

技术手册

解码生命
快乐科学

上海浦东解码生命科学研究院

上海浦东解码生命科学研究院

简介

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！

目 录

第一部分 多水平的核酸检测技术.....	4
一. DNA 水平检测.....	4
(一) 单核苷酸多态性, SNP.....	4
(二) 基因拷贝数变异, CNV.....	7
(三) 短串联重复序列, STR.....	9
(四) DNA 甲基化.....	11
二. RNA 水平检测.....	15
(一) 转录组.....	15
(二) 小 RNA.....	16
第二部分 测序平台.....	19
一. 3730XL 毛细管电泳测序系统.....	19
二. PGM/Ion Pronton 测序系统.....	21
三. Illumina Hiseq 3000/4000 测序系统.....	24
四. Pyrosequencing 焦磷酸测序.....	28
第三部分 芯片平台.....	31
一. Illumina 微阵列扫描仪系统.....	31
二. 安捷伦 DNA 微阵列扫描仪.....	32
三. Affymetrix 新型基因芯片扫描仪.....	33
四. 罗氏 MS200 扫描仪.....	34
第四部分 芯片技术服务.....	35
一. 表达谱芯片.....	35
二. 比较基因组杂交芯片.....	38
三. ChIP-on-chip 芯片.....	40
四. DNA 甲基化芯片.....	41
五. 微小 RNA 芯片.....	43
六. Sequenom MassArray 芯片.....	44
第五部分 生物信息学服务.....	46

第一部分 多水平的核酸检测技术

一. DNA 水平检测

(一) 单核苷酸多态性, SNP

单核苷酸多态性 (Single Nucleotide Polymorphism, SNP) 是指在基因组上单个核苷酸的变异, 包括置换、颠换、缺失和插入形成的遗传标记。SNP 数量众多, 多态性丰富, 是基因组学研究中的第三代遗传标志。人体的生理差异、对疾病的易感性、对药物反应等都与 SNP 有关。根据估算, 人类基因组大约每 1000 个碱基就有一个 SNP, SNP 的总量约为 3×10^6 。

1. SNP 的应用

➤ 致病基因定位

主要包括两个方面: 一是在疾病定位区域中寻找到的致病 SNP。这种 SNP 的出现可能直接导致了基因转录水平上和翻译水平上的变化, 即基因表达量的变化或者基因产物蛋白质的结构变化, 从而直接导致某种疾病或表型; 二是 SNPs 作为一个与疾病或表型紧密连锁的遗传标记, 可以用来绘制高密度的基因组图谱, 以便寻找疾病相关基因。

➤ 易感基因定位

在基因组关联分析中, SNP 已成为比其它序列变异 (如微卫星位点等) 更为理想的遗传标记。通过基于 SNP 的连锁不平衡研究, 有助于寻找疾病易感基因的区域甚至位点。

➤ 群体遗传学

SNP 可以协助对人类起源、人类进化和群体迁徙进行研究。国际基因地理计划就是基于 SNP 来绘制人类迁移路线图, 回答人类遗传多样性的国际合作研究计划。

➤ 药物基因组学

分析病毒、细菌基因的多态性, 有助于人们了解病毒、细菌的感染发病机制与耐药性形成机制。耐药基因检测的方法有 2 种: (1) 基因表达芯片检测药物代谢相关基因表达的改变; (2) 寡核苷酸芯片检测药物相关基因的 SNP 变化。目前已有乙肝耐药、结核杆菌耐药和艾滋病耐药的 SNP 检测基因芯片, 可指导临床对病人实施个体化治疗。

2. SNP 的检测方法

➤ 常规方法

包括 PCR-单链构象多态性 (PCR-SSCP) 分析、聚合酶链反应-限制性片段长度多态 (PCR-RFLP)、变性高压液相 (dHPLC) 和杂合双链分析法(HA) 等。这些分析方法速度慢、

通量低、难以判断碱基类型，无法进行大规模筛选。

➤ DNA Sanger 测序

Sanger 测序法又称双脱氧末端终止法，原理是采用核苷酸链终止剂—2,3-双脱氧核苷酸终止 DNA 的延长。根据碱基配对原则，每当 DNA 聚合酶需要 dNMP 参入到正常延长的 DNA 链中时，就有两种可能性，一是参入 ddNTP,结果导致脱氧核苷酸链延长的终止；二是参入 dNTP,使 DNA 链仍可继续延长，直至参入下一个 ddNTP。根据这一方法，就可得到一组以 ddNTP 结尾的长短不一的 DNA 片段。在尿素变性的 PAGE 胶上电泳，即可获得 DNA 序列。见图 1。

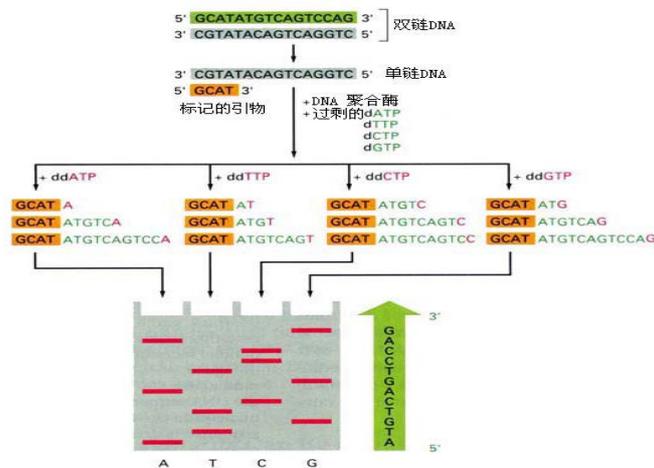


图 1. Sanger 法测序原理

➤ 深度测序

深度测序技术是可一次性对几十万到几百万条 DNA 分子进行并行测序的测序技术，可进行 SNP，cSNP（编码序列单核苷酸多态性）等研究。原理见图 2。



图 2. 深度测序的基本原理

➤ **SNP 芯片**

是一种高通量 SNP 检测方法，其原理是将大量探针分子固定于支持物上后，与标记的样品分子进行杂交，通过检测每个探针分子的杂交信号强度，获取样品分子的数量和序列信息。原理和流程见图 3。

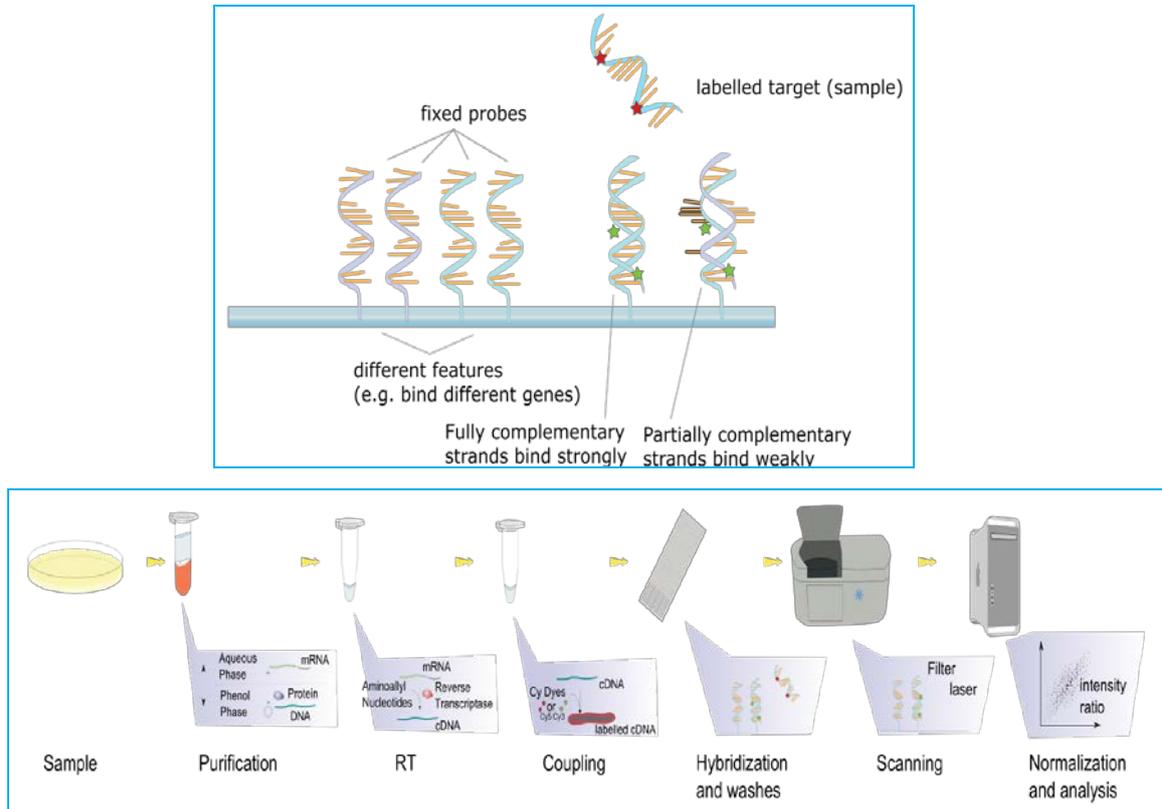


图 3. 芯片检测的原理及流程

3. 浦东解码生命科学研究院在 SNP 检测中的优势

- 浦东解码生命科学研究院拥有 ABI 3730xl 一代测序平台，能够准确迅速的对客户提供的样本进行 Sanger 测序分析，有着独有的检测体系。
- 浦东解码生命科学研究院拥有 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，能够准确迅速的对客户提供的样本进行各种个性化、针对性的多基因高通量测序服务。
- 浦东解码生命科学研究院拥有 Sequenom MassARRAY 时间飞行质谱，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，一张芯片可对 384 个样本进行多重检测，更可实现高达 25-30 重反应，是目前市场上中高通量 SNP 位点验证的最高性价比的检测系统。

- 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。
- 浦东解码生命科学研究院根据客户需要可以提供不同厂家的 SNP 基因分型芯片服务,包括: Illumina SNP 基因分型芯片服务和 Affymetrix SNP 基因分型芯片服务。

(二) 基因拷贝数变异, CNV

基因拷贝数变异 (copy number variations, CNVs) 是指 DNA 片段大小范围从 kb 到 Mb 的亚微观变化,包括了拷贝数的删除、插入、复制和复合多位点变异等。

1. CNV 的应用

➤ 遗传性疾病

研究发现 CNVs 可能是遗传病更重要的发病机制,CNVs 将成为研究遗传病的分子机制、寻找基因诊断的有效途径。

➤ 群体遗传学

CNVs 和 SNPs 是人类表型变异的两个重要潜在来源,对 CNV 的研究可以为人类起源、人类进化和群体迁徙的研究提供重要的信息。

➤ 疾病抗性性状的 QTLs 检测

CNV 可以作为遗传标记,与某种性状的数量性状基因组 (QTLs) 连锁,可能直接控制疾病的抗性或易感性状的表达,目前已作为一种有效的遗传标记用于畜禽抗病育种。

2. CNV 的检测方法

➤ 比较基因组杂交技术 (CGH)

CGH 是将消减杂交和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的一种方法。CGH 仅需少量 DNA,即可在整个基因组水平研究 DNA 拷贝数变异。其基本原理是将待检测组织 DNA 和正常对照用不同荧光标记后,一定比例杂交后,用荧光显微镜检测。但 CGH 技术所能检测到的最小的 DNA 扩增或丢失约 3-5Mb,故对于低水平的 DNA 扩增和小片段的丢失会漏检。此外,在染色体的拷贝数量无变化时,CGH 技术不能检测出平等染色体的易位。

➤ Fosmid 末端测序法

本方法是将基因组 DNA 随机打断成片段,然后连接到载体上,构建成一系列的重组子。研究时,只需筛选出感兴趣的重整子即可,见图 4。

➤ 深度测序

深度测序可以检测CNV。可先与基因组捕捉芯片结合（NimbleGen Sequence Capture array）或溶液杂交(Agilent Sure-Select System)，然后富集目标区域DNA，再行深度测序。

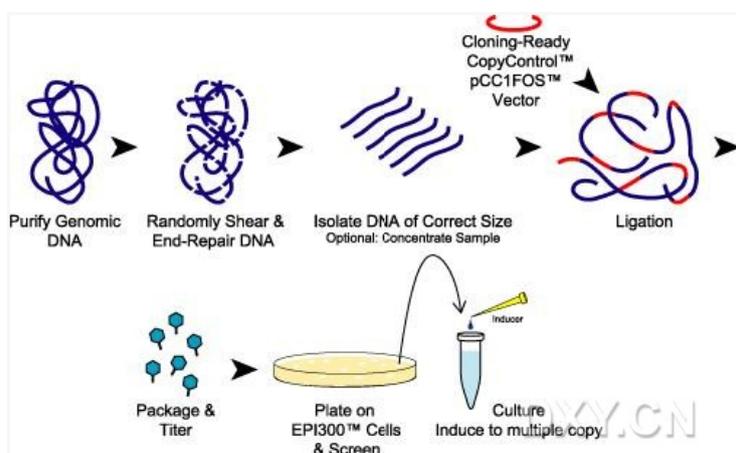


图 4. Fosmid 末端测序法原理

➤ 微阵列比较基因组杂交芯片（aCGH）

可以用于 CNV 的检测。通过将 DNA 与芯片杂交，检测信号强度，即可确定每个位点的基因组拷贝数。在拷贝数检测运算中，需要考虑探针的长度和 CG 含量，从而降低芯片的噪音，提高检测的精确度。图 5 为 aCGH 的基本原理。

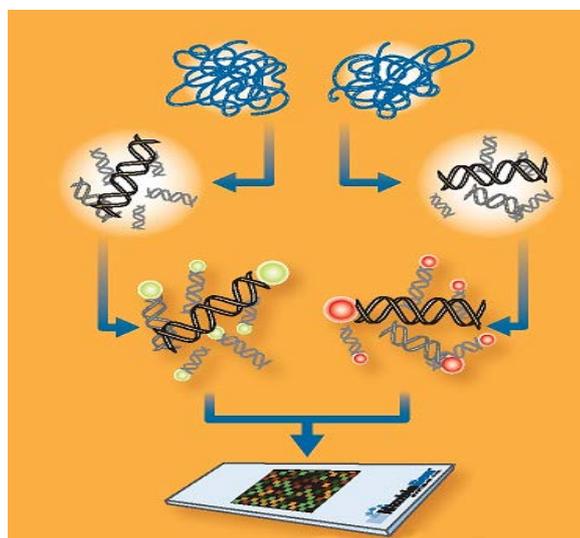


图 5. aCGH 的基本原理

3. 浦东解码生命科学研究院优势

- 浦东解码生命科学研究院一代测序平台可提供一套完整的 Fosmid 测序方案。
- 浦东解码生命科学研究院根据客户需要可以提供不同厂家的 SNP 基因分型芯片服务，

包括 Illumina SNP 基因分型芯片服务和 Affymetrix SNP 基因分型芯片服务。

- 浦东解码生命科学研究院拥有 Agilent 芯片平台, 可根据客户需要, 使用不同密度 aCGH 芯片检测低至数十 Kb 的片段扩增或缺失。
- 浦东解码生命科学研究院拥有的深度测序平台, 在 CNV 的检测中有自己独特的创新。

(三) 短串联重复序列, STR

短串联重复序列(short tandem repeat, STR)又称微卫星 DNA 标记(micro satellite DNA), 是一类广泛存在于人类基因组中的 DNA 多态性。STR 由 2-6 碱基对构成核心序列, 呈串联重复排列。STR 的长度一般在 100~300 bp 之间, 因个体间 DNA 片断长度或 DNA 序列差异而成高度多态性, 在基因传递过程中遵循孟德尔共显性方式遗传。因 STR 片段短、扩增效率高等特点, 已广泛应用于法医学中个体识别和亲子鉴定等领域。

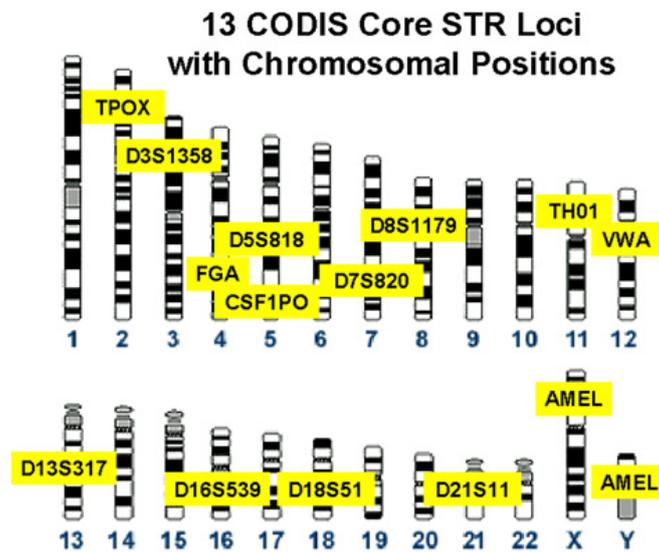


图 6. 用于个体识别的常用 STR 位点

1. STR 的应用

➤ 群体遗传学

STR在人类群体中具有高度的多态性和遗传稳定性。通过对不同人群的STR等位片段种类和频率分布进行检测, 可以发现STR的差异, 这些差异反映了不同人群间的遗传差异。运用STR 技术可以了解人类群体的基因流动及遗传差异。

➤ 个体识别和亲权鉴定

法医检案中, 经常会遇到极少量或降解严重的生物检材, 此时最好的方法是用PCR扩增STR。人体血液、精液、精斑、毛发、指甲、骨和牙齿均可作为分析STR的DNA来源。正是因为STR广泛存在于人类基因组中, 具有高度多态性、杂合性和稳定性, 当把一定数量的STR 位点联合分析后, 可以得到相当高的累积个体识别率和父权排除率。

➤ **寻找致病基因**

通过家系和对照研究，运用连锁分析和关联分析，可以找到与疾病高度相关STR位点。一般来讲，目的基因若与STR位点有连锁关系，则其位置可能与STR 位点邻近。对STR附近区域进行克隆测序，就有可能发现目的基因或致病基因。

2. 检测方法

➤ **PCR--STR 分型技术**

通过PCR或复合PCR扩增STR片段，然后用不同的电泳方法分离等位基因片段，最后经银染、溴化乙锭染色或荧光标记法检测STR分型结果，对照等位基因分型标准物(allelic ladder)判断基因型。

➤ **DNA 测序法分型**

常规的一代测序即可完成 STR 分型的检测，流程见图 7。

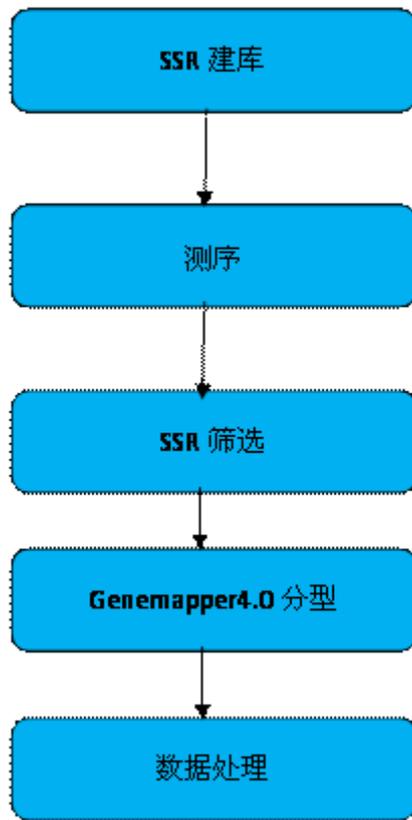


图 7. 应用测序方法进行 STR 分型

3. 浦东解码生命科学研究院优势

浦东解码生命科学研究院在国内率先开发建立了亚洲人特异 STR 数据库。将 3 个亚洲特异的 STR 序列与 15 个国际公认的 STR 序列结合，可使个体识别率超过千亿分之一，即 1000 亿人之间不会有二个个体发生重复。由此建立的独具特色的“基因身份证—身份鉴定

版”。

（四）DNA 甲基化

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。

1. DNA 甲基化检测的应用

➤ 遗传性疾病

某些遗传性疾病与基因的甲基化状态改变或甲基化相关基因的异常相关。例如甲基转移酶的变异导致的ICF综合征(immunodeficiency, centromeric instability and facial anomalies)是一种常染色体隐性遗传病,患者常有各种免疫缺陷和面部畸形的表征,且在其细胞中心粒周围染色体的稳定性下降。研究DNA甲基化与遗传病的关系可能成为揭示遗传病发病机制的重要方面。

➤ 恶性肿瘤

全基因组的低甲基化和局部性(CpG岛)高甲基化并存这种现象,既是癌症发生的重要原因之一,也是肿瘤良恶转化的重要标志。一般认为,癌症发生的机制主要包括癌基因的活化和抑癌基因的失活,目前研究发现二者都与基因甲基化异常有关。一方面,局部性的高甲基化使抑癌基因表达受到抑制,从而丧失抑癌的功能。另一方面,全基因组的低甲基化可促使癌基因活化、逆转座子的转录活性增强及基因的不稳定性增加。DNA甲基化的检测有望用于肿瘤的早期诊断以及预后的预测,而去甲基化药物在治疗肿瘤的前景更是令人瞩目。

➤ 动植物性状改良

近年来, DNA 甲基化作为一种新的分子遗传标记被用于动物遗传育种上,如预测禽畜杂交优势、检测动物生长性质等。在植物研究中,则内容更加广泛。DNA 甲基化作为高等植物中普遍存在的一种常见 DNA 共价修饰方式,参与了植物的许多重要生命过程,控制调节植物基因表达、生长发育、抵御逆境等生命活动。揭示 DNA 甲基化调控植物基因表达机理,对于改良植物表型性状,提高植物适应性等方面具有重要的作用。

2. 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 8。

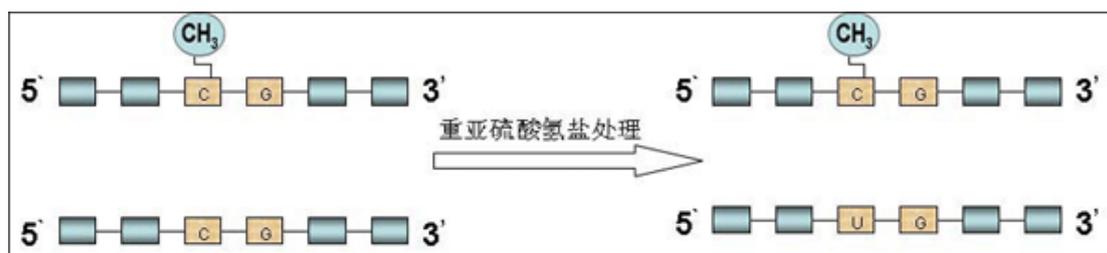


图 8. DNA 甲基化测序原理

➤ 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。

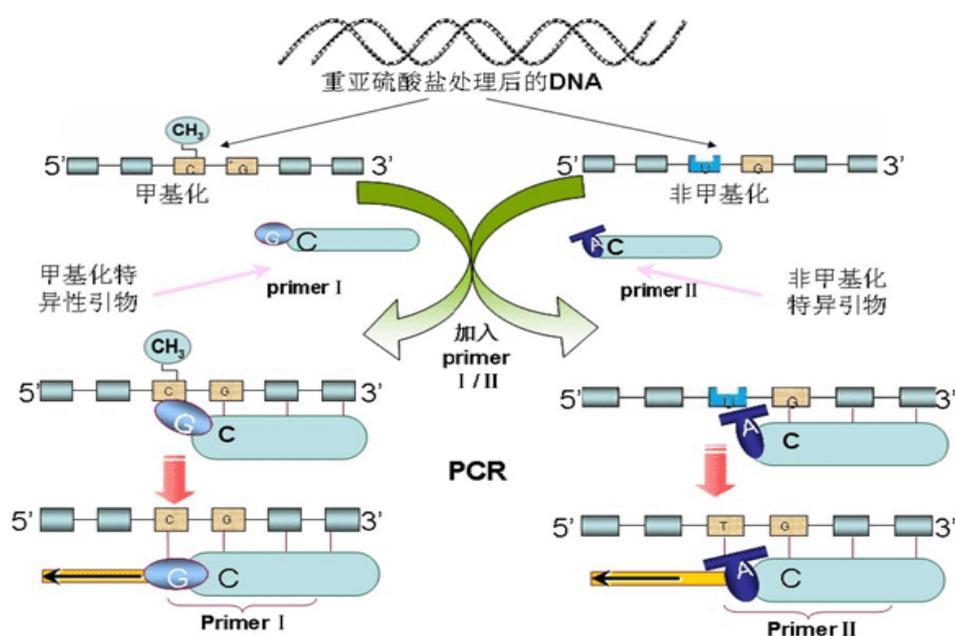


图 9. MS-PCR 原理

MS-PCR 的实验流程见图 9。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。

➤ **甲基化特异性多连接依赖性探针扩 (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 10。

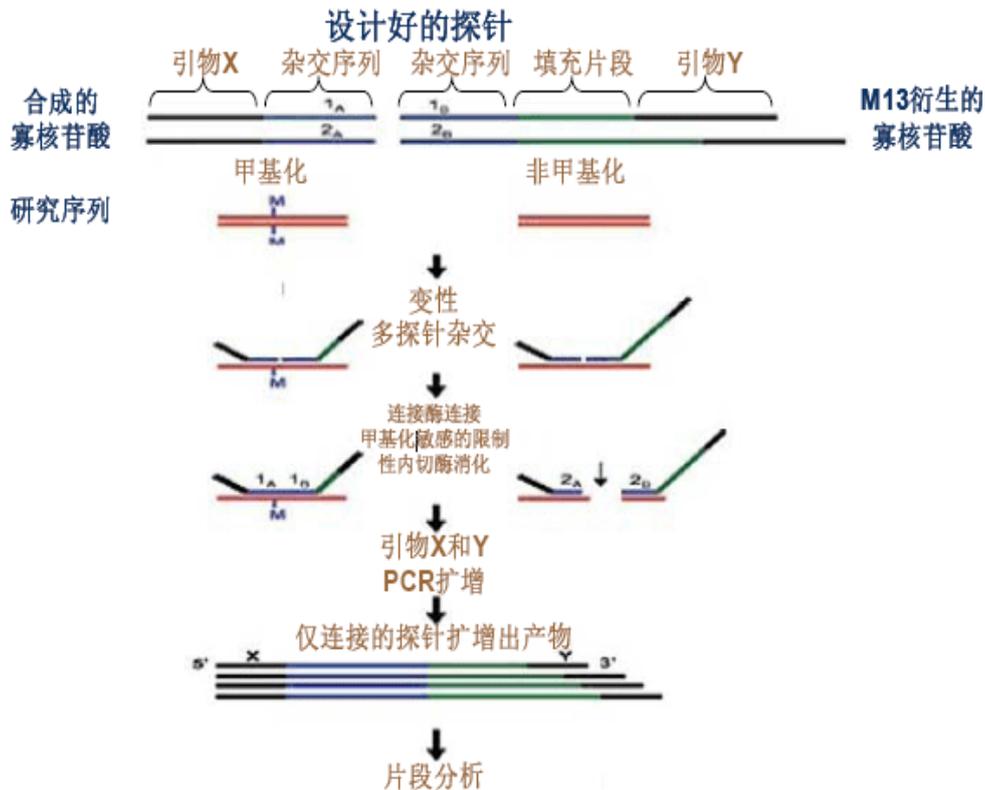


图 10. MS-MIP 原理

➤ 高通量测序

新一代测序仪的飞速发展，使得测序成本大幅度下降，也使得甲基化组（methyloome）的研究成为可能。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。

➤ 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、Illumina 和 Affymetrix。平台不同，过程也各异。Agilent 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）。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。

3. 浦东解码生命科学研究院的优势

DNA 甲基化生物标记物在疾病诊断、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 甲基化检测服务。

二. RNA 水平检测

(一) 转录组

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、核糖体 RNA、转运 RNA 及非编码 RNA;在狭义上是指所有 mRNA 的集合。

1. 检测方法

➤ 转录组测序

转录组测序是基因结构和功能研究的基础。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 11。



图 11. 转录组样品制备与测序流程

2. 应用领域

- 疾病分型研究；
- 分子标志和药物靶点的筛选；
- 药理与病理学研究；
- 功能基因组筛选；
- 生物代谢调控研究

3. 浦东解码生命科学研究院的优势

- 重复性好；
- 数字信号；
- 检测精度高，从几个到几十万个拷贝；
- 分辨率高，可以检测单个碱基差异。

(二) 小 RNA

小 RNA 是生物体内一类重要的特殊分子，以 RNA 干扰、翻译抑制、异染色质形成等

基因沉默方式调控生长发育、应激反应等多种多样的细胞进程。许多疾病的发生、发展都与小 RNA 的调控密切相关。

1. 检测方法

➤ 小 RNA 测序

二代测序省去了小 RNA 克隆文库的构建过程，可以一次性产生上百万条小 RNA 序列，能够快速鉴定特定条件下表达的已知小 RNA 并发现新的小 RNA，同时还可以研究不同条件下小 RNA 的表达差异。

➤ 微小 RNA (microRNA, miRNA) 芯片

miRNA 芯片的检测原理是：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 的表达量出现了异常。见图 12。

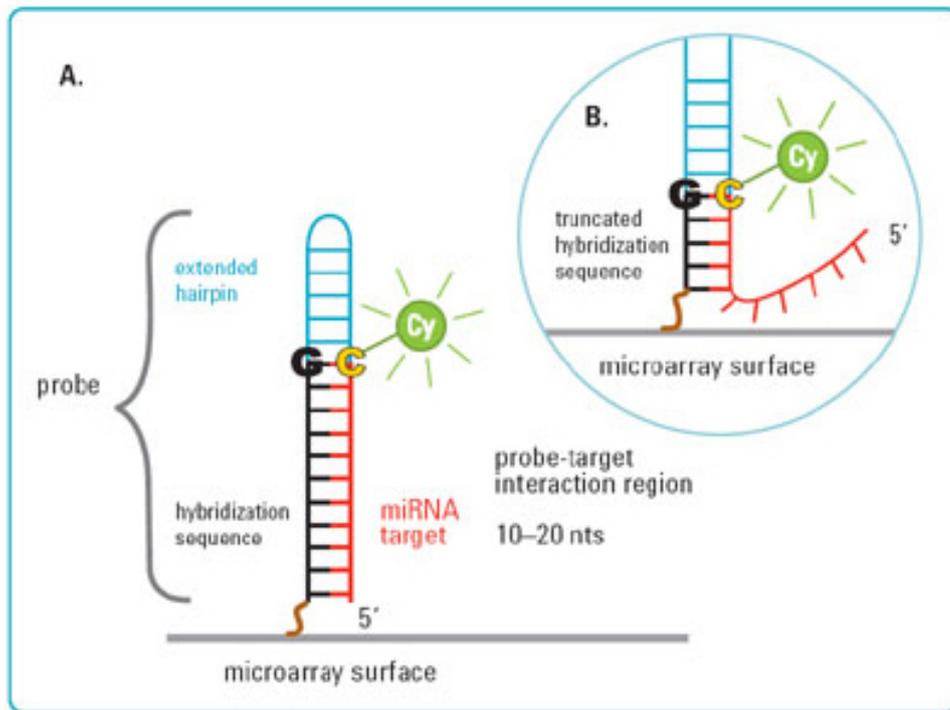


图 12. miRNA 的检测原理

2. 应用领域

- 疾病致病机制研究；
- 靶基因功能研究；
- 肿瘤研究；
- 药物药理研究；
- 生物标志物筛选；

- 发育进化研究。

3. 浦东解码生命科学研究院优势

- 高通量：一次测序得到的序列可以达到千万左右；
- 可拓展性：可以鉴定未知新小 RNA；
- 重复性良好：深度测序保证了检测的随机性，重复性好；
- 高分辨率：可以检测到单个碱基差异。

第二部分 测序平台

一. 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—ABI3730xl 测序仪（图 13），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 13. 3730XL 测序仪

1. 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° C，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。

2. 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 3730x1 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。

3. 一代测序服务流程（图 14）

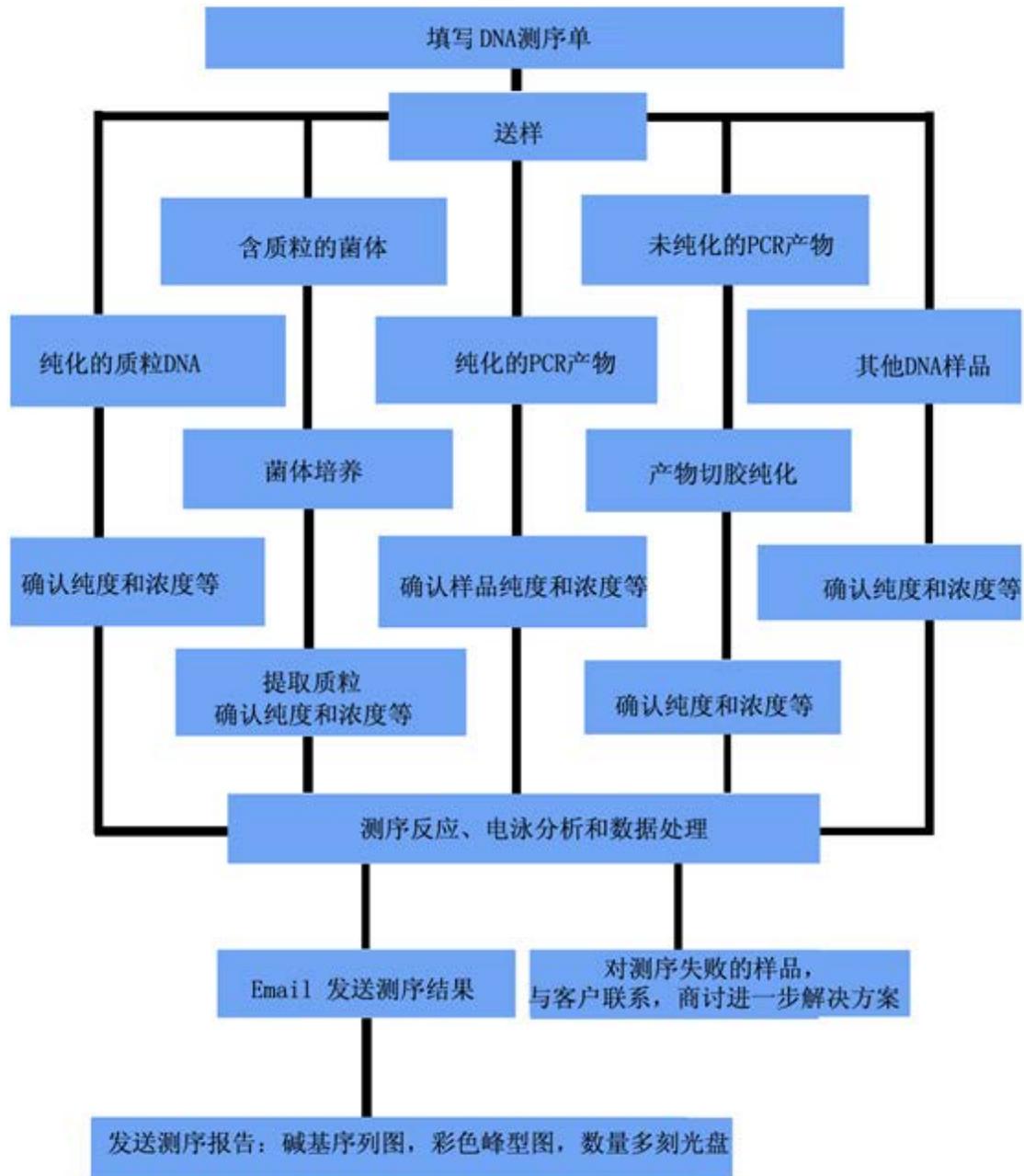


图 14. 一代测序服务流程图

二. PGM/Ion Pronton 测序系统

Ion Torrent（图 15），中文称个人化操作基因组测序仪(PGM)，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 15. 个人化操作基因组测序仪

1. 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。

2. 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 16）。

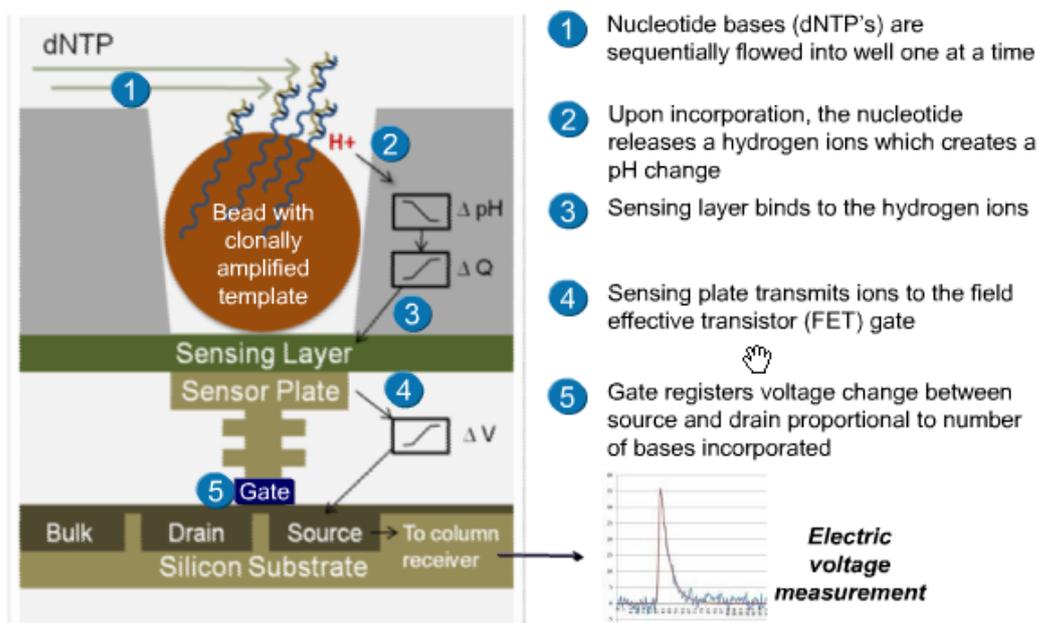


图 16. Ion Torrent 的测序机理及流程

3. 浦东解码生命科学研究院服务领域

- 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

4. 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。

5. 生物信息服务

- 基本生物信息学分析
 - A. 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。
 - B. 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
 - A. mapping 优化, 去除 PCR 重复子
 - B. 通过序列比对分析 SNP 和插入缺失 (INDEL)
 - C. 基于 peak 的多样品间差异分析
 - D. 基因表达分析
 - E. SNP 变异分析
 - F. 融合基因分析
 - G. 筛选突变

6. 应用实例

- Life Technology 公司在 nature 上发表了 Ion Torrent 技术特点及优势的文章。该研究展示了 Ion Torrent 的技术特点及原理。
- 论文: Rothberg JM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011; 475: 348–352

参考文献

- Howden BP, et al. Evolution of Multidrug Resistance during *Staphylococcus aureus* Infection Involves Mutation of the Essential Two Component Regulator WalKR. *PLoS Pathog*. 2011;7(11):e1002359.
- Alexander Mellmann, et al. Prospective Genomic Characterization of the German Enterohemorrhagic *Escherichia coli* O104:H4 Outbreak by Rapid Next Generation Sequencing Technology. *PLoS One*. 2011; 6(7): e22751
- Glenn TC, et al. Field guide to next-generation DNA sequencers. *Mol Ecol Resour*. 2011;11(5):759-69.

三. Illumina HiSeq 3000/4000 测序系统

Illumina 公司的 HiSeq3000/4000 测序系统（图 17）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 17. HiSeq 测序系统

1. 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。

2. 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生产 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 18。

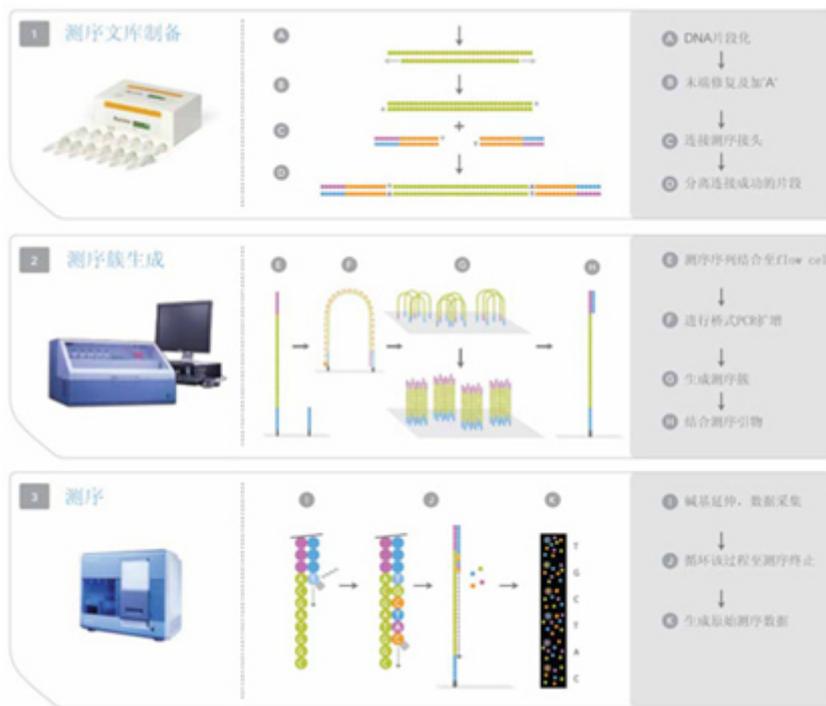


图 18. HiSeq 测序原理及流程

3. 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
- 小分子 RNA 测序
- ChIP-Seq

- 甲基化检测
- 羟甲基化测序
- 单细胞测序

4. 生物信息服务

- 基本生物信息学分析
 - A. 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.
 - B. 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。
 - C. RNA-seq 序列质量过滤
 - D. MeDIP-seq 序列质量过滤
 - E. MeDIP-seq 序列与参考序列的比对
- 高级生物信息学分析
 - A. mapping 优化，去除 PCR 重复子。
 - B. 通过序列比对分析 SNP 和插入缺失 (INDEL)
 - C. MeDIP-seq 序列数据在全基因组的分布趋势
 - D. 统计 MeDIP-seq 序列富集区域 (peak) 的信息
 - E. 基于 peak 的多样品间差异分析
 - F. RNA-seq 序列与参考基因组序列的比对
 - G. 基因表达分析
 - H. SNP 变异分析
 - I. 融合基因分析
 - J. 筛选突变

5. 浦东解码生命科学研究院的优势

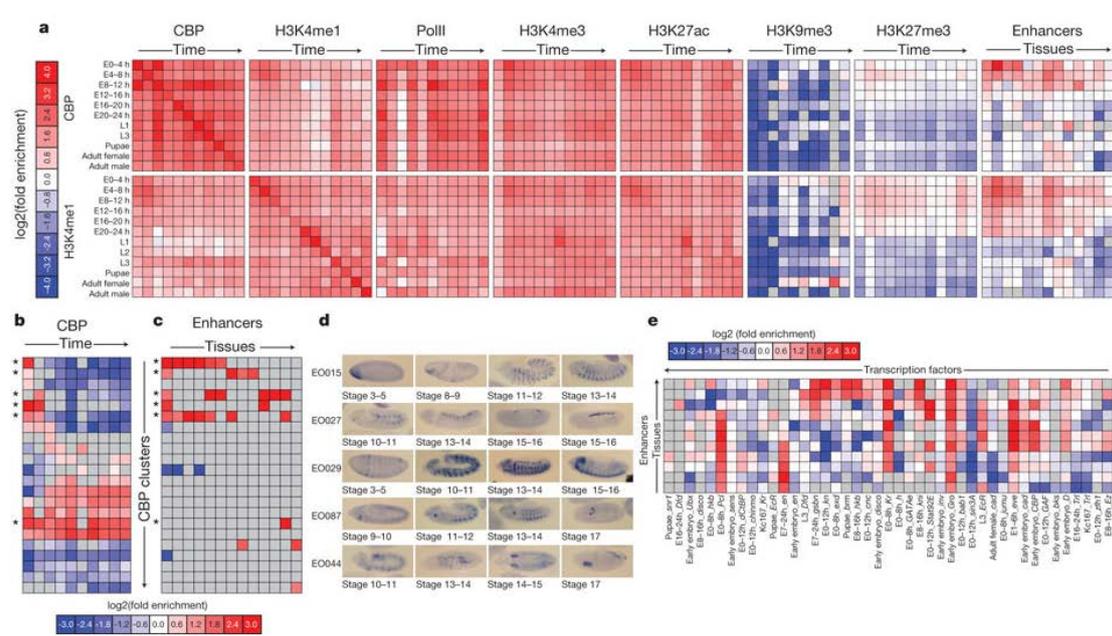
- 高质量的测序数据：在每个流动槽中对数百亿到数千亿碱基机型准确的碱基识别。在读测序时，模板可以重新生成。在两次读取间获得相同准确度，严格的功能检测确保检测稳定性和重复性。
- 应用广泛：所有的深度测序均可在该设备上完成。
- 高效的样本制备：高速自动化工作流程，低样品量要求，可在两周的时间内生成所有数据。
- 灵活的平台：稳定的 Solexa 测序化学反应支持多种读取长度，研究人员可量身定制每次所需读长，数据整体质量得以提高。

6. 应用实例

(1) 论文: Nègre N, et al. A cis-regulatory map of the Drosophila genome. *Nature*. 2011; 471: 527–531

(2) 研究亮点

- 基因调控元件的系统性诠释是基因组学研究中的一项重大挑战。染色质修饰的标志和全基因组转录因子结合位点的直接绘图成功地确定了具体亚型的调控元件。虽然在果蝇的转录因子结合位点、染色质位点等方面有着几项开创性的研究，但是全面诠释调控基因组仍然是一项重大挑战。本研究中通过 illumina 高通量测序方法制作了一张果蝇调控染色质基因图谱。
- 研究者使用这些数据推断出了超过 20,000 个的候选调控元件，发现了近 2,000 个基因组区域密集转录因子结合区域与染色质活动与辅助功能相关联。illumina 高通量测序手段为本研究绘制图谱提供了准确的数据。



参考文献

- Najmabadi H, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*, 2011;478: 57-63
- Amores A, et al. Genome Evolution and Meiotic Maps by Massively Parallel DNA Sequencing: Spotted Gar, an Outgroup for the Teleost Genome Duplication. *Genetics*. 2011; 188(4): 799–808
- Shinzato C, et al. Using the Acropora digitifera genome to understand coral responses to environmental change. *Nature*. 2011;476:320–323
- Wu G, et al. Statistical Quantification of Methylation Levels by Next-Generation Sequencing. *PLoS One*. 2011; 6(6): e21034
- Al-Balool HH, et al. Post-transcriptional exon shuffling events in humans can be evolutionarily conserved and abundant. *Genome Res*. 2011;21: 1788-1799

四. Pyrosequencing 焦磷酸测序

1. 技术特点

Pyrosequencing 焦磷酸测序（图 19）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。对于不同的应用来说，这意味着更经济省时。



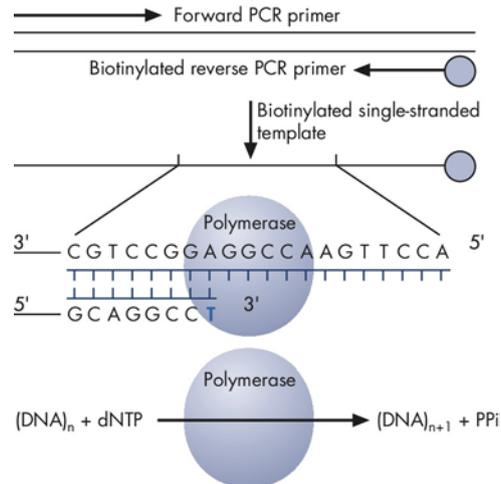
图 19. pyrosequencing 测序系统

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

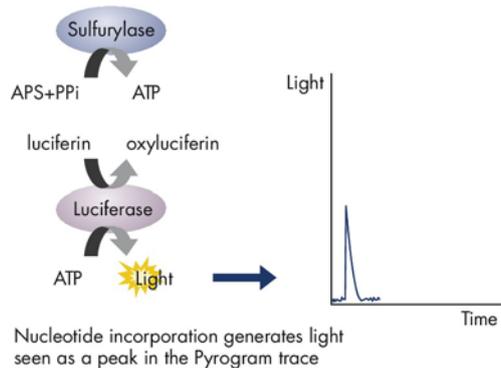
2. 检测原理

- 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。

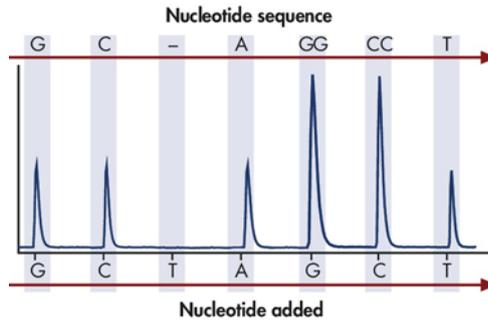


- 步骤 5: 腺苷三磷酸双磷酸酶是核苷酸的降解酶, 持续地对未结合的核苷酸和 ATP 进行分解。当分解完成时, 便会引入另一个核苷酸。



- 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。



3. 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

段片段重复序列分析

➤ 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

参考文献

- Guenin S, Mouallif M, et al. Aberrant promoter methylation and expression of UTF1 during cervical carcinogenesis. *PLoS One*. 2012;7(8):e42704.
- Dejeux, E., El abdalaoui, H., Gut, I.G., and Tost, J. (2009) Identification and quantification of differentially methylated loci by Pyrosequencing Technology. *Methods Mol. Biol.* 507, 189.
- Koontz, D.A., Huckins, J.J., Spencer, A., and Gallagher, M.L. (2009) Rapid detection of the CYP2A6*12 hybrid allele by Pyrosequencing technology. *BMC Med. Gen.* 10, 80.
- Goudie DR, et al. Multiple self-healing squamous epithelioma is caused by a disease-specific spectrum of mutations in TGFBR1. *Nat Genet.* 2011;43:365-369.
- Shen J, et al. Genome-wide DNA methylation profiles in hepatocellular carcinoma. *Hepatology*.2012. [Epub ahead of print]

第三部分 芯片平台

一. Illumina 微阵列扫描仪系统

Illumina 微阵列扫描仪系统（图 20）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 20. 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 2.x 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（图 21）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 21. 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（图 22）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 22. 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（图 23）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 23. 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。

第四部分 芯片技术服务

一. 表达谱芯片

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。

1. 应用领域

- 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。
- 临床诊断：可用于某些疾病的临床诊断。
- 药物筛选：药物药理机制分析，药物的毒性毒理研究。

2. 浦东解码生命科学研究院提供的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。

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。

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。

(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。

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。
- 芯片生产 100% QC 质控：利用生产过程中专利的解码技术，质控可深入到每张芯片每个微珠上，保证了数据的可靠性和重复性。

检测原理

微珠采用 3.1 um 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”。

(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。

二. 比较基因组杂交芯片

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。

1. 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。

2. 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。本公司提供后续的生物信息学服务。

(1) Agilent 比较基因组杂交 (CGH) 芯片服务

CGH 微阵列芯片 (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。
- 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。

- 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。
- 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。

检测原理

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 24。

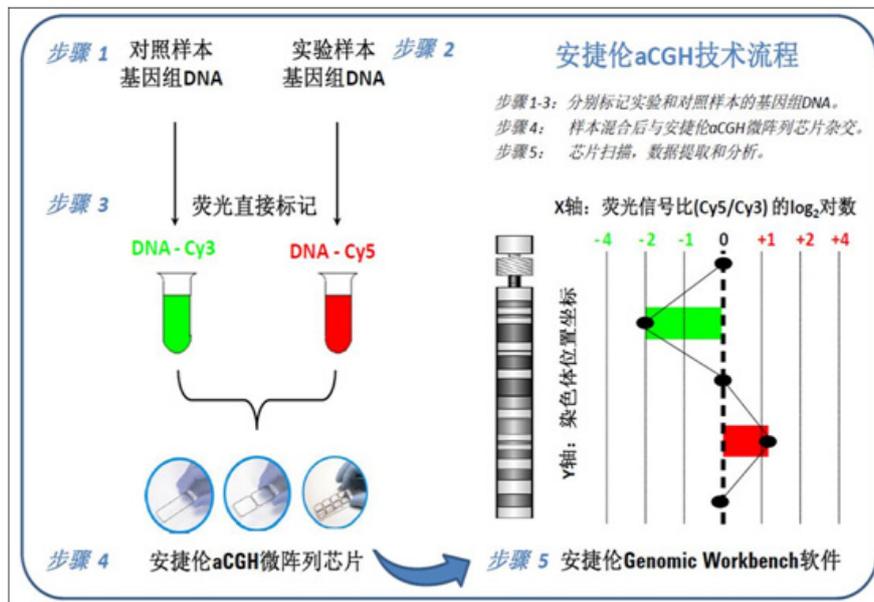


图 24. aCGH 的检测原理

三. ChIP-on-chip 芯片

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。

1. 应用领域

- 蛋白与 DNA 相互作用
- 转录因子和聚合酶的 DNA 结合位点研究

2. 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) 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。

四. DNA 甲基化芯片

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。

1. 应用领域

- 癌症研究
- 染色体结构稳定性研究
- 基因表达差异研究
- X-染色体失活
- 疾病发生及进展
- 细胞分化
- 发育

2. 浦东解码生命科学研究院提供的甲基化芯片服务

浦东解码生命科学研究院的 Agilent DNA 甲基化芯片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。

Agilent DNA 甲基化芯片服务

Agilent 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) 或者 MBD (Methyl Binding Domain Protein)，再将其用于芯片检测，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、大鼠等基因组的非重复区域。
- 高覆盖面，高解析度的探针：Agilent 高密度 244K 甲基化芯片集成有约 244,000 个 60-mer 的寡核苷酸探针，这些探针以 100-300bp 的平均间隔覆盖 UCSC 数据库中所有的 CpG 岛和 RefSeq 数据库中已研究清楚的约 17000 个转录本。
- 高重复性：Agilent 拥有多项芯片专利技术，芯片上探针均经过反复试验优化，另双色标记系统比单色标记系统得到灵敏度 and 精确度更高，重复性更好，既可有效检测处低频 DNA 甲基化水平改变，又可在同一张芯片上直接比较不同样本的差异。

- 支持灵活定制: Agilent 公司拥有完整的探针集可供选择, 通过 Agilent 在线设计平台 e-array, 研究者可根据感兴趣的组织及特定基因区域定制不同探针密度的芯片。
- 专业的数据处理: Agilent 公司 GeneSpring 套装可保证研究者更好检测出 DNA 甲基化状态并降低错判率, 通过该软件研究者更方便的将数据定位到基因组上并结合已有的基因组注释进行下一步的分析工作。

检测原理

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, 之后与甲基化芯片杂交来进行检测。

检测流程。

1. 将基因组 DNA 超声打断成 400-500bpDNA 片段;
2. 加热变性并将变性后的单链 DNA 样品分成两份;
3. 其中一份单链 DNA 样品加入抗 5'-甲基化胞嘧啶核苷抗体。使用免疫磁珠法分离样品中甲基化 DNA 片段的抗体复合物, 样品中其余的非甲基化 DNA 片段被洗脱;
4. 纯化免疫共沉淀的 DNA 片段; 对 MeDIP(Cy5)与 Input(Cy3)样品分别进行标记;
5. 标记后的 MeDIP 与 Input 样品混合、变性, 与芯片杂交;
6. 检测杂交信号并进行数据分析

五. 微小 RNA 芯片

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。

1. 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。

2. 浦东解码生命科学研究院为您提供小 RNA 芯片服务

Agilent microRNA 芯片

安捷伦高性能 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。

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 的表达量是否出现了异常。见图 25。

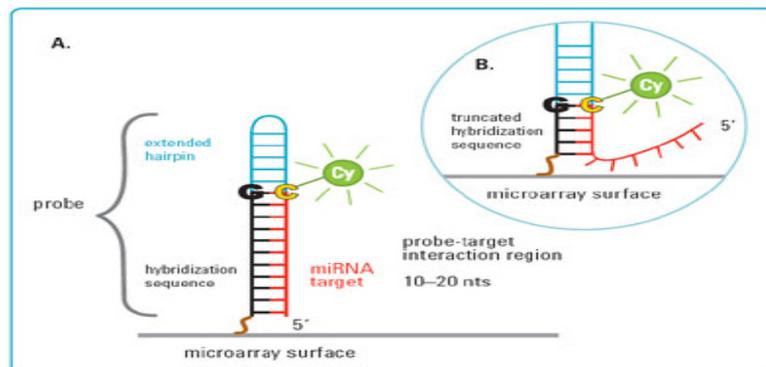


图 25. miRNA 表达谱芯片检测原理

六. Sequenom MassArray 芯片

Sequenom 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。

1. 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

2. 浦东解码生命科学研究院提供 Sequenom MassArray 芯片服务

Sequenom 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。

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。

检测原理

Sequenom MassArray 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 26。

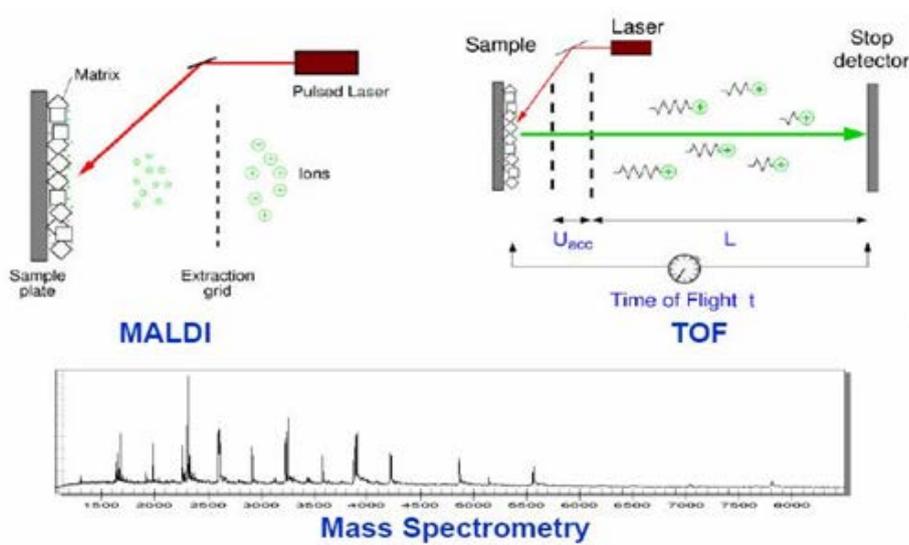


图 26. Sequenom MassArray 系统检测原理

第五部分 生物信息学服务

芯片结果的生物信息解码对于研究的深入至关重要。为了便于客户的研究，浦东解码生命科学研究院提供了人性化的后续生物信息学服务，不仅提供一目了然的芯片实验报告，同时提供高质量的生物信息服务。

1.实验报告

浦东解码生命科学研究院提供的芯片实验的实验报告，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。

(1) 芯片的原始数据

包括原始的.EXP、.DAT、.CEL 及report 文件。

(2) 芯片图像

由该图像，客户可方便的查看芯片各个部分的细节，了解芯片杂交，洗脱，扫描过程的效果。

(3) 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

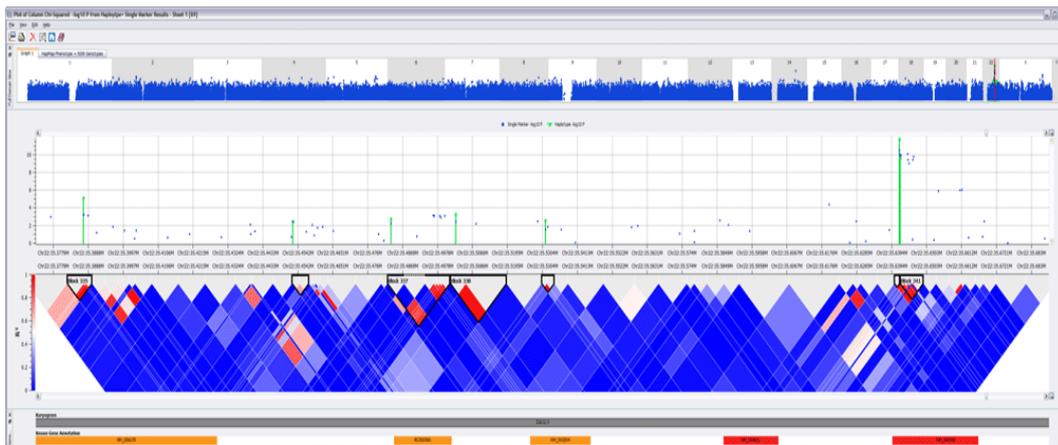
File	Bounds	Contrast QC	Contrast QC (Random)	Contrast QC (Nsp)	Contrast QC (Sty)	Contrast QC (Nsp/Sty Overlap)	QC Call Rate	QC Call Rate (Nsp)	QC Call Rate (Nsp/Sty Overlap)	QC Call Rate (Sty)	Computed Gender
22 F1.CEL	In	2.53	2.53	2.00	2.41	2.41	94.14	92.54	95.75	91.95	male
23 F2.CEL	In	2.67	2.67	2.63	2.67	2.76	95.57	93.67	97.17	95.01	female
24 G0.CEL	In	2.21	2.21	2.12	2.32	2.63	91.89	90.87	94.52	93.31	male
25 G1.CEL	In	2.96	2.96	3.19	2.89	3.22	96.56	95.76	97.84	94.20	male
26 G2.CEL	In	2.99	2.99	3.04	3.29	3.28	97.05	96.53	98.46	94.04	female
27 H0.CEL	In	2.22	2.22	3.27	3.24	3.52	96.23	94.47	97.29	95.65	male
28 H1.CEL	In	2.67	2.67	2.30	2.31	2.58	92.36	91.26	94.70	97.60	male
29 H2.CEL	In	2.16	2.16	2.53	2.53	2.89	93.65	92.90	95.50	89.86	female
30 I0.CEL	In	2.28	2.28	2.64	2.53	2.76	93.71	92.29	95.75	90.19	male
31 I1.CEL	In	2.98	2.98	3.01	2.89	2.99	96.59	96.02	98.03	93.56	male
32 I2.CEL	In	1.78	1.78	1.74	1.38	1.76	92.69	91.13	94.76	98.73	female
33 J0.CEL	In	2.07	2.07	2.36	2.35	2.71	94.11	93.44	96.00	90.02	male
34 J1.CEL	In	2.60	2.60	2.58	2.42	2.84	96.62	97.17	97.47	93.72	male
35 J2.CEL	In	1.92	1.92	1.50	1.55	2.20	90.21	88.56	92.85	85.35	female
36 JwT-0449.CEL	Out	0.36	0.36	-0.46	-0.64	-0.03	84.22	83.29	85.81	78.58	female
37 JwT-3048.CEL	Out	0.03	0.03	-0.34	-0.36	-0.02	87.59	88.69	89.59	81.00	male
38 M0.CEL	In	2.91	2.91	3.23	3.40	3.24	96.36	96.14	97.60	93.40	female
39 N0.CEL	In	2.72	2.72	2.72	2.66	2.92	97.39	96.40	98.89	94.69	male
40 N1.CEL	In	2.34	2.34	2.49	2.41	2.74	93.78	93.57	95.56	89.37	male
41 N2.CEL	In	2.09	2.09	2.18	2.13	2.17	92.39	90.36	94.95	88.24	female
42 P0.CEL	In	2.74	2.74	3.13	3.16	3.16	96.59	96.53	98.09	92.75	female
43 P1.CEL	In	1.26	1.26	0.79	1.39	1.40	89.15	88.43	90.88	85.51	male
44 P2.CEL	In	1.82	1.82	1.85	1.84	2.09	90.87	90.10	93.04	86.15	female
45 Q0.CEL	In	2.47	2.47	2.13	2.51	2.83	94.97	94.09	96.55	91.95	male
46 Q1.CEL	In	1.18	1.18	0.69	0.65	0.91	85.77	86.76	87.74	73.39	male
47 Q2.CEL	In	2.36	2.36	2.50	2.51	2.58	95.63	94.47	97.41	92.43	female
48 R0.CEL	In	2.72	2.72	3.06	3.18	3.27	95.67	93.57	97.72	92.91	male
49 R1.CEL	In	2.21	2.21	2.24	2.59	2.49	93.98	92.42	96.06	90.50	male
50 R2.CEL	In	1.87	1.87	1.63	1.74	2.00	88.19	85.73	91.37	82.93	female

➤ 样品 SNP call 结果总表

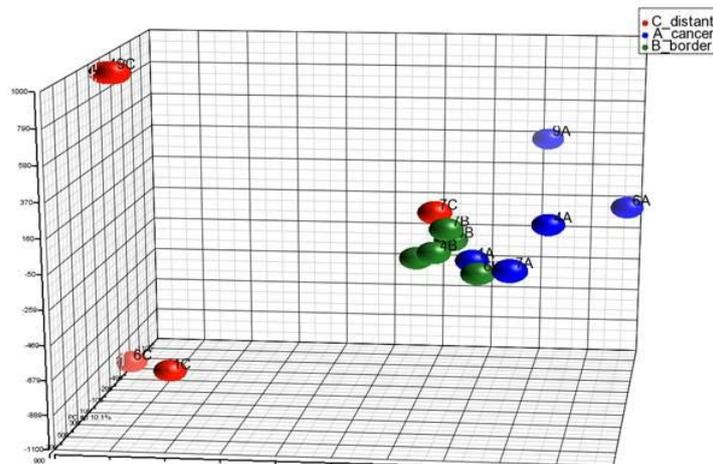
```
#GenomeWideSNP_6.na29.annot.db
##genome-version-ucsc=hg18
##genome-version-ncbi=36.1
```

Probe Set ID	Call Codes	Confidence	Forced Call Codes	Signal A	Signal B	Forward Strand
SNP_A-2131660	AB	0.001	AB	1830.388	1866.541	CT CT rs2887286 1 1145994
SNP_A-1967418	BB	0.006	BB	229.804	696.715	GG GG rs1496555 1 2224111
SNP_A-1969580	BB	0.004	BB	1161.019	4258.893	GG GG rs41477744 1 2319424
SNP_A-4263484	AB	0.001	AB	1555.889	1141.152	CT CT rs3890745 1 2543484
SNP_A-1978185	AA	0.002	AA	1727.128	386.225	GG GG rs10492936 1 2926730
SNP_A-4264431	AB	0.008	AB	1374.601	1140.755	GA GA rs10489588 1 2941694
SNP_A-1980898	BB	0.015	BB	392.428	970.943	CC CC rs2376495 1 3084986
SNP_A-1983139	AA	0.009	AA	2468.521	451.303	AA AA rs4648462 1 3155127
SNP_A-4265735	BB	0.003	BB	373.692	1277.006	AA AA rs10492939 1 3292731
SNP_A-1995832	BB	0.006	BB	597.921	1175.229	GG GG rs9424283 1 3695086
SNP_A-1995893	BB	0.007	BB	368.860	1841.182	GG GG rs2154068 1 3710825
SNP_A-1997689	BB	0.002	BB	648.964	3050.044	GG GG rs12060299 1 3753024
SNP_A-1997709	AA	0.003	AA	1605.028	242.587	TT TT rs10909802 1 3753427
SNP_A-1997896	AA	0.005	AA	1682.890	567.887	AA AA rs16824230 1 3756100
SNP_A-1997922	BB	0.002	BB	412.561	1198.438	GG GG rs17404435 1 3756146
SNP_A-2000230	AA	0.002	AA	3265.942	1298.384	AA AA rs12059199 1 4240737
SNP_A-2000332	AA	0.006	AA	2949.331	925.803	CC CC rs6702000 1 4243294
SNP_A-2000337	AA	0.000	AA	1980.041	283.237	TT TT rs16838547 1 4243405
SNP_A-2000342	AA	0.000	AA	4024.634	916.316	GG GG rs16838549 1 4243441
SNP_A-4268173	AA	0.012	AA	1326.251	278.357	CC CC rs3912752 1 4276892
SNP_A-2002663	AB	0.003	AB	2174.258	2285.586	GA GA rs505933 1 4371593
SNP_A-2004169	AA	0.006	AA	1472.315	218.324	CC CC rs4654438 1 4459761
SNP_A-2004249	BB	0.004	BB	424.713	1727.128	CC CC rs676853 1 4461025
SNP_A-4268681	AA	0.004	AA	1954.369	371.757	CC CC rs583027 1 4461905
SNP_A-2004332	AA	0.001	AA	1658.085	246.938	GG GG rs350165 1 4464544
SNP_A-4268770	BB	0.000	BB	338.034	2236.729	TT TT rs10489135 1 4482059
SNP_A-4268887	BB	0.005	BB	517.529	1560.505	CC CC rs12120353 1 4520856
SNP_A-2005859	NoCall	0.153	BB	1402.832	2317.540	--- GG rs241212 1 4531126

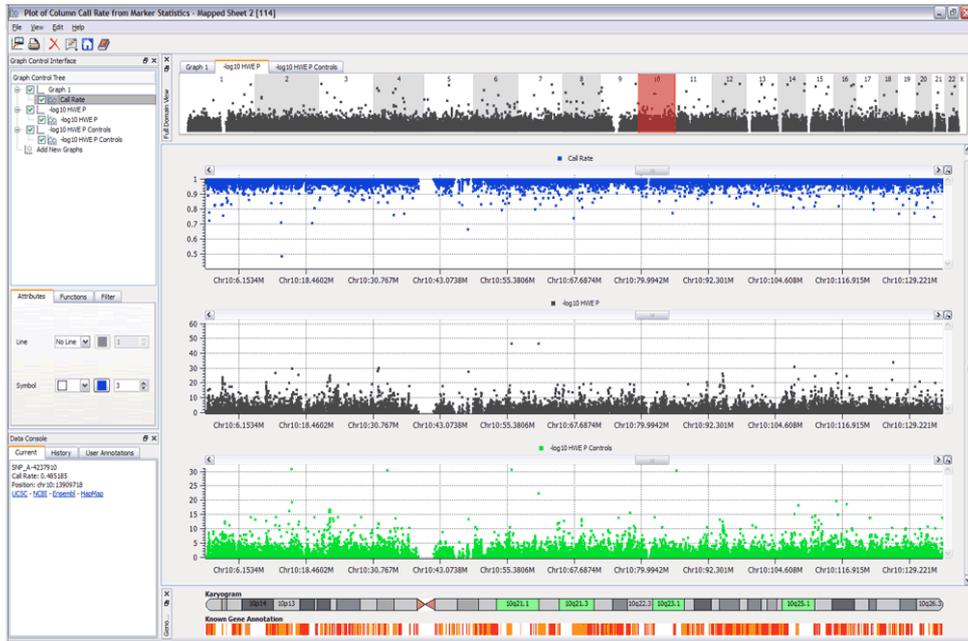
➤ 连锁不平衡性分析



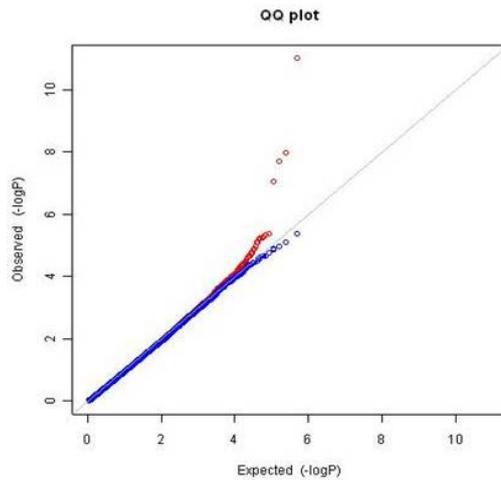
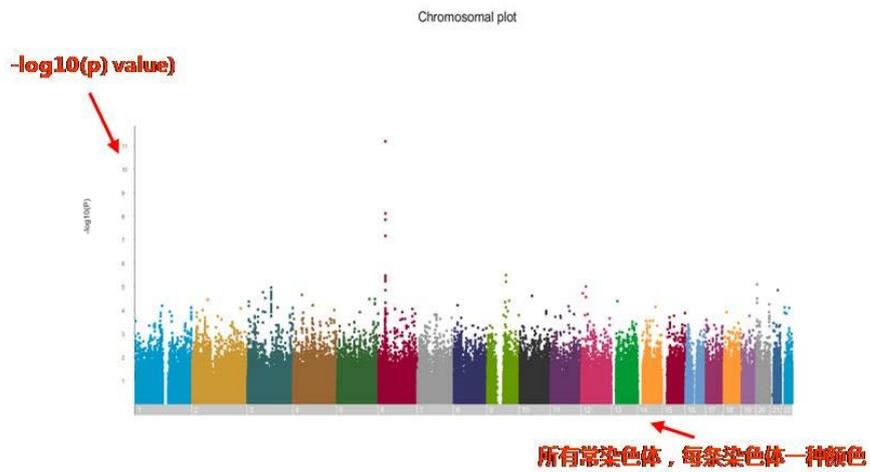
➤ 人群分层分析



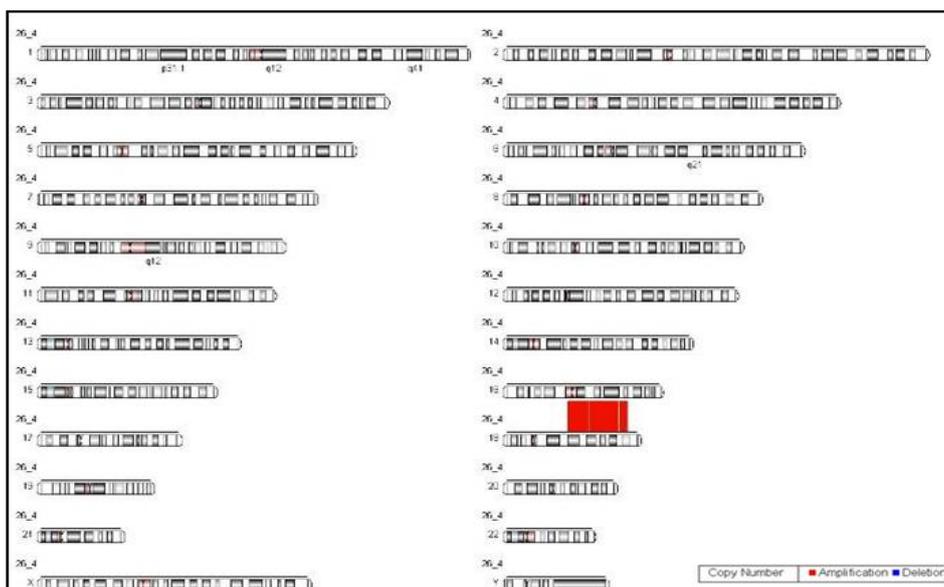
➤ 阳性位点深度过滤



➤ 单点关联分析



➤ 配对 CNV 分析、非配对 CNV 分析与区段筛选



2. 生物信息服务

浦东解码生命科学研究院还可根据客户需要提供相应的生物信息学分析服务。对于从基因芯片得到的实验数据，我们提供专业分析服务。

- 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。