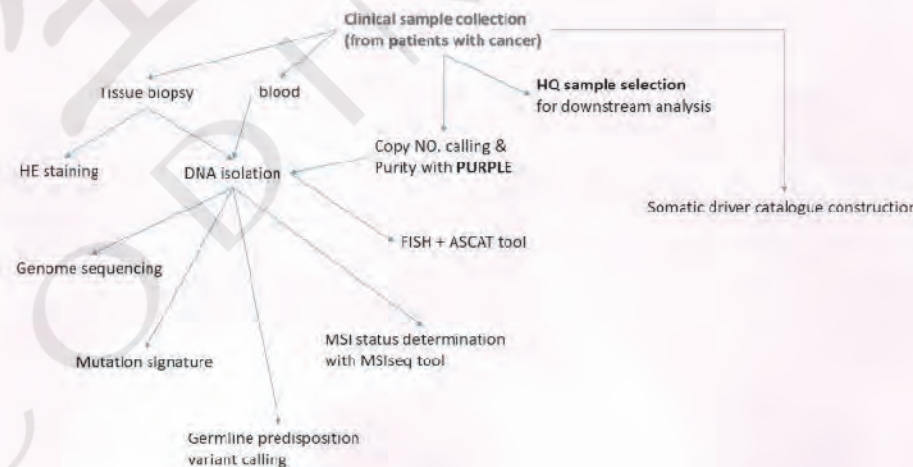


图2. 研究流程图

首先，作者根据组织的来源分析了每种癌症中的基因突变负荷。单核苷酸变异 (SNV) 的情况：黑色素瘤的中位SNV是44000个，肺癌是36000，肉瘤是4100，神经内分泌肿瘤是3500.这些突变特征与之前的研究基本匹配。多核苷酸变异 (MNV) 的情况：肺癌的中位MNV是821，皮肤癌是764，这两种癌症的MNV是其他癌症的5倍。插入和缺失 (InDel) 的中值只有SNV的十分之一左右，皮肤癌和肺癌的发生率相对较低。

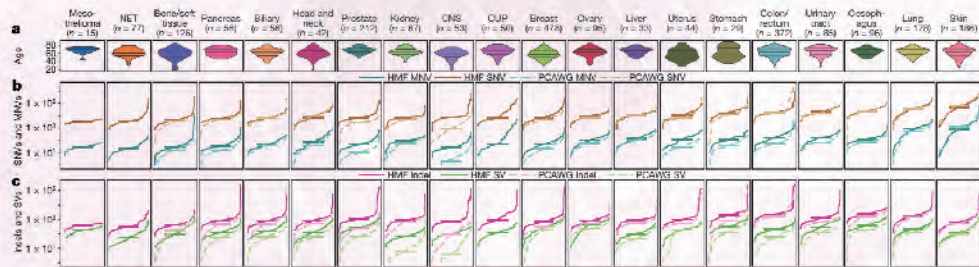


图3. 转移性肿瘤的突变负荷

之后，作者发现56%的肿瘤样本出现了全基因组加倍 (WGD) 的现象，从中枢神经肿瘤的15%到食管癌的80%。这个数据远高于之前报道的原发肿瘤的25%-37%。全基因组加倍在一定程度上增强了癌细胞对化疗的耐药性，并防止基因组不稳定、破坏性突变或者染色体片段丢失对癌细胞造成致命性影响。

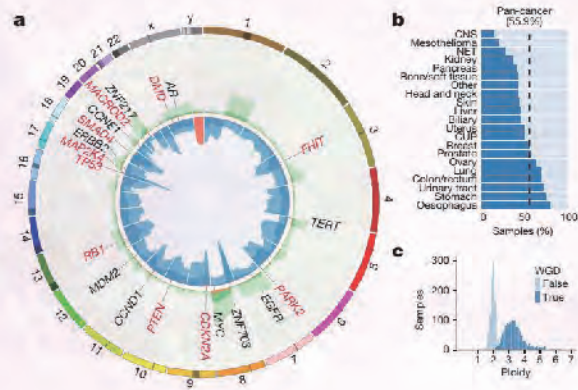


图4. 转移性肿瘤的拷贝数变化图谱

随后，作者给一些关键的癌症驱动突变基因列了个名录，结果与之前的发现类似。其中，TP53 (52%的样本)、CDKN2A (21%)、PIK3CA (16%)、APC (15%)、KARS (15%)、PTEN (13%) 和TERT (12%) 被确定为最常见的突变基因，它们共同构成目录中所有候选驱动突变的26%。

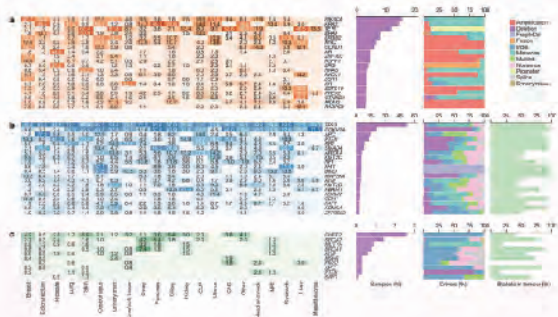


图5. 转移性肿瘤的最常见驱动基因突变

除了与原发肿瘤具有相似的驱动基因和基因组突变情况，作者还在转移性肿瘤发现了一些特殊的基因特征，可能有助于分析个别患者的治疗或耐药性反应。作者发现，在1480名患者（占有所有患者的62%）的转移灶肿瘤样本中，至少有一个与用药有关的变异。其中约一半患者的基因变异，有已经获批或者正在开展临床研究的药物可用，而且没有出现与耐药相关的基因变异。除了上面的50%有药物可用的患者之外，还有31%的患者需要的药物正处于实验室研发阶段。

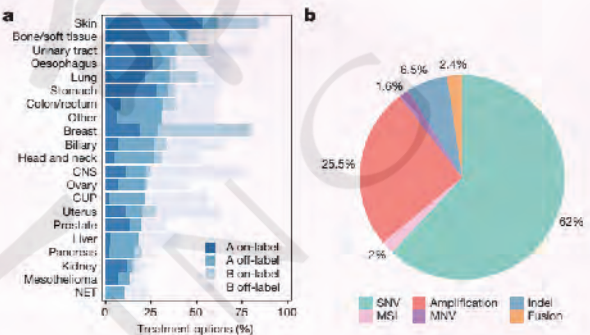


图6. 转移性肿瘤患者基因组变化的临床关联性和可操作性

总的来说，这个大型的队列研究揭示了一些癌症转移的小秘密，也证明了全基因组测序对癌症精准医疗的重要性。

参考资料

1. Pan-cancer whole-genome analyses of metastatic solid tumours. PMID: 31645765
2. Pan-cancer patterns of somatic copy number alteration. PMID: 24071852
3. Absolute quantification of somatic DNA alterations in human cancer. PMID: 22544022
4. Genome doubling shapes the evolution and prognosis of advanced cancers. PMID: 30013179

二.全外显子组测序

(一) 背景介绍

外显子是蛋白质的编码区域。外显子区域仅占人类基因组的1-2%，但却包含多达85%疾病相关的基因突变。全外显子组测序（Whole Exome Sequencing, WES），是指利用序列捕获技术将全基因组外显子区域DNA捕获富集后进行高通量测序。作为一项研究工具，结合大量的公共数据库提供的外显子数据，WES可以有效地调查在复杂的人类疾病中发挥重要作用的与蛋白质功能变异相关的常见遗传变异和稀有遗传变异。相比于WGS，WES具有如下优势：直接针对蛋白编码序列进行测序，找出影响蛋白结构的变异；测序深度高，可发现常见变异及频率低于1%的罕见变异；测序区域仅占基因组的1%，有效降低费用和工作量。总之，WES技术目标性强、性价比高。

(二) 测序原理和技术流程

全外显子捕获技术是以生物素标记的寡聚核苷酸为探针，与片段化后的基因组DNA文库杂交，通过碱基互补配对与外显子区域结合，再用链霉亲和素磁珠将外显子区域拉下来进行纯化富集，然后进行测序。

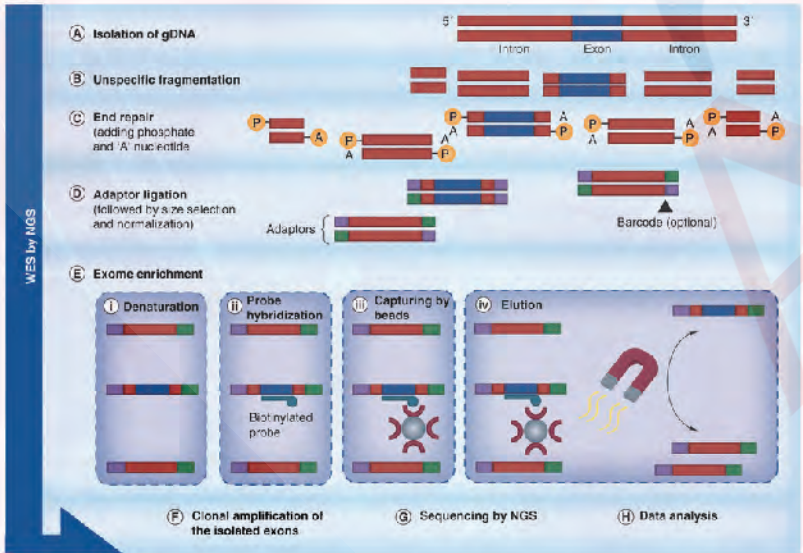


图1. WES测序原理和技术原理

(三) 生物信息学分析

疾病基因组学		
基本分析	◆ 数据质控：去除接头污染和低质量数据 ◆ 与参考序列比对、统计测序深度及覆盖度 ◆ SNP/InDel/CNV/SV检测，注释及统计	
单基因病 高级分析	◆ 突变位点筛选 ◆ 显/隐性遗传模式分析 - 显性遗传模式分析 - 隐性遗传模式分析	◆ 候选基因功能注释 ◆ 基因功能及通路分析 ◆ 家系连锁分析 ◆ 纯合子区域（ROH）分析
	◆ InDel检测与注释 ◆ 突变位点筛选 ◆ 显/隐性遗传模式分析 - 显性遗传模式分析 - 隐性遗传模式分析	◆ 共有突变筛选 ◆ 新生突变筛选 ◆ 共有突变筛选 ◆ 候选基因功能注释 ◆ 候选基因功能富集
复杂疾病 高级分析	◆ InDel检测与注释	◆ 基因显著性分析 (推荐30对case/control or trios样本)
个性化分析	◆ 药物效应多态性的遗传学机理研究 ◆ 生存分析 ◆ 疾病显著性关联位点分析 (建议基于150对以上的case/control)	◆ 蛋白互作网络分析（PPI） ◆ 疾病显著性关联基因分析 (建议基于100对以上case/control) ◆ 基于HLA型别的关联分析

(四) 应用范围

全外显子组测序主要用于识别和研究与疾病、种群进化相关的基因编码区变异，结合大量公共数据库提供的外显子组数据，有利于更好地解释基因变异和疾病之间的关联及致病机理。特别适合大样本量进行高深度测序，可发现低深度难以发现的基因变异，非常适合肿瘤研究，同时在其他疾病研究如遗传性疾病、复杂疾病，及群体研究中也应用广泛。

(五) 浦东解码生命科学研究院优势

拥有Agilent公司的多种外显子捕获平台，包括Agilent SureSelect Human All Exon V6/V7/ Clinical Research Exome V2和Agilent SureSelect Non-Human All Exon (Mouse/Bovine/Zebrafish/ Canine)等，能够准确迅速地对客户提供的样本进行各种个性化、针对性的全外显子组高通量测序服务。

具有强大的生物信息学分析能力，能够高效率地帮助用户完成海量数据分析，捕捉有用的DNA信息。同时，提供常规和高级信息分析，可根据客户需求提供个性化数据分析和作图。

根据样品的不同采用最优的实验流程。具备多年的分子生物学服务经验，可对gDNA、细胞、血液、新鲜组织、冷冻组织、石蜡切片 (FFPE)、制备好的文库等多种类型的样品进行最佳方案设计 & 检测流程规划。具备成熟稳定的微量或低起始量DNA (10 ng) 建库测序技术。

完善的售前、售后服务。经验丰富的技术支持和生物信息工程师可提供售前研究策略指导和完善的售后服务。

(六) 案例分析

案例1.全外显子测序在单基因遗传性疾病中的应用

研究背景：血浆中“低密度脂蛋白胆固醇 (LDL-C)”浓度升高是导致心脑血管疾病的主要风险。LDL-C水平受遗传和饮食双重控制，了解人体LDL-C水平的遗传调控机制是疾病诊治和医药研发的先决条件，而目前只有少数影响LDL-C的基因被鉴定出来。不同种族之间LDL-C的含量及冠心病的发病率有很大差异。

研究内容：为揭示新的胆固醇调控基因，该研究团队在针对新疆人群心脑血管疾病的风险调查中，发现了一个家族性低LDL-C的哈萨克族人家系，对三个LDL-C低浓度个体和一个正常个体进行全外显子测序，过滤后，只剩LDL-C个体的突变位点，接着通过过滤同义突变、dbSNP数据库、SIFT分值高而POLYPhen V2和Mutation Taster分值低的突变位点，发现LIMA1基因罕见移码突变 (K306fs) 与低LDL-C显著相关。深入研究发现，LIMA1特异性表达在小肠上，通过与NPC1L1蛋白互作将后者锚定到肌球蛋白Myosin Vb上，从而调控小肠胆固醇的吸收。

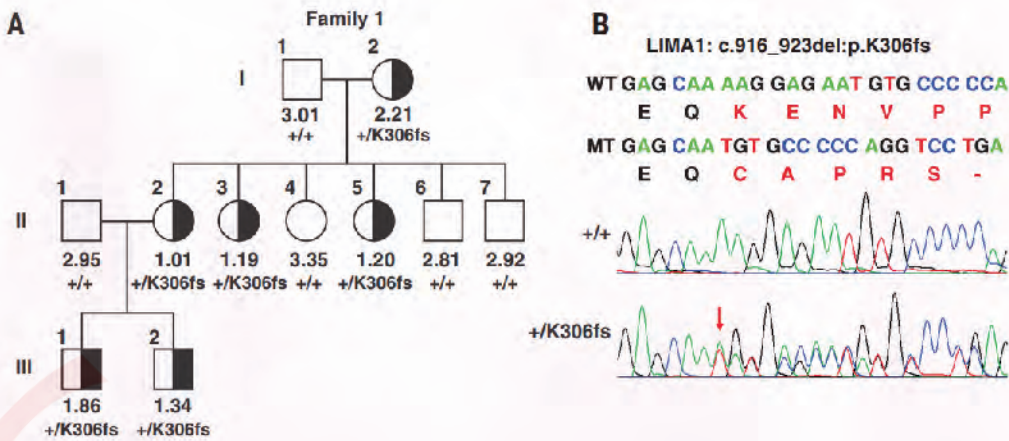


图2. 该低胆固醇的哈萨克族家系图和发现LIMA1基因变异

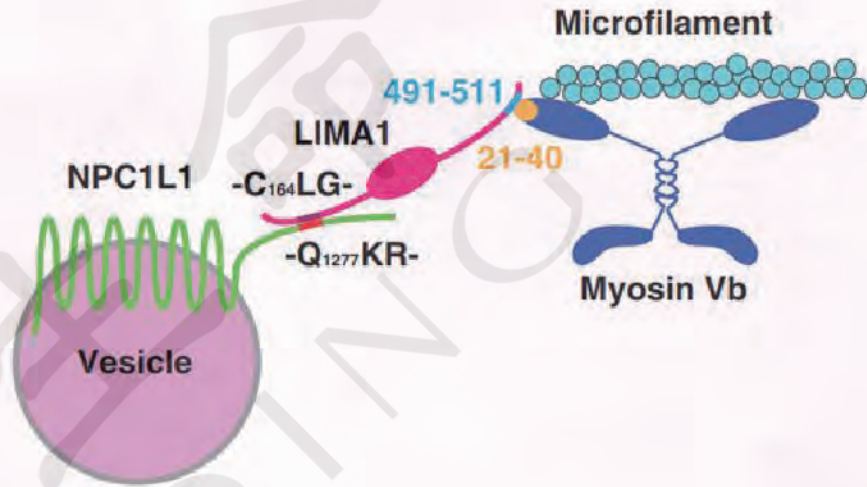


图3. LIMA1蛋白通过结合NPC1L1和Myosin Vb调控小肠胆固醇吸收

参考文献：

Zhang Y Y, Fu Z Y, Wei J, et al. A LIMA1 variant promotes low plasma LDL cholesterol and decreases intestinal cholesterol absorption[J]. Science, 2018, 360(6393): 1087-1092.

案例2.前列腺癌致癌基因的长尾效应分析

研究背景：前列腺癌具有高度的异质性。前期的前列腺癌深度基因组学分析已经鉴定出一些复发性相关的变异基因，这些基因参与雄激素信号传导，比如DNA修复和PI3K信号传导等基因。然而，更大规模的基因组学分析可以鉴定出一些其他低频的反复突变基因。

研究内容：研究人员汇总并统一分析来自1,013个前列腺癌的外显子组测序数据。鉴定和验证了一类新的由表观遗传调节因子中的突变所定义的E26转化特异性（E26 transformation-specific,ETS）-融合阴性肿瘤，以及在之前的前列腺癌研究中未涉及到的途径中的变异，如剪接体途径。同时作者还发现显著突变基因（significantly mutated genes, SMG）的突变率遵循长尾分布，许多基因的突变率不到3%。作者总共确定了97个SMG，包括70个在之前的研究中未报道的前列腺癌SMG，例如泛素连接酶CUL3和转录因子SPEN。最后，通过比较原发性和转移性前列腺癌中的突变位点信息，鉴定出一组可以预测前列腺癌危险分层的基因组标记。

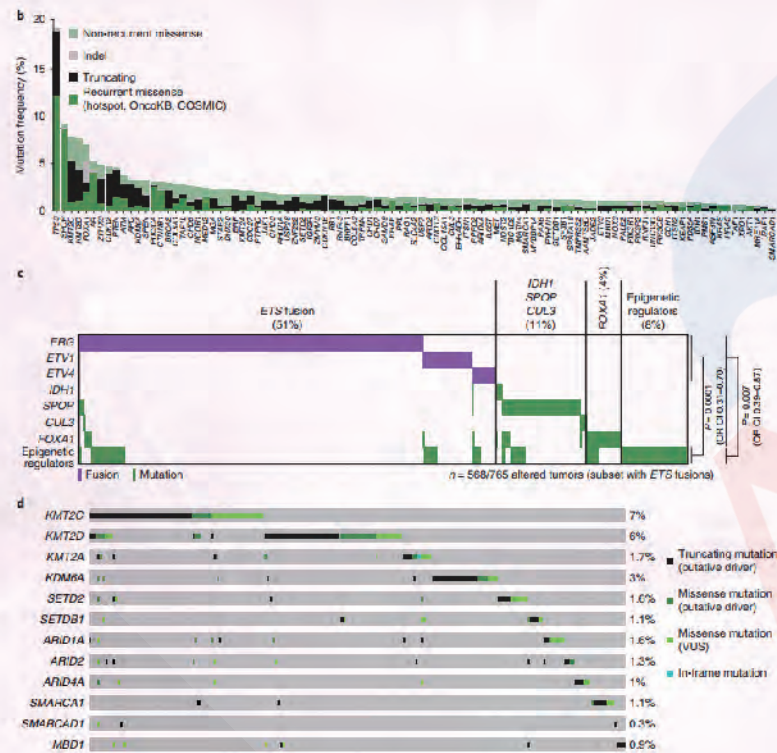


图4. 1013个前列腺癌的显著突变基因

参考文献：
Armenia J, Wankowicz S A M, Liu D, et al. The long tail of oncogenic drivers in prostate cancer[J]. Nature genetics, 2018, 50(5): 645.

三.目标区域测序

(一) 背景介绍

目标区域测序（Target Region Sequencing, TRS）是一种高度特异和靶向的方法，是指采用各种技术手段将待检测的目标区域富集之后，进行高通量测序的研究策略。目标区域富集常采用的技术手段有：探针捕获法和扩增子法。

探针捕获法是通过定制目标基因组区域的特异性探针，与基因组DNA进行杂交，将目标区域DNA片段进行富集后再利用第二代测序技术进行测序的研究策略。扩增子法是通过特异性设计的针对目标区域的引物对基因组DNA进行多重PCR扩增获得目标区域DNA，建库之后同样利用二代测序技术进行测序的研究策略。以上两种方法之间略有差别。

探针捕获法	扩增子法
富集策略	探针杂交
适用范围	100k<目标区域<100M
检测变异类型	SNP、大片段InDels、融合基因、基因扩增等
核酸起始量	100-200ng
优势	发现新的SNP、融合基因、大片段InDels等
制约因素	探针捕获效率

通过对大量样本的目标区域研究，有助于发现和验证疾病相关候选位点或基因，在临床诊断和药物开发方面有着巨大的应用潜力。

技术优势如下：

- 1.针对性强：比起全基因组及全外显子组水平的研究，目的区域测序更具有针对性，可以依赖大量的前期研究成果，获得候选染色体区域或者基于生物通路的大量候选基因。
- 2.费用低：目的待测区域较小，可对数百个样品进行快速测序，大大降低了研究成本。
- 3.信息量大：比起目的区域或者候选基因单倍型标签SNP分型的研究策略，目的区域测序可以完整覆盖整个基因区域，不仅可以获得高频SNP的分型数据，而且还可以发现低频的和个体特有的变异。
- 4.效率高：比起使用Sanger法的候选基因测序方法，基于二代测序技术的目的区域测序更加快速、高效。
- 5.高精度：目的区域的高测序深度保证了更准确的测序结果。

(二) 测序原理和技术流程

目标区域探针捕获技术是使用特异性定制的针对目标区域的生物素标记的寡聚核苷酸探针与片段化后的基因组DNA文库杂交，通过碱基互补配对与目标区域结合，再用链霉亲和素磁珠将目标区域拉下来进行纯化富集，然后进行高通量测序（如下图所示）。

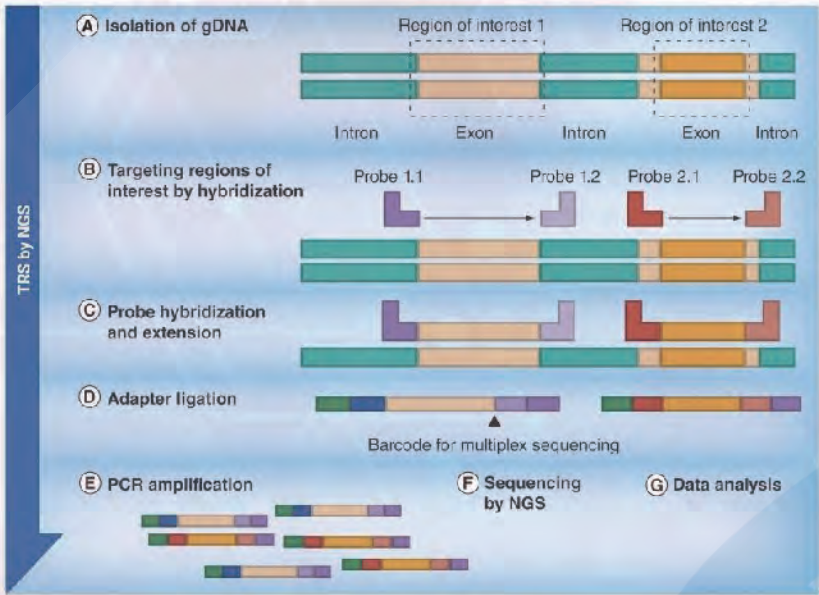


图1.TRS（探针捕获）测序原理和技术流程

扩增子测序是利用特异性引物来对感兴趣的DNA区域进行PCR扩增，形成高度富集的DNA库，将PCR产物纯化后再进行文库构建和高通量测序（如下图所示）。

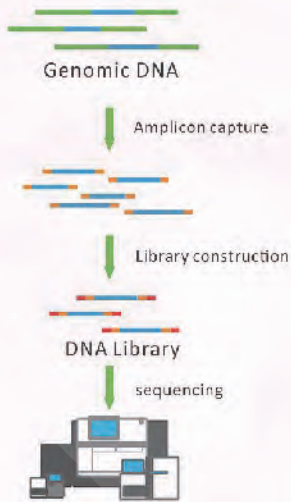


图2.TRS（扩增子）测序原理和技术流程

(三) 生物信息学分析

疾病基因组学	
基本分析	◆ 数据质控：去除接头污染和低质量数据 ◆ 与参考序列比对、统计测序深度及覆盖度 ◆ SNP/InDel检测，注释及在各基因功能元件上的统计
高级信息分析	◆ 突变位点筛选 ◆ Non-coding区突变位点筛选 ◆ 位点显著性分析（成对样本） ◆ 基因显著性分析（成对样本） ◆ 基于家系信息进行遗传模式分析 ◆ 候选基因功能注释和通路富集分析 ◆ 常见变异关联分析 ◆ 罕见变异基因负荷检测

癌症基因组学（肿瘤成对样本）	
标准分析	◆ 数据质控：去除接头污染和低质量数据 ◆ 与参考序列比对、统计测序深度及覆盖度 ◆ SNP/InDel检测、注释及统计 ◆ omatic mutation/InDel检测、注释及统计（成对样本）
高级分析	◆ 高频突变基因统计及通路富集分析（>2对样本） ◆ MRT高频突变基因相关性分析（>2对样本） ◆ 已知驱动基因筛选 ◆ 驱动基因预测（>2对样本） ◆ 突变位点分布情况分析 ◆ 高频突变基因靶向用药预测 ◆ 耐药性突变筛选

(四) 应用范围

目标区域测序适用于单基因遗传疾病的研究、多基因遗传病的研究、癌症研究、用药指导、群体研究等研究领域。

(五) 浦东解码生命科学研究院优势

- ◆ 可提供多种目标区域捕获试剂盒，来满足客户研究方向，如Agilent SureSelect Target Enrichment System、NimbleGen SeqCap EZ Choice等；也可提供扩增子法TRS测序，可根据客户需要设计特异性扩增引物；能够准确迅速地对客户提供的样本进行各种个性化、针对性的目标区域高通量测序服务。也可按客户需求个性化定制Panel，支持多个物种、任意区域的设计。
- ◆ 具有强大的生物信息学分析能力，能够高效率地帮助用户完成海量数据分析，捕捉有用的DNA信息。同时，提供常规和高级信息分析，也可根据客户需求提供个性化数据分析和作图。
- ◆ 根据样品的不同采用最优的实验流程。具备多年的分子生物学服务经验，可对gDNA、细胞、血液、新鲜组织、冷冻组织、石蜡切片（FFPE）、制备好的文库等多种类型的样品进行最佳方案设计及检测流程规划。具备成熟稳定的微量或低起始量DNA（10ng）建库测序技术。
- ◆ 完善的售前、售后服务。经验丰富的技术支持和生物信息工程师可提供售前研究策略指导和完善的售后服务。

(六) 案例分析

案例 杂合型RFX6蛋白截断突变与MODY型糖尿病的外显率降低有关
研究背景：寻找导致单基因糖尿病的新病因有助于理解人类血糖调节的机制。研究人员使用TRS对青少年发病的成人型糖尿病（maturity-onset diabetes of the young, MODY）患者进行测序（深度>10×），以期发现MODY的新致病因素。

研究内容：首先对38名表现出强MODY表型的欧洲（不包含芬兰人）先证者进行了TRS测序（含有29种已知单基因型糖尿病的Panel）。发现RTX6基因发生了杂合无义突变，也称为蛋白截断突变（PTV），这种突变会导致肽链合成提前终止。与ExAC数据库中的大量样本进行对比，结果表明，在MODY的发现队列（38人）中，RFX6 PTV的杂合突变频率显著较高（5.5% vs 0.045%）。在348个样本组成的MODY重复队列中，RFX6 PTV的杂合突变频率也显著高于gnomAD数据库中的频率（1.15% vs 0.045%）。这种突变在芬兰人中的频率更高，约为普通欧洲人群的10倍，并且均为p.His293Leufs型变异（第293位氨基酸由His转变为Leu）。

之后的研究发现，当RFX6发生PTV突变时，人群在25岁时的患病率为27%，51岁时的患病率为78%，显著低于HNF1A（25岁为70%，50岁为97%）和HNF4A（25岁为55%，50岁91%）突变，这暗示了RFX6杂合突变可以减少糖尿病的外显率。

综上，通过对未知病因的MODY病例进行目标区域靶向测序，并将PTV的频率与大型数据库进行了比较，从而确定了一种新的杂合RFX6 PTV突变与MODY发生相关。

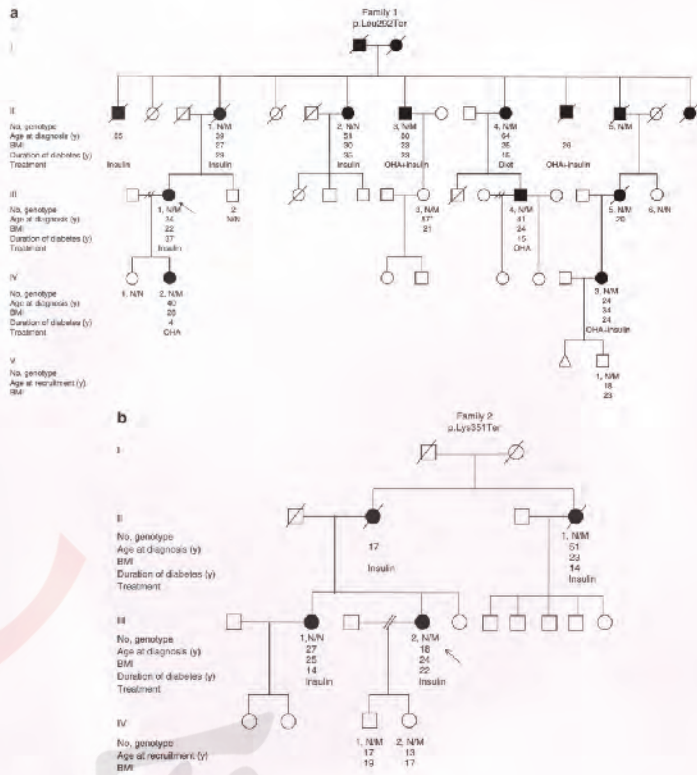


图3. MODY患者家系图

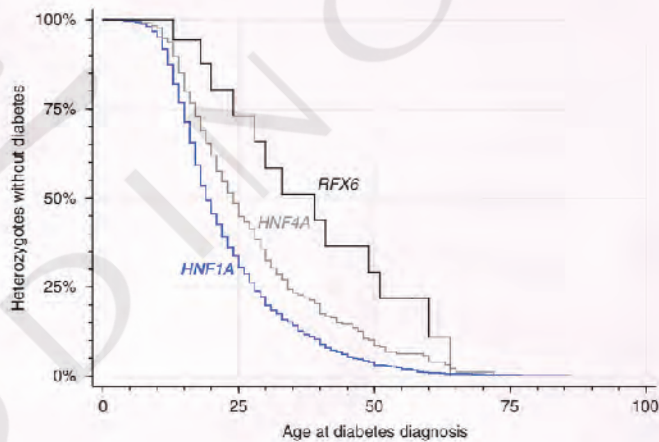


图4. MODY患者外显率变化

参考文献：

Patel K A, Kettunen J, Laakso M, et al. Heterozygous RFX6 protein truncating variants are associated with MODY with reduced penetrance[J]. Nature communications, 2017, 8(1): 888.

四.比较基因组杂交芯片 (CGH芯片)

(一) 检测原理

比较基因组杂交芯片是将消减杂交、荧光原位杂交相结合,用于检测DNA序列的变化(缺失、扩增、复制),并将其定位在染色体上,以此为基础开发的芯片技术就是比较基因组杂交芯片。aCGH的基本原理是用不同的荧光染料(常用Cyanine 3-dUTP和Cyanine 5-dUTP)分别标记对照样本基因组DNA与实验样本基因组DNA,然后与制备好的aCGH微阵列芯片进行杂交,通过扫描芯片检测染色体上两种荧光(红、绿)的相对强度比率,可以计算出DNA的缺失与扩增,从而了解实验样本基因组DNA拷贝数的改变情况,并能同时在染色体上进行精确定位。aCGH技术能在一次实验中实现全基因组扫描,检测出所有染色体的不平衡变化,并且具有较高的分辨率。见图1。

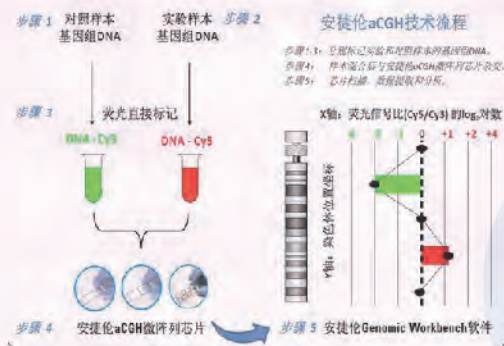


图1. aCGH的检测原理

(二) 应用领域

拷贝数变异 (CNV) 的检测:

- (1) 检测正常人中的CNV;
- (2) 检测和一些复杂疾病有关的常见或稀有的CNV。

遗传学/细胞遗传学研究:

- (1) 遗传疾病的产前诊断及产后筛查;
- (2) 检测研究病人的染色体变异情况;
- (3) 鉴定由CNV引发的遗传性疾病。

癌症研究:

- (1) 对肿瘤病人中的染色体异常进行全基因组范围研究或检测;
- (2) 寻找癌症易感性区域;

- (3) 寻找可用于肿瘤分类或分级的标志性CNV、寻找新的治疗靶向区域。

模式生物研究:

- (1) 鉴定正常的基因组变异(如进化和品系发育研究);
- (2) 在人类疾病的动物模型中鉴定疾病特异性的区域。

(三) 浦东解码生命科学研究院提供的比较基因组学芯片服务

根据客户需要可以提供全系列比较基因组芯片,以及后续标准化和个性化生物信息学服务。

(1) Agilent 比较基因组杂交 (CGH) 芯片服务

CGH芯片(aCGH)是安捷伦推出一种突破性的生物芯片平台,这种平台能够支持研究人员精确扫描所研究物种的全基因组染色体的微小变化。安捷伦提供多种规格的CGH芯片,除了人类CGH芯片,还提供其他多种物种的CGH芯片,如小鼠、大鼠及各种模式生物等。

技术特点

开放式的系统设计

安捷伦执行的是全球统一的芯片标准,既可以为客户提供全套解决方案,又可与其它技术平台或客户的自有技术衔接,为客户提供最大的自由空间。

灵敏的检测能力

60mer的探针设计赋予安捷伦芯片极佳的检测性能,这对低丰度基因的检测显得尤为重要。在美国FDA组织的全球基因芯片质量评估计划(MAQC)中,在所有参与评估的当今基因芯片主流产品中,安捷伦基因芯片能检测出最多的基因数。

独特的制作技术

无需“光掩膜”的喷墨化学原位合成专利技术提供了芯片制作极大的稳定性和灵活性,安捷伦提供高质量的各种密度和规格的芯片(1×244K, 1×1M, 4×44K, 2×105K, 8×15K, 2×400K, 4×180K, 8×60K等),满足客户的各种需求。

方便的定制服务

安捷伦推出唯一的免费芯片探针设计工具—eArray。科研客户只需简单登录即可访问安捷伦公司的各种物种、不同应用类型的基因芯片探针数据库,在其中选择符合自己要求的探针设计芯片。客户也可以上载自己专有的基因序列信息来设计芯片,安捷伦可根据提交的设计方案量身定制独有的芯片产品。

五.ChIP-on-chip芯片

(一) 背景介绍

ChIP-on-chip是研究体内DNA与蛋白质相互作用的方法。它的基本原理是在活细胞状态下把细胞内的蛋白质和DNA交联，超声波将其随机打断为一定长度范围内的染色质小片段，然后用所研究的目的蛋白质特异性抗体免疫沉淀蛋白质-DNA复合体，从而特异性地富集目的蛋白结合的DNA片段。

(二) 应用领域

(1) 蛋白与DNA相互作用； (2) 转录因子和聚合酶的DNA结合位点研

(三) 浦东解码生命科学研究院提供的ChIP-on-chip芯片服务

根据客户需要可以提供以下厂家的ChIP-on-chip芯片服务。

(1) Roche NimbleGen ChIP-on-chip芯片服务

Roche NimbleGen提供的ChIP-on-chip芯片具有极高的灵敏度。这类芯片能快速和精确地定位特定蛋白与DNA结合的区域，例如转录因子、组蛋白和聚合酶。

型号多样	高分辨率和高通量	高灵敏度和特异性	图形化显示结果
Roche NimbleGen提供覆盖全基因组、启动子和设计定制的ChIP-on-chip芯片。	2.1M芯片（单张芯片上210万探针），3×720K芯片（每张芯片3样本每样本72万探针），385K芯片（单张芯片38.5万探针），4×72K芯片（每张芯片4样本每样本7.2万探针）。	使用光敏合成50-75mer的长探针，使其具有高敏感性和特异性。双色荧光系统可消除芯片批次间的差异。可以稳定地检测到最低2倍丰度的目标结合位点。	Roche NimbleGen开发了SignalMap软件，可以非常直观地显示芯片结果，标识出DNA-蛋白结合区域。

检测原理

通过染色质免疫沉淀（chromatin immunoprecipitation）方法来富集DNA与蛋白结合区域，得到的DNA片段用Cy5和Cy3双色荧光标记分别标记IP Sample和Input Sample，之后与NimbleGen ChIP-on-chip芯片杂交来检测DNA与蛋白结合位点。见图1。

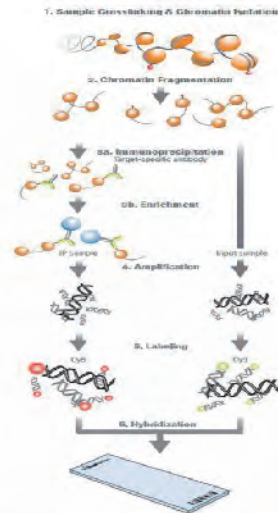


图1. ChIP-on-chip的检测原理

(2) Agilent ChIP-on-chip芯片服务

该项技术可对支配遗传表达的转录因子的结合区域进行高分辨率的精确分析。通过纯化染色质免疫共沉淀中的蛋白质-DNA结合物，可以富集DNA并进行标记。通过进行微阵列的杂交反应，可鉴别出蛋白质的结合位置。除了可探明转录因子的结合部位，利用这种技术也可以对乙酰化组蛋白、甲基化组蛋白和甲基化DNA进行分析。

技术特点

采用Agilent的Sure-Print技术生产的60-mer寡核苷酸探针阵列，玻片大小为“1×3”（25mm×75mm）。

探针设计依据UCSC Genome Browser hg17(NCBI Build 35)(human)，并可方便的进行数据更新。

产品内容灵活，既包括覆盖整个人类基因组ENCODE区域的Human ENCODE芯片（80,000 probes/2sides），也有针对目前已知的约17,000个转录调控区域，覆盖起始密码子上游—8.0kb到下游+2.0kb范围，密度高达25 probes/gene。

完善的数据处理软件（ChIP Analytics 1.1），具有图形化显示，强大的数据统计功能，简便易用。



第二章 表观遗传组学

一.全基因组甲基化测序

(一) 背景介绍

DNA甲基化是最早发现的基因修饰方式之一，可能存在于所有高等生物中，是表观遗传学(epigenetics)的重要组成部分。DNA甲基化能关闭某些基因的活性，去甲基化则可诱导基因的重新活化和表达。它参与了动物胚胎发育、基因印迹和X染色体失活等过程，在基因表达的调控中具有重要作用，异常甲基化可以导致肿瘤的形成。

全基因组甲基化测序 (Whole Genome Bisulfite Sequencing, WGBS) 其采用Bisulfite处理基因组DNA使未甲基化修饰的胞嘧啶C转化为尿嘧啶U，通过对处理后的DNA进行全基因组重测序，并与参考基因组进行比对，从基因组水平实现单碱基分辨率的、高精度度甲基化水平分析，从而构建精细的全基因组DNA甲基化图谱，广泛应用于细胞分化、组织发育等基础机制研究，以及动植物育种、人类健康与疾病治疗等研究领域。

(二) 测序原理和技术流程

WGBS原理是用Bisulfite处理，将基因组中未发生甲基化的C碱基转换成U，进行PCR扩增后变成T，与原本具有甲基化修饰的C碱基区分开来，再结合高通量测序技术，与参考序列比对，即可判断CpG/CHG/CHH位点是否发生甲基化。

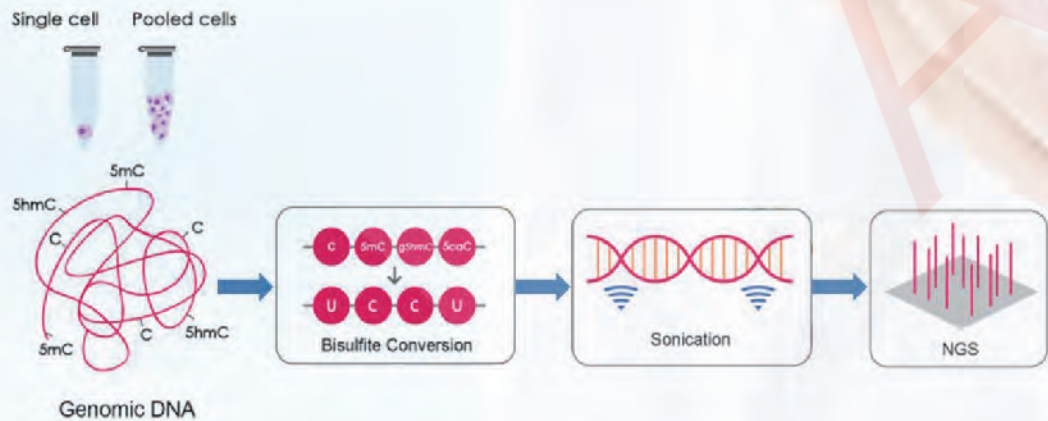


图1.WGBS测序技术流程

(三) 生物信息学分析

数据质控	◆ 去除接头、污染序列和低质量数据	
	◆ Clean reads数据量、测序质量统计	
标准分析	◆ 参考序列对比分析	◆ 甲基化位点calling
	◆ 测序数据去除重复序列分析	◆ 全基因组甲基化水平分布趋势分析
	◆ 碱基的有效深度和覆盖度分析	◆ 全基因组DNA甲基化图谱分析
高级分析	◆ 差异性甲基化区域 (DMRs) 分析	
	◆ -DMRs检查	
	-DMRs功能注释	
	-DMRs相关基因的GO富集分析	
个性化分析	◆ 药物效应多态性的遗传学机理研究	◆ 疾病显著性关联基因分析 (建议基于100对以上case/control)
	◆ 生存分析	
	◆ 疾病显著性关联位点分析 (建议基于150对以上的case/control)	◆ 基于HLA型别的关联分析

(四) 应用范围

WGBS适用于基因表达调控、肿瘤研究、染色质重塑、遗传印记、疾病研究、胚胎发育、环境与表观遗传、表观异质性等研究领域

(五) 送样要求

样本类型	适用检测项目	取样要求	保存及运输
血液样本	WGBS	3mL外周血 EDTA/柠檬酸钠抗凝 管采血，颠倒混匀； 禁用肝素抗凝管	液氮速冻， -80℃保存， 干冰运输
新鲜组织	WGBS	总量>2g 迅速取材，离体后 立即用0.9%的生理 盐水清洗	液氮速冻， -80℃保存， 干冰运输
细胞	WGBS	总量：106-107 细胞需用PBS快 速洗涤2次	-80℃保存， 干冰运输
核酸	WGBS	纯度OD260/280=1.8-2.2 浓度≥50ng/μL，总量≥5μg DNA完整纯净，无降解，无细胞凋亡， 无RNA、蛋白质、多糖、试剂成分等 残留或污染	-80℃保存， 干冰运输

(六) 浦东解码生命科学研究院优势

拥有ABI公司和Illumina公司的全系列二代测序平台,包括Nextseq 500、Hiseq 3000/4000、Miseq、Ion Torrent、Ion Proton等,能够准确迅速地对客户提供的样本进行各种个性化、针对性的Bisulfite处理后的全基因组高通量测序服务。

经过不断优化建库方案,我们的甲基化建库BS平均转化率高达99%以上,作为文库构建的质控指标进行严格控制,同时覆盖度和均一性显著提升,建库所需起始样本量不断降低。

浦东解码生命科学研究院具有强大的生物信息学分析能力,能够高效率地帮助用户完成海量数据分析,捕捉有用的全基因组甲基化信息。同时,具有丰富的多组学分析经验,也可根据项目的不同为客户提供量身定制的个性化数据分析服务。

根据样品的不同采用最优的实验流程。具备多年的分子生物学服务经验,可对gDNA、细胞、血液、新鲜组织、冷冻组织、制备好的文库等多种类型的样品进行最佳方案设计及检测流程规划。

(七) 案例分析

案例1.利用全基因组甲基化测序绘制肿瘤微环境表观遗传特征谱

研究背景: 实体瘤的发生与发展过程涉及肿瘤上皮细胞与其微环境之间的动态互作。目前,大部分关于肿瘤的表观基因组研究集中于上皮肿瘤细胞中。而有关维持肿瘤表型的肿瘤微环境的分子特征研究较少。肿瘤相关成纤维细胞(CAF)是肿瘤微环境组成的一类主要细胞。CAF可以促进前列腺癌细胞的迁移和增殖。有关CAF的表观基因组特征谱尚不清楚。

研究内容: 文章利用WGBS技术,对肿瘤相关成纤维细胞CAF和非恶性前列腺成纤维细胞NPF进行甲基化检测,筛选两者间的差异甲基化区域DMR,并进行DMR注释。结果显示CAF中的DMR富集发生于基因调控区域,通过与RNAseq的联合分析证明这些DMR会对基因的转录产生调控作用。另外CAF中DMR基因注释结果显示,DMR富集到与肿瘤相关的信号通路。CAF中筛选到的这些DMR可用于进行前列腺癌的诊断。

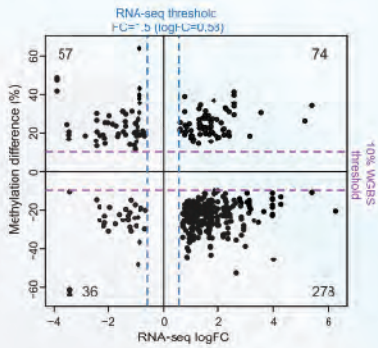


图2. 甲基化与转录组学联合分析展示DMR对DEG的调控关系

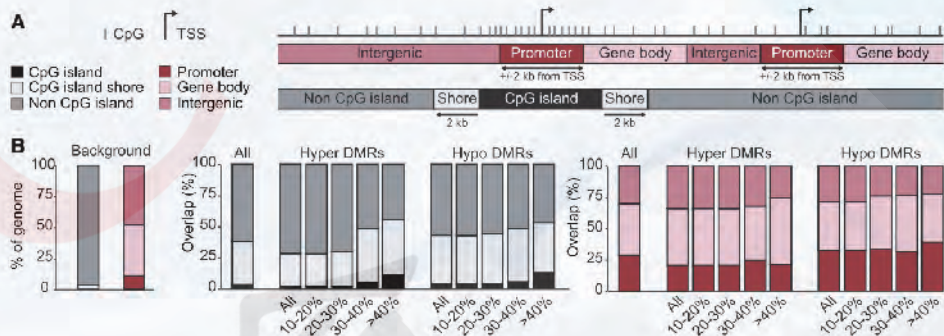


图3. DMR主要发生在基因组非CpG岛区域

参考文献:

Pidsley R, Lawrence MG, Zotenko E, et al. Enduring epigenetic landmarks define the cancer microenvironment. Genome Res 2018; 28(5): 625-638.

案例2. DNA甲基化应用于液体活检(血浆)鉴定早期肿瘤来源

研究背景: 近几年研究发现癌症无创检测技术通过液体活检可筛选出死亡癌细胞释放的ctDNA,从而能够检测癌症的早期发生,但却不能得知早期癌症发生的具体位置。

研究内容: 在该研究中,研究团队在血液中利用甲基化单倍型作为新的分子标记,不但准确检测出数量极少的癌细胞,而且还识别癌细胞在组织中的来源,为了开发该癌症早期检测方法,研究人员收集了10个不同组织(肝脏,肺等)的基因组数据,结合UCSD癌症中心癌症患者的肿瘤样品和血液样品数据进行分析,构建了不同组织癌症特异性甲基化标记数据库,该文章的方法是肿瘤ctDNA液体活检与甲基化检测的全新结合,新的技术能帮助人们更早发现肿瘤的发生位置,意义重大。

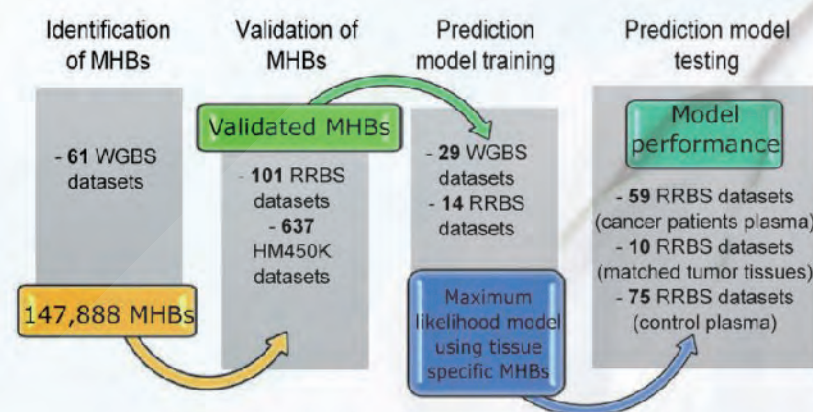


图4. 文章研究思路

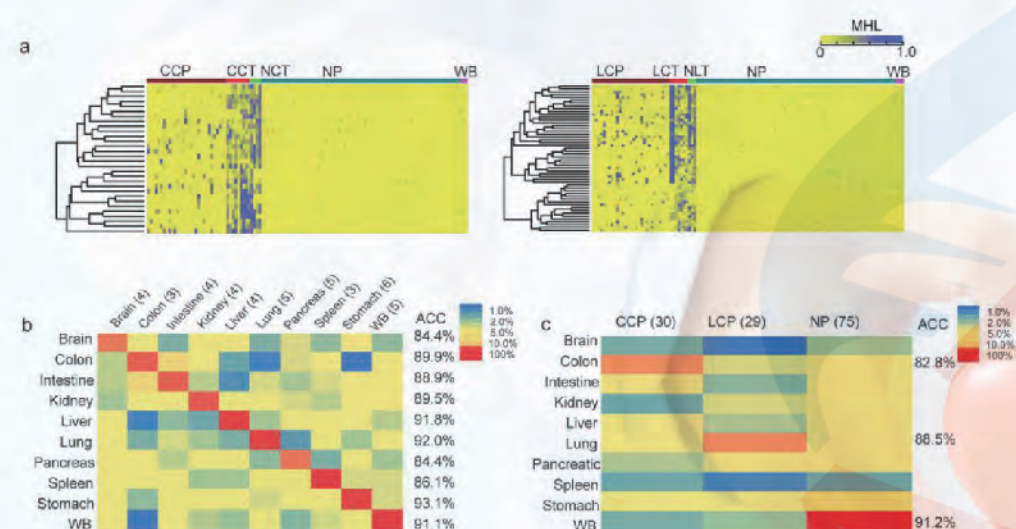


图5. 基于MHL方法进行ctDNA癌症组织来源预测

参考文献:

Guo S, Diep D, Plongthongkum N, et al. Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA.[J]. Nature Genetics, 2017, 49(4):635.

二.DNA 甲基化芯片

(一) 检测原理

DNA 甲基化检测芯片是一种将DNA 甲基化分析与基因芯片结合起来的高通量检测技术，可以对DNA 甲基化进行定性和定量分析。样本通过MeDIP或者MBD方法来富集甲基化区域，得到的DNA 片段用Cy5和Cy3双色荧光标记分别标记IP Sample和Input Sample，之后与Illumina DNA 甲基化芯片杂交来进行检测。见图1。

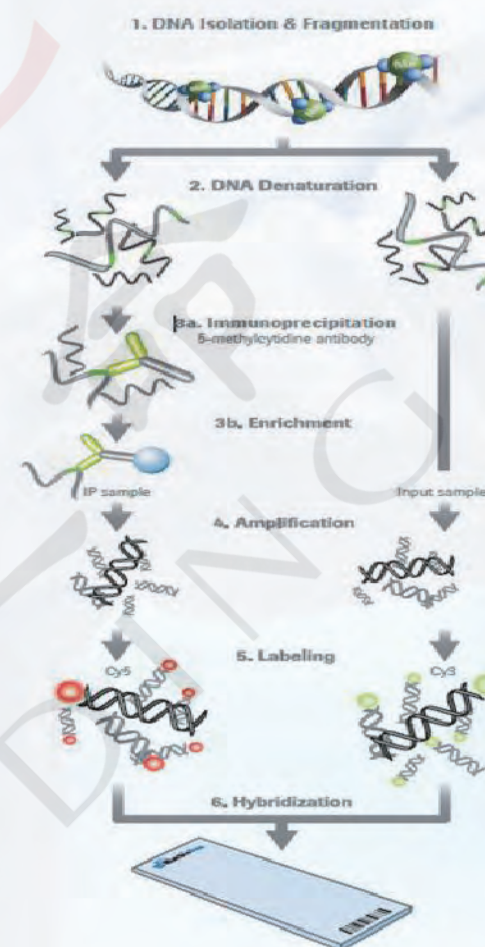


图1. DNA 甲基化检测芯片工作原理

(二) 应用领域

癌症研究

基因表达
差异研究

X-染色体失活

疾病发生
及进展

染色体结构
稳定性研究

细胞分化

发育

(三) 浦东解码生命科学研究院提供的甲基化芯片服务

解码生命科学研究院的Illumina DNA 甲基化芯片(Infinium Methylation EPIC BeadChip 850K)平台可提供甲基化检测服务, 以及后续生物信息学相关服务。

Illumina DNA 甲基化芯片服务

Illumina 850K 和 Roche NimbleGen DNA甲基化芯片具有极高的灵敏度和特异性。通过基于亲和原理的方法来富集甲基化DNA片段, 比如MeDIP (Methylated DNA immunoprecipitation), 再将其用于Illumina 850K DNA甲基化芯片检测, 可以得到包括启动子和CpG岛在内的全基因组范围或者目标范围里精确的甲基化DNA区域定位。

技术特点

💎 多种检测芯片型号: 包括全基因组、启动子和定制芯片。全基因组DNA甲基化芯片设计涵盖了人类、小鼠和植物基因组的非重复区域, 其分辨率在100bp-255bp。

💎 高分辨率和高通量: Roche NimbleGen具有目前市场上密度最高的DNA甲基化芯片平台, 提供四种格式的芯片: 2.1M (单张芯片上210万探针)、3×720K (每张芯片3样本每样本72万探针)、385K (单张芯片38.5万探针) 以及4×72K (每张芯片4样本每样本7.2万探针)。最高密度的2.1M芯片只需要10张就能以100bp的分辨率覆盖人类基因组所有非重复区域。

💎 高灵敏度和特异性: Roche NimbleGen芯片探针使用光敏合成50-75mer的长探针, 使其具有高敏感性和特异性。使用的双色荧光系统, 可消除芯片批次间的差异。

NimbleGen DNA甲基化芯片能稳定地检测最低500bp内的两个甲基化CpG岛。

💎 图形化显示结果: Roche NimbleGen开发了SignalMap软件, 可以非常直观地显示芯片结果, 标识出高甲基化区域。

三.DNA甲基化研究策略

(一) 背景介绍

从活性甲基化合物(如S-腺苷基甲硫氨酸)上催化其甲基转移到其他化合物的过程, 可形成各种甲基化合物, 或是对某些蛋白质或核酸等进行化学修饰形成甲基化产物, DNA碱基上添加甲基基团的化学修饰现象。细菌中的甲基化常发生在腺嘌呤的第6位氨基与胞嘧啶的5位碳原子上。高等生物中的甲基化主要是多核苷酸链的CpG岛上胞嘧啶的5位碳原子, 生成m5CpG。DNA的不同甲基化状态(过甲基化与去甲基化)与基因的活性和功能有关。

(二) DNA甲基化检测的应用

遗传性疾病/恶性肿瘤/动植物性状改良

(三) 检测方法

直接测序法

重亚硫酸盐可使DNA中未发生甲基化的胞嘧啶脱氨基转变成尿嘧啶, 而甲基化的胞嘧啶则保持不变。经PCR扩增所需片段, 则尿嘧啶全部转化成胸腺嘧啶, 最后对PCR产物进行测序, 与未经处理的序列比较, 即可判断CpG位点是否发生甲基化。此方法是一种可靠性及精确度很高的方法, 能明确目的片段中每一个CpG位点的甲基化状态, 但过程较为繁琐。见图1。

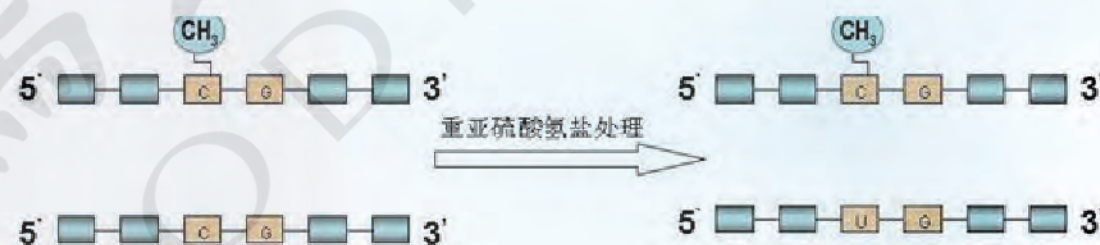


图1. DNA甲基化测序原理

甲基化特异性PCR (Methylation-specific PCR, MS-PCR)

将DNA先用重亚硫酸盐处理，未甲基化的胞嘧啶转变为尿嘧啶，而甲基化的不变。随后进行引物特异性的PCR。MS-PCR应按照要求设计两对引物：

- (1) 引物末端均设计至检测位点结束；
- (2) 两对引物分别只能与重亚硫酸盐处理后的序列互补配对，即一对结合处理后的甲基化DNA链，另一对结合处理后的非甲基化DNA链。检测MSP扩增产物，如果用针对处理后甲基化DNA链的引物能扩增出片段，则说明被检测的位点存在甲基化；若用针对处理后的非甲基化DNA链的引物扩增出片段，则说明被检测的位点不存在甲基化。

MS-PCR的实验流程见图2。MS-PCR的优点是：避免了使用限制性内切酶及其后续的相关问题，敏感性高，可用于石蜡包埋样本。其缺点是：要预先知道待测片段DNA的序列；引物设计至关重要；若待测DNA中5-甲基胞嘧啶分布极不均衡，则检测时较为复杂；MS-PCR只能作定性研究，即只能明确是否存在甲基化。若要求定量，则需用其它方法进一步检测；存在重亚硫酸盐处理不完全导致的假阳性。

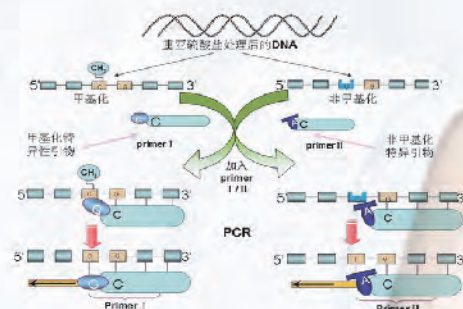


图2. MS-PCR原理

甲基化特异性多连接依赖性探针扩增 (Methylation-Specific multiplex ligation-dependent probe amplification, MS-MIPA)

本方法针对甲基化和非甲基化的位点分别设计两个探针，每个探针都包括两个寡核苷酸部分，其中短的部分由合成产生，长的部分来源于phageM13的衍生物，后者有一个非杂交填充片段，不同探针其长度不同。探针的两个寡核苷酸杂交序列均与目标序列互补，其末端都连有相同的PCR引物。要求设计的MS-MIPA探针具有甲基化敏感的限制性内切酶Hha I (GCGC) 或Hpa II (CCGG) 的识别位点。将探针和目标序列杂交后，降低反应体系温度，并向其中加入连接酶Hha I。若原样本DNA中含有Hha I识别的非甲基化位点，则非甲基化探针的两个寡核苷酸部分不能连接，而甲基化的探针顺利连接而不被切割。在随后的PCR反应中，只有两个寡核苷酸部分连接后的探针才能被扩增。最后分析扩增片段，确定待研究位点的甲基化情况。见图3。



图3. MS-MIPA原理

高通量测序

近两年，多个研究小组将传统的甲基化工具（如DNA的亚硫酸氢盐转化）与目标基因组捕获技术和高通量测序相结合，绘制出了多张甲基化图谱。而新一代测序技术的出现，更是让甲基化的直接测定成为可能。全基因组甲基化测序已在上文中具体介绍，这里不再赘述。

基于芯片的甲基化图谱分析

就甲基化图谱分析而言，目前流行的分析方法是甲基化芯片。很多公司都提供了这种工具，包括Agilent、Roche NimbleGen、Illumina和Affymetrix。平台不同，过程也各异。Agilent和NimbleGen的甲基化芯片，将基因组DNA分成两份，一份用来做MeDIP，另一份作为对照。两个样品都标记荧光（样品用Cy5标记，对照用Cy3标记），然后与芯片杂交。芯片上每个探针的Cy5/Cy3比例显示出该区域的甲基化程度。

Agilent的244,000-element array覆盖了27000多个人CpG岛和甲基化不足区域 (UMR)。NimbleGen的Human 2.1M Deluxe Promoter array也覆盖了27000多个CpG岛，还包括了27000多个启动子区域。以上产品都具有100 bp的分辨率。然而这两个平台都不能以单核苷酸的分辨率报告甲基化状态，但是Illumina的Infinium和GoldenGate的甲基化芯片可以做到。

Illumina的Infinium HumanMethylation27 BeadChip芯片覆盖了27,578个CpG位点。利用两个位点特异探针可检测出位点上的甲基化差异。一个探针是为甲基化位点 (M磁珠类型) 设计的，而另一个是为未甲基化位点 (U磁珠类型) 设计的。探针的单碱基延伸掺入了一个标记的ddNTP，它随后被荧光试剂染色。通过计算甲基化与未甲基化位点的荧光信号比例，可确定确定位点的甲基化水平。

(四) 浦东解码生命科学研究院的优势

DNA甲基化在疾病诊断、预后以及新药研发中有着巨大的应用前景。浦东解码生命科学研究院能够提供所有的DNA甲基化检测和数据个性化分析服务。

第三章 转录组学

一.转录组测序

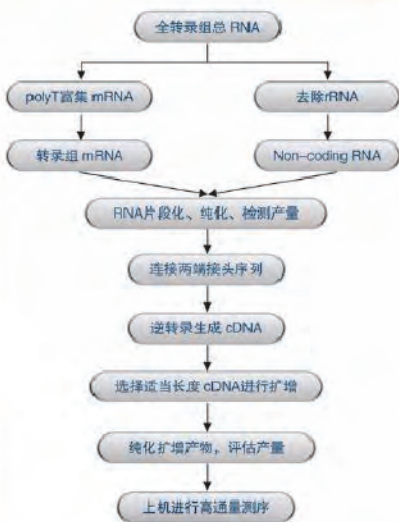
(一) 背景介绍

转录组在广义上指某一生理条件下,细胞内所有转录产物的集合,包括信使RNA (mRNA)、核糖体RNA (rRNA)、转运RNA (tRNA) 以及非编码RNA (ncRNA); 在狭义上是指所有mRNA的集合。转录组研究是基因功能及结构研究的基础和出发点。通过二代测序,能够全面地获得某一物种的特定组织或器官在某一状态下的几乎所有转录本序列信息,从而准确地分析基因表达差异、发现新转录本、基因结构变异、基因融合、RNA编辑等生命科学的重要基础问题,转录组测序已经广泛应用于基础研究、临床科研和药物研发等领域。

RNA-Seq是通过PolyT捕获真核生物的mRNA的3'端PolyA尾巴结构进行建库的测序方法,它能够检测所有的差异表达或结构变异的mRNA,以及筛选与疾病或性状相关的分子标志物。RNA-Seq可以揭示生物转录组的复杂性,确定基因以及转录本结构、可变剪接、RNA编辑、带PolyA尾的非编码RNA和新型转录本。

(二) 转录组测序流程

针对不同种类的RNA,可采用不同的方法进行分离和富集,然后进行测序。见图1。



1. 转录组样品制备与测序流程

(三) 测序模式

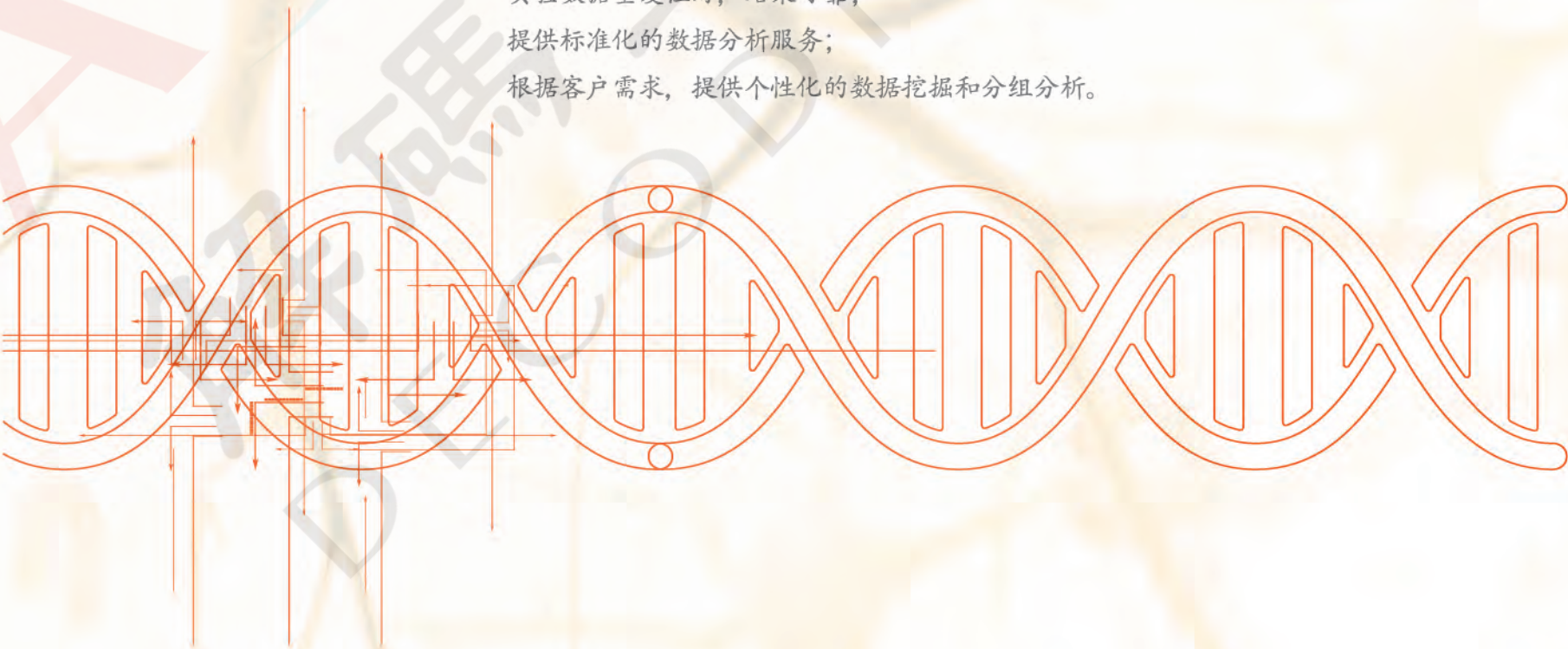
Illumina HiSeq, PE150, 大于等于6G clean reads.

(四) 应用领域



(五) 浦东解码生命科学研究院的优势

- 一次获取已知和新的转录本信息, 并实现可变剪切定量分析;
- 实验数据重复性好, 结果可靠;
- 提供标准化的数据分析服务;
- 根据客户需求, 提供个性化的数据挖掘和分组分析。



二.lncRNA测序

(一) 背景介绍

长链非编码RNA (Long non-coding RNA, lncRNA) 是一类长度在200nt以上、不编码蛋白质的RNA分子。lncRNA在发现之初,被认为是RNA聚合酶II转录的副产物,不具备生物学功能。然而,近几年的研究表明,lncRNA具有保守的二级结构,可以与蛋白、DNA和RNA相互作用,参与多种生物学过程的调控,如染色质修饰、转录激活或抑制、转录后调控以及作为miRNA的吸附海绵而干扰靶基因表达等。

通过lncRNA测序,可以发现与特定疾病相关的潜在lncRNA,并对后续的功能研究有重要的指导意义。上海浦东解码生命科学研究院通过去除rRNA,并富集lncRNA和mRNA,建库测序,可以同时分析lncRNA和mRNA的表达情况,并发现大量新的lncRNA。

(二) 研究思路



(三) lncRNA测序流程

- (1) 提取样本Total RNA;
- (2) 去除rRNA;
- (3) 构建测序文库;
- (4) 上机测序;
- (5) 原始测序数据过滤清除;
- (6) 比对、去除rRNA数据;
- (7) 比对参考基因组;
- (8) 数据标准化和个性化分析。

(四) 测序模式:

Illumina HiSeq, PE150, 12G clean reads.

三.miRNA测序

(一) 背景介绍

Small RNA主要包括miRNA、piRNA、snoRNA、和tsRNA等微小RNA分子,微小RNA参与基因表达与调控、RNA加工与剪切、蛋白质翻译、遗传“入侵”抑制、配子生成等重要生物学过程,是生命活动中不可少的调控因子,miRNA是微小RNA中最主要的一类,与癌症、心血管疾病、免疫疾病等密切相关。

(二) 技术特点

- (1) 能检测单个碱基差异,精准区分相同家族或序列相似度极高的miRNA;
- (2) 延伸mRNA之外的非编码RNA研究领域,实现转录组功能研究的全覆盖;
- (3) 可以鉴定已知的miRNA,也能检测和预测新的miRNA。

(三) 测序模式

Illumina HiSeq, SE50, 10M clean reads.

四.circRNA测序

(一) 背景介绍

环状RNA (circular RNA, circRNA) 是近些年发现的一类特殊的非编码RNA, 它大量存在于真核细胞胞质内, 主要由pre-mRNA通过可变剪切加工产生, 具有很多重要的调控功能。circRNA可作为竞争性内源RNA (ceRNA) 结合胞内miRNA, 阻断miRNA对靶基因表达的抑制作用。此外, circRNA也具有调控其他类型的RNA、调节蛋白质活性等功能。circRNA在不同物种中具有保守性且存在组织表达差异, 环状结构对RNase不敏感, 因而比线性RNA更加稳定。研究表明, circRNA在开发新型疾病诊断与治疗上具有重大潜力和研究价值。

(二) 技术特点

- (1) 去rRNA、去线性RNA, 针对环状RNA建库, 排除了线性RNA干扰;
- (2) 有效提高数据可靠性和检测效率, 可检测到低丰度的稀有环状RNA;
- (3) 针对环状RNA开发了相应的算法和分析方法, 深度挖掘环状RNA生物学功能、发现新的环状RNA。

(三) circRNA测序流程

- (1) 提取样本Total RNA;
- (2) 去除rRNA;
- (3) RNase R消化
- (4) 构建测序文库;
- (5) 上机测序;
- (6) 原始测序数据过滤清除;
- (7) 高质量测序数据比对参考基因组;
- (8) 数据标准化和个性化分析;
- (9) 筛选和研究目标circRNA。

(四) 测序模式

Illumina HiSeq, PE150, 10~12G clean data.

(五) 浦东解码生命科学研究院的优势

- (1) 一次获取已知的和新的circRNA;
- (2) 实验数据重复性好, 结果可靠
- (3) 提供标准化的数据分析服务;
- (4) 构根据客户需求, 提供个性化的数据挖掘分析。

五.全转录组测序

(一) 测序原理

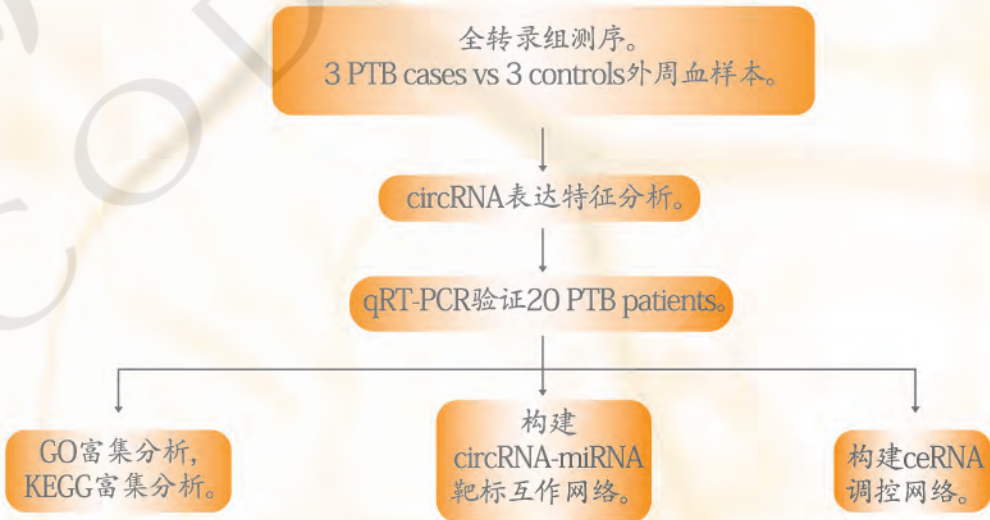
通过全转录组测序可以快速地获取某一条件下组织和细胞中大部分转录本的表达情况。与常规转录组测序相比, 全转录组测序可以同时获得mRNA、lncRNA和circRNA三种类型RNA的序列和丰度信息。既可研究已知基因, 也能发掘新基因。

(二) 案例分析

案例1. 鉴别肺结核中circRNA的表达特征及构建ceRNA调控网络

研究背景: 肺结核 (PTB) 是由结核杆菌引起的慢性疾病, 对人类健康构成严重威胁。circRNA被认为在许多疾病的发病机理和发展过程中具有潜在的调控作用。然而, 肺结核中circRNA的表达模式和调控机理有待于研究。

研究内容: 通过全转录组测序探究肺结核中circRNA的表达特征。阐明肺结核中circRNA的表达特征, 通过GO和KEGG富集分析对circRNA进行了功能富集分析, 构建了肺结核中circRNA的ceRNA调控网络。



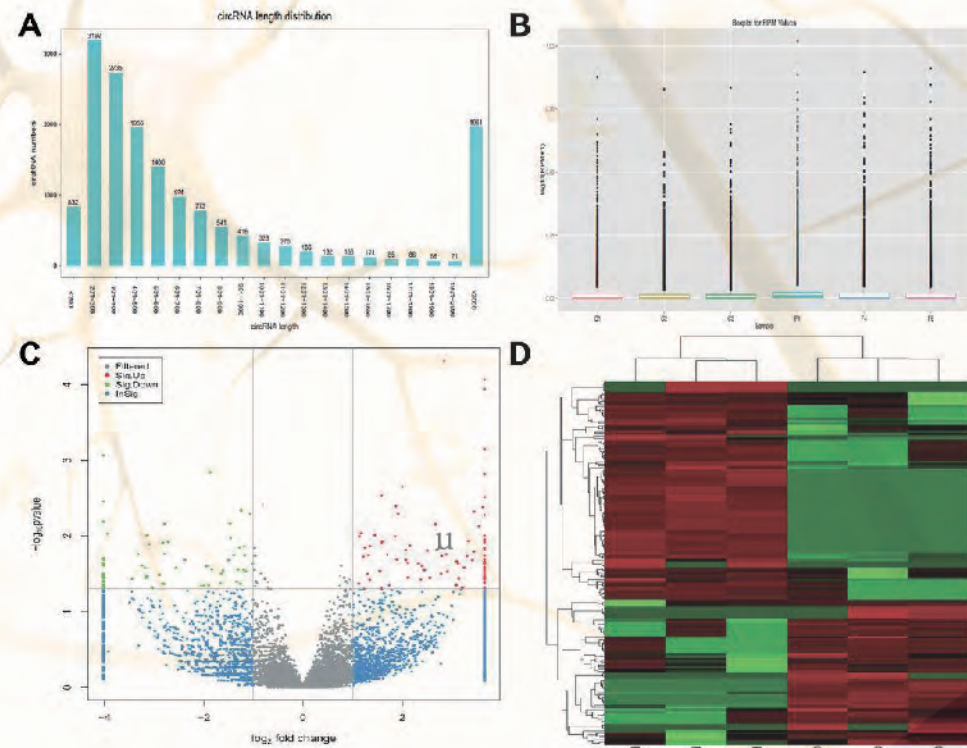


图1.利用全转录组测序检测肺结核(PTB)患者和健康个体的环状RNA表达模式

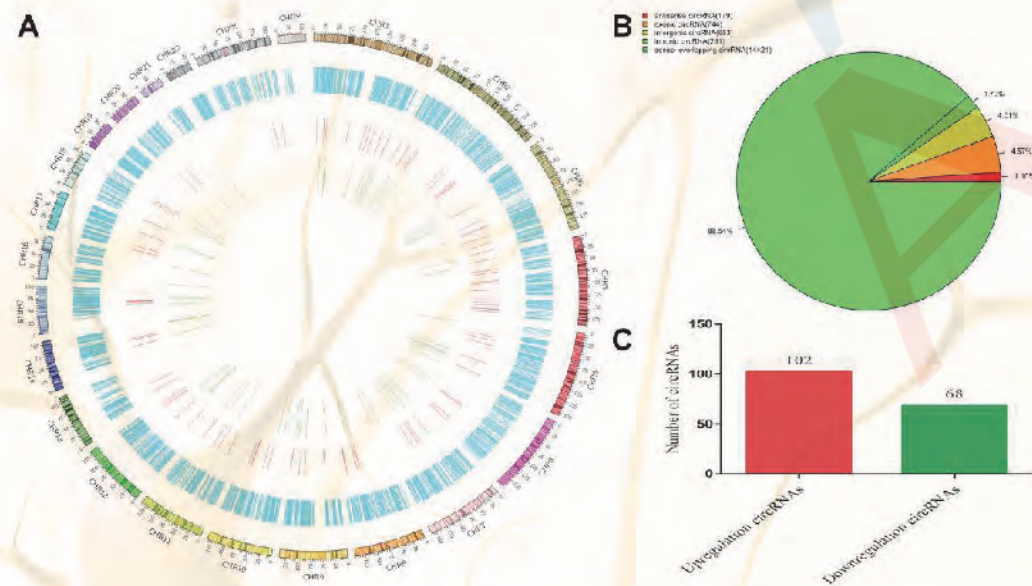


图2.鉴定肺结核(PTB)患者和健康个体中差异性表达的circRNAs

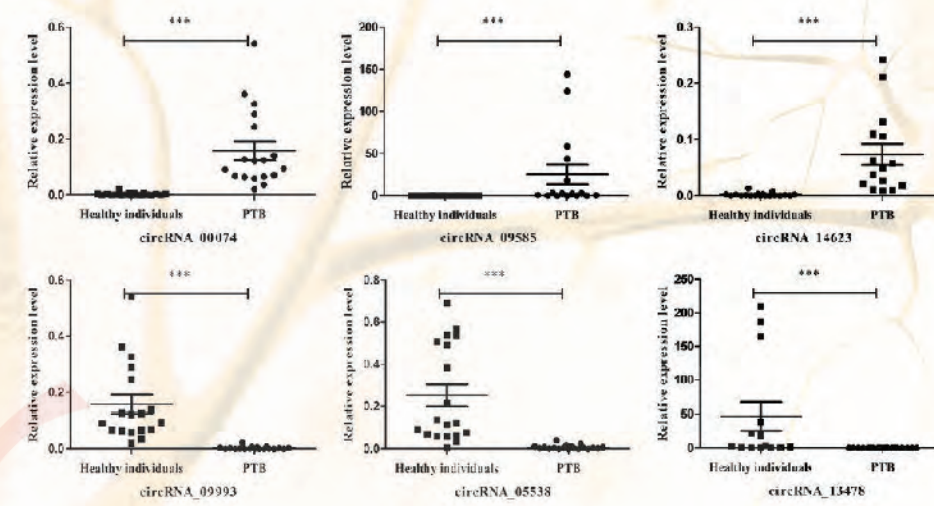


图3. RT-qPCR验证候选circRNAs表达水平

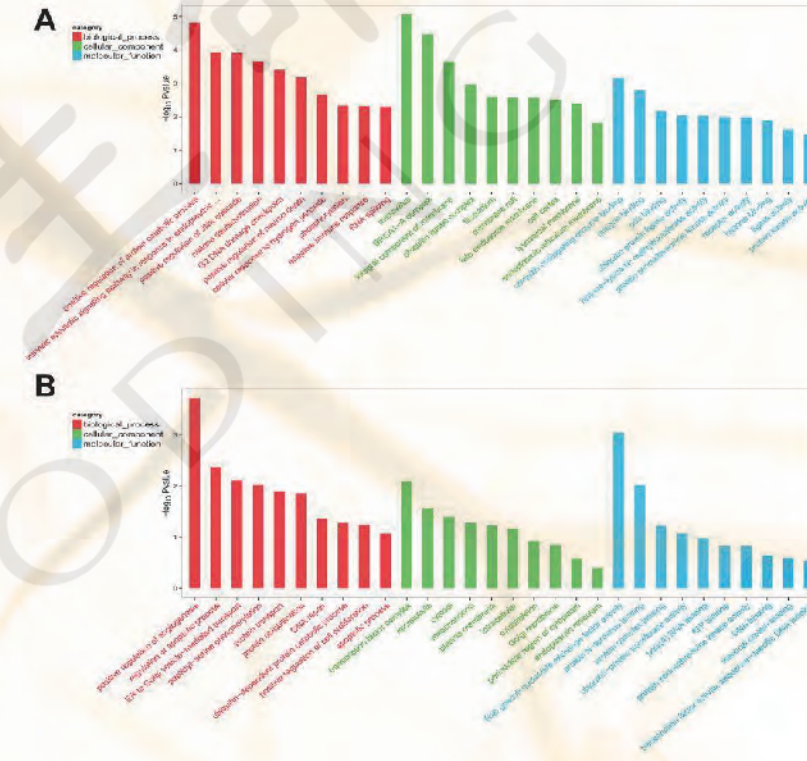


图4.肺结核(PTB)患者和健康个体中差异性表达的circRNAs富集的前20个GO注释

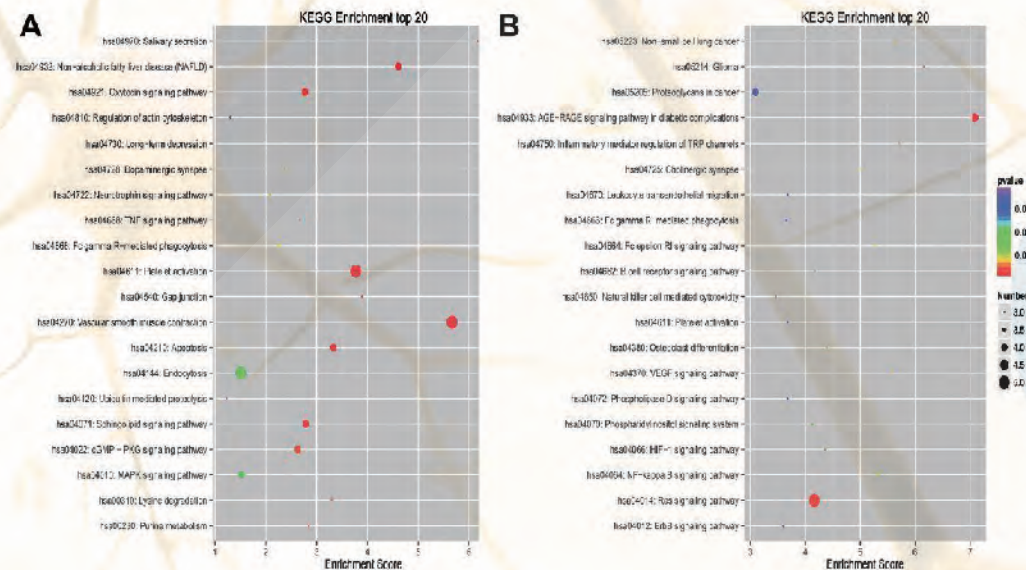


图5. 肺结核(PTB)患者和健康个体中差异性表达的circRNAs富集的前20个KEGG通路

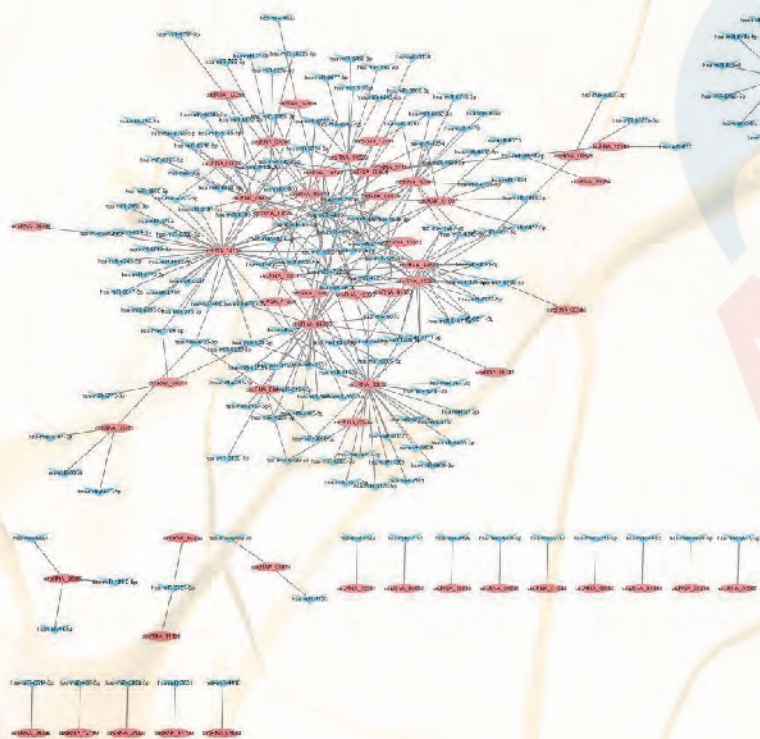


图6. circRNA-miRNA共表达网络

研究结果:

- 1.采集3例肺结核患者和3个正常人的全血样品进行全转录组测序, 筛选发现170个circRNA的表达量发生了显著性变化;
- 2.在20例患者中对6个circRNA的表达进行验证, 结果与全转录组测序一致;
- 3.GO和KEGG功能分析发现, 差异表达circRNA在一些免疫通路如MAPK信号传导通路、癌症中的内吞作用通路出现富集;
- 4.构建肺结核中miRNA介导的circRNA-mRNA的相互网络, 即ceRNA调控网络。

参考文献:

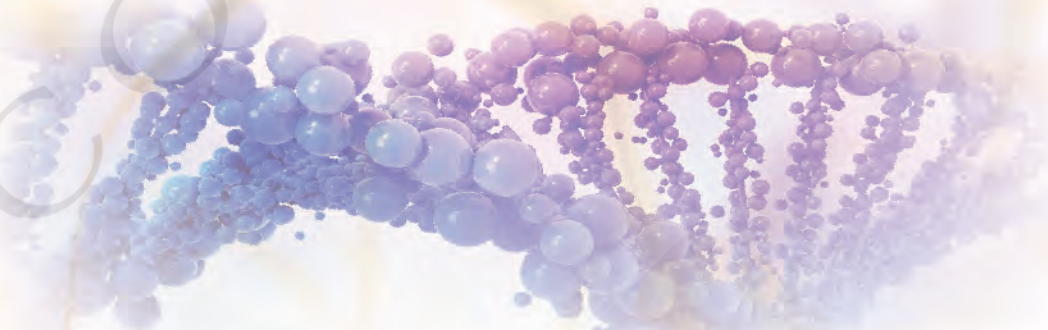
1.Zhang X, Zhu M, Yang R, et al. Identification and comparison of novel circular RNAs with associated-expression and competing endogenous RNA networks in pulmonary tuberculosis. Oncotarget,2017;8(69):113571-113582.(IF:5.168).

2.Zhang S, Zhu D, Li H, et al. Analyses of mRNA Profiling through RNA Sequencing on a SAMP8 Mouse Model in Response to Ginsenoside Rg1 and Rb1 Treatment. Frontiers in Pharmacology,2017,8. (IF:3.920).

3.Patterson AM, Mulder IE, Travis AJ, et al. Human Gut Symbiont Roseburia hominis Promotes and Regulates Innate Immunity. Frontiers in Immunology, 2017,8(8). (IF:4.710).

4.Lagarde J, Uszczyńska-Ratajczak B, Carbonell S, et al. High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. Nat Genet,2017,49(12):1731. (IF:27.125).

5.Jiang M, Zhang S, Yang Z, et al. Self-Recognition of an Inducible Host lncRNA by RIG-I Feedback Restricts Innate Immune Response. Cell, 2018.(IF:31.398).



六.表达谱芯片检测

(一) 检测原理

表达谱芯片是采用cDNA或寡核苷酸片段作探针，固化在芯片上，将待测样品（处理组）与对照样品的mRNA以两种不同的荧光分子进行标记，然后同时与芯片进行杂交，通过分析两种样品与探针杂交的荧光强度的比值，来检测基因表达水平的变化。

(二) 应用领域

01

疾病研究：

研究疾病的基因表达时空特征，揭示疾病的分子基础；研究疾病的基因表达差异，寻找疾病分型的分子指标；研究疾病治疗前后的基因表达差异，从基因组学上评估其时效性。

02

临床诊断：

可用于某些疾病的临床诊断。

03

药物筛选：

药物药理机制分析，药物的毒性毒理研究。

(三) 浦东解码生命科学研究院提供全基因组表达谱芯片检测服务

客户根据需要，可以选择不同厂家的全基因组表达谱芯片，包括Agilent、Illumina、Affymetrix等平台，可为您提供全基因组表达谱芯片分析以及后续的生物信息学检测分析服务。

(1) Agilent全基因组表达谱芯片

安捷伦高性能的表达谱芯片为客户的基因表达研究提供了灵活以及高性价比的研究平台，产品芯片覆盖了30多种不同物种，配备以高品质的60-mer寡核苷酸针，设计出以4×44K规格为主的表达谱芯片。此外客户可以根据自己的需要在安捷伦的eArray在线服务平台设计任何需要的芯片，不受目录芯片限制。

检测原理：用两种不同的荧光分别标记实验组和对照组扩增后的cRNA，调整好两组的杂交用量，然后与表达谱芯片杂交，通过扫描芯片上两种荧光（红、绿）的相对强度比率，两组RNA有差异的部分会出现颜色偏移，从而计算出相应基因表达量是上调还是下调。如果选择做单色荧光标记的表达谱芯片检测，只需用一种荧光标记一个样本进行杂交即可。

技术特点

独特的SurePrint原位合成技术：SurePrint技术依靠强大的探针设计运算法则和准确的确认方法，无需“光掩膜”，并利用工业规模的喷墨打印技术和独特的质控程序来合成芯片。

双色芯片检测：可以根据客户的需要，将相互比较的两个样本分别用不同的荧光标记，杂交在同一张芯片上。相较于到传统的单色芯片检测，此方法可以发现更微小的表达差异，并有效地控制实验的批次差异。

用量少：只需200ng 总RNA 即可用来进行标记实验（不包括质检所消耗的量），大大降低了RNA抽提的难度，能让客户利用剩下的总RNA进行验证。

灵活方便的定制服务：安捷伦公司除了提供成品芯片外，还开放了一个免费的eArray芯片设计平台，让客户可以根据自己的需求设计专有的芯片，同时亦可让客户分享其他客户的定制芯片，满足客户不同的研究需要。

(2) Illumina全基因组表达谱芯片

Illumina全基因组表达芯片是利用独特的微珠阵列BeadArray技术生产的，具有非常高的灵敏度、精确性和信息含量，并大幅度降低了成本。

检测原理：微珠采用3.1 μm无孔无荧光硅珠，每个微珠均单独质控，为芯片杂交提供了有效的反应表面积和体积比。探针在合成后与微珠进行偶联，每个微珠可偶联近百万个寡核苷酸探针，每个微珠分别代表不同的检测信息，如基因序列、SNP 位点等。每种类型的探针由特异的地址序列和探针序列组成，地址序列用于对每种微珠进行解码，探针序列对基因信息通过专利的扫描技术进行确定。每种类型的微珠（微珠+探针）单独反应，连接后混合到一起形成“微珠池”。

技术特点

- 💎 高密度：每平方厘米面积上可有约400万个点。
- 💎 低上样量：每个芯片在一轮反转录的情况下仅需50-100ng RNA。
- 💎 高重复性：其设计采取了“无序自组装”方式，每种类型微珠30倍重复，“片中片”的设计特点保证了它的高重复性。
- 💎 高特异性探针：Illumina通过生物信息学分析，针对每个基因设计2-3个候选探针，这些探针被用来和26种不同组织样本进行杂交实验，所得数据通过专利算法分析后，选择特异性和反应性能最好的唯一探针制成芯片。这种设计充分保证了其特异性及高灵敏度。
- 💎 实验过程全程监控：芯片设计了几百个内参微珠对实验进行监控，其中包括：样品质量、标记反应成功率、杂交严谨性、信号产生等。
- 💎 高均一性：芯片采取了微珠和光纤的设计，芯片点大小、点与点之间间距完全均一。

(3) Affymetrix基因表达谱芯片

Affymetrix是原位合成寡核苷酸技术的首创者，是基因芯片的早期开发者。Affymetrix专利的寡核苷酸原位光刻合成技术使得芯片上的信息储量最大化，使研究工作者们可以通过全基因组研究对遗传学与健康之间的关系进行分析。其产品具有种类丰富、信息量大的特点。

技术特点

- 💎 光蚀刻原位合成的寡核苷酸芯片：该方法的最大优点在于用很少的步骤可以合成大量的DNA阵列。
- 💎 高密度的点阵技术：1个平方厘米的面积上至少可以排列几十万到一百多万个探针合成点。
- 💎 独特的PM-MM探针设计：使得特异性和灵敏度大大提高，芯片结果准确可靠；
- 💎 极其丰富的数据信息。



七.microRNA芯片检测

(一) 原理介绍

微小RNA (miRNA) 芯片将寡核苷酸先固定在芯片介质上, 再通过杂交方式与样品中的miRNA结合, 以鉴定miRNA的表达与否及表达高低的高通量检测技术。

(二) 应用领域

肿瘤研究: 比较癌症和癌旁组织中miRNA表达差异; 寻找抑癌药物靶点。

其它疾病研究: 疾病组织miRNA的表达特点。

药物研发: 寻找药物靶点; 研究药物途经。

(三) 检测原理

用Cy-3对总RNA样品进行标记, 然后与miRNA芯片杂交。通过扫描芯片上绿光信号的强度, 计算出该样品中miRNA的表达量, 从而了解该样品中哪些miRNA的表达量出现异常。见图1。

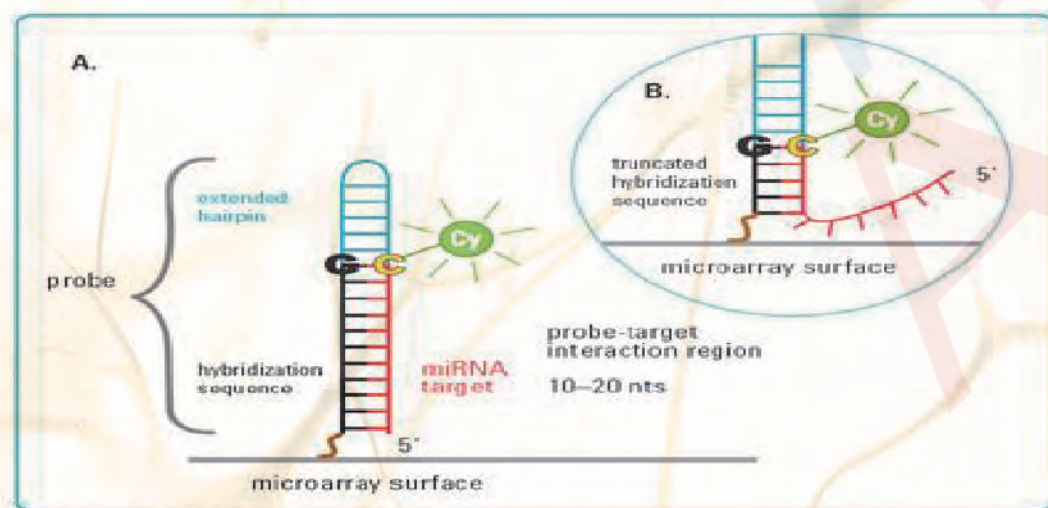


图1. miRNA表达谱芯片

(四) 浦东解码生命科学研究院为您提供的miRNA芯片服务

Agilent高性能miRNA芯片平台, 配备以最优化的试剂和硬件设备, 为miRNA的表达研究提供了一个动态的检测范围。利用SurePrint技术合成的探针, 其优越的敏感性和独特性使得近亲相关的miRNA依然能够有效地被分辨出来。

技术特点

独特的SurePrint原位合成技术: SurePrint技术依靠强大的探针设计运算法则和准确的确认方法, 利用工业规模的喷墨打印技术和独特的质控程序来合成芯片。

最新的miRNA数据库: 以最新的Sanger miRBase数据库为基础, 为客户提供最前沿的miRNA研究信息。

优越的灵敏性与特异性: 结果可靠。

解码生命
造福人类



第四章 微生物组学

一.16S/18S/ITS测序

(一) 背景介绍

微生物多样性测序（又称扩增子测序）是利用二/三代测序平台，对16SrDNA/18S rDNA/ITS/功能基因等特定区段PCR产物进行高通量测序，突破传统微生物不可培养的缺点，获得环境样本中的微生物群落结构、进化关系以及微生物与环境相关性等信息。

(二) 产品种类和测序原理

(1) 16S rDNA测序：

16S rDNA为编码原核生物核糖体小亚基rRNA的DNA序列，存在于所有细菌染色体基因中，共包括9个可变区和10个保守区，保守序列区域反映了生物物种间的亲缘关系，而高变序列区域则能体现物种间的差异。采用HiSeq或最新的MiSeq测序仪，对16S rDNA某个高变区进行测序，进一步分析环境微生物中细菌或古细菌的多样性，可以了解环境样品中群落多样性信息；



图1. 16S rDNA序列及常用扩增引物序列

(2) 18S rDNA测序：

18S rDNA为编码真核生物核糖体小亚基rRNA的DNA序列。在结构上分为保守区和高变区，保守区反映生物物种间的亲缘关系，高变区反映物种间的差异。在系统发育研究中较适用于种级以上阶元的分类；

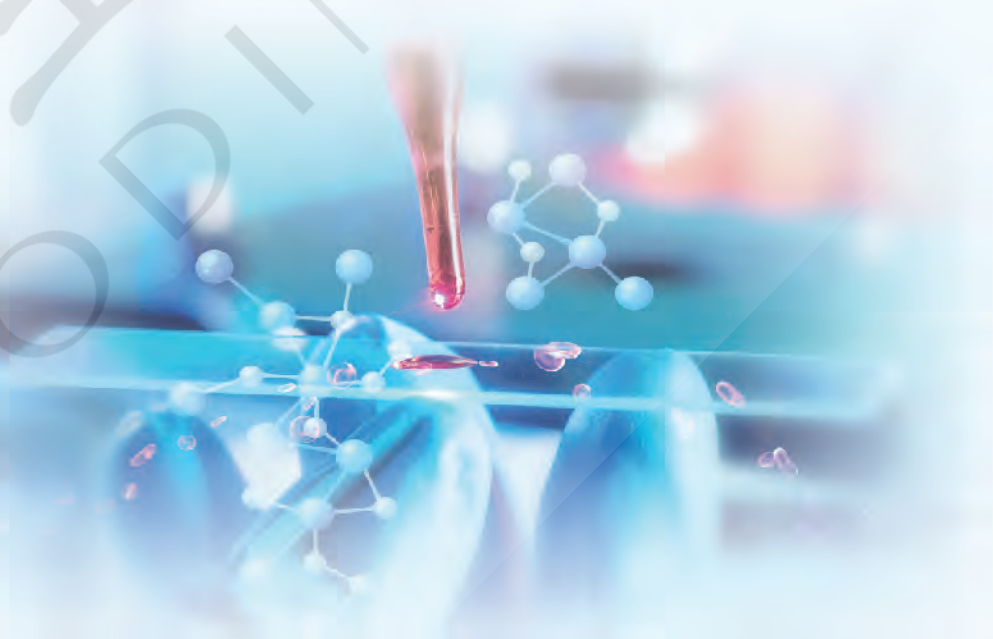


图2. 18S rDNA序列及常用扩增引物序列

(3) ITS测序：ITS分为两个区域：ITS1/ITS2，其中ITS1位于真核生物核糖体rDNA序列18S和5.8S之间，ITS2位于真核生物核糖体rDNA序列5.8S和28S之间。ITS序列由于自然选择压力小，因此在进化过程中能够承受更多的变异，通过对ITS1或者ITS2进行测序，可分析环境微生物中真菌的多样性。（Transcribed Spacer）被广泛应用在真菌分类鉴定中。ITS属于中度保守区域，利用它可研究种及种以下的分类阶元。

产品种类和研究对象信息如表下所示：

产品类型	研究对象
16s rDNA/rRNA	细菌或古菌的多样性
18s rDNA/rRNA	真核生物的多样性
ITS	测序真菌的多样性
功能基因扩增子测序	对特定基因的PCR产物进行测序



(三) 测序流程

包括细菌基因组DNA提取、16SrDNA特异引物PCR扩增、扩增产物纯化、DNA测序、序列比对等步骤。是一种快速获得细菌种属信息的方法。



图3. 实验操作流程

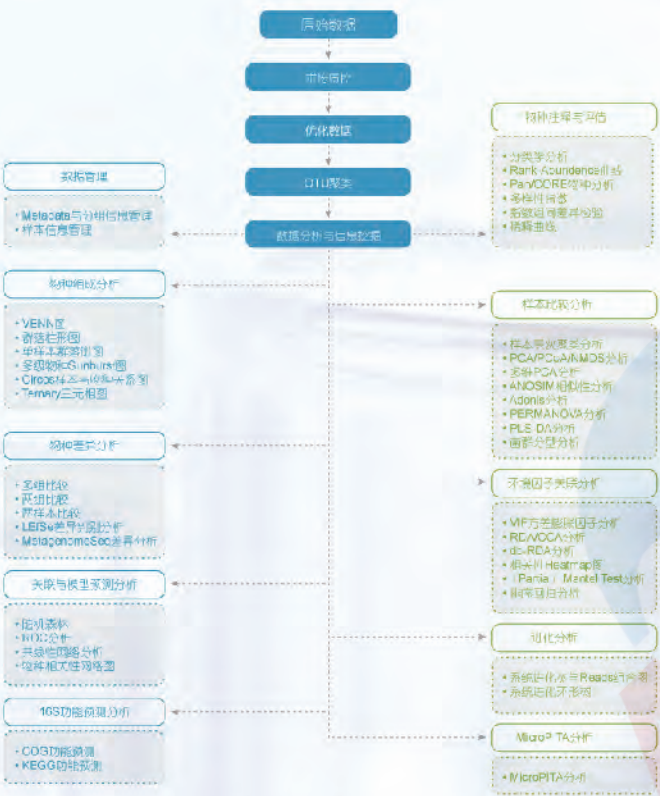


图4. 数据分析流程

(四) 应用范围

包括疾病诊断、疾病机理、药效评价、药物机理等医药方面；环境微生物差异研究、生物地理学研究等环境生态学方面；植物保护、连作障碍、作物生长与根际微生物关系等植物科学方面；动物肠道微生物研究等动物科学方面；污水处理、油污利用、酿酒微生物研究等工业应用方面；还有某些重要次生代谢物相关的环境微生物多样性研究；极端环境条件下的微生物种群研究等。

(五) 浦东解码生命科学研究院优势

- (1) 精准鉴定：菌群多样性精确到“种”，无偏差覆盖低丰度物种；
- (2) 先进的分析及算法：提供精准的微生物样品组间差异、菌落组成及群落进化等分析；
- (3) 高效率低成本：相比传统菌落鉴定方法，通量提高，成本降低。

(六) 案例分析

案例1. 对于个体和群体出生一年内婴儿的肠道微生物群落研究

研究背景：人类出生时肠道理论上是无菌的状态，到了3岁左右，肠道的微生物群落逐渐形成与成年人类类似的状态。大量研究表明婴儿的肠道微生物群落与人体健康息息相关，像肥胖、糖尿病和肠胃炎等，均与肠道微生物密切相关。目前已有一些基于粪便样品的不规律采样，来监测个体婴儿肠道微生物的变化，但从群体的角度出发，从出生开始高频率持续收集粪便样品来分析婴儿的肠道微生物变化的研究还没有。

研究内容：本研究收集了12个婴儿在出生一年内的粪便样品，收集频率几乎是每日一次，从个体和群体的角度出发，研究婴儿在出生的一年内肠道微生物变化。研究样本包括一对异卵双胞胎和一对出生16个月后分开的同胞兄弟。实验采用16S rRNA高通量测序的方法，测所有样本的细菌群落结构及丰度，结合婴儿的饮食用药等环境因素，探究时间尺度上婴儿肠道微生物的演替和变化规律。

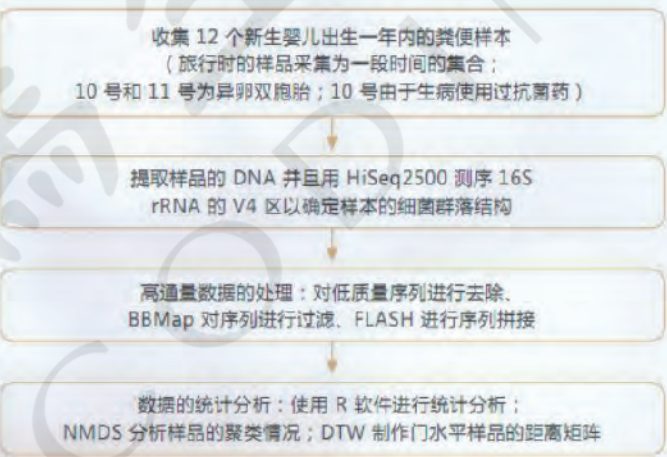


图5. 课题研究思路

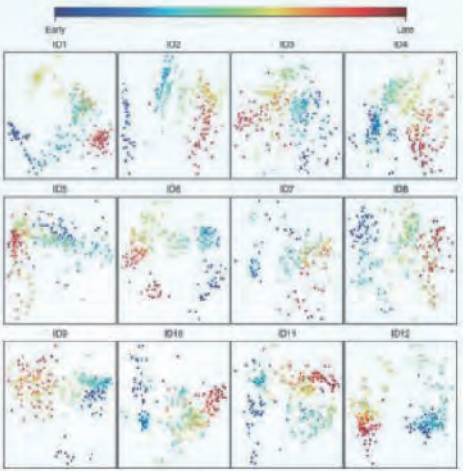


图6. 婴儿肠道微生物在时间尺度上的变化

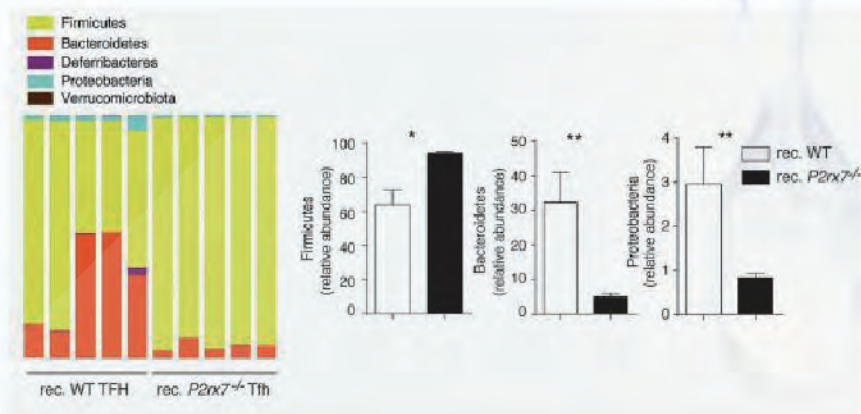


图8.将缺乏P2X7与正常小鼠的Tfh细胞转移到cos^{-/-}小鼠体内后的微生物变化

研究结论

Tfh细胞中的P2X7受体对于塑造肠道微生物群是很重要的，缺乏P2X7导致肠道中致肥胖的微生物选择性增加，Tfh细胞控制IgA的分泌来促进产生健康的肠道微生态，且P2X7受体识别细菌产生的ATP，从而促进形成有利于宿主代谢稳态的肠道生态系统。

参考文献

1. Individuality and convergence of the infant gut microbiota during the first year of life. Nat Commun, 2018 (IF: 12.353).
2. T Follicular Helper Cells Promote a Beneficial Gut Ecosystem for Host Metabolic Homeostasis by Sensing Microbiota-Derived Extracellular ATP. Cell Reports, 2017 (IF: 8.28).

二.宏基因组测序

(一) 基本概念

宏基因组学 (Metagenomics)，也称元基因组学，利用新一代高通量测序技术 (NGS) 以特定环境下微生物群体基因组为研究对象，在分析微生物多样性、种群结构、进化关系的基础上，可进一步探究微生物群体功能特性、相互协作关系及与环境之间的关系，发掘潜在的生物学意义。

(二) 适用范围

- (1) 医学领域：代谢病、肿瘤癌症等；
- (2) 畜牧领域：肠道、瘤胃（如产甲烷菌类群）与动物健康/营养消化研究等；
- (3) 农业领域：根际微生物与植物互作、农业耕作/施肥处理与土壤微生物群落等；
- (4) 环境领域：雾霾处理、污水治理、石油降解、酸性矿水处理及海洋环境等；
- (5) 生物能源：特殊功能的菌株、基因挖掘、工程菌的开发；
- (6) 特殊极端环境：极端环境条件下的微生物类群研究。

(三) 技术优势

- (1) 自然界中，有许多微生物是不能在实验室条件下进行培养繁殖的，宏基因组学研究不要求对微生物进行分离培养，从而大大扩展了微生物研究范围；
- (2) 宏基因组学引入了宏观生态的研究理念，对环境中微生物菌群的多样性、功能活性等宏观特征进行研究，可以更准确地反映出微生物生存的真实状态。因此，宏基因组测序研究摆脱了微生物分离纯培养的限制，扩展了微生物资源的利用空间，为环境微生物群落的研究提供了有效工具。
- (3) 通过宏基因组测序可以进行功能基因的发掘，分析微生物群体基因组成及功能，解读微生物群体的多样性与丰度，发掘和研究新的、具有特定功能的基因。

(四) 实验流程



图1.实验操作流程

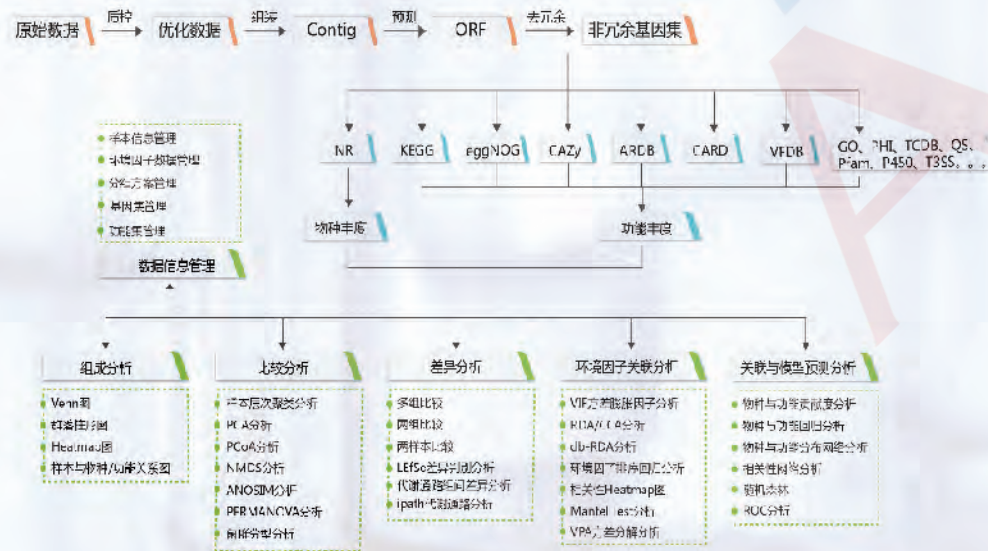


图2.生信分析流程图

其中，数据分析部分主要包括：

- (1) 物种鉴定：将所得序列（通常为16S/18S rRNA等兼具保守及高变特性的序列）与专业数据库（Silva、RDP等）进行比对，得出样品中所含物种的信息；
- (2) 多样性统计学分析：将所得序列（通常为16S/18S rRNA等兼具保守及高变特性的序列）进行聚类，得到相应的OTUs（分类操作单元）。通过统计学手段，分析出环境样品中的主要成分及不同样品间的明显差异因素。结合物种鉴定，可以得到关键菌群；
- (3) 宏基因组拼接：对环境样品DNA进行大规模测序后，通过严格的拼接方式，可获得较长的DNA片段。当样品的生物多样性较低，且达到一定测序通量后，很有可能直接获得一个或多个微生物基因组草图；
- (4) 功能分析：将所得序列与已有的数据库进行比对，进行基因功能的注释。其中，常用的数据库包括NT、NR、GO、COG、KEGG、SEED、Swiss-Prot等；
- (5) 微生物群落结构及功能：通过大量测序，可以获得样品的群落结构信息，如微生物物种在该环境下的分布情况及成员间协作关系等。通过实验还可以确定一些特殊的主要基因或DNA片段。对于多个样品，还可做相应的比较分析，发掘样品间的相同点与不同点。

(五) 案例

案例1.全球海洋微生物结构与功能

研究背景：微生物是生物化学主要动力，但对它们的微生物种群结构、功能多样性以及生态因素进行总体分析还存在很大挑战。

方法流程：

取材：全球主要代表性海洋区域（除北极外）68个采样点分三层（表层SRF，最低光合作用层DCM，中层MESO）不同海水中采集243份样品。

建库：海水样品总DNA提取（NEBNext Sample Reagent Set）

测序：illumina PE100

分析：

- (1) OTU聚类和分析；
- (2) 基因预测；
- (3) 菌群结构和功能分析；
- (4) 统计模型和相关性分析

研究结果:

(1) 海水中微生物菌群

本研究建立了一个全球海洋微生物内参基因目录 (OM-RGC), 包含>4,000万个来自病毒、原核生物、微型真核生物的非冗余, 且大部分为新发现的基因, 约为人体肠道微生物内参基因目录数目的4倍, 大部分为细菌中的基因。

(2) 海洋的生态环境与微生物菌群的相关性

PCoA分析显示, 跟光照、温度和营养相比, 海水深度对微生物种群构成的贡献率为73%, 随海水深度的增加, 微生物物种和功能多样性增加, 而细胞密度和细胞最大增长速率降低。

(3) 海洋微生物vs人体肠道微生物

人体肠道微生物核心直系同源家族中仅有9%是未知的, 而海洋微生物中这一比例是40%, 分析两个微生物生态系统 (海洋与人体肠道) 发现, 尽管这两种生态系统理化因素差别很大, 但海洋微生物核心功能与人体肠道微生物相似性>73%。

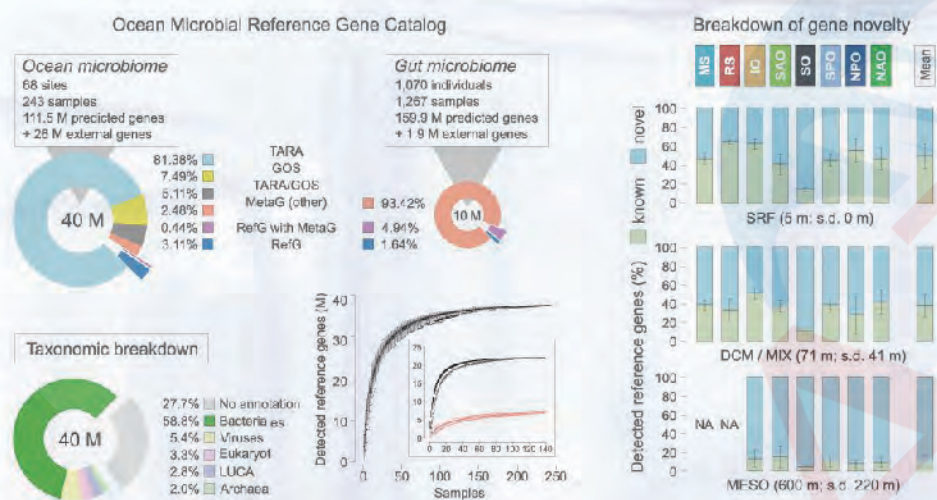


图3.全球海洋发现的微生物新基因多样性

研究结论:

Tara Ocean研究团队, 结合生态学、系统生物学和海洋学来研究海洋中浮游生物, 研究得到很多珍贵资源如海洋微生物内参基因列表、浮游生物 (病毒、原核生物、微小真核生物) 多样性普查表, 为研究海洋微生物之间及微生物与环境互作提供方法理论基础。研究发现温度是微生物物种多样性最大的驱动力, 海洋微生物中广泛的功能预示其影响全球气候变化的潜力。

案例2. 结直肠癌患者肠道微生物宏基因组和代谢组学研究

研究内容: 该课题对616名接受结肠镜检查参与者的样本进行了粪便宏基因组和代谢组学研究, 以评估肠道微生物群和代谢物的分类学和功能特征。在多发性息肉样腺瘤和粘膜内癌的病例中以及更严重的病变中, 微生物组和代谢组的变化明显。研究人员发现了两种不同的微生物群升高模式。首先, 具核梭杆菌的相对丰度 ($P<0.005$) 从粘膜内癌到更晚期持续升高。其次, 仅在多发性息肉样腺瘤和或粘膜内癌中, 同时出现在粘膜内癌中的特异性小瘤和放线菌显著增加 ($P<0.005$)。代谢组学分析显示, 多发性息肉样腺瘤和或多发性息肉样腺瘤中支链氨基酸和苯丙氨酸显著增加 ($P<0.005$), 胆汁酸 (包括脱氧胆酸) 显著增加 ($P<0.005$)。研究人员鉴定了宏基因组和代谢标记物, 以区分粘膜内癌病例与健康对照组。课题组的大型队列多组学数据表明, 微生物组和代谢组的变化发生在结直肠癌发展的早期阶段, 这对可能病因的判定以及诊断具有重要意义。

据悉, 在大多数的散发性结直肠癌中, 肿瘤发生是一个多步骤的过程, 包括基因组的改变与形态学变化。此外, 越来越多的证据表明, 人体肠道微生物组与结直肠癌的发展有关。

参考文献:

1. Structure and function of the global ocean microbiome, Science, 2015, IF=31.477.
2. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer, Nature Medicine, 2019, IF=30.641.

解码生命 快乐科学



第五章 蛋白质组学

一.蛋白质组定性检测

背景介绍

上世纪90年代, 澳大利亚科学家Wilkins和Williams首次提出蛋白质组 (Proteome) 的概念。蛋白质组是指一种基因组所表达的所有对应蛋白质, 即包括细胞、组织或机体内全部蛋白质的存在。蛋白质是生命活动的直接执行者, 从蛋白质水平研究生命现象和规律, 以及疾病防治和病理研究, 已逐渐成为生命科学研究的发展趋势。与基因组相对应, 蛋白质组也是一个整体。基因组内各个基因的转录表达会受到多个水平的调控, 因此其表达的产物蛋白质在不同的时空和环境条件下也不尽相同。

之后, Peter James在此基础上提出了蛋白质组学 (Proteomics) 的概念, 即以蛋白质组为研究对象, 研究细胞、组织或生物体蛋白质组成及其变化规律。它是在细胞和组织的整体水平上, 研究蛋白质的种类、表达, 进而探究蛋白质表达的调节控制以及蛋白质之间的相互作用。蛋白质组学的研究不仅能为生命活动规律提供物质基础, 也能为众多疾病机理的阐明及攻克提供理论根据和解决途径, 不仅是探索生命奥秘的必须工作, 也能为人类健康事业带来巨大的利益。

蛋白质组学研究大多需要对不同时期、不同条件下, 细胞、组织或器官中所含的蛋白质进行定性和定量。生物质谱技术的不断发展为蛋白质鉴定和差异表达分析提供了可靠的研究技术。本章节将会介绍现有的主流的蛋白质鉴定、定量、及蛋白质表达修饰分析的检测方法。

蛋白质定性分析

蛋白质定性分析是指对蛋白质进行鉴定和序列分析。如今使用质谱技术进行蛋白质鉴定越来越广泛, 质谱技术具有较好的灵敏度、分辨率、准确度。实验流程一般先将蛋白质样本经过蛋白酶解消化, 将酶解后的肽段按需求使用不同的质谱技术进行分析, 最后将得到的结果与数据库进行比对, 通过对肽段序列的鉴定推断蛋白质序列。



图1.蛋白质组定性检测流程

(一) 蛋白质全谱分析 (shotgun)

技术原理

蛋白质全谱分析也可称为质谱shotgun分析, 研究对象是完整的组织、体液或其提取物, 其目的在于鉴定出尽可能多的肽和蛋白质分子, 并由此衍生出宏蛋白质组学, 即研究与宏基因组或者宏转录组对应的复杂样品中蛋白质的表达。

将溶液内蛋白质分子或SDS-PAGE条带的复杂混合物酶解成肽段混合物, 通过液相色谱分离, 串联质谱测试, 最后用相应的数据库进行检索匹配, 可同时鉴定成百上千种蛋白质。

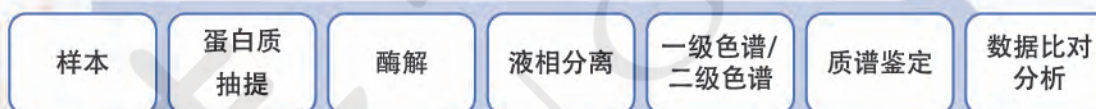
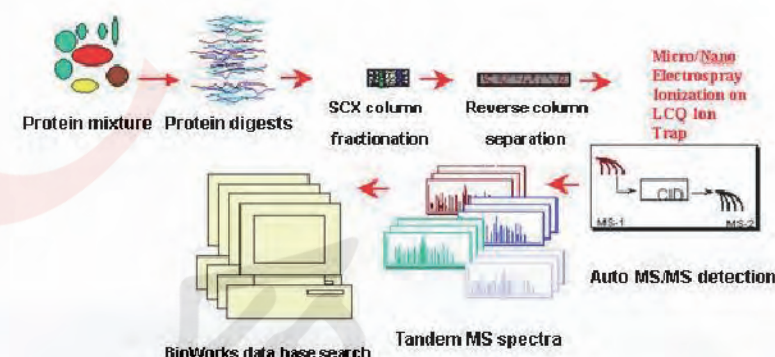


图2. Shotgun技术流程图

技术优势

- (1) 可全面鉴定整个细胞、组织或复杂混合样品的大量蛋白质;
- (2) 高通量, 分离能力强;
- (3) 兼容性强, 可分析蛋白质条带、组织提取液、全细胞裂解液、亚细胞分离组分等;
- (4) 自动化程度高。

技术平台

ABSciexTriple TOF 5600; Thermo Scientific Q-Exactive; Thermo Scientific LTQ-Velos.

应用领域

针对某一生理状态下组织、细胞或者细胞器中所有表达蛋白质的鉴定，或者对一些环境样本及肠道微生物等复杂样品中的蛋白质进行鉴定。

(二) 蛋白质胶条/混合液鉴定

技术原理

利用LC-ESI-MS/MS可对胶条样本（SDS-PAGE样本）、IP、co-IP、Pull-down等纯化溶液等中等复杂样本进行蛋白鉴定。

电喷雾电离质谱（electrospray ionization mass spectrometry, ESI-MS）是在毛细管的出口处施加高压电场，使从毛细管流出的液体雾化成带电荷的微滴，随着液滴中挥发性溶剂的蒸发，液滴表面的电荷密度随体积减少而逐渐增大，达到临界点后，分析物以单电荷或多电荷离子的形式进入质谱分析器。电喷雾离子化产生的高电荷离子，使在较小的质量电荷比范围内检测到更大分子量的分子，扩展了分子量的分析范围，离子的真实分子质量可根据质荷比及电荷数算出。电喷雾质谱可以与多种分离技术联合使用，一般与液相色谱（LC）联用，对复杂样品进行质谱分析，根据样品在色谱中不同的保留时间对其进行分离后鉴定，大大提高了分析的效率和精确度。

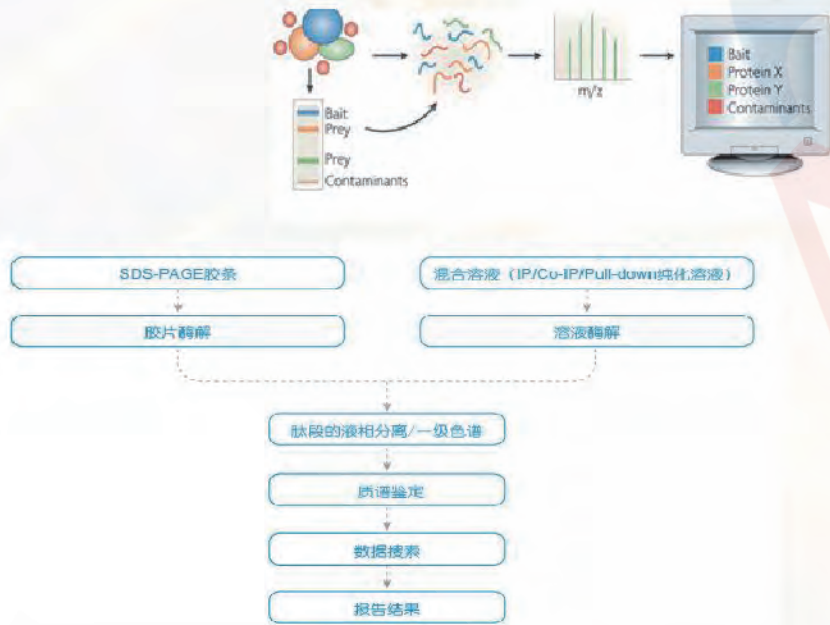


图3. 蛋白质胶条/混合液检测流程图

技术优势

- (1) 适合中度复杂的蛋白质样品，如相互作用的蛋白质条带；
- (2) 高准确性，高灵敏度。

技术平台

Thermo Scientific LTQ-Velos; Thermo Scientific Q Exactive.

应用领域

蛋白质或多肽鉴定实验。

样品要求

- (1) SDS-PAGE条带：考染、银染（质谱兼容）条带清晰可见；。
- (2) 蛋白质溶液：蛋白质总量>5 μg，浓度>0.1 μg/μl；
缓冲液不含去污剂NP40、Triton X-100等。

(三) 蛋白质胶点鉴定

技术原理

利用MALDI-TOF/TOF技术可对考/银染的2D或DIGE胶点样本或者较纯的蛋白样本进行蛋白鉴定。

基质辅助激光解吸质谱技术（Matrix-Assisted Laser Desorption/Ionization Time Of Flight, MALDI-TOF）由两部分组成：基质辅助激光解析电离离子源（MALDI）和飞行时间质量分析器（TOF）。MALDI是将分析物分散在基质分子中并形成晶体，在吸收激光提供的能量后晶体升华，使基质和分析物进入气相并离子化。随后通过TOF使离子在电场作用下加速飞过飞行管道，根据到达检测器的飞行时间不同而被检测，即测定离子的质荷比（M/Z）与离子的飞行时间成正比，从而检测离子。MALDI-TOF所产生的质谱图多为单电荷离子，因而质谱图中的离子与多肽和蛋白质的质量有一一对应关系。

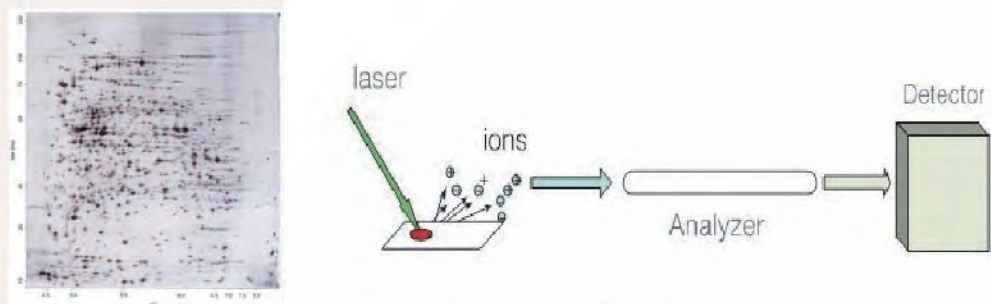


图4. MALDI-TOF质谱原理示意图和技术流程图

技术优势

- (1) 适合较纯或比较单一的蛋白质样品，如双向凝聚电泳差异点；
- (2) 速度快，通量高。

技术平台

ABI 4800 MALDI-TOF/TOF

应用领域

蛋白质或多肽鉴定。

二.蛋白质组定量检测

背景介绍

相对定量的目的是测定目标蛋白在两个或多个样本中的表达量的相对比例，而不需要知道它们在每个样本中的表达量。比如设置处理组和对照组样本，将对照组的样本定位基准，规定其目标蛋白浓度为100%，将处理组样本的定量结果与对照样品相比较，就可以计算各个处理样本的蛋白含量相对于对照样本的百分比。

目前有多种技术可用于蛋白质相对定量，下表总结了几种主流技术的原理及技术优势，后会详细介绍每种技术的原理、检测流程及应用。

技术名称	技术原理	技术优势
iTRAQ	采用4种或8种同位素的标签，通过特异性标记多肽的氨基基团，然后进行串联质谱分析，可同时比较4种或8种不同样品中蛋白质的相对含量	可检测膜蛋白、核蛋白等多种蛋白类型；以及低丰度蛋白质、强酸（碱）性蛋白质、<10KD或>100KD的蛋白质分子。
TMT	该技术采用2种、6种或10种同位素的标签，通过特异性标记多肽的氨基基团，然后进行串联质谱分析，可同时比较2组、6组或10组不同样品中蛋白质的相对含量。	与iTRAQ检测原理相似，可同时检测10组样本。
Label-free	待测蛋白质样本不需要任何标记，直接酶解后采用液质联用方法对蛋白质酶解肽段进行质谱分析，通过软件对谱图进行归一化，比较对应质谱峰峰强和峰面积定量蛋白质，分析蛋白质在几组样本中的相对表达量变化，对分析组数无限制。	蛋白质不需要进行标记，所需样品总量少，耗费低。
2D双向电泳	将不同种类蛋白质按等电点和分子量差异进行高分辨率分离的分析方法。检测时先将蛋白质根据其等电点在pH梯度胶内进行等电聚焦，然后按照其相对分子量大小进行SDS-PAGE第二次电泳分离。电泳后对凝胶进行染色，并用相关软件对图像进行分析处理。	可同时将上千种蛋白质同时分离和展示的方法；蛋白质差异变化直观可见；检测成本低。
DIGE 荧光差异 双向电泳	荧光差异双向凝胶电泳（DIGE）在2D电泳之前标记蛋白质样品，保证所检测到的蛋白质丰度变化是真实的，极大地提高结果的准确性、可靠性和重复性。对分析组数无限制。	检测灵敏度更高，节约样品；采用内标对数据进行归一化处理，消除胶与胶之间差异造成的误差，准确度高。
SILAC	在细胞培养基中加入轻、中或重型稳定同位素标记的必需氨基酸，细胞经培养后，新合成的蛋白带上稳定同位素标签。等量混合各类型蛋白质，酶解后进行质谱分析。	体内标记，更接近样品真实状态，适用于培养细胞或者小鼠蛋白质表达动态变化研究。

(一) iTRAQ (isobaric Tags for Relative and Absolute Quantitation)

技术原理

iTRAQ是由ABSciex公司在2004年研发推出的用于相对和绝对定量的同重同位素标记技术, 可实现对多种不同样品中蛋白的比较。iTRAQ试剂用于标记酶解后的肽段, 来自不同样品的同一肽段经标记后具有相同的质量数, 并在质谱图中表现为同一个质谱峰。当此质谱峰被选定进行碎裂后, 在串联质谱检测中, 报告基团、质量平衡基团和多肽反应基团之间的键发生断裂, 质量平衡基团丢失, 带不同同位素标签的同一多肽释放出不同的报告基团, 根据报告离子的信号强度值可获得样品间相同肽段的定量信息, 再经过软件处理得到蛋白质的定量信息。目前iTRAQ可同时比较8组样品的蛋白质表达差异。

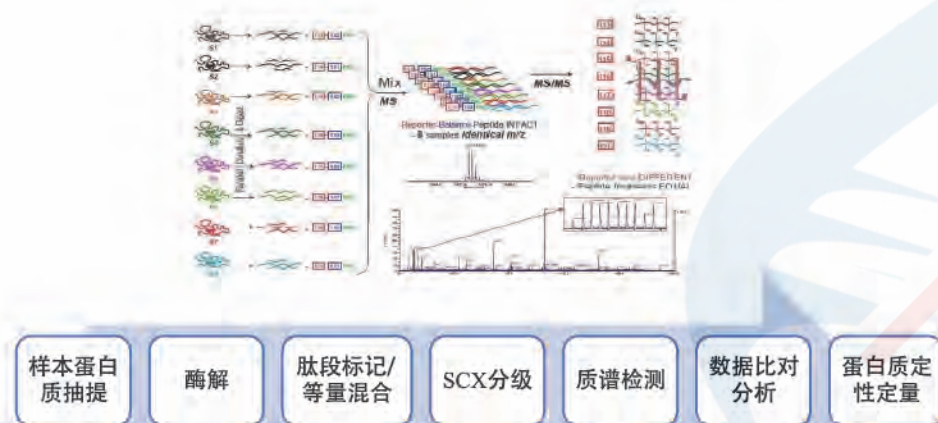


图1.iTRAQ技术流程

技术平台

ABSciexTripleTOF 5600;
ThermoScientific Q-Exactive;
Thermo Scientific Orbitrap FusionTribrid.

技术优势

- (1) 可同时对2-8种不同条件样品进行蛋白质差异分析;
- (2) 适合广泛样品类型: 胞浆蛋白、膜蛋白、核蛋白、胞外蛋白等;
- (3) 与凝胶分离的蛋白质定量技术相比, iTRAQ质谱定量技术可检测到更多类型的蛋白质分子, 如低丰度蛋白质、强酸(碱)性蛋白质、<10KD或>100KD的蛋白质分子。

应用领域

机理和调控机制研究, 动态过程中机制研究, 疾病标志物的筛选, 植物抗逆研究, 用药及预后标志物筛选, 特殊功能蛋白质筛选。

(二) TMT (Tandem Mass Tag)

技术原理

TMT技术是由美国Thermo Scientific公司研发的一种多肽体外标记技术。其原理与iTRAQ技术类似, 但最多可采用10种同位素的标签, 因此可同时比较10组不同样品中蛋白质的相对含量。TMT标签由三部分组成: 分子量报告子、分子量标准化部分和反应基团, 形成2种、6种或10种相对分子质量均等量的异位标签。

TMT试剂通过反应基团标记酶解后的肽段。在一级质谱图中, 分子量标准化部分使任何一种TMT试剂标记的不同样本中的同一肽段表现为相同的质荷比。在串联质谱中, 可剪切臂能够优先断裂以便释放分子量报告子。每个报告子都有各自独特的分子量, 并且能够在MS/MS分析中反应出所标记的多肽的样品丰度, 根据报告离子的信号强度值可获得样品间相同肽段的定量信息, 再经过软件处理得到蛋白质的定量信息。

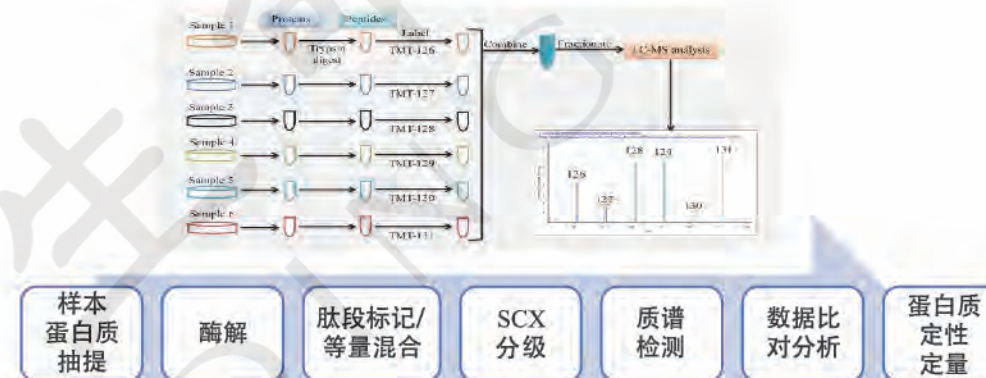


图2.TMT技术流程图

技术平台

Thermo Scientific Q Exactive.

技术优势

- (1) 可同时对多达10种不同条件样品进行蛋白质差异分析;
- (2) 适合广泛样品类型: 胞浆蛋白、膜蛋白、核蛋白、胞外蛋白、分泌蛋白等;
- (3) 与凝胶分离的蛋白质定量技术相比, TMT质谱定量技术可检测到更多类型的蛋白质分子, 如低丰度蛋白质、强酸(碱)性蛋白质、<10KD或>100KD的蛋白质分子。

应用领域

疾病标志物筛选；分子机制研究；植物抗逆研究；药物作用靶点研究；特殊功能蛋白质筛选。

(三) Label-free

技术原理

非标记定量技术 (Label Free)，待测蛋白质样本不需要任何标记，直接酶解后采用液质联用方法对蛋白质酶解肽段进行质谱分析，通过软件对谱图进行归一化，比较对应质谱峰峰强和峰面积定量蛋白质，分析蛋白质在几组样本中的相对表达量变化，对分析组数无限制。

相比于iTRAQ和TMT等标记定量技术，Label Free减少了前处理过程中的样品损失，无需使用昂贵的稳定同位素标签做内部标准，只需分析大规模鉴定蛋白质时所产生的质谱数据，比较不同样品中相应肽段的信号强度，从而对肽段对应的蛋白质进行相对定量。单个样品单独分析，没有样品数目的限制。

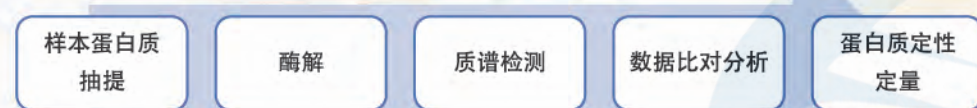


图3.Label-free技术流程图

技术优势

蛋白质不需要进行标记，所需样品总量少，耗费低。

技术平台

ABSciexTripleTOF 5600;

Thermo Scientific Q-Exactive;

Thermo ScientificOrbitrap FusionTribid.

应用领域

机理和调控机制研究；疾病标志物的筛选；用药及预后标志物筛选；疾病的分子分型；植物抗逆研究；药物作用靶点研究；特殊功能蛋白质筛选。

(四) SILAC (Stable Isotope Labeling By Amino Acids In Cell Culture)

技术原理

SILAC，细胞培养氨基酸稳定同位素标记技术，其原理是在细胞培养基中加入轻、中或重型稳定同位素标记的必需氨基酸（主要是赖氨酸Lys和精氨酸Arg），比如在A组培养基中加入正常氨基酸（light），而B组培养基则含有“重型（heavy）”氨基酸，即稳定同位素标记的氨基酸。细胞经5至6代倍增周期后，新合成的蛋白带上稳定同位素标签。随后等量混合各类型蛋白质，酶解后进行质谱分析。通过比较一级质谱图中同位素峰型的面积大小进行相对定量，同时二级谱图对肽段进行序列测定从而进行蛋白鉴定。该技术适用于细胞和动物的体内标记。

细胞实验流程如下：

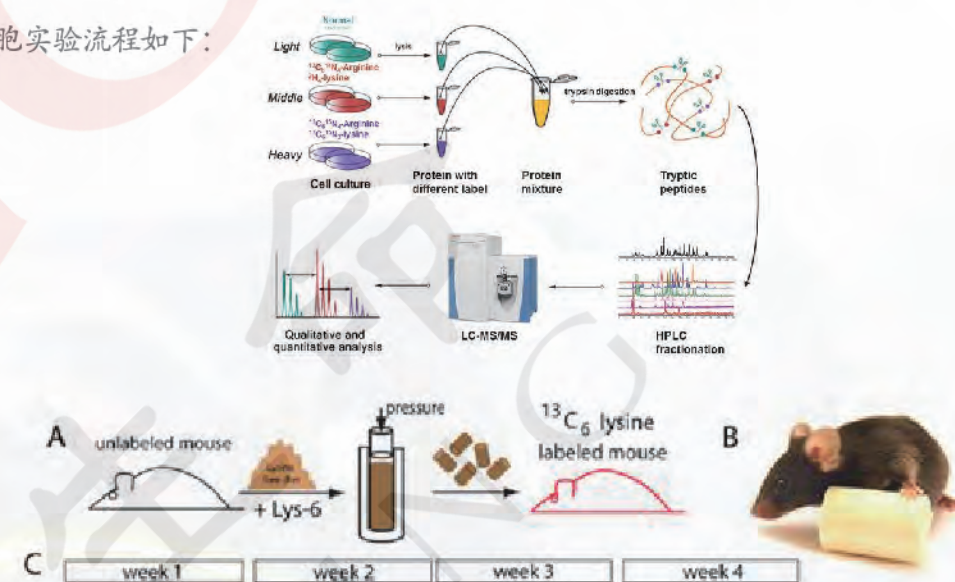


图4.SILAC技术流程图

(五) 2D双向凝胶电泳

技术原理

双向凝胶电泳是将不同种类的蛋白质按照等电点和分子量差异进行高分辨率分离的分析方法。成功的二维电泳可以将2000到3000种蛋白质进行分离，是唯一能同时将上千种蛋白质同时分离和展示的方法，在蛋白质组学中发挥重要的作用。

检测时先将蛋白质根据其等电点在pH梯度胶内进行等电聚焦，然后按照它们的相对分子量大小进行SDS聚丙烯酰胺凝胶（SDS-PAGE）第二次电泳分离。电泳后对凝胶进行染色，并用相关软件对图像进行分析处理。

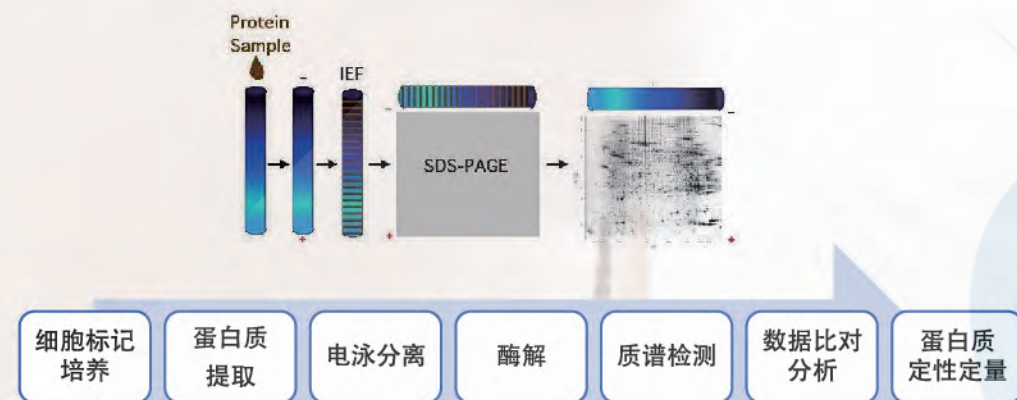


图5.2D双向凝胶电泳技术流程图

技术优势

经典蛋白质组学方法，蛋白质差异变化直观可见；检测成本低。

技术平台

第一向IEF: EttanIPGphor Isoelectric Focusing System (GE Amersham)；

第二向SDS-PAGE: Hofer SE 600 (GE Amersham)；

扫描仪: UMaxPowerlook 2110XL (GE Amersham)。

应用领域

疾病标志物筛选，作用机制研究，植物抗逆研究，药物作用靶点研究，特殊功能蛋白质筛选。

(六) DIGE荧光差异双向凝胶电泳

技术原理

DIGE荧光差异蛋白质表达分析系统在传统2D双向凝胶电泳技术的基础上，结合了多重荧光分析的方法。荧光染料采用分子量和等电点相互匹配的三种荧光标记分子Cy2、Cy3、Cy5，其特殊设计的NHS脂活性基团通过氨基酸桥连接作用于蛋白质的赖氨酸上，使检测灵敏度达到125pg。DIGE实验中，在同一块胶上共同分离多个分别由不同荧光标记的样品，根据胶点灰度值和面积定量蛋白质。同时将Cy2标记所有样品的等量混合样作为内标，理论上包含了需要分析样品的所有蛋白质点，即每个蛋白质点都有它自己的内标，存在于每一块需要被分析的胶中。软件全自动根据每个蛋白质点的内标对其表达量进行校准，保证所检测到的蛋白质丰度变化是真实的，极大地提高结果的准确性、可靠性和重复性。对分析组数无限制。

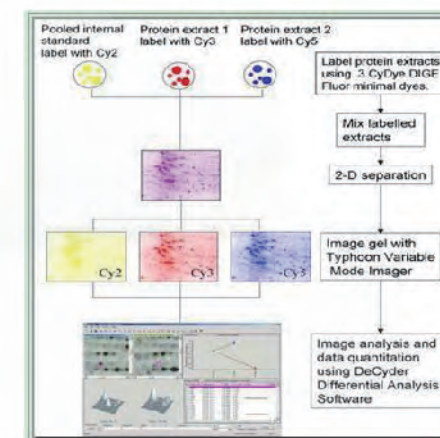


图6.DIGE荧光差异双向凝胶电泳流程图

技术优势

检测灵敏度相对应传统的双向电泳技术有很大提高；

采用内标对数据进行归一化处理，消除胶与胶之间差异造成的误差，准确度高。

技术平台

第一向IEF: EttanIPGphor Isoelectric Focusing System (GE Amersham)；

第二向SDS-PAGE: Hofer SE 600 (GE Amersham)；

扫描仪: UMaxPowerlook 2110XL (GE Amersham)。

应用领域

疾病标志物筛选；作用机制研究；植物抗逆研究；药物作用靶点研究；特殊功能蛋白质筛选。

三.蛋白质化学修饰组学检测

背景介绍

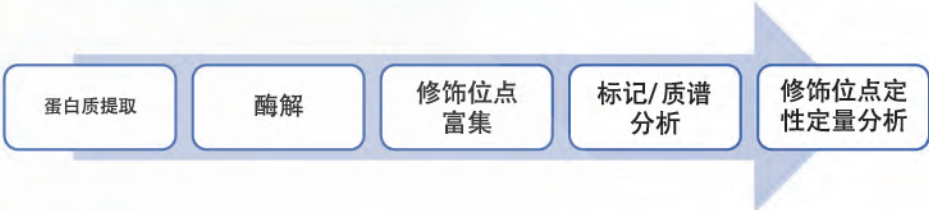
生命进程不仅由蛋白质含量控制，也受可逆的翻译后修饰所调控，因此研究翻译后修饰的内在规律是解析复杂多样的生物功能的重要内容。蛋白质修饰后的分子质量会发生相应的改变，使用质谱能够精确测定修饰后的蛋白质或多肽分子的质量变化。由于翻译后修饰的蛋白质在样本中含量低且动态范围广，检测前先对修饰蛋白质或肽段进行富集，然后进行质谱鉴定。

技术原理

检测的修饰位点包括磷酸化、糖基化、甲基化等研究热点。磷酸化涉及到细胞信号转导，细胞的增殖，发育和分化过程；糖基化在细胞的免疫保护等生物过程中起重要作用。具体修饰位点及富集方法见表。

翻译后修饰	修饰位点	富集方法	生物学功能
磷酸化	Ser, Thr, Tyr	二氧化钛；酪氨酸磷酸化基序抗体	信号转导、细胞周期、生长发育及癌症机理等
N-糖基化	Asn	凝集素	蛋白质折叠、更新以及免疫应答等
乙酰化	Lys	赖氨酸乙酰化基序抗体	基因表达调控、细胞凋亡、细胞代谢、蛋白质稳定性以及神经退行性病变等
泛素化	Lys	赖氨酸泛素化基序抗体(K-ε-GG)	细胞周期、细胞凋亡、转录调控、DNA修复以及信号转导等
甲基化	Lys, Arg	赖氨酸/精氨酸mono/di/tri-甲基化基序抗体	染色质结构及转录调控等
丙酰化	Lys	赖氨酸丙酰化基序抗体	主要存在于线粒体中，与细胞代谢、表观遗传调控等相关
丙二酰化	Lys	赖氨酸丙二酰化基序抗体	主要存在于线粒体中，与细胞代谢、表观遗传调控等相关
戊二酰化	Lys	赖氨酸戊二酰化基序抗体	主要存在于线粒体中，与细胞代谢密切相关
琥珀酰化	Lys	赖氨酸琥珀酰化基序抗体	主要存在于线粒体和细菌中，与代谢过程密切相关

将蛋白质样本进行酶解后，首先根据目标修饰位点，使用相应的富集方法进行富集，再使用质谱检测鉴定修饰位点。也可在修饰位点富集后，选择蛋白质定量技术（iTRAQ、TMT、Label-free等）进行蛋白质修饰的定量分析。



技术平台

ABTriple TOF 5600;
Thermo Q-Exactive;
Thermo Orbitrap Fusion Tribrid.

技术优势

基于高精度的质谱分析，对修饰位点精确鉴定。

应用领域

疾病机理研究；信号转导途径分析。

四.蛋白质组学检测平台

(一) AB Sciex TripleTOF 5600

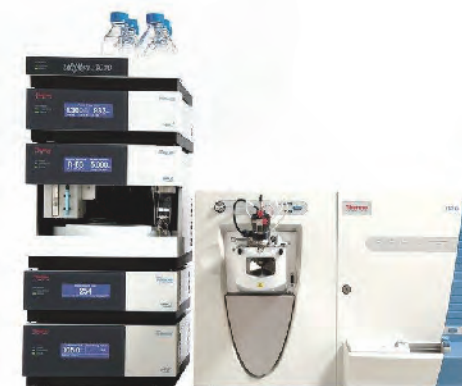
此系统标准配置了用于蛋白质组样本的高通量分析的最新技术—SWATH™，具有准确质量数和高分辨进行同时定量和定性分析的LC-MS/MS系统，该系统具有快速扫描速度和串联四极杆质谱系统的定量灵敏度能力，集质谱系统的高分辨、准确质量数稳定性、高灵敏度和高速度扫描特点于一个质谱平台。

(二) Thermo Scientific Q-Exactive



使用Thermo Scientific Q-Exactive组合型四极杆Orbitrap质谱仪可以快速可靠地识别、定量和确认更多化合物。本台式LC-MS/MS系统将四极杆母离子选择性与高分辨率和准确质量数 (HRAM) Orbitrap检测相结合，提供出色性能和多功能性。Q-Exactive质谱仪特别适用于非目标或目标化合物筛查，也能够实现广泛的定性和定量应用，可广泛用于药物发现、蛋白质组学、环境和食品安全、临床研究和法医毒理学。

(三) Thermo Scientific LTQ XL 线性离子阱质谱仪



为蛋白质组学和代谢组学应用获得更多结构信息Thermo Scientific™ LTQ XL™ 离子阱质谱仪为常规结构解析提供多种解离技术 (PQD、ETD和CID)。PQD是消除低质量截止效应的专利技术。广泛应用于可预测和不可预测代谢物的检测，并可以使用同重元素标记法进行肽定量分析。

(四) Orbitrap Fusion™ Tribrid™ 质谱仪



使用Thermo Scientific™ Orbitrap Fusion™ Tribrid™ 质谱仪能够实现最具挑战性的生物学研究目标。本质谱仪在革命性的Tribrid架构中结合了四极杆、离子阱和Orbitrap质量分析的优点，能够提供前所未有的分析深度和易用性。创新的软件和仪器设计可以使您专注于科学研究而非方法建立和仪器使用。

一.非靶向代谢组学 (Untargeted metabolomics)

非靶向代谢组学分析一般包括:

- (1) 代谢谱分析 (也称为差异表达分析): 在一组实验和对照样品中, 寻找丰度改变有统计学意义的感兴趣代谢物;
- (2) 鉴定: 进行代谢谱分析后, 测定这些代谢物的化学结构;
- (3) 解释: 研究流程的最后一步, 解释所发现的代谢物与生物过程或生物状态之间的关联。

数据质控与校准体系:

非靶向代谢组学平台设置了严格的质控体系，以保证每批次数据的质量和批次间数据的可比性。在每轮测试中，均设置有12个质控样品孔（4种Calibrators，3种QCs和1个NIST血清样品）进行测试稳定性监控。Calibrators监测分析过程中的可能污染，控制实验本底干扰。QCs确保每批数据准确性，控制技术误差。NIST血清（全球唯一国际标准化代谢组学首选的独立外部验证样本）用于不同批间检测数据的校准，保证多批次检测数据的可比性。

背景介绍

代谢组学是效仿基因组学和蛋白质组学的研究思想，对生物体内所有代谢物进行定量分析，并寻找代谢物与生理病理变化的相对关系的研究方式，是系统生物学的组成部分。其研究对象大都是相对分子质量1000以内的小分子物质。先进分析检测技术结合模式识别和专家系统等计算分析方法是代谢组学研究的基本方法。

相比基因组和蛋白组的研究, 代谢组学研究具有独特的应用优势:

基因和蛋白表达在功能水平上的微小变化会在代谢物上得到放大,使代谢物检测更容易;

代谢物的变化除了反应基因组变化外，还受到环境因素、肠道菌群的影响，其动态性更强，对生物体变化的反映更加灵敏：

代谢反应及终产物在各个物种的生物体系中都是类似的，因此，代谢组学方法学通用性更强；

代谢组学的技术不需建立全基因组测序及大量表达序列数据库，直接对几乎可以对所有样本类型进行检测，包括全血、血浆/血清、组织、细胞、细胞培养上清、尿液、粪便、食物、唾液、脑脊液、脂肪等。

基因组、转录组、蛋白质组等表型组广泛应用于临床疾病等的研究中，期望发现常见某种疾病的共性和突破口，代谢组表型针对其共性的代谢途径进行治疗可能是一个更好的精准医疗战略，代谢组恰好是表型组的生物学终点，是未来精准医疗中值得期待的重要战略。

For FTN & FL6 well plate (Waters 700 uL round well, Mfr#: 186005837; 7 mm round plug cap-mat, pre-slit, Mfr#: 186006332)												
	1	2	3	4	5	6	7	8	9	10	11	12
Port A 01	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12
B 13	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24
C 25	S25	S26	S27	S28	S29	S30	S31	S32	S33	S34	S35	S36
D 37	S37	S38	S39	S40	S41	S42	S43	S44	S45	S46	S47	S48
E 49	S49	S50	S51	S52	S53	S54	S55	S56	S57	S58	S59	S60
F 61	S61	S62	S63	S64	S65	S66	S67	S68	S69	S70	S71	S72
G 73	S73	S74	S75	S76	S77	S78	S79	S80	S81	S82	S83	S84
H	CAL0	CAL1	CAL2	CAL3	CAL4	CAL5	CAL6	CAL7	QC1	QC2	QC3	QC4
	85	86	87	88	89	90	91	92	93	94	95	96

Calibrators

QCs

NIST

技术路线:



- (1) 检测生物样本种类齐全: 血 (血浆、血清、全血、血片)、尿液、组织、粪便、唾液、精液、卵泡、乳汁、羊水、胎盘、细胞、细胞培养后基质、微生物代谢物、植物样本等;
- (2) 高灵敏度的时间飞行质谱能够检测代谢物类群丰富: 包含有机酸、脂肪酸、核苷酸、氨基酸、糖类、脂类等; 1500+内源性小分子代谢物, 覆盖95%以上小分子代谢网络 (KEGG);

二.定量代谢组学 (Quantitative metabolomics)

定量代谢组学 (Quantitative metabolomics), 是在同一代谢网络或者具有特定功能代谢的一类小分子代谢物进行绝对定量检测的代谢组学方法。这种方法通过选择稳定同位素标记的标准品作为基质效应校正, 和一系列已知浓度的标准品来定量样本中代谢物的未知浓度。通过选择合适的样本前处理方法, 优化色谱分离、质谱条件等对代谢物进行最优化检测, 大大提高检测的灵敏度、准确性、特异性及重现性, 从而精确检测出每一个指定的代谢物在生物样本中的浓度, 该类研究在临床转化代谢组学中有着极其重要的意义, 它可以给出临床上诊断的临界值。



Laboratory Developed Test Kits

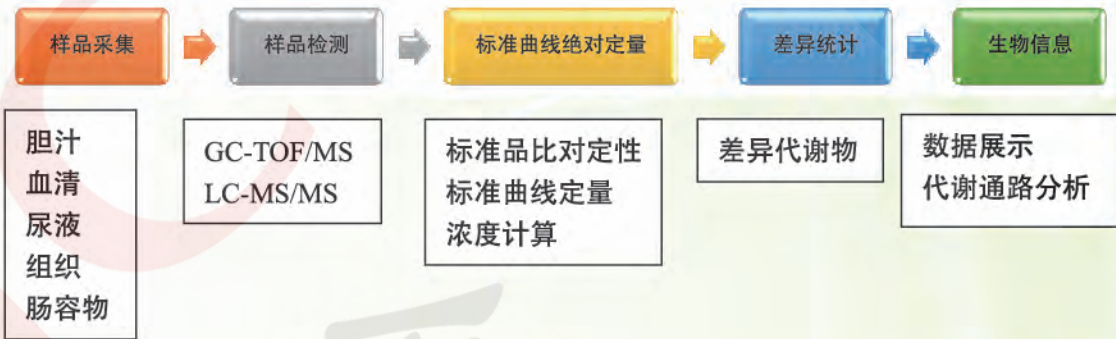
For FTN 84: 96 well plate (Waters 700 ul round well, MFC 136/05552, 7 mm round plug cap-mat, vial-06, MFC 136/06352)											
	1	2	3	4	5	6	7	8	9	10	11
A01	801	522	523	524	525	526	527	528	529	530	531
B13	532	533	534	535	536	537	538	539	540	541	542
C25	543	544	545	546	547	548	549	550	551	552	553
D37	554	555	556	557	558	559	560	561	562	563	564
E49	565	566	567	568	569	570	571	572	573	574	575
F61	576	577	578	579	580	581	582	583	584	585	586
G73	587	588	589	590	591	592	593	594	595	596	597
H85	598	599	600	601	602	603	604	605	606	607	608
I97	609	610	611	612	613	614	615	616	617	618	619
J09	620	621	622	623	624	625	626	627	628	629	630

Calibrators QC's NIST SRM 1950

同时, 通过定量代谢组学研究方法检测机体收到外界扰动后体内所有代谢物的量及

其量的变化, 可为疾病诊断和防治、药物开发, 药效筛选、药物的独立药理等提供更加精准、可靠的量化指标。该方法具有极高通量、卓越的可靠性和信息可挖掘性, 使得生物标志物的筛选过程变得简单、可靠。

技术路线:

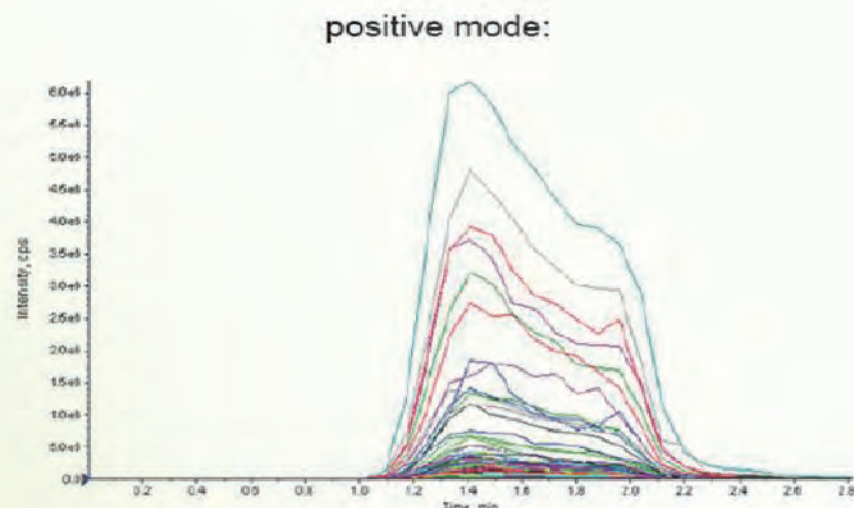


(一) 脂质检测

包括八大脂质 (脂肪酰类、甘油脂类、甘油磷脂类、鞘脂类、固醇类、异戊烯醇类、糖脂类、聚酮化合物) 276个亚类大于40000种脂质。

脂质组学是对生物体、组织或细胞中的脂质以及与其相互作用的分子进行全面系统的分析、鉴定, 了解脂质的结构和功能, 进而揭示脂质代谢与细胞、器官乃至机体的生理、病理过程之间的关系的一门学科, 属于代谢组学的一部分。通过比较不同生理、病理状态下脂代谢网络的变化, 揭示脂质在蛋白质表达和基因调控过程中的作用及在生命活动中所发挥的重要作用。

LC-MS/MS平台可准确定性定量检测各类样本中的脂质代谢物达140+种, 关注有生物学功能解释的酰基肉碱类、甘油磷脂类、鞘脂类、游离脂肪酸等, 该方法灵敏度高、准确性高且通量高。



LC-MS/MS 正离子模式检测脂质的流动注射分析 (FIA) 色谱图

(二) 胆汁酸检测

胆汁酸是一大类胆烷酸的总称，在脂肪代谢中起着重要作用。胆汁酸的生理功能可概括为：影响胆汁分泌、促进肠道对脂类物质的吸收并影响肠道功能。胆汁酸作为消化液的组成部分之一是由肝合成并随胆汁排入肠内，促进对脂类物质的消化和吸收。主要包括胆酸、鹅去氧胆酸、脱氧胆酸、石胆酸、熊去氧胆酸等及其甘氨酸和牛磺酸结合型。肝分泌到胆汁中的胆汁盐是强有力的乳化剂，胆汁流经小肠上部后，胆汁盐乳化脂肪，助其消化吸收。在乳化脂肪小滴的脂肪酸和甘油酯被小肠下部吸收以后，胆汁盐也被重吸收。它们回到肝并被重新使用，因此，胆汁盐总在肝和小肠间循环。肠道中大量的微生物会对这种肝肠循环产生影响。主要的影响菌类有乳酸菌、双歧杆菌、肠杆菌、拟杆菌和梭状芽孢杆菌。这些菌通过影响宿主胆汁酸的种类和含量，进而对诸如膳食脂肪和脂溶性维生素的吸收，促进脂质吸收，保持肠道屏障功能，信号系统内分泌功能，并能够在一定程度上调节甘油三脂、胆固醇、葡萄糖和能源体内平衡。

(三) 游离脂肪酸检测

脂肪酸是指一端含有一个羧基的长的脂肪族碳氢链的有机物。脂肪酸在有充足氧供给的情况下，可氧化分解为二氧化碳和水，释放大量能量，是机体主要能量来源之一。脂肪酸 β -氧化是体内脂肪酸分解的主要途径，脂肪酸氧化不但为机体提供了大量的能量，同时也生成了大量的体内化合物合成的中间原料乙酰CoA。脂肪酸在大部分组织中可以彻底氧化生成二氧化碳和水，但在肝脏中，脂肪酸的氧化很不完全，经常生成一些中间代谢产物，如乙酰乙酸、 β -羟基丁酸和丙酮等酮体。脂肪酸不仅为营养所必需，而且与儿童生长发育和成长健康有关，还具有降血脂、防治冠心病等治疗作用，且与智力发育、记忆等生理功能有一定关系。

(四) 肠道菌群-宿主代谢物检测

肠道菌群是哺乳动物体内重要的代谢“器官”，影响着宿主的整体代谢，宿主代谢受自身基因和肠道菌群基因的双重影响，宿主和菌群之间有着“共代谢”的过程。通过代谢组学方法来研究肠道菌群和宿主之间关联的复杂代谢体系，作为一种相对全面、连续动态并无刺激的分析手段，它可客观地检测肠道微生物的代谢组分及浓度变化，展示肠道菌的代谢状态。

数据平台特点：

探讨研究内容，设计方案实验——填写检测分析要求、签订合同并寄样——样品检测及数据分析——获取数据分析报告、探讨实验结果并验收。

案例分析

案例1.肝脏胆汁酸失调联合推动肝癌的发生(Dysregulated hepatic bile acids collaboratively promote liver carcinogenesis)

期刊名: International Journal of Cancer, IF= 5.5(2016).

研究对象: 肝癌小鼠

技术平台: UPLC-TQMS、Histological、16S rDNA、RT-PCR

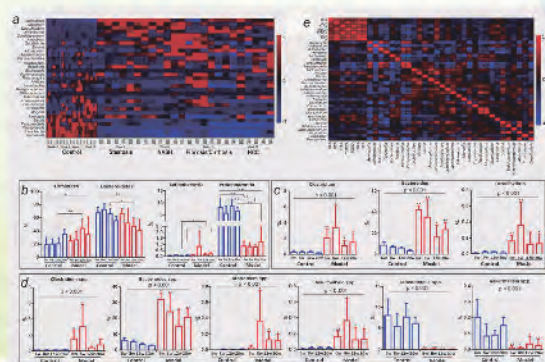
方案设计: 为了研究肝脏内肠道菌群代谢的胆汁酸的堆积是否会导致肝细胞的持续损伤密集而引发肝纤维化和肝癌的科学假说。

1.使用链脲佐菌素结合高脂饮食诱导非酒精性脂肪肝-肝癌小鼠模型, 6周-脂肪肝, 8周-脂肪性肝炎, 12周-肝纤维化, 16周-肝癌。

2.另一组小鼠高脂饮食中添加离子交换树脂以清除肠道经细菌代谢的胆汁酸, 收集小鼠血清、肝脏、粪便进行胆汁酸靶向代谢组学研究。

研究结论:

- (1) 高脂可以诱导肝脏胆汁酸的堆积, 而胆汁酸在肝内的滞留促进了肝癌的发生。
- (2) 胆汁酸代谢平衡的恢复伴随着肠道菌群平衡的恢复及肝脏病理状态的改善, 菌群代谢产物—疏水性胆汁酸很可能是肝癌形成的促发剂。
- (3) 胆汁酸一方面能显著改变肠道细菌的结构, 而肠道细菌结构的改变也伴随着胆汁酸水平的变化。



参考文献:

- (1) Guoxiang Xie, Xiaoning Wang, Fengjie Huang, Aihua Zhao, Wenlian Chen, Jingyu Yan, Yunjing Zhang, Sha Lei, Kun Ge, Xiaojiao Zheng, Jiajian Liu, Mingming Su, Ping Liu, Wei Jia*. Dysregulated hepatic bile acids collaboratively promote liver Carcinogenesis. Int J Cancer, 2016, 139(8): 1764-75. (IF=5.5)

- (2) Guoxiang Xie, Xiaoning Wang, Ping Liu, Runmin Wei, Wenlian Chen, Cynthia Rajani, Brenda Y. Hernandez, Rosanna Alegado, Bing Dong, Defa Li, Wei Jia*. Distinctly altered gut microbiota in the progression of liver disease. Oncotarget, 2016, 7(15): 19355-66. (IF=5.0)

案例2.肠道菌群代谢和三聚氰胺肾毒性相关性研究 (Melamine-Induced Renal Toxicity is Mediated by the Gut Microbiota)

期刊名: ScienceTranslationalMedicine, IF= 16.3 (2013).

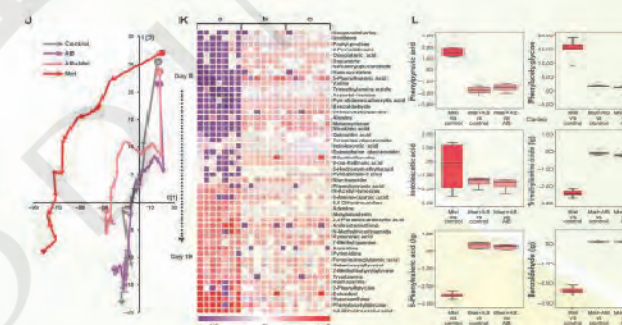
研究对象: 雄性Wistar鼠尿液样本

技术平台: UPLC-QTOF-MS、16S rDNA

方案设计: 假设哺乳动物体内肠道细菌能够代谢三聚氰胺产生三聚氰酸, 这两个物质在体内结合形成大分子物质产生肾毒性。三聚氰胺致大鼠肾毒性模型通过代谢组学平台观察肠道菌-宿主代谢物是否发生变化; 体内肠道菌抑制剂观察毒性是否下降, 通过体外肠道菌的培养说明体内肠道菌的代谢机制找到关键作用菌, 通过体内关键菌的添加考察肾毒性是否增加。

研究结论:

发现并验证了三聚氰胺产生肾毒性的机制—体内肠道菌群对三聚氰胺的代谢产生三聚氰酸, 造成肾小管中三聚氰胺-三聚氰酸的结晶沉淀; 证实了Kiebsiella菌在肠道菌群中为主要转化三聚氰胺的关键菌 代谢组学差异代谢物和通路的发现, 提出机制研究的假设, 对假设因素的抑制和添加, 体内动物实验和体外细胞实验的验证, 为临床肾毒性代谢组学的机制性研究提供了思路。



参考文献

- Melamine-Induced Renal Toxicity is Mediated by the Gut Microbiota. Sci Transl Med, 2013, 5(172): 172ra22. (IF=16.3)

案例3. 小分子代谢物诊断结直肠癌 (A Distinct Metabolic Signature of Human Colorectal Cancer with Prognostic Potential)

期刊名: Clinical Cancer Research, IF= 8.7 (2014)

研究对象: 376例结直肠癌和癌旁组织 (上海、杭州、北京、加州)

技术平台: GC-TOFMS 非靶向代谢组学检测、RT-PCR

方案设计: 对上述4个不同地区收集的样本进行非靶向代谢组学研究筛选出差异代谢物, 进而对差异代谢物所在代谢通路上的基因表达水平的变化进行研究并对结果进行验证。

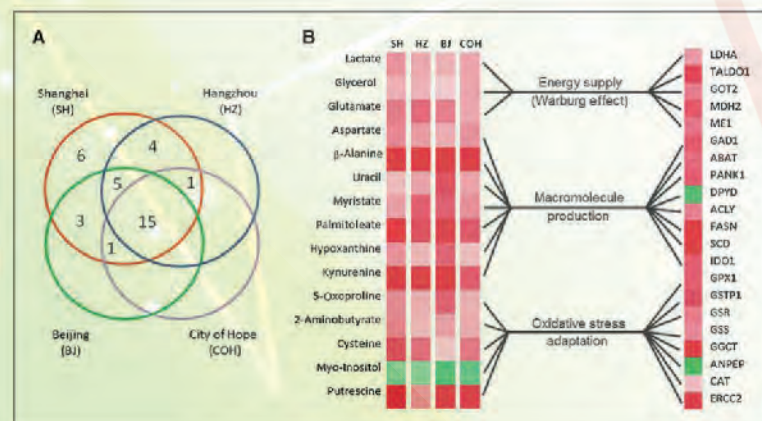
研究结论:

1. 对376例结直肠癌和癌旁组织分别进行非靶向代谢组学研究, 筛选出4组样本中检测到的变化水平完全一致的15个小分子代谢物。
2. 基于差异代谢物关联到的代谢通路上的基因进行RT-PCR分析, 筛选到21个差异表达的基因, 结果显示差异代谢物的变化和差异表达的基因变化趋势高度一致。
3. 结合代谢物和基因表达变化发现的结直肠癌的代谢物模式和基因表达模式特点主要从三个方面阐释其生物特性:

(1) “瓦伯格效应” (Warburg Effect): 是肿瘤细胞能量代谢的典型特征, 表现在大量地摄取葡萄糖进行有氧糖酵解, 生成大量的乳酸, 同时为不断生长的肿瘤细胞提供生物合成原料;

(2) 伴随着糖酵解的上升, 肿瘤细胞的代谢会产生大分子中间体来支持细胞生长, 导致某些特定的游离脂肪酸 (豆蔻酸、棕榈油酸) 和核酸 (次黄嘌呤) 的浓度上升。在肿瘤细胞中高表达的ACLY、FASN和SCD同样提示了脂肪酸合成的增强;

(3) 肿瘤细胞内维持较高的氧化应激水平, 氧化应激的生物标志物视晶酸、2-氨基丁酸在肿瘤细胞中上升。且与谷胱甘肽相关的基因GPX1、GSR、GGCT、GSTP1也在肿瘤组织中显著升高。



案例4. 游离脂肪酸预测肥胖者发展为糖尿病的风险 (Circulating Unsaturated Fatty Acids Delineate the Metabolic Status of Obese Individuals)

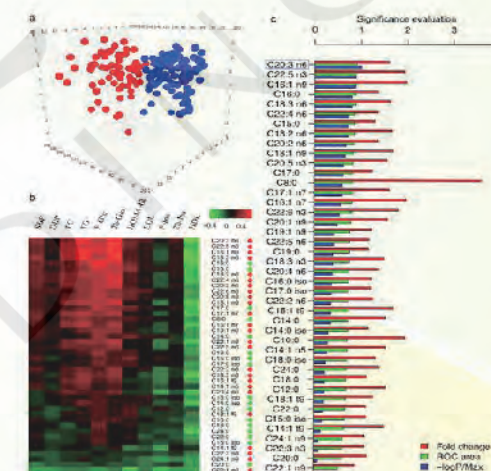
期刊名: EBioMedicine.

研究对象: 452份血清

技术平台: UPLC-QTOF-MS

方案设计: 一个横断面对照 (体重正常的健康者、健康肥胖者、肥胖伴糖尿病患者) 样本集、一个健康肥胖者长达10年随访的样本集、一个代谢术后随访2年的样本集、一个8周低热量饮食干预的样本集, 收集452份临床血清样本, 为寻找“代谢健康型肥胖”, 和“代谢非健康型肥胖”的代谢特征, 研究人员采用液质联用仪对其血清样本中的40余种游离脂肪酸进行了靶向代谢组学研究, 全面分析了游离脂肪酸与肥胖人群代谢状态的关系。

研究结论: 体重正常的健康者与健康肥胖者血清游离脂肪酸谱非常相似, 但肥胖伴糖尿病患者血清游离脂肪酸明显升高, 尤其是一些不饱和脂肪酸如DGLA (C20:3 n6) 升高更为突出, 且与糖尿病的临床指标明显相关; 对健康肥胖者的10年随访中发现, 那些未来发生糖尿病的肥胖者, 在基线时血清中的不饱和脂肪酸就出现了显著变化, 这些变化远远早于血糖、胰岛素等临床指标的变化; 而接受代谢手术治疗的肥胖者, 其不饱和脂肪酸在术后明显下降, 且术前这些不饱和脂肪酸水平的高低可以预测术后2年内糖尿病的复发情况。



参考文献

Yan Ni, Linjing Zhao, et al. Circulating Unsaturated Fatty Acids Delineate the Metabolic Status of Obese Individuals. EBioMedicine, 2015, 2(10): 1513-2

三.代谢组学检测平台

(一) GC-TOF/MS

仪器性能特点:

- ◆ 业界领先的采集速率, 最优的数据完整性;
- ◆ 自动峰识别和自动解卷积算法;
- ◆ 动态信号跟踪技术 (DSL), 优异的动态线性范围;
- ◆ 先进的数据挖掘算法, 提高被基质或共流干扰物遮盖化合物的识别效率;
- ◆ 全自动检测流程: 样品衍生化、进样、检测、数据挖掘、统计。



应用方向

疾病标志物开发; 肿瘤/癌症; 代谢疾病/心血管疾病; 营养学; 炎症/免疫; 感染性疾病; 神经科学; 植物学; 药物开发; 毒理研究; 微生物研究

(二) LC-MS/MS

仪器性能特点:

- ◆ 优异的耐用性, 实现复杂基质中低含量分析物的重现性检测;
- ◆ 采用高速采集速率, 可靠定量更多分析物;
- ◆ 快速极性切换可涵盖单次进样中的正负模式下离子化化合物;
- ◆ 配备Xtended Dynamic Range检测器, 实现简单易用的方法转换与高灵敏度;
- ◆ 可同时获取全扫描 (MS) 和MRM (MS/MS) 数据, 揭示样品复杂性。



第七章 测序平台

一.ABI 3730XL毛细管电泳测序系统

DNA测序是重要的分子生物学分析方法之一，它不仅为基因表达、基因调控等生物学基础研究提供重要数据，而且也在疾病诊断学、基因治疗等应用研究中起着重要的作用。浦东解码生命科学研究院拥有多台国际先进的DNA测序设备—ABI 3730xl测序仪（图1），提供高通量、高质量基因测序和片段分析服务，提供基因突变、多态性分析和基因克隆等配套服务，以满足客户的不同需求。



图1. ABI 3730XL测序仪

技术特点

- 自动化操作。
- 快速电泳，3小时左右即可达到测序要求。
- 新型液体分离胶使读序长度达1100bp，精确读序达800bp。
- 新型碱基识别与质量评分软件提高了测序准确性。
- 电泳温度可达70℃，有助于去除二级结构的影响。
- 高灵敏度提高了测序模板DNA的浓度适应范围。
- 良好的温控装置保证了片段分析准确性及重现性。
- 毛细管内荧光检测，灵敏度高。
- 双光束双侧激光激发，荧光信号强度高度均一。

浦东解码生命科学研究院的优势

- 成功率高。平均成功率在90%以上，大单成功率在98%以上。
- 测序准确率高、质量好，600bp内序列信噪比达98%以上。
- 价格实惠，一次送检100个样本以上价格特惠。差错率低。多台3730xl测序仪，一流的管理，一流的服务，认真负责的态度，将差错率降至最低。
- 保密性佳，保证您的成果不泄露、不被使用、不被剽窃。

一代测序服务流程如下：

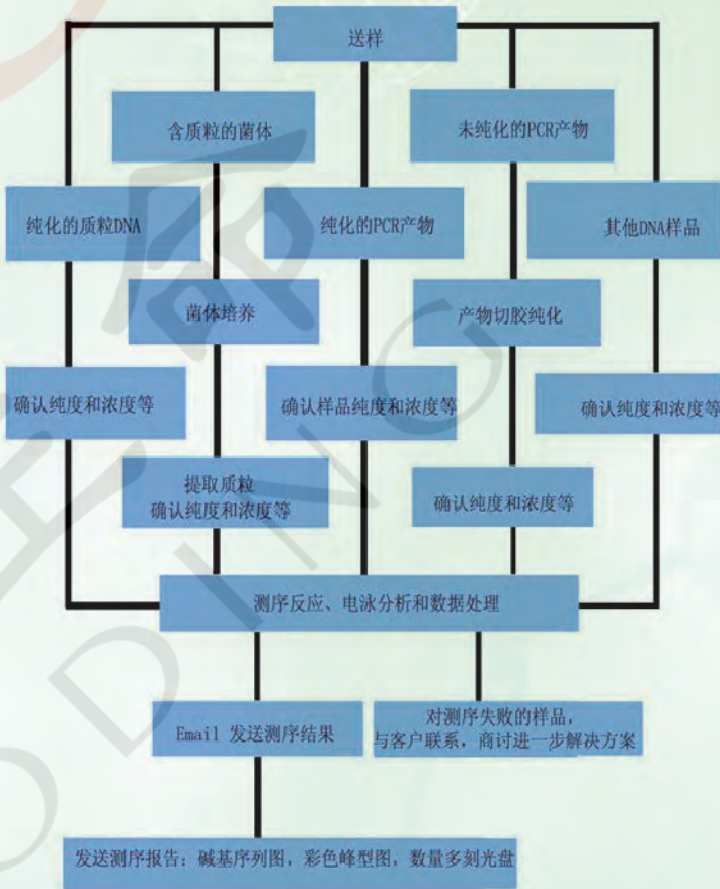


图2. 一代测序服务流程图

二. PGM/Ion Torrent测序系统

Ion Torrent, 即个人化操作基因组测序仪(PGM) (见图3), 其设计配合了革新性的半导体芯片技术, 使整个平台具有极高的扩展性和快速测序的性能。该技术使用了一种布满小孔的高密度半导体芯片, 一个小孔就是一个测序反应池。每次DNA聚合酶把核苷酸聚合到延伸中的DNA链上时, 都会释放出一个氢离子, 反应池中的pH也会随之发生改变, 此时位于池下的离子感受器感受到信号, 把化学信号直接转化为数字信号, 从而读出DNA序列。与其它新一代测序仪相比, 它不需要激发光、CCD成像仪或荧光标记, 能直接并快速“读”出DNA序列。



图3. 个人化操作基因组测序仪

技术特点

扩展性强: 大规模可扩展的半导体芯片技术, 比拟半导体技术升级的强大扩展性。技术的基础以超过千亿美元规模的半导体产业为支持, 芯片密度的提升以40年来累积验证的摩尔定律为证据, 芯片产量的扩展以四百家标准化半导体生产企业为选择。

技术简洁: 无需光学部件, 简单的生物化学原理, 真实的离子流测序原理, 无修饰的核酸, 无需酶化学级联, 无需任何荧光和化学发光经济的标准试剂。

快速: 2个小时完成测序工作。

检测原理

Ion Torrent 技术通过专有的大规模并行半导体感应器, 对于DNA 扩增时产生的离子流, 实现直接和实时的检测。当试剂通过集成的流体通路进入Ion Torrent半导体芯片中, 密布于芯片上的反应孔立即成为上百万个微反应体系。这种独特的流体体系、微体系机械设计和半导体的技术组合, 使得该系统得以快速直接地将遗传信息翻译成数码的DNA测序结果, 从而获得大量高质量的测序数据 (图4)。

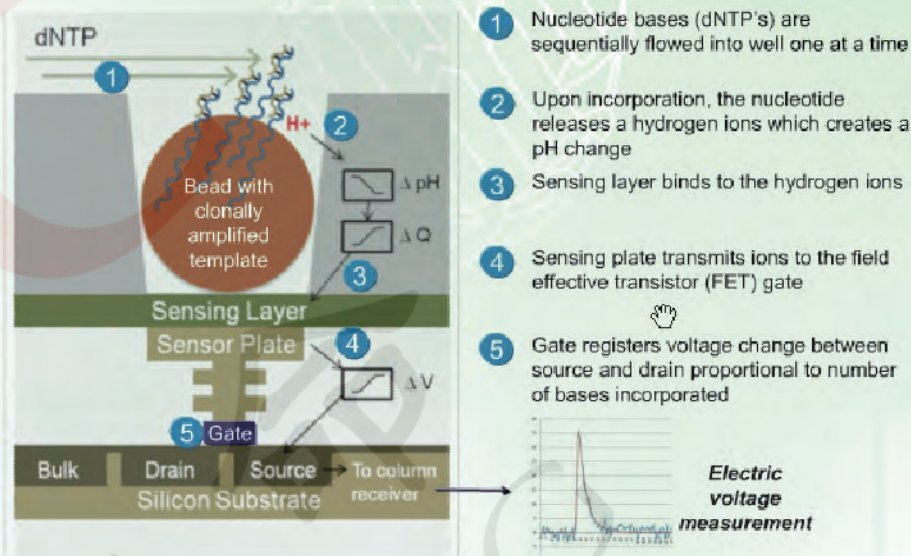


图4. Ion Torrent的测序机理及流程

应用范围

基因组验证; 全外显子验证; 扩增子重测序; 小基因组测序; 靶向重测序。

浦东解码生命科学研究院的优势

最大化地提高Ion Torrent 技术本身的优势特点, 独创了一套操作体系, 能够显著提高客户样本的反应成功率。

通过专业的团队迅速为客户提供最优的实验方案, 以及备用实验方案, 保证了客户珍贵样本的数据准确率和实验成功率。

拥有先进的生物信息学分析团队, 开发出多种强大的分析软件, 为客户提供最全面、最优秀的数据分析。

生物信息分析服务

标准生物信息学分析

- (1) 序列与参考序列的mapping比对, 生成bam文件, 包含序列在基因组的位置信息;
- (2) 同时对数据进行质控检测, 包括测序深度、覆盖度等分析的检测报告。

高级生物信息学分析

- (1) apping优化, 去除PCR重复子;
- (2) 通过序列比对分析SNP和插入缺失 (InDel) ;
- (3) 于peak的多样品间差异分析;
- (4) 基因表达分析;
- (5) SNP变异分析;
- (6) 融合基因分析;
- (7) 筛选突变。

三. Illumina Hiseq 3000/4000测序系统

Illumina公司的Hiseq 3000/4000测序系统 (见图5) 是一次运行就能产生数百亿到数千亿碱基数据的高通量测序平台, 该技术的成本不足毛细管测序的万分之一。Hiseq测序平台运用了微阵列技术和专有的可逆终止子技术, 可快速而准确的进行大规模测序。



图5. Illumina Hiseq测序系统

技术特点

💎 快速, 高通量——HiSeq使用一或两个流动槽和一种新颖的双表面成像方法, 能够在单次运行中产生多至1.5TB的数据及50亿条序列, 使测序通量及实验灵活性提升到新的水平。

💎 操作简便, 省时——HiSeq 具有最简化和直观的操作流程, 预先配置的即插即用试剂, 足够300个循环使用, 启动运行手工操作时间短。触摸屏界面, 以及实时的远程运行监控, 让仪器操作更简单。文库制备时96个样品的平行处理显著降低了手工操作时间和整体的费用。

💎 费用低廉, 灵活应用——对人类基因组进行测序时, 假设人基因组覆盖度为30倍, 可同时检测多达12个人全基因组, 每个基因组的测序费用更低。HiSeq能在单个或两个流动槽模式下操作, 独立操控的流动槽让需要不同读长的应用能同时运行, 达到了实验灵活性和仪器可扩展性的较好结合。

HiSeq测序系统目前可应用在多个领域, 包括: DNA测序、基因调控分析、转录组分析、SNP发现与结构变异分析、细胞遗传学分析、DNA-蛋白质相互作用分析 (ChIP-Seq)、测序法甲基化分析以及小RNA的发现及分析。

检测原理

Hiseq测序平台可以生成高密度、海量平行测序反应, 可对每个流动槽中成百上千万的模板进行单端或对读测序。全自动的illumina簇生成工作站在恒温条件下在流动槽表面扩增产DNA簇, 每个簇含有单个模板分子的500-1000克隆拷贝。用全自动基因分析仪在流动槽表面对生成的高密度序列模板进行测序。采用专有的荧光标记的可逆终止核苷酸, 可对样本边合成边测序。完成首次读取后, DNA簇进行原位改良, 产生二次测序的模板。之后应用第二个测序引物对同一DNA簇测序, 生成第二次数据。见图6。

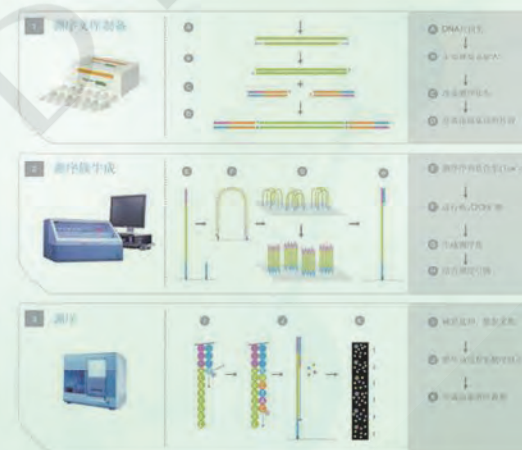


图6. Hiseq测序原理及流程

浦东解码生命科学研究院服务领域

SNP和结构变异检测；全基因组重测序；全外显子测序；转录组测序；微小RNA测序；ChIP-Seq；甲基化检测；羟甲基化测序；单细胞测序。

生物信息服务

标准生物信息学分析

- (1) 序列与参考序列的mapping比对，生成bam文件，包含序列在基因组的位置信息；
- (2) 对数据进行质控检测，包括测序深度、覆盖度等分析的检测报告；
- (3) RNA-seq序列质量过滤；
- (4) MeDIP-seq序列质量过滤；
- (5) MeDIP-seq序列与参考序列的比对。

高级生物信息学分析

- (1) mapping优化，去除PCR重复子；
- (2) 通过序列比对分析SNP和插入缺失(InDel)；
- (3) MeDIP-seq序列数据在全基因组的分布趋势；
- (4) 统计MeDIP-seq序列富集区域(peak)的信息；
- (5) 基于peak的多样品间差异分析；
- (6) RNA-seq序列与参考基因组序列的比对；
- (7) 基因表达分析；
- (8) SNP变异分析；
- (9) 融合基因分析；
- (10) 筛选突变。

浦东解码生命科学研究院的优势

高质量的测序数据：在每个流动槽中对数百亿到数千亿碱基机型准确的碱基识别。在读测序时，模板可以重新生成。在两次读取间获得相同准确度，严格的功能检测确保检测稳定性和重复性。

应用广泛：所有的深度测序均可在该设备上完成。

高效的样本制备：高速自动化工作流程，低样品量要求，可在两周的时间内生成所有数据。

灵活的平台：稳定的Solexa测序化学反应支持多种读取长度，研究人员可量身定制每次所需读长，数据整体质量得以提高。

四.焦磷酸测序(Pyrosequencing)系统

焦磷酸测序(Pyrosequencing)系统(见图7)通过合成反应进行测序，是一项独特的检测技术，能够提供定量的实时数据，并且无需凝胶、探针或标记物。在对单核苷酸多态性(SNP)，插入和缺失(indel)及未知的序列变异进行鉴定时，该技术还能对等位基因频率及CpG与非CpG(CpN)位点的DNA甲基化水平进行灵敏的定量检测。强大的PyroMark平台同时集成了遗传变异的检测和定量功能，该系统上的焦磷酸测序技术在靶向分析短DNA序列时优于其他基于序列的解决方案。



图7. Pyrosequencing系统

技术特点

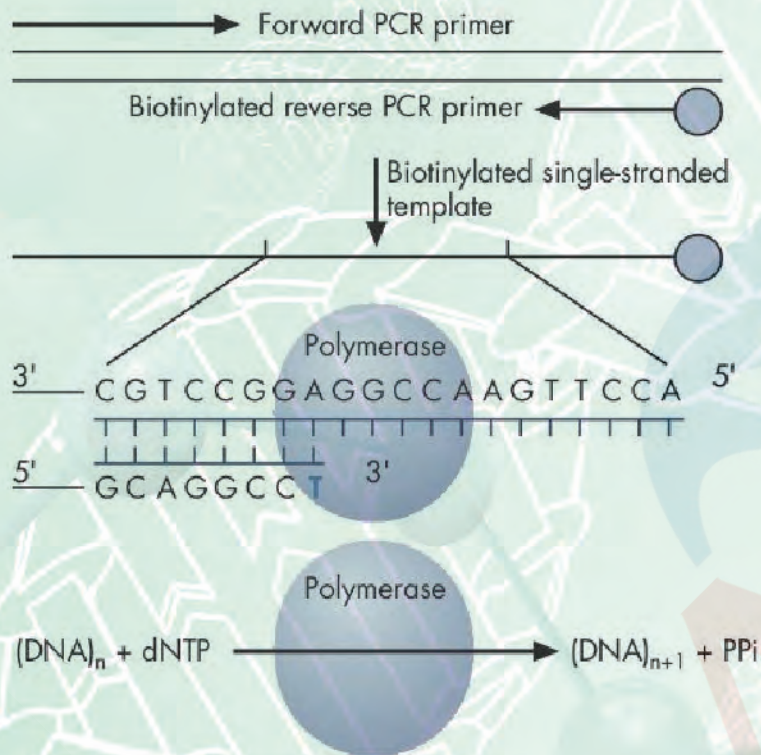
- (1) 操作简单。
- (2) 在同一台仪器上对同一批样品进行多功能分析。
- (3) Shinzato 可靠的定量分析等位基因和甲基化状态。
- (4) 序列信息帮助发现稀有突变。
- (5) 15分钟内平行分析1-24个样本。
- (6) 紧凑的检测平台，占用空间最小。

检测原理和实验步骤

步骤1：扩增DNA片段并对焦磷酸模板链进行生物素化。通过变性，分离生物素化的单链PCR扩增子并使其与一条测序引物进行杂交。

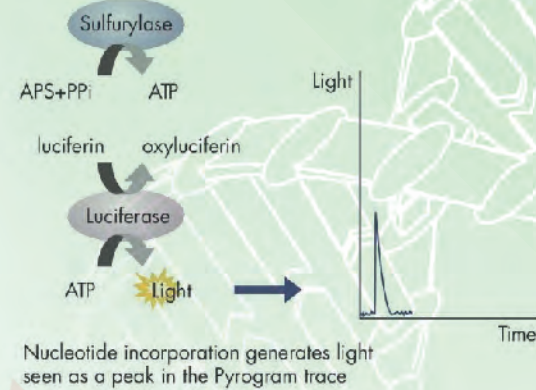
步骤2：将杂交引物和单链模板与DNA聚合酶、ATP硫酸化酶、荧光素酶和腺苷三磷酸双磷酸酶，以及底物腺苷-5'-磷酸硫酸酐（APS）和荧光素共孵育。

步骤3：第一个脱氧核苷酸三磷酸(dNTP)被引入反应。DNA聚合酶催化dNTP，使其在互补模板链的情况下对测序引物进行延伸。每个核苷酸的添加都对应着等摩尔数焦磷酸(PPi)的释放。

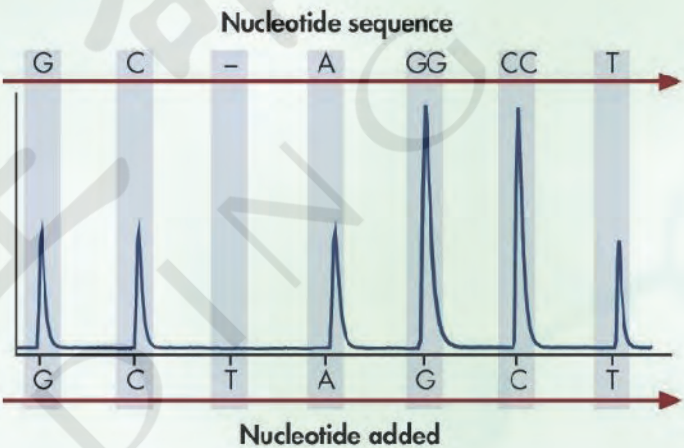


步骤4：ATP硫酸化酶在腺苷-5'-磷酸硫酸酐（APS）存在的情况下将PPi转化为ATP。生成的ATP为荧光素酶介导的荧光素氧化反应提供能量，并生成与ATP的数量成比例的可见光。电荷耦合器(CCD)摄像头检测到荧光素酶催化反应所生成的光，并将其作为原始输出数据中的一个峰值（热解图Pyrogram）。每个峰值（光信号）的高度与核苷酸结合的数目呈比例关系。

步骤5：腺苷三磷酸双磷酸酶是核苷酸的降解酶，持续地对未结合的核苷酸和ATP进行分解。当分解完成时，便会引入另一个核苷酸。



步骤6：dNTP的添加反应在持续进行着。应注意脱氧腺苷三磷酸α-硫代三磷酸（dATP α S）能够作为天然脱氧腺苷三磷酸（dATP）的替代物，被DNA聚合酶有效利用，却并不会被荧光素酶所识别。当反应进行时，逐渐合成了互补的DNA链，其核苷酸序列能够凭借热解图（Pyrogram）轨迹中的信号峰值来确定。



浦东解码生命科学研究院服务领域

序列分析: 序列鉴定；微生物分型鉴定；线粒体DNA鉴定/法医；转基因序列鉴定；片段重复序列分析。

突变检测：已知序列寻求突变位点；未知序列含有热点序列；点突变；插入/删除。

定量分析：等位基因频率检测；甲基化检测。

五. Agena MassArray系统

(一) 背景介绍

Agena MassArray时间飞行质谱生物芯片系统是为基因组学研究提供兼顾灵敏度和特异性服务的中高通量技术平台，广泛地应用于遗传突变检测、SNP分型以及DNA甲基化定量分析研究，是目前唯一采用质谱法进行直接检测的方法。

(二) 应用领域

SNP分型以及SNP位点等位基因频率计算。

体细胞突变检测和分析。

甲基化定量分析。

基因表达定量分析。

CNV检测分析。

寡核苷酸质量控制和检测。

(三) 技术特点

高通量：一张芯片可对384个样本进行多重检测；每个体系最多可实现40重反应；通量可根据客户要求个性化调整。

高性价比：无需荧光标记，仅需合成普通引物，大大降低成本；单个分析成本低；适用范围广，几十到成千上万个样本，同时检测几十到成百上千个位点。

高灵敏度：分析所需样本量少（10ng）；检测精度高；可进行定量分析。

高灵活度，功能多样：适用于SNP分型以及DNA甲基化定量分析；一张芯片上样本数量和位置可随意选择；一张芯片上样本和位点检测匹配可随意选择。

(四) 检测原理

Agena MassArray质谱检测系统主要是利用基质辅助激光解吸电离飞行时间质谱 (MALDI-TOF MS) 进行分析，即PCR扩增产物或者预处理样本在延伸单碱基后，将制备的样品分析物与芯片基质共结晶，将该晶体放入质谱仪的真空管，而后用瞬时纳秒 (10-9s) 强激光激发。由于基质分子经辐射所吸收的能量，可导致能量蓄积并迅速产热，从而使基质晶体升华，核酸分子就会解吸附并转变为亚稳态离子，产生的离子多为单电荷离子，这些单电荷离子在加速电场中获得相同的动能，进而在非电场漂移区内按照其质荷比率得以分离，在真空小管中飞行到达检测器。见图1。

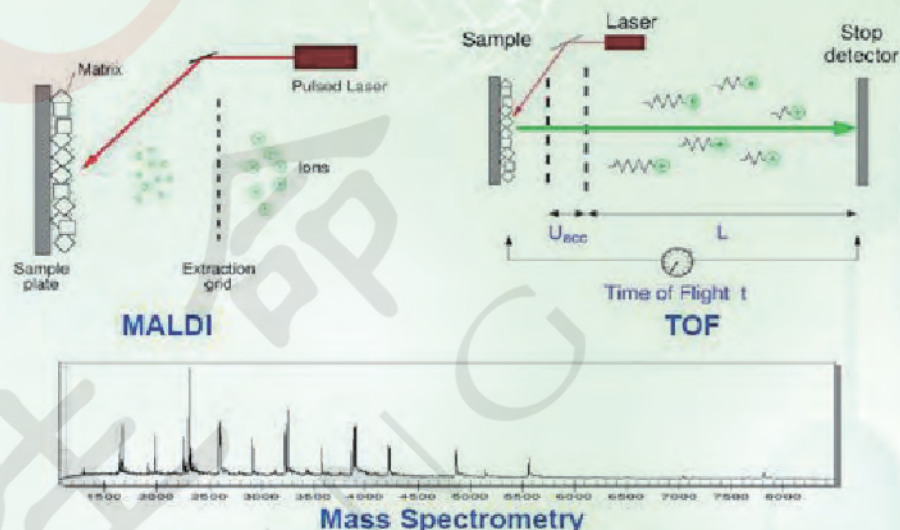


图1. Agena MassArray检测系统检测原理

(五) 浦东解码生命科学研究院提供Sequenom MassArray芯片服务

Agena MassArray系统反应体系为非杂交依赖性，不需要各种标记物，实验设计灵活，更可实现高达40重反应，是目前市场上最高性价比的产品。浦东解码生命科学研究院提供多种后续的生物信息分析服务。

第八章 芯片平台

一. Illumina微阵列扫描仪系统

Illumina微阵列扫描仪系统（见图8）的核心是iScan 阅读器，它包含了高性能激光器，光学仪器及其检测系统，能够提供亚微米分辨率和非匹配的数据输出率，甚至最高密度的芯片亦可在几分钟内完成扫描，每天可以处理多达96张芯片，数百个样品。



图8. Illumina微阵列扫描仪系统

技术特点

这些成像特性尤其有利于进行高密度基因型分型，CNV分析，DNA甲基化和基因表达分析。该扫描系统支持各种多样品芯片格式的infinium、GoldenGate、DASL、基因表达（直系杂交）和甲基化分析。

AutoLoader 2x 可被加入到单或双扫描系统，从而能每天高通量输出成千上万的样品数据。

二. 安捷伦DNA微阵列扫描仪

安捷伦DNA微阵列扫描仪（见图9）是一个具有48片装片器的扫描系统，它能够读取任何1" x 3" (25mm x 75mm) 芯片，厚度和宽度可有一定的差别，可以利用安捷伦图像分析采集软件无缝连接进行图象分析，采用SureScan高分辨技术提高基因组的覆盖率。无论是进行基因表达研究、aCGH、miRNA表达谱，还是其它新应用，安捷伦的扫描技术都能够提供很高质量的数据。



图9. 安捷伦DNA微阵列扫描仪

技术特点

图像稳定、灵敏度强。在扫描过程中，动态自动聚焦功能可以大大提高稳定性，增强灵敏度。超低限检测，最佳的精确光学系统、宽动态范围以及最低的光谱干扰使得设备能够检测微弱的信号，并且获得快速、准确和可靠的结果。

快速扫描：可在2、3、5、10微米分辨率下同时进行双色扫描，迅速得出结果。

三. Affymetrix新型基因芯片扫描仪

Affymetrix基因芯片扫描仪（见图10）具有扫描各种Affymetrix专属芯片的功能，范围2.5-0.51 μm ，支持最新的高密度GeneChip芯片，可进行从tiling、全外显子到单核苷酸多态性基因分型的研究。这种新型扫描仪设计用于各种操作的峰扫描，包括RNA表达和DNA分析。



图10. Affymetrix新型基因芯片扫描仪

技术特点

灵敏度高，可检测小于0.5 个藻红蛋白/ μm^2 。

扫描分辨率高，可达2.5 微米。

扫描速度快，2分钟检测一张标准芯片。

系统可靠性高，每年可以保证完成1 万次检测。

更新的Auto-Zero 系统，大幅降低背景信号。

四. 罗氏MS200扫描仪

罗氏MS200扫描仪（见图11）是首台经过优化能揭示阵列全部潜力并捕获所有 DNA 微阵列完整图片的扫描仪。这台高灵敏扫描仪能获取最小为2微米像素分辨率的双色荧光阵列图像数据，并生成高质量的图像文件。



图11. 罗氏MS200扫描仪

技术特点

具有全自动系统的多项重要功能，能进行无人自动扫描和通宵运行。系统包括一个芯片盒，能自动加载最多48个芯片。

通过高级动态自动对焦，可清楚地扫描数据图片，并且自动获取清晰的扫描效果。

内置QC内参芯片，可对芯片结果进行轻松校正，保证扫描结果准确可靠。

高信噪比，内部校准，集成条码读取器，可减少臭氧功能，这些功能可确保每次扫描数据的高质量及可靠性。

第九章 生物信息学平台

一. 平台简介

目前, 解码的生信平台主要由基因组分析平台、转录组分析平台、表观遗传组学分析平台、微生物组学分析平台、以及蛋白质组和代谢组分析平台组成。其中, 基因组分析平台主要包括全基因组测序分析、全外显子组测序分析和扩增子测序分析等。转录组分析平台主要包括mRNA测序分析、lncRNA测序分析、sRNA和 microRNA测序分析、以及10×Genomics单细胞测序分析等。表观遗传组学分析平台主要包括全基因组甲基化测序分析、全基因组甲基化芯片检测分析等。蛋白质组检测分析平台主要包括蛋白质组定性检测、蛋白质组定量检测、以及蛋白质化学修饰组学检测分析等。代谢组检测分析平台主要包括非靶向代谢组学检测和靶向代谢组学检测分析。

此外, 为了便于客户的研究, 浦东解码生命科学研究院提供了标准化和个性化的生物信息学服务, 不仅提供一目了然的高通量测序和芯片检测结果报告, 同时提供高质量的数据深度挖掘分析服务, 以助力每一位临床科研人员的科研事业。

二. 测序分析流程

(一) 基因组测序常规分析流程

基因组测序数据的常规分析流程只要给出样本中检测到的SNP和InDel的注释信息。

该流程分析从测序下机数据开始, 主要得到以下结果:

- (1) 各样本中注释后的SNP和InDel信息;
- (2) SNP、InDel与表型的关联分析结果;
- (3) 家系分析。包括寻找de novo变异以及复合杂合变异。

(二) mRNA、lncRNA和circRNA常规分析流程

常规分析流程中包括了mRNA、lncRNA和circRNA三部分内容。该流程分析从测序下机数据开始, 主要得到以下结果:

- (1) 各样本中mRNA、lncRNA和circRNA的表达谱;
- (2) 比较组中差异表达的RNA;
- (3) 比较组中差异表达的mRNA的GO、KEGG富集结果;
- (4) 预测新的lncRNA;
- (5) 预测已知lncRNA的靶基因;
- (6) 以上结果的相关图表包括基因表达热图、通路散点图等。

(三) sRNA测序常规分析流程

常规分析流程中包括了microRNA和sRNA两部分内容。该流程分析从测序下机数据开始, 主要得到以下结果:

- (1) 样本中miRNA和sRNA的表达谱;
- (2) 比较组中差异表达的RNA;
- (3) 预测新的miRNA;
- (4) 预测已知miRNA的靶基因;
- (5) 以上结果的相关图表。

(四) 16S测序常规分析流程

16S测序分析的常规流程主要给出各样本中微生物的分布情况。该流程分析从测序下机数据开始, 主要得到以下结果:

- (1) 各样本中微生物丰度;
- (2) Alpha多样性分析;
- (3) Beta多样性分析;
- (4) ANCOM(analysis of composition of microbiomes)。这是一种微生物群落差异分析的方法, 能给出组间具有差异的微生物。

(五) 个性化分析与绘图

除上述常规流程外, 客户可以选择多种个性化的分析手段, 定制分析流程。可选项包括:

- (1) 回归分析;
- (2) 相关性分析;
- (3) PCA/PLS-DA分析;
- (4) ROC曲线分析;
- (5) 个性化图表绘制。

三. 结果报告

浦东解码生命科学研究院提供的二代测序和芯片检测的结果报告, 内容涵盖了芯片的原始数据和芯片图像。原始数据包含了各种格式的文件, 包含了客户所需要的全部信息。所提供的芯片图像, 便于客户查看芯片的各个细节和结果。

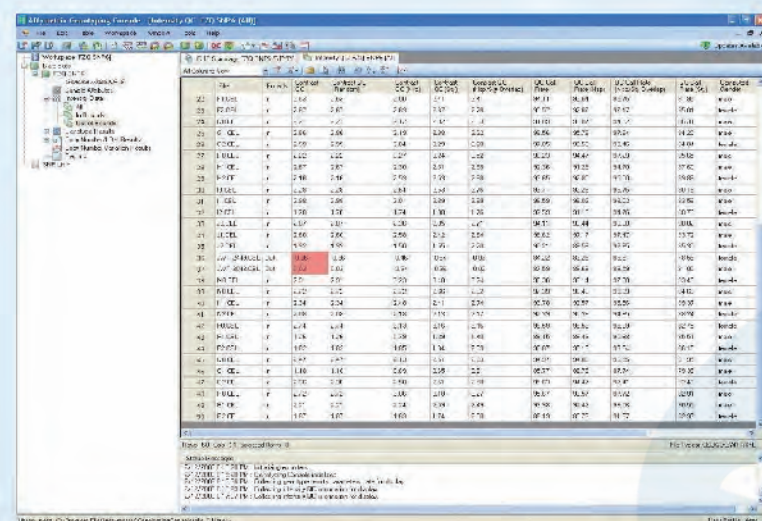
连锁不平衡性分析

包括原始的.EXP、.DAT、.CEL 及report 文件。

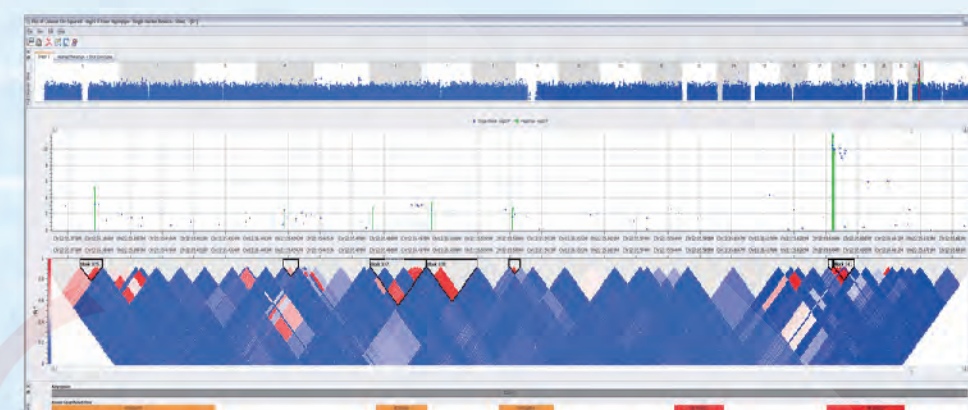
(二) 各种分析结果

以文本文件格式提供给客户所有样品SNP位点的基因分型(Genotype Call)和置信度评分(Confidence Score)等。

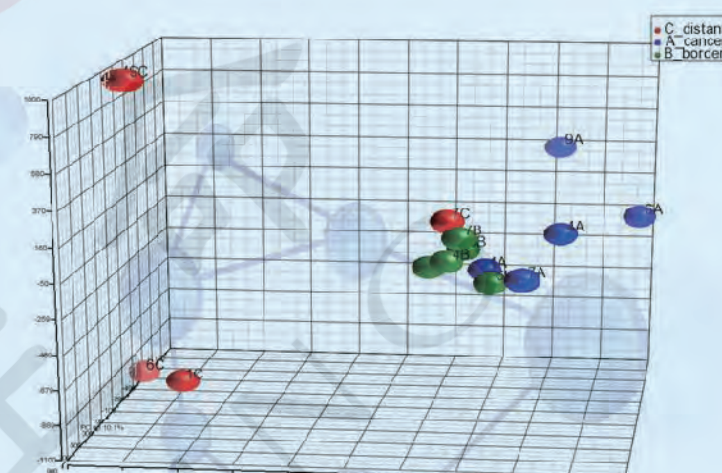
Genotyping Console软件给出的质控报告



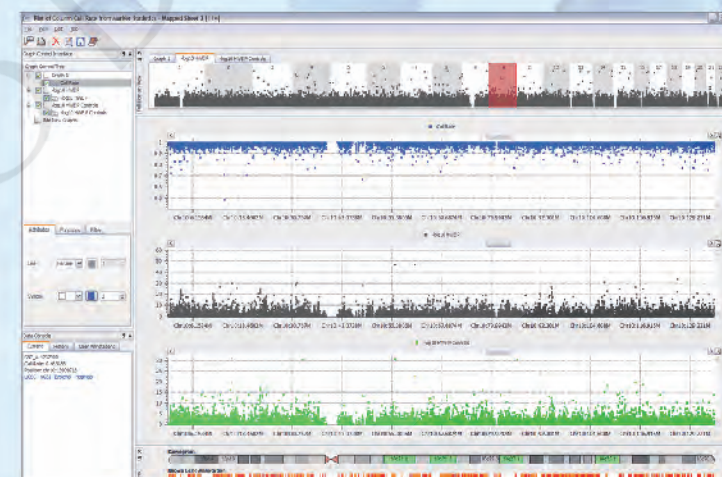
样品SNP call结果总表

[illegible]

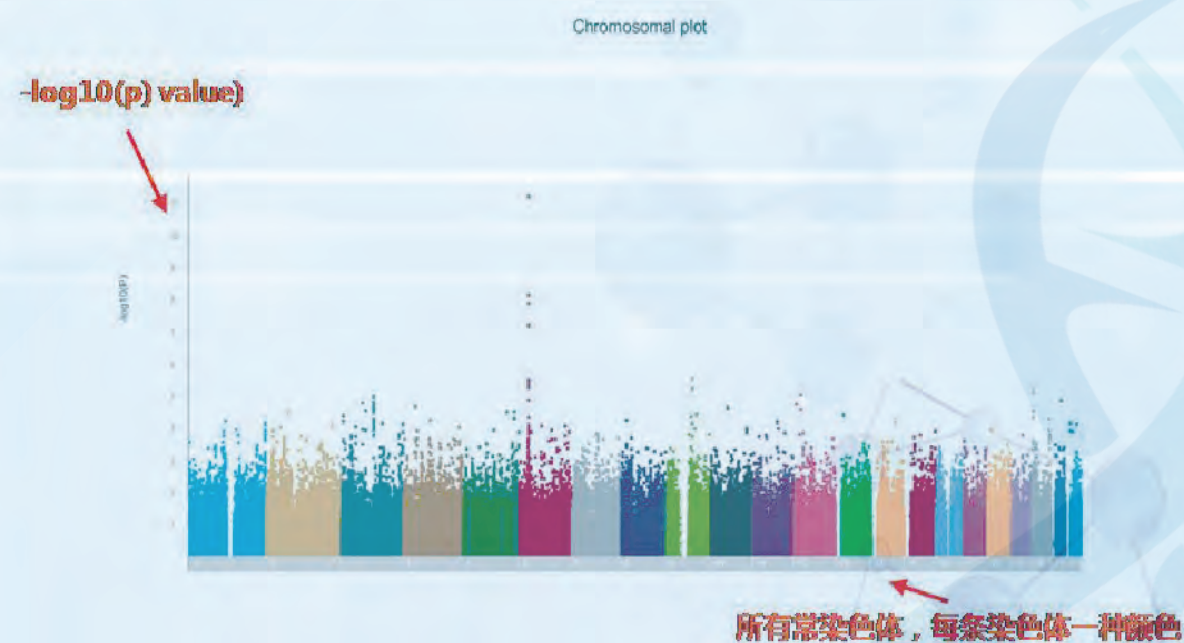
人群分层分析



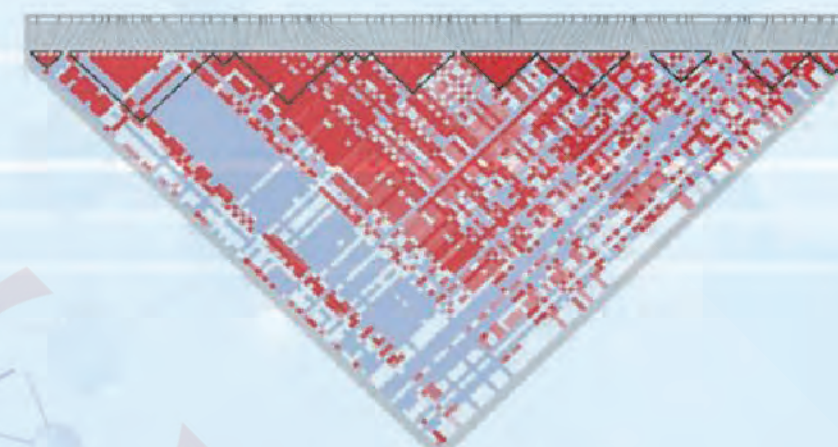
阳性位点深度过滤



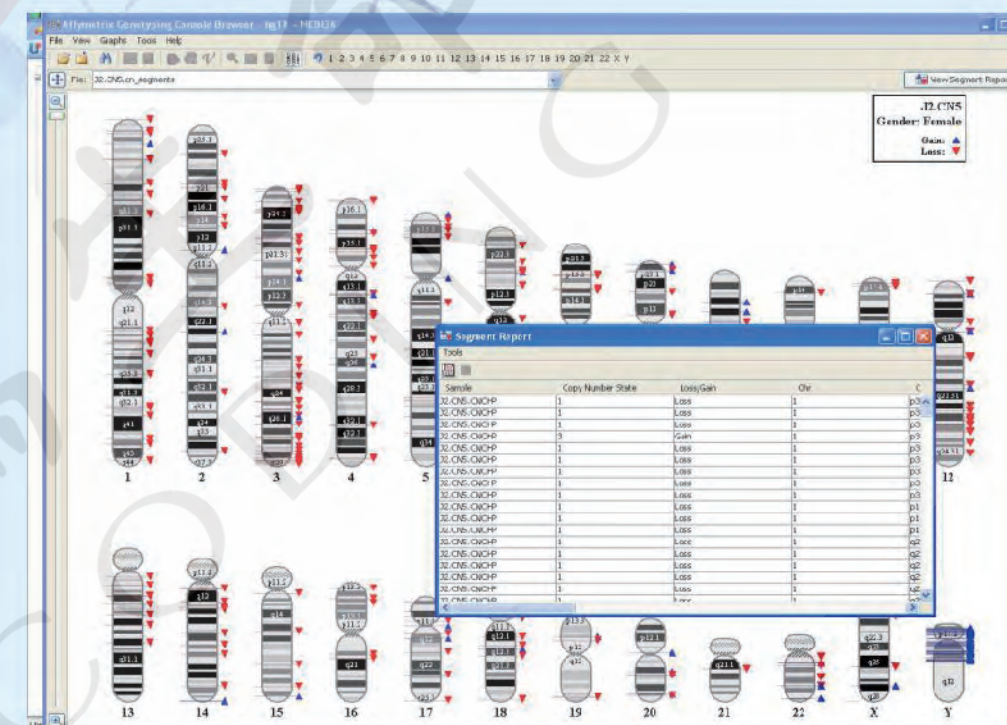
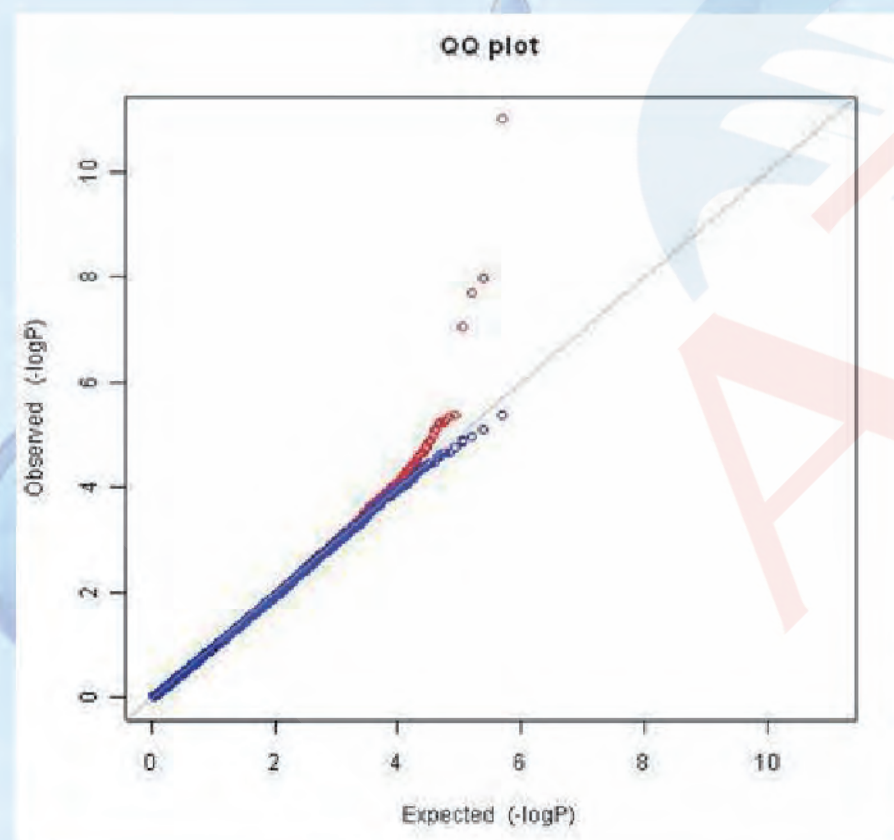
单点关联分析



精确定位分析



与客户指定对照组样品间的CNV比对结果总图



四.生信平台优势

基因功能分类注释：针对差异基因进行生物学功能分类，采用GO数据库中的功能聚类注释结果，并统计P-value值。

代谢通路分析 (Pathway分析)：建立信号通路和生物功能网络，将差异基因与相关信号通路整合和比较，根据统计检验方法 (P-value) 筛选出显著差异的代谢通路。

转录因子分析 (TF分析)：利用相关转录因子数据库，针对每个转录因子分析其在差异基因中的分布情况。统计寻找有差异的转录因子。

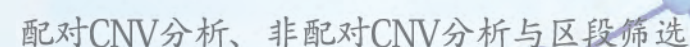
基因网络构建：采用贝叶斯方法，对表达数据进行机器学习，构建差异基因间的动态网络。

网络构建：利用文本挖掘、同源预测、PPI和基因融合等，进行网络构建。

疾病分型：主要是针对复杂疾病，采用芯片数据对这些疾病的子类进行区分定型。

聚类分析：主要是进行芯片间及差异表达基因间的双向聚类，对新基因进行功能预测，分类。聚类方法有层次聚类，K-means聚类及SOM等。

针对客户研究特点, 设计并提供个性化分析服务。



第十章 分子生物学平台

一. 基因编辑(CRISPR-Cas9)

(一) 背景介绍

CRISPR-Cas9系统(Clustered Regularly Interspaced Short Palindromic Repeats-Cas9 Systems)是细菌及古细菌进化出来用以抵御病毒和质粒入侵的适应性机制。CRISPR-Cas9系统的高效基因组编辑功能已被应用于多种生物,包括斑马鱼、小鼠、大鼠、秀丽隐杆线虫、植物及细菌。多个科研小组的研究都显示,与锌指核酸酶(ZFNs)和转录激活样效应核酸酶(Transcription activator-like effector nucleases, TALEN)相比较,CRISPR-Cas9系统介导的基因组靶向实验在细胞或斑马鱼中显示出相似甚至更高的效率。

(二) 技术原理

在CRISPR-Cas9系统中,CRISPR RNA(crRNA)与转录激活crRNA(Trans-activating crRNA, tracrRNA)退火形成的复合物能特异识别基因组序列,并引导Cas9核酸内切酶在靶生成DNA双链断裂(DSB),诱导基因敲除、基因敲入(配合供体克隆)、序列修饰等基因组编辑操作。识别复合物可以通过融合crRNA与tracrRNA形成sgRNA(single-guided RNA)进行简化。基因组靶点序列上上长有约20bp与sgRNA识别区序列一致。该序列末端紧接着三核苷酸区域PAM(5'-NGG-3')为Cas9识别位点,是实现剪切功能的关键。sgRNA识别并结合该序列的互补链;随后Cas9核酸酶剪切被识别的靶点DNA双链。

CRISPR-Cas9体系的RNA-DNA识别机制为选择性基因组编辑提供了一个简便而强大的工具。该体系其中一个最重要的优势是Cas9蛋白可在多个不同的sgRNA的引导下同时修饰多个基因组靶点。

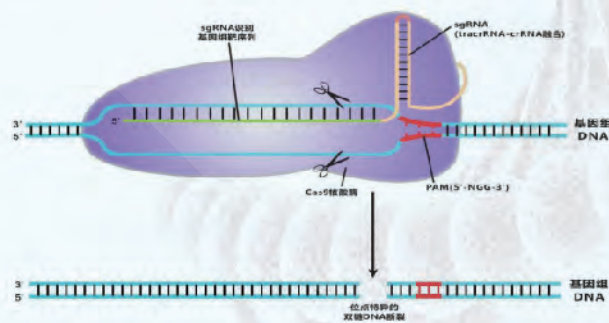


图 1. CRISPR-Cas9介导基因组编辑原理图

优势: RNA导向的基因组DNA识别,无需考虑DNA甲基化;较ZFN和TALEN,有相同或更高的基因编辑效率;可同时编辑多个基因(多重靶向编辑);设计简单快速,无需重复构建核酸内切酶。

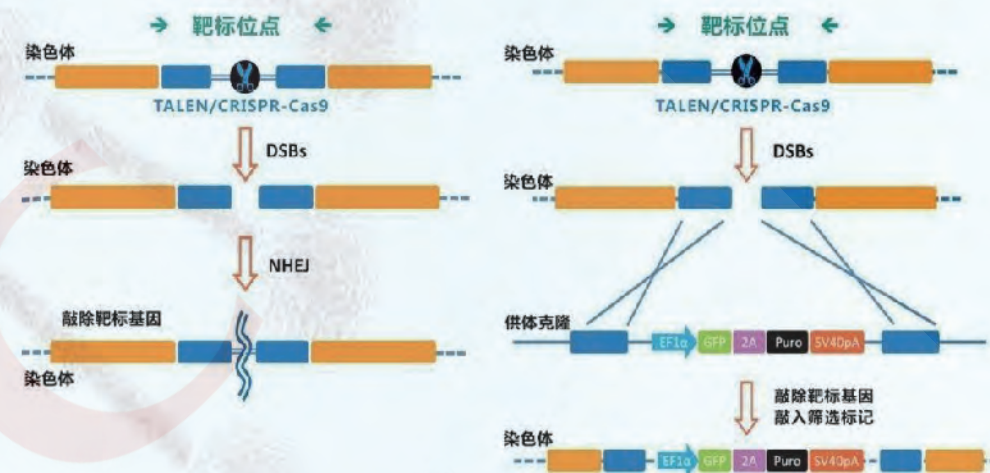


图2. CRISPR-Cas9介导的基因编辑。(左): sgRNA引导Cas9核酸酶作用于基因组,形成的DSBs被非同源末端连接(NHEJ)机制修复;(右): sgRNA引导Cas9核酸酶作用于基因组,形成DSBs,同源重组(HR)作用下,供体质粒上的目标基因及筛选标记(或其他遗传元件)在断裂处被整合进基因组。

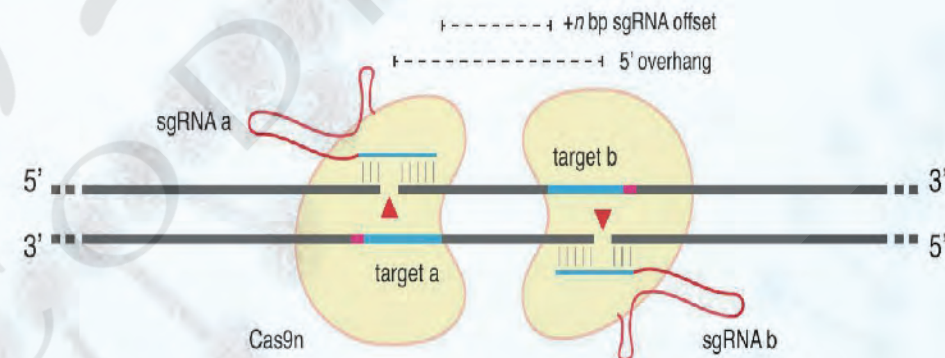


图3. 切口酶在结合链生产单链切口的原理图

(三) 相关服务

服务项目		内容描述
验证服务	错配酶切检测	染色体水平功能验证。通过检测CRISPR-Cas9介导的非同源末端连接 (NHEJ) 在染色体靶点上生成的插入缺失突变验证其功能。
供体克隆服务	供体克隆设计与构建	我们会依据您的实验需求, 提供供体克隆客户定制服务。供体克隆能将您的目的基因、筛选标记或遗传因子通过CRISPR-Cas9介导的同源重组 (HR) 整合到宿主细胞的基因组靶点。我们提供一系列含有不同筛选标记及遗传因子的供体载体设计以满足您的实验需求。
稳转细胞株服务	单克隆细胞株	含有CRISPR-Cas9介导的基因组修饰的单克隆细胞系。
	细胞库	为含有CRISPR-Cas9介导的基因组修饰的单克隆细胞系建库。
个性化分析	转基因小鼠	含有CRISPR-Cas9介导的基因组修饰的转基因小鼠。

(四) CRISPR-Cas9 基因编辑流程概览



(五) CRISP Cas9 核酸酶及 D10A 切口酶表达克隆

Cas9核酸酶克隆表达按照人类密码子优化过的Cas9核酸酶基因。Cas9切口酶克隆表达一个Cas9 D10A 切口酶基因。该切口酶是一个Cas9突变体, 第10位氨基酸从野生型的天冬氨酸突变成了丙氨酸。

CRISPR-Cas9体系是由sgRNA和Cas9核酸酶构成的DNA靶向识别剪切复合体。sgRNA识别一段20bp、3端紧接着5'-N-G-G-3' PAM位点的DNA靶序列, Cas9核酸酶在该靶序列上生成DNA双链断裂 (DSB)。DNA双链断裂 (DSB) 会被细胞的非同源末端连接 (NHEJ) 或同源重组 (HR) 机制修复。NHEJ机制容易在修复的断裂位点引入插入/缺失突变, 造成阅读框移码而产生基因敲除的效果。同源重组 (HR) 需要细胞内同时存在一段带有靶点同源序列的供体DNA作为修复模板, 我们可以通过对供体DNA进行设计, 在众多实验体系中实现基因敲除或者敲入突变修饰、报告基因等等一系列基因组修饰。

Cas9 D10A切口酶跟野生型的Cas9一样, 也使用sgRNA识别靶序列。但Cas9 D10A切口酶在靶序列上生成的是单链切口而非双链断裂。单链切口通常会由NHEJ和HR以外的机制完美地修复。但通过设计一对定向正确的sgRNA分别识别靶点两侧的DNA双链, 我们可以诱导Cas9 D10A分别在靶点两侧的DNA双链上各生成一个单链切口, 构成一个双链断裂 (DSB), 进而诱导NHEJ或HR修复机制。这种双切口策略常被用在要求低脱靶率的实验设计中。

产品	启动子	报告基因/筛选标记
Cas9 核酸酶表达克隆	CMV	mCherry / Neomycin
Cas9 核酸酶慢病毒表达克隆	CMV	Neomycin
Cas9 核酸酶慢病毒表达克隆	CMV	eGFP / Neomycin
Cas9 D10A 切口酶表达克隆	CBh	N/A
Cas9 D10A 切口酶表达克隆	CMV	mCherry / Neomycin

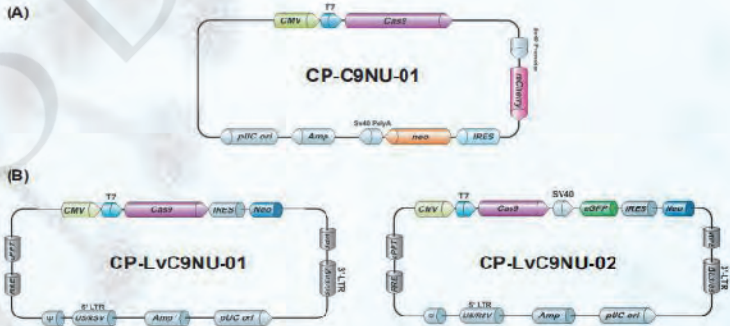


图4. (A) Cas9 核酸酶表达克隆质粒图谱; (B) Cas9 核酸酶慢病毒表达克隆质粒图谱。



图 5. Cas9 切口酶表达克隆质粒图谱。

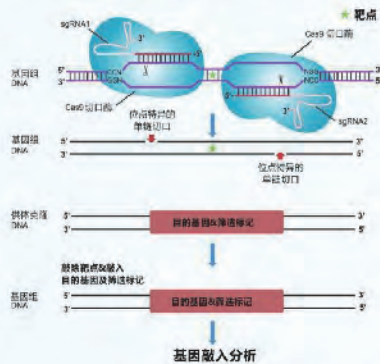


图6. Cas9 D10A切口酶的“双切口”设计原

(六) CRISPR sgRNA设计与克隆定制服务

sgRNA克隆表达一个单链的融合sgRNA（由crRNA和tracrRNA融合而成）。当Cas9核酸酶存在时，sgRNA能识别靶标DNA序列并引导Cas9核酸酶剪切靶标位点，形成DNA双链断裂（DSBs），进而实现基因敲除、敲入及突变等基因组编辑。多个sgRNA克隆与一个Cas9克隆共转染可同时在基因组中编辑多个位点，使实验设计更高效、更灵活。

载体	启动子	sgRNA 及 Cas9	筛选标记/报告基因
pCRISPR-SG01	U6	只表达 sgRNA*	Hygromycin
pCRISPR-LvSG03	U6	只表达 sgRNA*	Puromycin/mCherry
pCRISPR-CG01	U6	载体含有 CMV 启动子启动表达的Cas9 核酸酶基因	Neomycin / mCherry
pCRISPR-CG02	U6	载体含有 CBh 启动子启动表达的Cas9 核酸酶基因	N/A
pCRISPR-CG04	U6	载体含有 CMV 启动子启动表达的Cas9 核酸酶基因	Neomycin / copGFP
pCRISPR-CG07	U6	载体含有 CBh 启动子启动表达的Cas9 核酸酶基因	GFP
pCRISPR-CG08	U6	载体含有 CBh 启动子启动表达的Cas9 核酸酶基因	mCherry

*配合 Cas9 核酸酶 (Cat.No. CP-C9NU-01, CP-LvC9NU-01, CP-LvC9NU-02) 或 Cas9 D10A 切口酶 (Cat.No.CP-C9NI-01, Cat.No.CP-C9NI-02) 表达克隆使用。

图 7. sgRNA 载体图谱

(七) 供体克隆设计与构建服务

提供供体克隆设计及构建服务。供体克隆装载客户定制的目的基因、筛选标记及其他遗传因子，通过位点特异性的基因组编辑工具介导的同源重组（HR）将所装载的片段整合到基因组上的靶点。提供含有不同筛选标记及遗传因子的多种供体载体供客户选择。

载体	启动子	报告基因	筛选标记	LoxP 位点
pDonor-D01	EF1a	copGFP	Puromycin	N/A
pDonor-D02	CMV	copGFP	Neomycin	N/A
pDonor-D03	CMV	N/A	Neomycin	N/A
pDonor-D04	CMV	N/A	Puromycin	N/A
pDonor-D05	EF1a	N/A	Neomycin	N/A
pDonor-D07	EF1a	copGFP	Puromycin/TK	Loxp
pDonor-D08	CMV	copGFP	Neomycin/TK	Loxp
pDonor-D09	EF1a	N/A	Puromycin/TK	Loxp
pDonor-D10	CMV	N/A	Neomycin/TK	Loxp
pDonor-D11	PGK	copGFP	Puromycin/TK	Loxp
pDonor-D12	EF1a	copGFP	Hygromycin/TK	Loxp
pDonor-D13	PGK	copGFP	Neomycin/TK	Loxp
pDonor-D14	PGK	N/A	Puromycin/TK	Loxp

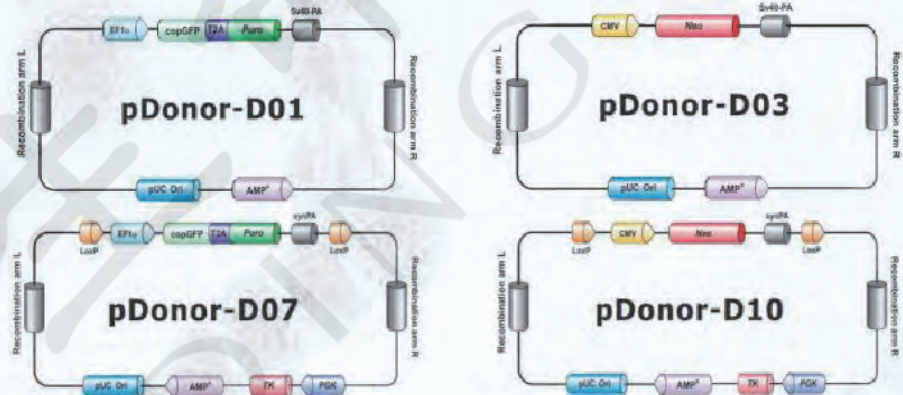


图8. 部分供体载体图谱

(八) 稳转细胞株服务

为含有CRISPR-Cas9介导的基因组修饰的细胞提供单克隆稳转细胞株构建服务及细胞建库服务。

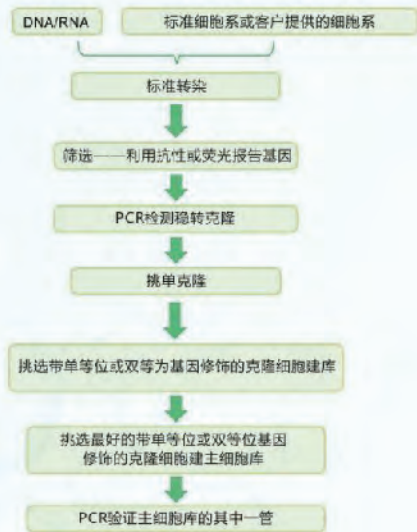


图9. 利用CRISPR构建稳转细胞系技术路线图

(九) 转基因小鼠服务

转基因小鼠模型是研究基因产物的分子机制、其相互作用以及对细胞进程影响的理想工具。为客户定制带有 CRISPR-Cas9 介导的基因组修饰的转基因小鼠。

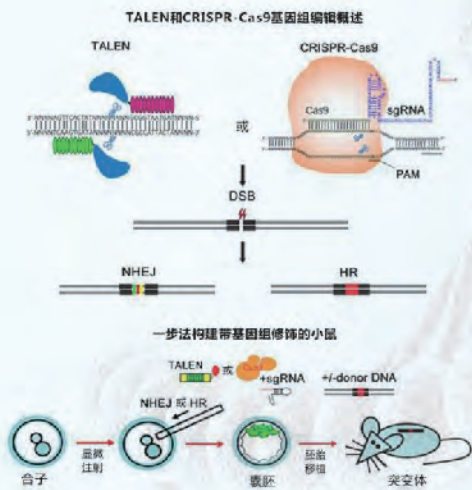


图10. 利用CRISPR-Cas9技术构建小鼠模型

(十) CRISPR-Cas9在基因治疗中的应用

CRISPR-Cas9基因编辑技术能够方便、准确的对人类遗传DNA进行修饰。目前，已有实验室使用 CRISPR-Cas9技术治愈了小鼠的白内障遗传疾病，也有实验室修复了囊性纤维症患者的细胞缺陷疾病等。遗传性高酪氨酸血症是一种致命性遗传疾病，由于Fah基因的8号外显子最后一个核苷酸G-A突变，导致FAH蛋白含量降低，从而导致有毒代谢物质积累，引起肝脏严重损伤。利用CRISPR-Cas9技术对Fah致病位点进行修正，可以有效治疗遗传性高酪氨酸血症。CRISPR-Cas9技术对遗传性致病特异位点的纠正，可以实现对由基因缺陷所引起的疾病的有效治疗，这也为人类基因治疗提供了一种新思路。

设计3个sgRNA位点，将每一个sgRNA和Cas9同时克隆到pX330载体上，同时为了促进同源重组修复，设计一个199nt长度的包含突变Fah基因正确序列的单链DNA模板作为修复模板，将Cas9、sgRNA和修复模板分别注射到细胞中（见图11）。

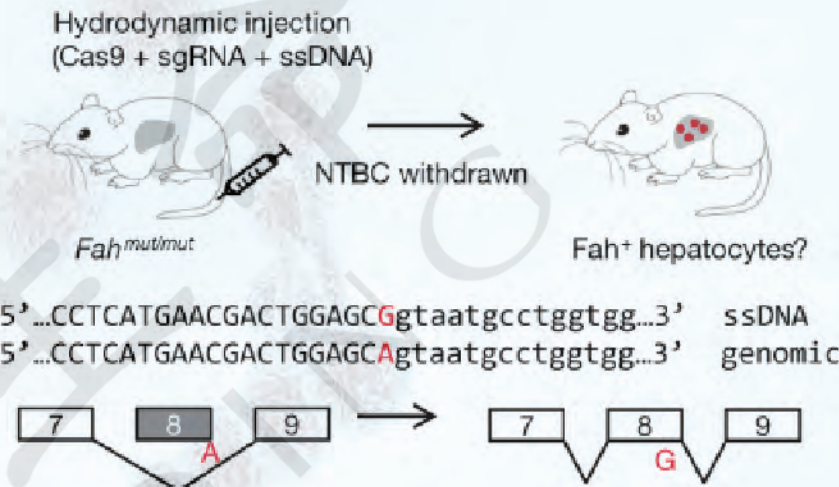


图11.运用CRISPR-Cas9技术修正小鼠初始细胞中的Fah基因

CRISPR-Cas9基因组编辑技术可以有效治疗成年动物的基因疾病，也有潜在在人类基因疾病治疗中得到应用。未来，基于CRISPR-Cas9技术的遗传病矫正将会在临床基因治疗中发挥强大的力量。

二. MLPA

(一) 背景介绍

MLPA(Multiples Ligation-dependent Probe Amplification), 即多重连接依赖探针扩增, 是一种在同一反应管内检测多达50个核苷酸序列的拷贝数变化的方法。MLPA可以快速同时鉴定几十个基因的缺失和插入, 可用于血液, 肿瘤样本的DNA, mRNA的表达谱分析。而且, MLPA方法可用于甲基化分析。

(二) 技术原理

MLPA的基本原理包括探针和靶序列DNA进行杂交, 之后通过连接、PCR扩增, 产物通过毛细管电泳分离及数据收集, 分析软件对收集的数据进行分析最后得出结论。每个MLPA探针包括两个荧光标记的寡核苷酸片段, 两个寡核苷酸片段都与靶序列进行杂交, 之后使用连接酶连接两部分探针。只有当连接反应完成, 才能进行随后的PCR扩增并收集到相应探针的扩增峰。如果检测的靶序列发生点突变或缺失、扩增突变, 那么相应探针的扩增峰便会缺失、降低或增加, 因此, 根据扩增峰的改变就可判断靶序列是否有拷贝数的异常或点突变存在。

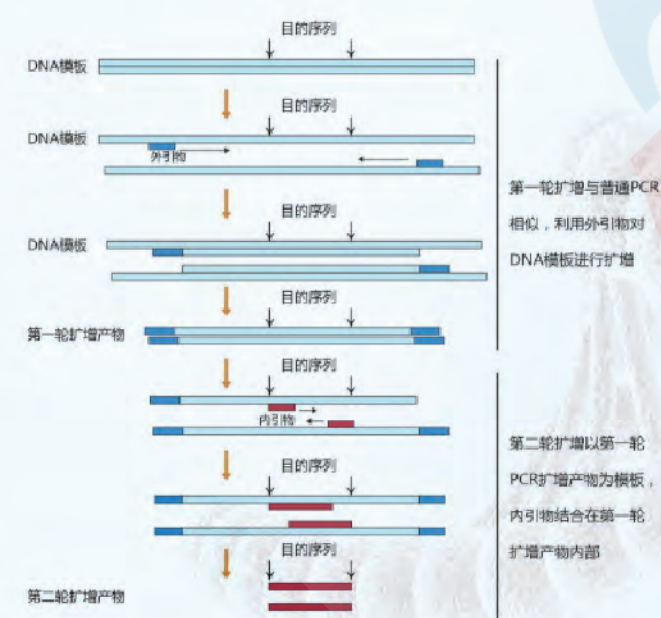


图1. MLPA原理图

(三) 技术优势

- (1) 仅通过选用不同探针即可以检测多种疾病;
- (2) 为多重反应, 一个反应可提供近50个靶位点;
- (3) 反应费用较低;
- (4) 重复性好, 简单易操作, 可以同时检测多个样本;
- (5) 敏感性高, 仅需20 ng 人DNA, 结果不受样品DNA量的影响;
- (6) 能区分单个碱基的差异, 能检测少量的拷贝数改变, 如2个, 3个拷贝数改变;
- (7) 可以检测单一外显子的缺失和插入突变。

(四) 应用范围

- (1) 检测小的重排: BRCA1, BRCA2, MSH2等;
- (2) 检测大范围的染色体重排: williams syndrome, prader-willi/angelman syndrome等;
- (3) 检测亚端粒区的拷贝数改变;
- (4) 检测染色体非整倍体改变, (包括羊水样本);
- (5) 肿瘤诊断: ALL, CLL, Oligodendrogliomas, melanomas, neuroblastomas等病的拷贝数改变;
- (6) 甲基化定量检测: Parader-willi、angelman syndrome, Fragile X等;
- (7) mRNA分析细胞凋亡和炎症反应。

三. RT-qPCR

(一) 基本原理

qPCR全称为实时荧光定量PCR技术, 随着PCR循环数的递增, PCR产物不断积累, 荧光信号也会相应的增加, 这样就可以通过荧光强度的变化来对PCR反应进行实时的监测。通过对DNA或RNA的荧光定量PCR监测, 实现基因的相对定量、绝对定量和定性分析。

(二) 技术优势

- (1) 实时监测: 通过扩增曲线能够实时监测PCR产物的积累。
- (2) 特异性强: 实验完成后, 溶解曲线的分析可以验证扩增产物的特异性, 降低假阳性。
- (3) 精确定量: 利用扩增进入指数增长期的CT值来定量测定起始模板量, 从而实现了极为精确的核酸定量检测。

(三) 服务内容

该技术主要用于:

- (1) RNA抽提与定量
- (2) 引物设计及荧光定量PCR、数据分析

(四) 注意事项

- (1) 在实验过程中要防止RNA的降解, 保持RNA的完整性。在总RNA的提取过程中, 注意避免mRNA的断裂;
- (2) 为了防止非特异性扩增, 必须设阴性对照;
- (3) 内参的设定: 主要为了用于靶RNA的定量。常用的内参有G3PD (甘油醛-3-磷酸脱氢酶)、 β -Actin (β -肌动蛋白) 等。其目的在于避免RNA定量误差、加样误差以及各PCR反应体系中扩增效率不均一各孔间的温度差等所造成的误差;
- (4) PCR不能进入平台期, 出现平台效应与所扩增的目的基因的长度、序列、二级结构以及目标DNA起始的数量有关。故对于每一个目标序列出现平台效应的循环数, 均应通过单独实验来确定;
- (5) 防止DNA的污染: 采用DNA酶处理RNA; 在可能的情况下, 将PCR引物置于基因的不同外显子, 以消除基因和mRNA的共线性。

四. TaqMan检测

(一) 背景介绍

aqMan方法通过采用荧光探针在PCR循环过程中随着特定PCR产物的累计而对其进行检测。

(二) TaqMan探针技术原理

TaqMan荧光探针是一种寡核苷酸探针, 荧光基团连接在探针的5'末端, 而淬灭剂则在3'末端。PCR扩增时在加入一对引物的同时加入一个特异性的荧光探针, 探针完整时, 报告基团发射的荧光信号被淬灭基团吸收; PCR扩增时, Taq酶的5'-3'外切酶活性将探针酶切降解, 使报告荧光基团和淬灭荧光基团分离, 从而荧光监测系统可接收到荧光信号, 即每扩增一条DNA链, 就有一个荧光分子形成, 实现了荧光信号的累积与PCR产物形成完全同步。常用的荧光基团是FAM, TET, VIC, HEX。

而新型TaqMan-MGB探针使该技术既可进行基因定量分析, 又可分析基因突变(SNP), 有望成为基因诊断和个体化用药分析的首选技术平台。

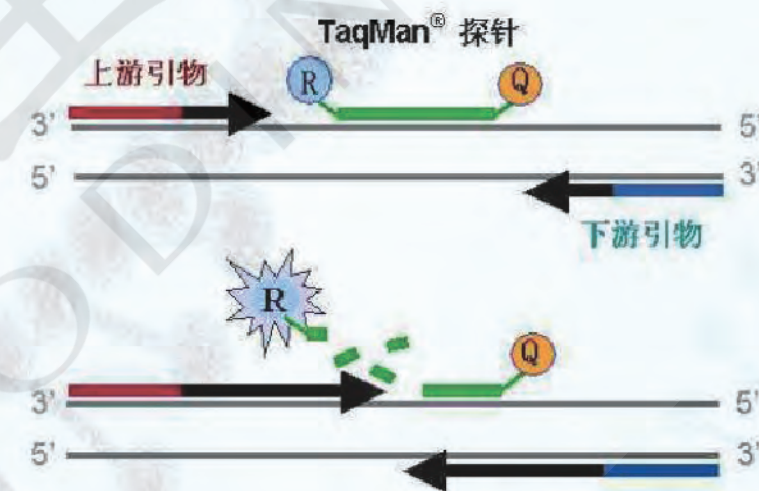


图1. TaqMan探针的荧光信号产生机制



图2. TaqMan MGB探针

(三) 技术平台:



图3. ABI 7500检测仪

(四) 应用范围

- (1) 基因表达 (mRNA) 分析, 包括基因产物定量、芯片验证、通路分析、发育生物学、质量控制和检测验证、siRNA/RNAi 实验、低拷贝数区分、使用OpenArray平台进行中到高通量表达谱分析;
- (2) microRNA与非编码RNA分析;
- (3) 遗传变异分析, 包括SNP基因分型、药物代谢酶 (DME) 基因分型以及拷贝数变异 (CNV) 分析;

(五) 服务流程

- (1) 根据目的基因设计引物及引物合成;
- (2) DNA抽提;
- (3) DNA定量及质量分析;
- (4) 实时定量PCR;
- (5) 数据处理及Figure制作。

五.基因定点突变

(一) 背景和原理

定点突变是指通过PCR等方法向目的DNA片段中引入所需变化,包括碱基的添加、删除、点突变等。定点突变能迅速、高效的提高DNA所表达的蛋白的性状及表征,是基因研究工作中一种非常有用的手段。

蛋白质的结构决定其功能,二者之间的关系是蛋白质组研究的重点之一。对某个已知基因的特定碱基进行定点改变、缺失或者插入,可以改变对应的氨基酸序列和蛋白质结构,对突变基因的表达产物进行研究有助于人类了解蛋白质结构和功能的关系,探讨蛋白质的结构/结构域。而利用定点突变技术改造基因:比如野生型的绿色荧光蛋白(wtGFP)是在紫外光激发下能够发出微弱的绿色荧光,经过对其发光结构域的特定氨基酸定点改造,GFP能在可见光的波长范围被激发(吸收区红移),而且发光强度比原来强上百倍,甚至还出现了黄色荧光蛋白,蓝色荧光蛋白等等。

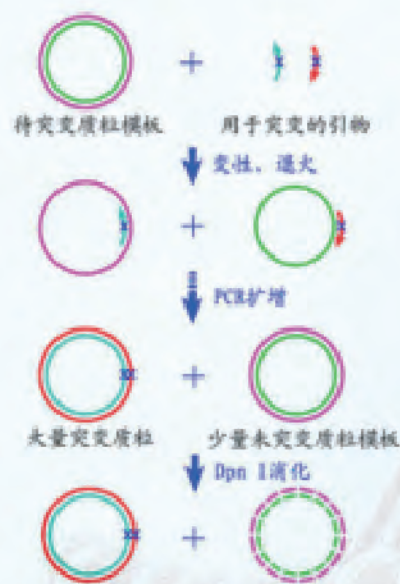


图1. 定点突变原理图

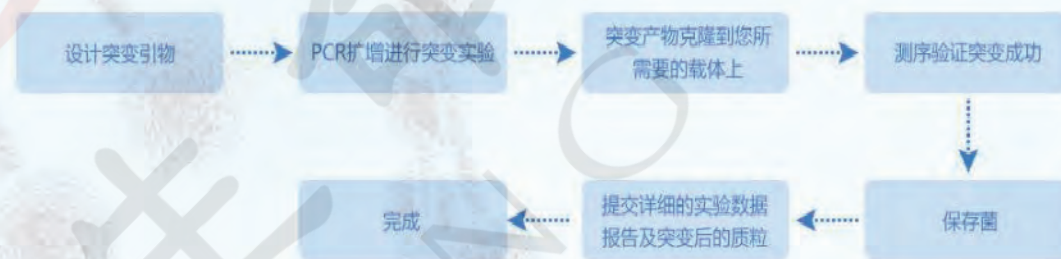
(二) 应用范围

定点突变技术的潜在应用领域很广,比如研究蛋白质相互作用位点的结构、改造酶的不同活性或者动力学特性,改造启动子或者DNA作用元件,引入新的酶切位点,提高蛋白的抗原性或者是稳定性、活性、研究蛋白的晶体结构,以及药物研发、基因治疗等等方面。

(三) 送样要求

- (1) 含有待突变基因的高纯度质粒,不少于5 μ g,电泳图清晰,达酶切及测序要求;
- (2) 待突变基因原始序列文件及原始测序图片文件、突变目标序列及详细信息、基因两端酶切位点,并提供待变基因的生物特性,如是否影响细菌增殖等特性;
- (3) 相关载体大小、抗性及图谱等信息;若需要克隆至目的载体,请提供目的载体及其详细信息。

(四) 实验流程



六. 载体构建

(一) 背景介绍

载体构建(vector construction)是分子生物学研究常用的手段之一,主要包括已有载体多克隆位点MCS的改造和已有载体启动子、增强子、筛选标记等功能元件的改造。理想运载体是质粒(plasmid),载体构建即是构建含外源DNA的质粒,构建完成后可利用PCR进行测序验证,提供多种类型的高品质载体构建服务。

(二) 技术特色

- ① 我们提供的载体均为非病毒载体；
- ② 过表达载体有 NeoR 或EGFP 等不同标记的骨架可选；
- ③ 可根据客户的不同需求构建不同的表达载体，能够保证所克隆序列的正确性及完整性；
- ④ 在哺乳动物细胞内高效过表达；
- ⑤ 提供对应的质粒 (2*5 μ g)、甘油菌 (100 μ L) 和实验报告书。

(三) 载体类型

1. lncRNA/基因过表达载体

对lncRNA/基因进行过表达，进行功能获得性实验 (gain-of-function) 是研究基因功能的重要手段，与以RNAi为代表的功能缺失性实验 (loss-of-function) 相互补充以说明lncRNA/基因在相关细胞功能中的作用。pEXP-RB-Mam系列过表达载体，在多克隆位点上游有CMV强启动子，可高效的表达插入到多克隆位点中的lncRNA/基因。

2. miRNA靶基因3' UTR报告基因载体

miRNA是一类具有调控功能的非编码小RNA，主要通过与其靶基因mRNA 3' UTR完全或不完全互补配对而调控靶基因的表达。pmiR-RB-Report™载体是专门用来检测miRNA作用的检测工具，将基因的3' UTR区克隆至载体海肾萤光素酶基因(hRLuc)下游，作为海肾萤光素酶基因的3' UTR序列随海肾萤光素酶转录表达，故可以海肾萤光素酶基因作为报告基因，以萤火虫萤光素酶基因(hLuc)作为内参基因，通过检测海肾萤光素酶活性的相对变化情况来看鉴定miRNA对该基因是否有调控作用。

3. miRNA Sensor报告基因载体

miRNA Sensor报告基因载体主要用于检测细胞内目标miRNA的活性变化，譬如内源竞争性RNA (ceRNA) 对miRNA活性的影响。miRNA Sensor报告基因载体以pmiR-RB-Report™载体骨架为基础，将目标miRNA的结合序列克隆至海肾萤光素酶基因(hRLuc)下游，目标miRNA可通过该结合序列调控海肾萤光素酶的表达，故可以海肾萤光素酶基因作为报告基因，以萤火虫萤光素酶基因(hLuc)作为内参基因，通过检测海肾萤光素酶活性的相对变化情况来判断细胞内miRNA的活性变化，如可以调控miRNA的ceRNA高表达，则miRNA被ceRNA结合，活性下降，海肾萤光素酶的相对活性上升。

4. circRNA过表达载体

环状RNA (circRNA) 是一种闭环的单链RNA分子，在细胞中发挥着多种重要作用，如调控转录，或者以ceRNA的方式调控miRNA的活性等等。pCE-RB-Mam环状RNA过表达载体，多克隆位点两翼高效促进RNA成环剪切的反向重复序列可使插入到多克隆位点中的序列转录后成环剪切，从而实现circRNA的过表达。

5. 启动子报告基因载体

转录因子与启动子的结合是基因转录调控的重要事件，启动子报告基因载体可用于检测特定转录因子通过启动子上特定位点调控基因转录的检测。pPro-RB-Report载体，海肾萤光素酶基因上游多克隆位点可用于插入目的基因的启动子，故可以海肾萤光素酶基因作为报告基因，以萤火虫萤光素酶基因作为内参基因，通过检测海肾萤光素酶活性的相对变化情况来看确定该段启动子的转录活性及转录因子通过该段启动子特定位点调控基因表达的活性。

6. 信号通路报告基因载体

细胞通过胞内胞外的受体与配体 (信号分子) 结合，产生一系列的信号传导，使细胞对环境变化做出一系列反应。信号通路非常多样，而每个信号通路往往也由数十个基因参与调控，信号通路研究是目前生物学及药物基础性研究的重点。信号通路报告基因载体pREG-RB和pREL，分别以EGFP和海肾萤光素酶为报告基因，在将信号通路关键转录因子的结合序列 (recognition motif或binding motif) 克隆至报告基因上游，一旦细胞感受到特定的信号活化该信号通路，则激活报告基因的转录表达，从而可通过信号通路报告基因载体检测信号通路的活化情况。

以EGFP为报告基因非常适用于高内涵平台筛选对特定信号通路有活化或抑制作用的药物筛选，或者筛选在信号通路中起作用的基因，或者鉴定基因对信号通路活化的影响；以海肾萤光素酶为报告基因则适合使用多功能酶标仪或萤光照度计进行简易的检测。

7. shRNA载体

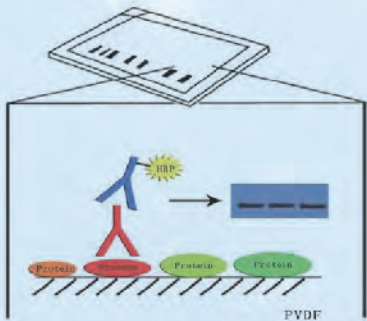
short hairpin RNA (shRNA) 是化学合成siRNA之外的另一种RNAi实验方法常用工具。shRNA载体骨架为pRNAT-U6.1/Neo，转染细胞后，shRNA在U6启动子的调控下表达，并最终剪切形成siRNA duplex起作用。载体上另外GFP标签，可用于检测确认质粒的转染情况。

第十一章 蛋白质验证和细胞研究平台

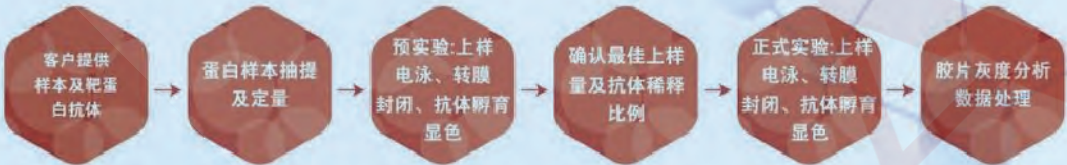
一. Western Blot(WB)

(一) 实验原理

免疫印迹(immunoblotting)又称蛋白质印迹(Western blotting), 它是根据抗原抗体的特异性结合检测复杂样品中的某种蛋白的方法。免疫印迹常用于鉴定某种蛋白, 并能对蛋白进行定性和半定量分析。凝胶迁移或电泳迁移率实验 (EMSA-electrophoretic mobility shift assay) 是一种研究DNA 结合蛋白和其相关的DNA 结合序列相互作用的技术, 可用于定性和定量分析。这一技术最初用于研究DNA 结合蛋白, 目前已用于研究RNA 结合蛋白和特定的RNA 序列的相互作用。原理图如下:



(二) 实验流程

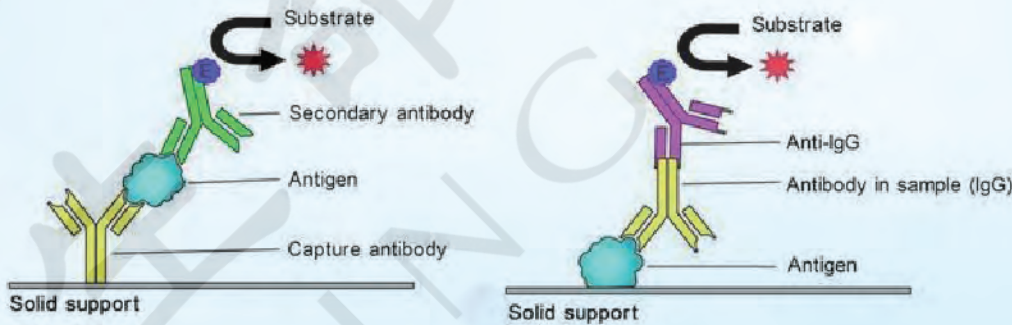


- (1) 蛋白质抽提;
- (2) 蛋白质定量;
- (3) 变性聚丙烯酰胺不连续凝胶电泳 (SDS-PAGE) ;
- (4) 蛋白质转移到PVDF膜;
- (5) 膜的封闭和抗体孵育;
- (6) 结果检测。

二. ELISA

(一) 实验原理

酶联免疫吸附剂测定(Enzyme Linked Immunosorbent Assay, ELISA), 是以免疫学反应为基础, 将抗原、抗体的特异性反应与酶对底物的高效催化作用相结合起来的一种敏感性很高的实验技术。基本原理是: 1) 使抗原或抗体结合到某种固相载体表面, 并保持其免疫活性。2) 使抗原或抗体与某种酶连接成酶标抗原或抗体, 这种酶标抗原或抗体既保留其免疫活性, 又保留酶的活性。在测定时, 把受检标本 (测定其中的抗体或抗原) 和酶标抗原或抗体按不同的步骤与固相载体表面的抗原或抗体起反应, 再用洗涤的方法使固相载体上形成的抗原抗体复合物与其他物质分开, 最后结合在固相载体上的酶量与标本中受检物质的量成一定的比例。加入酶反应底物后, 底物被酶催化为有色产物, 产物的量与标本中受检物质的量直接相关, 故可根据反应后颜色的深浅进行定性或定量分析。



(二) 实验流程



(三) 样品要求

新鲜标本尽早检测，对收集后当天就进行检测的标本，及时储存在4℃备用，如有特殊原因需要周期收集标本，请造模取材后，将标本及时分装后放在-20℃或-70℃条件下保存。因冰室与室温存在一定温差，蛋白极易降解，直接影响实验质量，所以避免反复冻融。

(1) 液体类标本：标本必须为液体，不含沉淀。包括血清、血浆、尿液、胸腹水、脑脊液、细胞培养上清、组织匀浆等。

(2) 血清：室温血液自然凝固10-20分钟后，离心20分钟左右（2000-3000转/分）。收集上清。如有沉淀形成，应再次离心。

(3) 血浆：应根据试剂盒的要求选择EDTA、柠檬酸钠或肝素作为抗凝剂，加入10%（v/v）抗凝剂（0.1M柠檬酸钠或1%heparin 或2.0%EDTA.Na2）混合10-20分钟后，离心20分钟左右（2000-3000转/分）。仔细收集上清。如有沉淀形成，应再次离心。

(4) 尿液、胸腹水、脑脊液：用无菌管收集。离心20分钟左右（2000-3000转/分）。仔细收集上清。如有沉淀形成，应再次离心。

(5) 细胞培养上清：检测分泌性的成份时，用无菌管收集。离心20分钟左右（2000-3000转/分）。仔细收集上清。检测细胞内的成份时，用PBS（PH7.0-7.4）稀释细胞悬液，细胞浓度达到100万/ml左右。通过反复冻融，以使细胞破坏并放出细胞内成份。离心20分钟左右（2000-3000转/分）。仔细收集上清。保存过程中如有沉淀形成，应再次离心。

(6) 组织标本：切割标本后，称取重量。加入一定量的PBS，缓冲液中可加入1 μg/L蛋白酶抑制剂或50U/ml的Aprotinin（抑肽酶）。用手工或匀浆器将标本匀浆充分。离心20分钟左右（2000-3000转/分）。仔细收集上清置于-20度或-70度保存，如有必要，可以将样品浓缩干燥。分装后一份待检测，其余冷冻备用。

三. 免疫荧光 (IF)

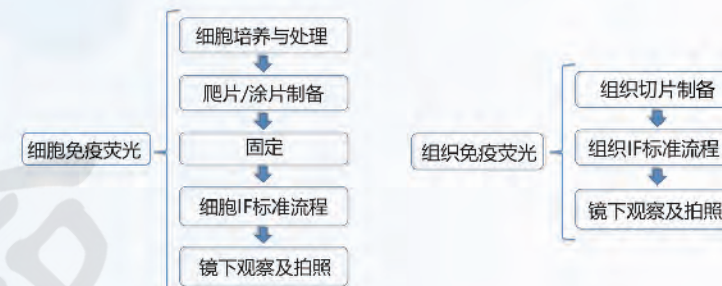
(一) 实验原理

根据抗原抗体反应的原理，先将已知的抗原或抗体标记上荧光素制成荧光标记物，再用这种荧光抗体（或抗原）作为分子探针检查细胞或组织内的相应抗原（或抗体）。在细胞或组织中形成的抗原抗体复合物上含有荧光素，利用荧光显微镜观察标本，荧光素受激发光的照射而发出明亮的荧光（黄绿色或桔红色），可以看见荧光所在的细胞或组织，从而确定抗原或抗体的性质、定位，以及利用定量技术测定含量。

实验流程免疫荧光单标记方法免疫荧光单标记是指只标记一种蛋白质分子，方法比较简单，只要按照染色步骤去做，通常不存在太多的问题。但要注意固定液的选择，固定液选择的合适与否，可能会直接影响染色结果。

免疫荧光双标记方法免疫荧光的双标记是指同时标记细胞内两种蛋白质分子，当怀疑某种配体与已知受体结合后可用此方法加以证明。此方法稍微复杂一些，首先应注意所用的一抗必须是来自不同种属动物的两种特异性抗体（例如：A抗体为多克隆抗体，来自家兔；B抗体为单克隆抗体，来自小鼠）。其次，两种二抗所带荧光素的发射光不应重叠，且尽量远离，通常可以选择FITC和TRITC、Alex488和TRITC、FITC和Cy5，或Cy3和Cy5等来组合。通常情况下，染色时两种一抗可以同时孵育，然后可以同时孵育两种二抗。但当染色结果一种颜色非常弱，而另一种颜色比较强时，应考虑先孵育颜色较弱的二抗。其他步骤同免疫荧光单标记。

(二) 实验流程



四. 免疫组织化学(IHC)

(一) 实验原理

免疫组织化学（Immunohistochemistry）是一项以免疫学的抗原-抗体反应为理论基础，应用理化方法显示细胞结构的物质成分，进而分析研究细胞和组织的代谢、功能及其形态变化的实验技术。该方法将抗原-抗体反应的特异性与酶的高效催化作用有效结合起来，即用酶标记抗体，以酶与底物的显色反应对抗原进行示踪，并对呈色反应后的免疫阳性产物做半定量分析。

(二) 实验流程



(三) 技术优势

特异性强免疫学的基本原理决定抗原与抗体之间的结合具有高度特异性，因此，免疫组化从理论上讲也是组织细胞中抗原的特定显示，如角蛋白(keratin)显示上皮成分，LCA显示淋巴细胞成分。只有当组织细胞中存在交叉抗原时，才会出现交叉反应。敏感性高在应用免疫组化的起始阶段，由于技术上的限制，只有直接法、间接法等敏感性不高的技术，那时的抗体只能稀释几倍、几十倍；现在由于ABC法或SP三步法的出现，使抗体稀释上千倍、上万倍甚至上亿倍仍可在组织细胞中与抗原结合，这样高敏感性的抗体抗原反应，使免疫组化方法越来越方便地应用于常规病理诊断工作。定位准确、形态与功能相结合该技术通过抗原抗体反应及呈色反应，可在组织和细胞中进行抗原的准确定位，因而可同时对不同抗原在同一组织或细胞中进行定位观察，这样就可以进行形态与功能相结合的研究，对病理学领域开展深入研究是十分有意义的。

以上三种方法的比较：

对于IHC来说，定位是免疫组化的优势。该方法可在组织和细胞中进行抗原的准确定位，因而可同时对不同抗原在同一组织或细胞中进行定位观察，这样就可以进行形态与功能相结合的研究，相比于其他蛋白检测方法，免疫组化具有定位较直接准确、定性灵敏度高，是定位检测分析首选方法对病理学领域开展深入研究是十分有意义，尤其对于有些因子的转位研究十分有用。

WB与IHC技术相比，定量可能更加准确。当然WB也可定性和定位（通过提取膜蛋白或核蛋白、胞浆蛋白分别检测其中抗原含量，进而间接反映它们的定位），但敏感性远远低于IHC。

ELISA与免疫组化技术相比，定量准确，是分泌性蛋白检测首选方法之一。尤其对于微量分析多个样本非常有用，所用试剂和样品量很少，结果精确。IHC针对性比较强，而且一般是以组织或者细胞为样本，而ELISA一般是使用血清或者组织磨碎液。免疫组化可以快速检测样本中抗体的阳性或者阴性，一般不用于定量检测。

(四) 浦东解码生命科学研究院的优势

解码研究院具备完善的蛋白检测和细胞生物学平台，可以为客户提供高质量的WB、ELISA、IF和IHC数据，所有原始数据可查询、可追溯、可获取。

第十二章 外泌体研究

一. 外泌体研究

(一) 背景介绍

外泌体是指包含了复杂 RNA 和蛋白质的小膜泡 (30-150nm)，现今，其特指直径在40-100nm的盘状囊泡。1983年，外泌体首次于绵羊网织红细胞中被发现，1987年Johnstone将其命名为“exosome”。多种细胞在正常及病理状态下均可分泌外泌体。其主要来源于细胞内容酶体微粒内陷形成的多囊泡体，经多囊泡体外膜与细胞膜融合后释放到胞外基质中。所有培养的细胞类型均可分泌外泌体，且外泌体天然存在于体液中，包括血液、唾液、尿液、脑脊液和乳汁中。有关他们分泌和摄取及其组成、“运载物”和相应功能的精确分子机制刚刚开始研究。外泌体目前被视为特异性分泌的膜泡，参与细胞间通讯，对外泌体的研究兴趣日益增长，无论是研究其功能还是了解如何将其用于微创诊断的开发。

(二) 提取方法

(1) 超速离心法 (UC)：

超速离心法是外泌体提取中最早建立并使用的一种方法。这种方法原理就是通过外泌体与其他细胞器的沉降系数不同，将细胞、细胞碎片以及其他细胞器等，分离出去，从而得到纯净的外泌体。这种方法的优点就是外泌体提取方法简单，提取过程中不会引入其他的标记物，适合大剂量的样品处理。缺点是，操作繁琐，费时费力，操作过程中需要超高速离心。此外，由于离心力的因素，可能会导致外泌体结构破坏从而影响下游实验。

(2) 密度梯度离心法：这是一种普通超速离心法（方法01）与密度梯度离心法结合的方法。由于外泌体的密度范围是1.13g/ml-1.21g/ml，可以使用梯度介质去除非囊泡颗粒。最开始使用的梯度介质为蔗糖，所以又称蔗糖密度梯度超速离心，在后来经过方法的不断改进，可以使用碘克沙醇作为新的一种高效介质。相比于普通的超速离心法，它可以通过使用蔗糖等介质更好的去除样品中大的蛋白质集体等杂质，从而得到更加纯净的外泌体。但是这种方法操作起来更加繁琐。

(3) 聚合物沉降法:

这种方法是通过高分子的疏水聚合物,如聚乙二醇(PEG),它可以改变外泌体的溶解性和分散性,使其能够在比较低的离心力下沉降出来。其原理可能与竞争性结合游离水分子和PEG本身形成的网状结构有关。这种方法的优点就是操作起来比较简单,不需要超速离心机,得到的外泌体数量较多,也适合大剂量的样品处理。但是这种方法提取得到的外泌体,杂质含量会比较多,尤其是一些杂蛋白。

(4) 分子排阻法(凝胶过滤层析):

这种方法主要是利用多孔的凝胶球,通过分子大小不同将外泌体纯化出来,其原理类似层析法纯化蛋白。在分离提取过程中,大分子物质不能进入凝胶孔,很快沿多孔凝胶间的缝隙被流动相洗脱出来,而小分子物质能进入凝胶孔,滞留时间更长,会更慢地被洗脱出来。外泌体的分子粒径大于一般的蛋白分子,所以外泌体会先被洗脱下来,而粒径小的蛋白质会被后洗脱下来,从而分离出高纯度的外泌体。优点就是纯度相对较高,且外泌体结构不易受到破坏。缺点就是如果有一些大小相近的颗粒也会被纯化出来,比如乳糜微粒、低密度脂蛋白等。

(5) 超滤法:

这种方法是将溶剂及小分子物质过滤到膜的另一侧,而将相对大分子物质截留在超滤膜上,以达到分离的目的。外泌体直径范围40~200 nm,我们用超滤膜经过长时间的低速离心就可以分离样本中的外泌体。这种方法的优点是操作方便,富集效率比较高,易于搭配其他膜或操作,并且成本较低。缺点是过膜易损耗变形,导致膜堵塞,纯度较低,大颗粒蛋白污染严重,并且在洗脱过程中,附着于膜上的外泌体难以回收,导致提取损失。

(6) 磁珠特异性捕获法:

外泌体表面带有许多特殊的膜蛋白质,如CD9、CD63、CD81和ANNEXIN等,可以将这些外泌体标志蛋白作为分离外泌体的特异性标记。将抗体固定于磁珠等基质上,利用相应性抗体与这些标志蛋白之间的特异性相互作用,以实现对外泌体的特异性富集。这种方法的优点是外泌体纯度较高,并且对外泌体膜结构影响较小。缺点就是成本较高,无法进行大剂量的提取,并且磁珠保存难度也比较高。

综上所述,一般采用超速离心法/试剂盒法/磁珠法。

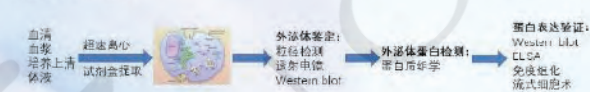
(三) 后续研究实验

- (1) 外泌体鉴定(NTA粒径、透射电镜、WB、流式检测);
- (2) 外泌体示踪/细胞对外泌体摄取行为;
- (3) 外泌体流式检测、外泌体载药(siRNA、质粒等);
- (4) 外泌体高通量测序(miRNA/lncRNA)、外泌体lncRNA/circRNA芯片检测;
- (5) 外泌体miRNA/lncRNA/mRNA/circRNA定量PCR检测;
- (6) 外泌体蛋白质组学检测(Label-free/iTRAQ/TMT);
- (7) 外泌体动物功能研究(外泌体注射动物、皮下荷瘤、活体成像、免疫荧光等);
- (8) 外泌体细胞功能检测(细胞增殖、细胞凋亡、细胞周期、细胞迁移/侵袭等);
- (9) 外泌体miRNA靶基因验证:双荧光素酶报告基因检测;

基因研究路线图:



蛋白研究路线图:



(四) 注意事项

- 1.分离exosome 前的样品不能加入任何RNA保护剂(如Trizol, RNA later);外泌体血液样本提取时,请大家注意血清和血浆都可以正常提取。如果提供全血,是无法完成的,必须提前做好血清或者血浆的分离,并且采血后尽快进行分离,期间不能经过冷冻,离心时转速不能太高,以800-1000转离心2-5分钟左右为宜,离心后取上清即可。
- 2.样本储存——冷冻储存
 - (1) 样本采集后,样本迅速放入进口高质量冻存管中,拧紧后立即到-80℃中保存。
 - (2) 胞上清样本处理后储存于-80℃冰箱中。在-20℃冰箱中储存不宜太久。
 - (3) PAXgene管(BD)保存。
 - (4) RNA/蛋白样本储存于-80℃冰箱中。
- 3.分离好的外泌体如需进行粒径检测或流式鉴定,需放4℃保存,并尽快进行实验。

(五) 送样要求

样本种类包括细胞培养上清、血清、血浆、脑脊液、腹水、胸水、尿液、组织样本等。

样本类型		所需样本参考量 (体积)
表征检测 (电镜、粒径、Marker蛋白质检测)		组学研究 (NSG测序/蛋白质组)
细胞上清*	20mL	50-60mL
血清	1mL	2mL
血浆	1mL	2mL
尿液	20mL	50-60mL
脑脊液	10mL	20-30mL
腹水	10mL	20-30mL
羊水	10mL	20-30mL
乳液	5mL	10-20mL
精液	1mL	2-4mL
唾液	5mL	10-20mL
*细胞上清液		指常见的肿瘤细胞, 如HeLa、A549、HepG2等; 然而神经细胞、干细胞分泌的外泌体量是肿瘤细胞的1/5~1/10左右, 因此所需培养基上清量是肿瘤细胞

(六) 应用范围

- (1) 疾病早期诊断和预后:
- (2) 外泌体新型治疗剂:
- (3) 外泌体作为药物治疗的递送系统。

(七) 案例分析

案例1.干细胞来源的囊泡小RNA在与年龄相关的人类疾病和长寿的不同疗法中的新兴作用。

研究摘要: 有机体衰老涉及器官功能的逐步下降和对与年龄有关的疾病的易感性增加。这与体内干细胞群体的衰老有关, 衰老降低了干细胞自我更新, 分化和再生受损组织

和器官的能力。这篇综述旨在探讨衰老如何与干细胞衍生的细胞外囊泡 (SCEV) 及其相应的miRNA货物 (SCEV-miRNA) 的失调相关。最近的证据表明, 在衰老的干细胞中, SCEV-miRNA可能在调节各种导致衰老的过程中起着至关重要的作用: 细胞衰老, 干细胞衰竭, 端粒长度和昼夜节律。因此, 进一步阐明SCEV-miRNA通过其下游效应发挥作用的年龄依赖性分子机制可能有助于人们更好地了解衰老生物学, 阐明其在干细胞功能中的作用, 并确定未来再生疗法的重要靶标。另外, 还讨论了评估SCEVs和SCEV-miRNA在治疗几种与年龄相关的疾病中的治疗作用的最新研究。

参考文献:
Emerging Role of Stem cell-derived Extravesicular microRNAs in Age-associated Human Diseases and in Different Therapies of Longevity. Ageing Res Rev, IF=10.39 doi: 10.1016/j.arr.2019.100979.

案例2.揭示癌细胞抑制抗肿瘤免疫反应新机制-从远处释放携带PD-L1的外泌体

研究内容: 癌细胞并不仅仅是一群生长失去控制的细胞; 为了自身的生存, 它们积极地参与与免疫系统之间的斗争。能够逃避免疫系统检测是癌症的一种特征。在一项新的研究中, 来自美国宾夕法尼亚大学的研究人员发现癌细胞释放生物“无人机”——在血液中循环的被称作外泌体 (exosome) 的小囊泡, 这些小囊泡携带着PD-L1蛋白, 这种蛋白导致T细胞在到达肿瘤并进行战斗之前精疲力竭——来协助这种斗争。尽管这项研究主要针对转移性黑色素瘤, 但是这些研究人员发现乳腺癌和肺癌也会释放携带着PD-L1的外泌体。相关研究结果于2018年8月8日在线发表在Nature期刊上。

这项研究以一种颠覆现有观念的方式展示了癌症如何采取一种系统性方法来抑制免疫系统。此外, 它还指出了一种新方法预测哪些癌症患者会对通过破坏免疫抑制来抵抗肿瘤的抗PD1疗法 (anti-PD1 therapy) 作出反应, 并提供一种追踪这类疗法治疗效果的方法。

Guo说: 对许多转移性黑色素瘤患者来说, 免疫疗法可以挽救生命, 但是大约70%的这些患者没有作出反应。这类疗法费用昂贵并且具有毒副作用, 因此了解哪些患者会作出反应将是有帮助的。鉴定血液中的生物标志物可能有助于及早预测哪些患者会作出反应, 以及随后可能为患者及其医生提供一种监测治疗效果的方法。

参考文献:
Exosomal PD-L1 contributes to immunosuppression and is associated with anti-PD-1 response. Nature, 560(7718), 382 - 386. doi:10.1038/s41586-018-0392-8.