中国科学 C辑:生命科学 2008 年 第 38 卷 第 10 期: 900 ~ 906
www.scichina.com life.scichina.com
《中国科学》杂志社
SCIENCE IN CHINA PRESS

专题

基因组分型技术的发展和应用

施锦绣, 王莹, 黄薇*

国家人类基因组南方研究中心, 上海市-科技部共建疾病与健康基因组学重点实验室, 上海 201203

* 联系人, E-mail: huangwei@chgc.sh.cn

收稿日期: 2008-10-03; 接受日期: 2008-10-08

国家杰出青年基金(批准号: 30625019)和上海市青年科技启明星计划(批准号: 08QA14053)资助项目

摘要 伴随着人类基因组计划、人类单倍型图谱和人类拷贝数变异图谱的完成, 基因分型技术获得了长足的发展, 同时促进了对复杂性状疾病的研究. 本文结合国家人类基因组南方研究中心的发展和研究项目, 介绍了目前国际上广泛应用的几种基因分型技术, 及国内外复杂性状疾病遗传研究的突破和进展.

关键词
单核苷酸多态性
拷贝数变异
复杂性状疾病

从古至今, 人类一直在探索疾病的相关知识. 1986 年, 人类基因组的第一个完整的限制性片段长度多态(restriction fragment length polymorphism, RFLP)图谱绘制完成, 这是人类的第一张分子遗传图谱, 标志着人类疾病发生机制的遗传学研究进入了一个新的阶段[1]. 而后, 微卫星位点(short tandem repeat , STR)代替了RFLP[2]. 此时, 基于家系连锁分析的疾病遗传学机制研究占据了统治地位. 1996 年, Risch和Merikangas指出, 通过对一般人群进行病例-对照研究, 即关联分析, 对捕获变异基因具有更高的效力[3], 因而提出了构建人类基因所有变异图谱的全球计划, 人类单倍型图谱计划得以启动. 随着人类单倍型图谱的完成[4,5], 单核苷酸多态性(single nucleotide polymorphism, SNP), 即单个核苷酸的变异所引起的DNA序列多态性, 因其数量众多和易检测性, 成为新的主要的遗传标记. 目前, 公共数据库内已有超过九百万的SNP[6]. 人类拷贝数变异图谱的完成[7], 标志着拷贝数变异(copy number variation, CNV)正在逐渐成为另一个越来越受到关注的遗传标记[8]. CNV被定义为超过 1 kb的DNA序列, 相对于参考基因组, 具有不同的拷贝数, 其中包括缺失、插入及其他较为复杂的变化.

与此同时, 基因分型的方法也在不断发展. 从繁琐低通量的单链构象多态性(single strand conformation polymorphism, SSCP)[9], 变性梯度凝胶电泳(denaturing gradient gel electrophoresis, DGGE)[10], 到自动化高通量的生物芯片技术, 使得我们发现和鉴别DNA序列上多态位点的工作变得快速而准确. 不同遗传标记中, 适合高通量、自动化分型的多态主要是SNP和CNV, 因而近年来迅猛发展的基因分型技术主要是基于SNP检测的分析技术.

国家人类基因组南方研究中心(简称南方中心)成立之时, 正是人类基因组计划STR遗传图谱绘制完成并开始被世人所认识的重要时期. 利用微卫星的全基因组分型技术, 南方中心遗传学研究部联合国内各临床学科, 广泛开展了中国人群疾病易感基因的全基因组定位研究, 在鼻咽癌[11]、肝癌、高血压、糖尿病、自身免疫性疾病、皮肤病等复杂性状疾病的遗传易感性研究中取得了令人鼓舞的成果. 与此同时, 南方中心遗传学研究部还积极参与国际合作项目, 承担了人类单倍型图谱计划中 21 号染色体的图谱构建工作, 并籍此建设、发展和完善了我国基因分型技术平台, 包括Illumina生物芯片技术, TaqMan技术和SNPstream基因分型系统等. 在此, 本文将结合

国家人类基因组南方研究中心的发展和研究项目, 简要介绍几种近年来发展的检测技术及复杂性状疾病遗传研究的突破和进展.

1 基因分型技术的革新与发展

目前基因分型技术主要是针对SNP的检测分析, 其原理包括目标DNA片段的扩增及扩增产物通过质谱、荧光或化学的方法进行测定, 从而获得SNP的位点信息[12]. 但是, 即便同样是对SNP的检测, 根据样本的多少和每个样本上所需检测SNP的多少, 人们必须采用不同的技术平台, 以获得最佳的"投入产出比".

1.1 高通量的基因分型技术

高通量的基因分型技术主要为基于各种基因芯片的技术; 其制作是将大量已知序列的 DNA 探针, 采用特殊方法固定在厘米见方的硅芯片或玻片上, 以此获得高密度的 DNA 探针阵列, 然后通过杂交获得 SNP 分型的信息. 具体的检测技术包括: DNA 探针阵列的构建、待测样品的准备、杂交及杂交信号的检测分析. 代表性产品主要有下列两大类:

(1) Illumina Beadarray的SNP分型技术系统包括全基因组分型的Infinium技术和自定义芯片的GoldenGate技术. DNA探针阵列的构建是利用微珠芯片技术[13,14]. Infinium技术无需PCR, 不受内切酶的位点限制, 在单管反应中获得高得率和精确度的实验数据. GoldenGate 技术具有高度灵活的特点, 它根据实验者的需要在一个样品里同时检测 96, 384, 768 或 1536 个SNP, 是最早生产的用户特制芯片代表性产品. 两种技术的原理见图 1(http://www.illumina.com).

(2) Affymetrix公司是利用原位光刻专利技术, 生产高密度寡核苷酸基因芯片[15]. 其GeneChip系列芯片最新产品是Affymetrix Genome-Wide SNP Array 6.0. 这张芯片上富集超过 906600 个SNP和超过 946000 个用于检测拷贝数变化的探针, 探针高密度且均匀地覆盖整个基因组. 样品准备及检测原理见图 2.

1.2 中等通量的基因分型技术

与高通量基因分型技术主要在基因组中心使用, 用于大量样本的"全基因组基因分型"不同, 中等通量的基因分型技术可在各个层次使用, 目的一方面是候选基因验证的相关分析, 另一方面多数是在走向临床前的大规模样本检测这样的转换研究. 为此目的, 很多公司致力于该技术的改进, 目前已有较多很好的发展. 其中较常用的主要有 ABI 公司的 TaqMan 技术, Beckman 公司的 SNPstream, Sequenom 公司的时间飞行质谱等.

(1) TaqMan技术的原理是合成一对能与PCR产物杂交的探针. 这对探针序列仅区别于多态性位点. 这对探针的 3′端连有荧光粹灭剂, 5′端分别有 2 种不同的荧光染料. 在PCR扩增过程中, 利用Taq 酶 5′核酸外切酶活性降解与目标序列完全互补的探针, 使荧光剂与粹灭剂分离而发出荧光. 如果探针与目标序列间存在错配, 就会大大减少荧光的释放量. TaqMan探针技术最突出的优点是不需要分离或洗脱等PCR后处理过程, 从而提高了检测速度[16,17].

(2) SNPstream基因分型系统是一套自动化的多重SNP分析系统, 在特制 384 孔板的每个微孔中可分析 12 个或 48 个SNP位点. 检测SNP位点的引物由 2 部分组成, 3′端是能与模板互补的部分, 约 20～25 个碱基的长度, 5′端是 20 个碱基的尾段序列, 与微孔中的核苷酸序列互补. 当引物与待检测的PCR片段杂交后, 在靶SNP位点处单碱基延伸. 延伸后的产物随后被带有与尾段互补的核苷酸序列的微矩阵芯片捕获并检测[18].

(3) 时间飞行质谱系统是目前唯一采用质谱法直接检测SNP的设备, 主要步骤包括碱基延伸, 样品离子化和离子化样品的质量分离三个部分. 该系统反应体系为非杂交依赖性, 不存在潜在的杂交错配干扰, 不需要各种标记物, 因而其突出特点是能以极高的精确度快速进行基因型识别, 直接测出带有SNP或其他突变的目标DNA[19,20].

基因分型技术仍然在向着更高通量, 更高准确性和更低廉成本的方向不断发展. 在全基因组测序技术的费用下降到足够低廉之前, 基因分型技术将是研究复杂性状疾病的主要技术手段之一. 与此同时, 如 SSCP, DGGE, DHPLC 等技术, 在筛查特定基因的突变或多态中仍旧有其重要作用.

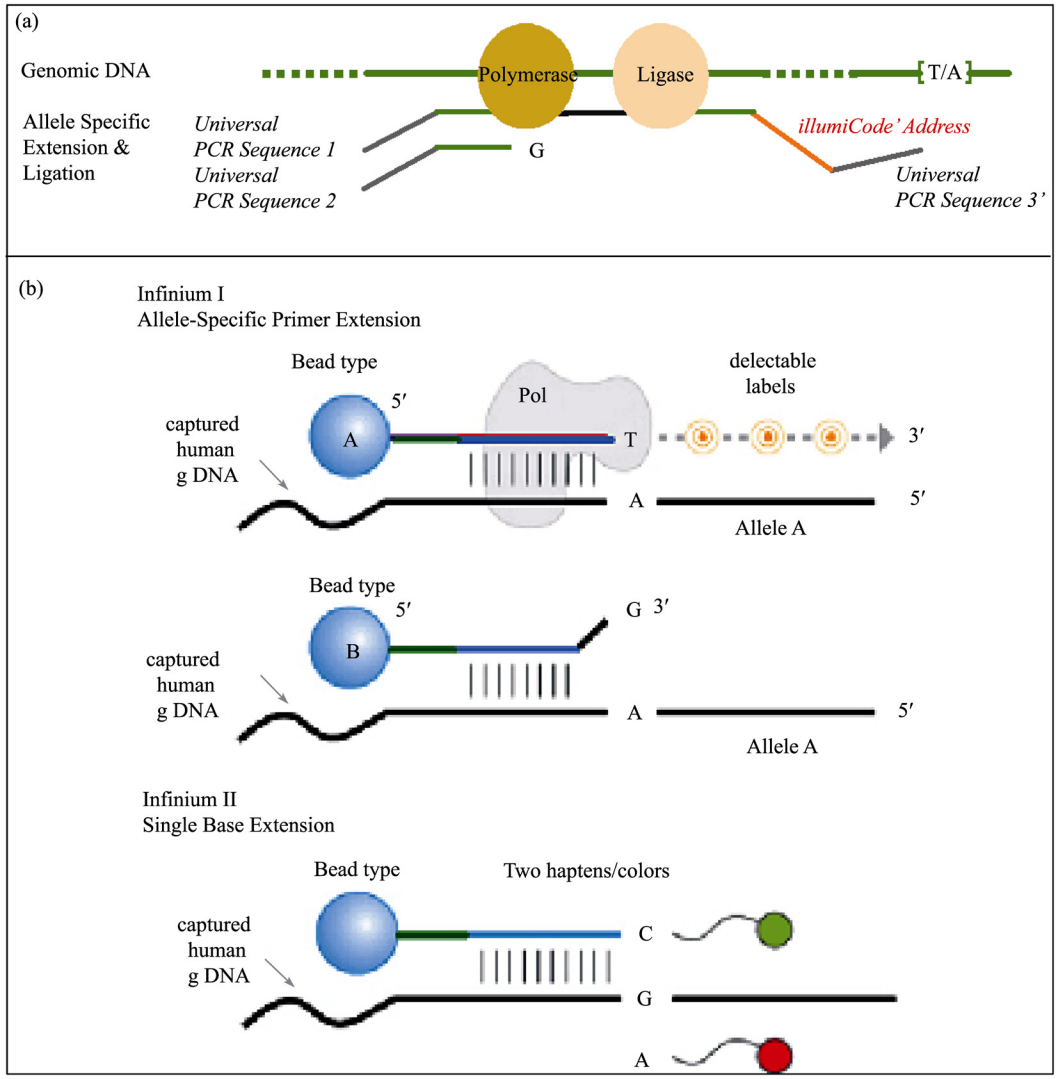

图 1　Illumina 芯片技术原理

(a) GoldenGate 原理图. 该方法的原理是根据 SNP 两侧已知 DNA 序列设计上下游引物, 检测每个 SNP 需要 3 条寡核苷酸引物：两条上游引物, P1 和 P2，覆盖同一位点, 分别代表二态 SNP 中的一种等位基因型; 一条下游引物, P3．这 3 条寡核苷酸都包括与基因组 DNA 互补的区域和与通用 PCR 引物配对的序列. 下游探针还包含一段特异性“标志”序列, 能与 BeadArray 中特定类型珠子上的探针互补结合. 当 3 条寡核苷酸引物与少量的基因组 DNA 样品杂交时, 上游引物 P1 和 P2 的 3 末端碱基与 SNP 位置上碱基互补后, 随即被 DNA 聚合酶有效地延伸到与下游引物 P3 的 5 磷酸基团毗邻, 然后通过连接反应与之连接起来. 这样产生的全长拷贝即成为下一步 PCR 循环扩增的模板. 在第二步的 PCR 扩增中, Cy3 和 Cy5-标记通用上游引物 P1 和 P2, 生物素标记通用下游引物 P3. PCR 扩增之后, 产物中由 P3 延伸的链被带有抗生物素蛋白的磁珠吸附从而去除, 留下荧光标记的单链. 然后荧光标记的单链与微阵列杂交, 通过特异性“标志”序列与 BeadArray 中特定类型珠子上的探针互补结合. (b) Infinium 技术原理. 变性的 DNA 片段与芯片微珠上位点特异的 50 个碱基长度的探针退火复性. 过夜杂交后, 洗去芯片上未杂交的和非特异结合的 DNA, 以捕获到的 DNA 为模板, 掺入特殊标记的核苷酸, 进行碱基延伸反应, 再通过标记物与荧光基团的免疫结合, 将样本的 SNP 位点信息转换成可以用扫描仪检测的荧光信号

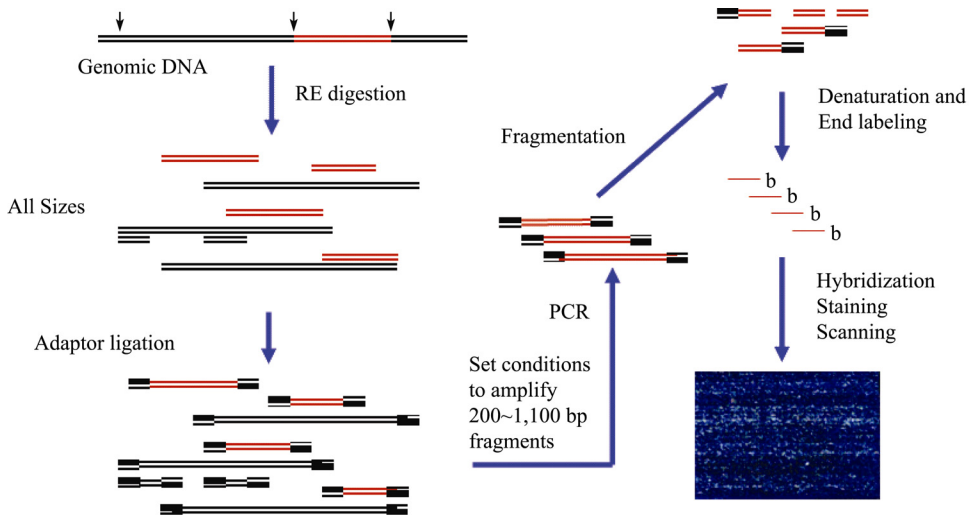

图 2　Affymetrix 芯片技术原理

转自GeneChip® Mapping 500K Assay Manual

2　复杂性状疾病遗传研究的突破和展望

复杂性状疾病是受遗传和环境因素共同作用和/或相互作用所引起. 它的遗传模式往往很难用一种固定的遗传模式来描述, 且具有遗传异质性和表型异质性[21]. 因而, 如何抽丝剥茧, 寻找到真正的致病基因一直是遗传学上的难点和重点.

连锁分析早在 20 世纪 70 年代, 就成功应用于疾病基因的鉴定[22]. 第二代人类遗传图谱的完成, 使得基于家系的全基因组连锁策略在单基因疾病的基因定位方面取得了巨大的成功. 如南方中心参与的恒齿缺失这一孟德尔常染色体显性遗传病的研究, 取得了致病基因的成功定位[23]. 在一个中国人家族性房颤家系的研究中, 南方中心课题组与国内临床合作, 通过全基因组扫描、分型以及连锁分析, 在 11p15.5 上定位了这个房颤家系的致病基因; 进一步在该区域内检测到了致病基因突变位点——一个编码钾离子通道蛋白的基因KCNQ1. 这是在我国成功定位、克隆的迄今为止第一个引起家族性房颤的致病基因[24], 标示了我国科学家在疾病基因组学研究上进入了一个新的高度.

20 世纪 90 年代, 全球遗传学家将全基因组连锁分析广泛应用于一些复杂性状疾病的研究, 例如精神分裂症[25], I 型糖尿病[26] 等. 南方中心遗传学研究课题组利用我国丰富的疾病遗传资源, 在上世纪末本世纪初定位了一批重要疾病, 包括肝癌、鼻咽癌、原发性高血压、II 型糖尿病、甲状腺疾病、皮肤病等的遗传易感基因位点和区域[11,27,28]. 然而, 对于大多数常见疾病而言, 连锁分析所取得的成功非常有限, 对于那些具有较低外显率的等位基因的检测能力不强, 而且一项研究的结果很难在不同人群被其他的研究重复.

因而比连锁分析具有更强检测能力的关联分析应运而生. 关联分析的基本原理是检验一个遗传标志物在病例组的频率是否高于对照组. 如果得到阳性关联的结果, 排除各种混杂之后, 可以推断该标志物存在于疾病易感基因座内或者与易感基因座连锁不平衡.

候选基因的关联分析, 即对有限的高风险候选基因内的遗传标记进行分型, 来分析其与疾病或某种表型之间的关联. 候选基因方法因其经济、省时成为众多研究者的共同选择. 例如, 我们采用候选基因策略, 在脑部神经发育和主要神经生化通路的相关基因中挑选了 112 个基因的 1536 个标签 SNP, 利用 Illumina 的 GoldenGate 技术平台, 在 265 个自闭症核心家系中, 完成了共 1019 个 DNA 样本的基因分型. 数据分析结果发现 7 号染色体存在疾病易感基因. 然而, 采用候选基因关联分析方法最多只能确定遗传

危险因素的部分原因, 一些基因可能与所研究的疾病关联, 却因为功能未知而没有被选为候选基因. 此外, 虽然研究发现了大量候选基因与疾病有关联, 但很难综合评价某个基因对疾病的作用.

随着基因分型技术的发展, 检测SNP的速度更快, 成本更低, 全基因组关联研究(Genome-Wide Association Study, GWAS)已成为可能, 并正在逐渐成为人类复杂性状疾病基因定位研究的重要手段[29]. 本质上, GWAS就是从人类全基因组范围内的序列变异中筛选出那些与疾病性状关联的SNP. GWAS的优势在于分析了基因组上所有基因, 不再受预先设定的候选基因限制, 使得众多功能不明的基因及大量基因间区域的SNP都为疾病的研究提供了线索. 自2005 年*Science*杂志报道了第一项有关年龄相关性(视网膜)黄斑变性GWAS研究以来[30], 截止 2008 年的 9 月 17 日, 已有 180 篇GWAS陆续被报道, 研究内容涉及肥胖、糖尿病、精神病等威胁人类健康的常见疾病[31], 表 1 统计了这些年来发表在*Nature*, *Science*等共 6 份杂志的GWAS文章篇数, 并列举了一些重要的疾病或性状及相应的检测所用平台. 这些研究不但很好地重复发现了过去已证实的关联信号, 而且还产生了很多新的候选基因. 目前, 我国多个研究小组亦采用最新的GWAS策略, 对心血管疾病、代谢性疾病、神经、精神性疾病、皮肤、免疫系统疾病、

表 1　*Science*, *Nature* 等杂志发表的全基因组关联研究文章统计

杂志	年份	论文篇数	疾病或性状	部分论文 PubMedID — 检测平台
Nature	2008	2	Lung cancer/Nicotine dependence	18385738-Illumina/18385739-Illumina
	2007	6	Bipolar disorder/Coronary disease/ Crohn's disease/Hypertension/…	17554300-Affymetrix/…
Science	2008	3	Fasting plasma glucose/Schizophrenia/ Recombination rate (females and females)	18451265-Illumina /18369103-Illumina/ 18239089-Illumina
	2007	8	Coronary disease/Type 2 diabetes/ Body mass index/Exfoliation glaucoma/…	17478681-Perlegen/17463246-Affymetrix/ 17434869-Affymetrix/17690259-Illumina/…
	2006	4	Inflammatory bowel disease/Memory performance/ Wet age-related macular degeneration/ Obesity	17068223-Illumina/17053149-Affymetrix/ 17053108-Affymetrix/16614226-Affymetrix
	2005	1	Age-related macular degeneration	15761122-Affymetrix
Nature Genetics	2008	30	Schizophrenia/Body mass index/Bipolar disorder/ Systemic lupus erythematosus/…	18677311-Affymetrix/18454148-Affymetrix/ 18711365-Affymetrix/18677312-Affymetrix/…
	2007	20	Rheumatoid arthritis/Colorectal cancer/ Prostate cancer/Breast cancer/…	17982456-Affymetrix/17618284-Illumina/ 17603485-Illumina/17529973-Illumina/…
	2006	1	QT interval prolongation	16648850-Affymetrix
PLoS Genet	2008	5	Soluble ICAM-1/Protein quantitative trait loci/ Psoriasis/Schizophrenia/…	18604267-Illumina/18464913-Illumina/ 18369459-Illumina/18282107-Affymetrix/…
	2007	3	Serum urate/Obesity-related traits/Crohn's disease	17997608-Affymetrix/17658951-Affymetrix/ 17447842-Illumina
N Engl J Med	2008	5	Myopathy/Neuroblastoma/Bone mineral density (hip)/ Systemic lupus erythematosus/…	18650507-Illumina/18463370-Illumina/ 18445777-Illumina/18204098-Illumina/…
	2007	5	Rheumatoid arthritis/Amyotrophic lateral sclerosis/ Multiple sclerosis/Restless legs syndrome/…	17804836-Illumina/17671248-Affymetrix and Illumina/ 17660530-Affymetrix/17634447-Illumina/…
Am J Hum Genet	2008	4	Knee osteoarthritis/C-reactive protein/Iris color/ cardiovascular disease	18471798-Illumina/18439552-Illumina/ 18252221-Affymetrix and Illumina/ 18179892-Affymetrix
	2007	3	Skin pigmentation by reflectance spectroscopy/ Type 2 diabetes/Progressive supranuclear palsy	17999355-Perlegen/17668382-Illumina/ 17357082-Affymetrix
	2005	1	Parkinson's disease	16252231-Perlegen

以及重要肿瘤等一直受到广泛关注的复杂性状疾病展开研究[29,32]. 可以预见, 在未来的 5～10 年内, GWAS的研究将引领遗传学、临床医学等各个相关学科的研究, 对生物医学的发展带来新的契机.

2006 年 11 月发布的第一张拷贝数变异图谱, 使我们逐渐认识到拷贝数变异在人类基因组上是普遍存在的. 这种基因变异可以通过破坏基因编码蛋白的活性部分, 改变一个基因的表达"量"[33]; 或者破坏基因组控制基因活性的调节区域, 影响基因活性[7,34,35]. 现在, 关于CNV的检测主要是利用芯片技术. 例如Illumina human 1M芯片上有 206665 个探针位于已报道的CNV区域; Affymetrix Genome-Wide SNP Array 6.0的芯片上有超过946000个用于检测拷贝数变异的探针. 此外, 仅仅利用SNP数据, 也可以计算出可能的CNV, 第一张拷贝数变异图谱就是利用GeneChip Human Mapping 500K early access的数据计算所得[7]. 拷贝数变异的概念, 使我们对复杂性状疾病的认识得到了拓展. 我们开始在疾病中寻找CNV区域, 例如CNV与缺血性脑卒中[36]、CNV与自闭症[37,38]、 CNV与精神病[39]等等.

20 世纪 80 年代遗传图谱的构建, 90 年代人类基因组计划的启动, 本世纪 2002 年人类单倍型图谱计划的启动; 遗传标记从 RFLP, STR, 到 SNP, CNV; 研究策略从连锁分析, 到关联研究, 再到全基因组的关联研究. 科学的发展促进了我们对疾病的了解. 目前正在进行中的癌症基因组计划, 还有蓬勃兴起的各类组学, 如药物基因组学、肿瘤基因组学等等, 引领着我们继续探索有关人类的生、老、病、死以及人类自身与自然环境和谐发展过程中大量的未知.

3 总结

基因分型技术, 是多基因复杂性状疾病的遗传易感性研究的重要技术手段之一. 目前, 在整个课题的研究中, 相对于样品的获得, 候选基因的选择, 数据的分析, 基因分型技术已经成为最容易实现的步骤, 甚至可以让其他拥有技术平台的实验室或科研服务公司去获得数据. 为了能够更好地利用所获得的海量数据, 我们需要加强对实验的设计和数据的分析, 促进疾病研究策略的发展, 从而遴选出真正与疾病相关的基因, 增加我们对疾病的认知.

参考文献

1 Donis-Keller H, Green P, Helms C, et al. A genetic linkage map of the human genome. Cell, 1987, 51(2): 319—337[DOI]

2 Weissenbach J, Gyapay G, Dib C, et al. A second-generation linkage map of the human genome. Nature, 1992, 359(6398): 794—801[DOI]

3 Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science, 1996, 273(5281): 1516—1517[DOI]

4 The International HapMap Consortium. A haplotype map of the human genome. Nature, 2005, 437(7063): 1299—1320

5 The International HapMap Consortium, Frazer K A, Ballinger D G, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature, 2007, 449(7164): 851—861

6 Rocha D, Gut I, Jeffereys A J, et al. Seventh international meeting on single nucleotide polymorphism and complex genome analysis: ever bigger scans and an increasingly variable genome. Hum Genet, 2006, 119(4): 451—456[DOI]

7 Redon R, Ishikawa S, Fitch K R, et al. Global variation in copy number in the human genome. Nature, 2006, 444(7118): 444—454[DOI]

8 Freeman J L, Perry G H, Feuk L, et al. Copy number variation: new insights in genome diversity. Genome Res, 2006, 16(8): 949—961[DOI]

9 Hayashi K. PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. PCR Methods Appl, 1991. 1(1): 34—38

10 Fodde R, Losekoot M. Mutation detection by denaturing gradient gel electrophoresis (DGGE). Hum Mutat, 1994, 3(2): 83—94[DOI]

11 Feng B J, Huang W, Shugart Y Y, et al. Genome-wide scan for familial nasopharyngeal carcinoma reveals evidence of linkage to chromosome 4. Nat Genet, 2002, 31(4): 395—399

12 Kim S, Misra A. SNP genotyping: technologies and biomedical applications. Annu Rev Biomed Eng, 2007, 9: 289—320[DOI]

13 Oliphant A, Barker D L, Stuelpnagel J R, et al. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. Biotechniques, 2002, Jun, Suppl: 56—58, 60—61

14 Shen R, Fan J B,Campbell D, et al. High-throughput SNP genotyping on universal bead arrays. Mutat Res, 2005, 573(1-2): 70—82
15 Matsuzaki H, Dong S, Loi H, et al. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods, 2004, 1(2): 109—111[DOI]
16 Wang W P, Ni K Y. Zhou, G H. Approaches for SNP genotyping. Yi Chuan, 2006, 28(1): 117—126
17 Holland P M, Abramson R D, Watson R, et al., Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. Proc Natl Acad Sci U S A, 1991, 88(16): 7276—7280[DOI]
18 Bell, P.A., et al., SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. Biotechniques, 2002. Jun, Suppl: 70—72, 74, 76—77
19 Braun A, Little D P, Köster H. Detecting CFTR gene mutations by using primer oligo base extension and mass spectrometry. Clin Chem, 1997, 43(7): 1151—1158
20 Jurinke C, Van den Boom D, Cantor C R, et al. The use of MassARRAY technology for high throughput genotyping. Adv Biochem Eng Biotechnol, 2002, 77: 57—74
21 Yan W L, Gu D F. Issues on association studies on complex diseases. Yi Chuan Xue Bao, 2004, 31(5): 533—537
22 Kan Y W, Dozy A M. Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. Proc Natl Acad Sci USA, 1978, 75(11): 5631—5635[DOI]
23 Liu W, Wang H, Zhao S, et al. The novel gene locus for agenesis of permanent teeth (He-Zhao deficiency) maps to chromosome 10q11.2. J Dent Res, 2001, 80(8): 1716—1720
24 Chen Y H, Xu S J, Bendahhou S, et al. KCNQ1 gain-of-function mutation in familial atrial fibrillation. Science, 2003. 299(5604): 251—254[DOI]
25 Stefansson H, Sigurdsson E, Steinthorsdottir V, et al. Neuregulin 1 and susceptibility to schizophrenia. Am J Hum Genet, 2002, 71(4): 877—892[DOI]
26 Nisticò L, Buzzetti R, Pritchard L E, et al. The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Belgian Diabetes Registry. Hum Mol Genet, 1996, 5(7): 1075—1080[DOI]
27 Zhang X J, He P P, Wang Z X, et al. Evidence for a major psoriasis susceptibility locus at 6p21(PSORS1) and a novel candidate region at 4q31 by genome-wide scan in Chinese hans. J Invest Dermatol, 2002, 119(6): 1361—1366[DOI]
28 Jin Y, Teng W, Ben S, et al. Genome-wide scan of Graves' disease: evidence for linkage on chromosome 5q31 in Chinese Han pedigrees. J Clin Endocrinol Metab, 2003, 88(4): 1798—1803[DOI]
29 The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature, 2007, 447(7145): 661—678
30 Klein R J, Zeiss C, Chew E Y, et al. Complement factor H polymorphism in age-related macular degeneration. Science, 2005, 308(5720): 385—389[DOI]
31 Hindorff L A, Junkins H A, Manolio T A. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/26525384. Accessed 2008/9/18
32 Samani N J, Erdmann J, Hall A S, et al. Genomewide association analysis of coronary artery disease. N Engl J Med, 2007, 357(5): 443—453[DOI]
33 Jakobsson M, Scholz S W, Scheet P, et al. Genotype, haplotype and copy-number variation in worldwide human populations. Nature, 2008, 451(7181): 998—1003[DOI]
34 Perry G H, Ben-Dor A, Tsalenko A, et al. The fine-scale and complex architecture of human copy-number variation. Am J Hum Genet, 2008, 82(3): 685—695[DOI]
35 Stranger B E, Forrest M S, Dunning M, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science, 2007, 315(5813): 848—853[DOI]
36 Matarin M, Simon-Sanchez J, Fung H C, et al. Structural genomic variation in ischemic stroke. Neurogenetics, 2008, 9(2): 101—108[DOI]
37 Autism Genome Project Consortium, Szatmari P, Paterson A D, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. Nat Genet, 2007, 39(3): 319—328
38 Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. Science, 2007, 316(5823): 445—449[DOI]
39 Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. Nature, 2008, 455(7210): 232—236[DOI]