

DNA 限制性片段长度多态性

党进军 孙志贤

(军事医学科学院放射医学研究所, 北京)

限制性内切酶的发现和应用, 给生命科学研究带来了深刻变化。通过测定 DNA 分子中核苷酸排列顺序, 研究基因的结构与功能, 从而揭示生命现象的本质, 是一个具有十分重要意义的课题。其中利用 DNA 重组和分子探针技术在基因水平上进行限制性酶切图谱分析, 确定 DNA 结构中具有多态性的核苷酸位点, 进行基因定位和异常基因的检测, 是近年来分子遗传学研究中颇受重视的一个新领域。Kan 和 Dozy^[1]1978 年首先在 β -球蛋白基因中发现第一个 DNA 多态性位点以来, 人们已将注意力从基因表达产物——酶和蛋白质水平的多态性研究转移到 DNA 分子水平。随着这方面研究的普遍开展, 证明 DNA 多态性研究为基因分析提供了一个新的方法, 这不仅具有重要的理论意义, 而且对遗传病的早期诊断和预防也具有重要的实际意义。

一、DNA 多态性的分子基础

DNA 多态性是指染色体 DNA 等位基因中核苷酸排列顺序的差异性。在生物进化过程中, 由于种种原因引起的基因突变和 DNA 分子结构重排, 造成了 DNA 分子内核苷酸排列顺序的改变, 当这种改变影响到基因(编码基因或非编码基因)中的限制性内切酶识别位点时, 利用限制性酶切图谱分析, 就可以检测出这种存在于群体之内的 DNA 限制性片段长度多态性(Restriction Fragment Length Polymorphisms, RFLPs)。

1. 基因突变产生 DNA 多态性

根据基因突变的结果, 可分为三类: 中性

突变、有利突变和不利突变^[2]。突变引起限制酶识别位点的移动、产生或消失, 通常, 基因结构中某些核苷酸位点可以发生中性突变, 其结果并不影响基因的表达, 而往往只是作为一种遗传标记而已。有利突变促进机体对环境的适应和生物进化, 不利突变则是遗传性疾病的来源。这种存在于正常基因与异常基因之间的差异性, 正是我们利用特定基因的多态性特征, 进行遗传病基因检测的基础。

基因突变的发生具有多向性, 同一基因可以发生多次突变形成复等位基因, 如 β -球蛋白 β^A 、 β^S 和 β^C 三个等位基因^[3]。由于基因突变的发生是随机的, 所以整个人基因组中任何位点均有可能发生突变而产生 DNA 变异体。

2. DNA 多态性的类型

DNA 多态性根据其产生的方式基本上有两种: 一是碱基对突变类型, 即基因点突变(碱基对的转换或颠换)造成限制酶识别位点产生或消失, 构成 DNA 多态性, 叫做碱基取代多态性。另一种为 DNA 分子结构重排型, 由于重排造成部分 DNA 片段的缺失、插入或倒位引起识别位点的移动, 结果形成染色体重排多态性^[4]。

以上两种类型以碱基取代多态性为多见^[5]。

二、DNA 多态性的检测

DNA 多态性检测一般分为三个步骤:

1) 样品 DNA 制备和限制性酶切片段的分离及 Southern^[6]转移。2) 利用 DNA 重组和分子克隆技术制备标记探针^[7]。3) 通过分子杂交

鉴定载有特定基因的 DNA 片段的多态性。用于基因定位的探针应是单拷贝基因,同时 DNA 片段不宜过长以避免含有重复顺序 DNA,导致结果混乱^[8]。

关于变异基因的检测大致可分为三种^[3]: 1)间接测定法。根据基因突变位点与 RFLPs 位点的连锁进行检测。在不完全连锁情况下,可通过家系研究确定在多态性限制片段中基因分离的图谱,但这常常因需要第二个 RFLP 位点而受到限制。2)直接测定法。由于点突变影响到限制酶识别位点,就可以用对该位点专一性的限制酶直接检测。但因点突变直接发生在识别位点上的频率很小,此方法也受到一定的限制。3)寡核苷酸探针测定法。在已知基因突变位点碱基顺序时,可人工合成寡核苷酸探针准确地检出突变基因。

三、DNA 多态性研究的应用

1982 年第六次国际人类基因制图会议上报告已发现 25 个 RFLPs 位点^[4],到 1984 年第

七次会议时增加到 159 个^[9],最近见报道的有 200 多个 RFLPs 被检测^[10]。其中有些 RFLPs 作为遗传标记提供 DNA 结构信息已经显示出重要的应用价值。目前, DNA 多态性研究主要应用于基因突变分析、基因定位和遗传病基因的早期检测等方面。

1. 基因定位

家系研究证明, RFLPs 的遗传方式符合孟德尔遗传规律^[11,12]。RFLPs 构成了一类新的遗传标记,为连锁研究提供了新的实验手段。根据 RFLPs 位点与特定基因的密切连锁,可以进行基因定位的研究。Fearon 等人^[13]通过 RFLPs 分析,研究确定了位于人 11 号染色体短臂上几个已知基因位点的线性排列顺序和重组距离,证明 c-Ha-ras-1 基因与球蛋白和胰岛素基因密切连锁,并估算出它们在染色体 11p 上的排列顺序为:着丝粒-甲状旁腺激素- β -球蛋白-c-Ha-ras-1-胰岛素(见图 1)。图中各位点之间的距离以 1cm 表示,1cm 等于 0.01 θ 的遗传距离。

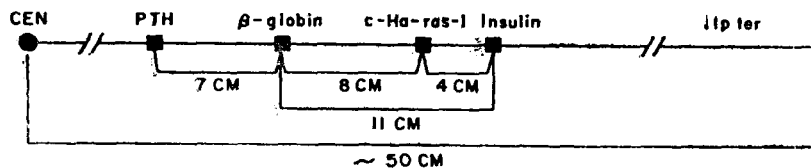


图 1 人染色体 11p 的连锁图谱^[13]

DNA 多态性位点普遍存在于整个人基因组中。Botstein 认为^[14],如果建立一套以 20 分摩 (centimorgan) 间隔平均分布于整个基因组的 RFLPs 位点,并找出这些遗传标记与其它遗传特征的连锁关系,通过 RFLPs 在人染色体 DNA 中的精确定位,就有可能绘制出完整的人类基因连锁图谱。由于缺乏大量合适的 RFLPs 位点,这一工作进展缓慢,但人们从重组 DNA 文库中抽提出任意 DNA 片段来检测 DNA 多态性,可以大大扩展 RFLPs 的数目,随着这些任意 DNA 多态性位点在基因组中定位,必将促进基因连锁分析。目前许多实验室正在共同努力组建一个以 20 分摩间隔分布于整个人基

因组中的 RFLPs 位点基因定位图谱^[15],以便进一步绘制出完整的人类基因图。

Bell 等人^[16]用限制酶 BglI 对人胰岛素基因分析,在侧翼序列的非编码基因中检测到三种 BglI 酶切多态性片段(图 2)。由于 RFLPs 不仅存在于编码基因中,也存在于非编码基因中,这大大扩展了遗传标记的范围。人染色体中含有大量重复顺序 DNA,它们分散在单拷贝或低拷贝基因之间。因此,人们也试图利用重复顺序 DNA 作为探针进行基因定位的研究。Gusella 等人^[17]利用 11 号染色体上重复 DNA 探针确定出位于该染色体上 19 个 DNA 片段及 5 个基因位点的顺序。重复顺序 DNA 不涉及结构

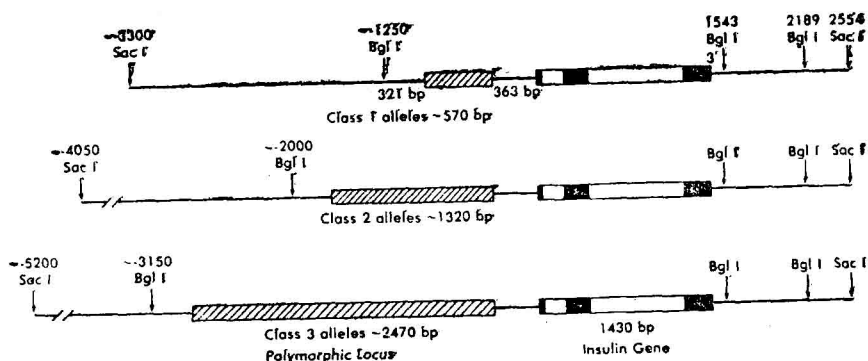


图2 人胰岛素基因侧翼序列的 DNA 多态性图谱^[16]

基因,并间隔分散于各结构基因之间,所以该方法有一定的应用价值,能够进行基因组结构测定以及研究由于 DNA 分子结构的重排、缺失和插入等产生的 DNA 多态性。

2. 基因突变的研究

通过 DNA 多态性分析,人们对基因突变的性质及规律有进一步的了解。Barker^[18]和 Cooper 等人^[5]利用多种限制酶作了大量的 DNA 多态性调查,均发现较多的 RFLPs 与含有 CpG 二核苷酸的限制酶识别位点有关,即这类识别位点产生多态性的频率较高。这是由于人基因组内的 CpG 二核苷酸中胞嘧啶碱基发生高度甲基化所致,胞嘧啶常因甲基化引起脱氨最后被胸腺嘧啶所取代。由于这种 C → T 转变的高频率造成了人基因组 DNA 结构中缺乏 CpG,所以识别位点中含有 CpG 的限制酶 MspI (CCGG) 和 TaqI (TCGA) 检出 DNA 多态性的频率就高于其它限制酶。作者认为,甲基化的胞嘧啶是哺乳动物 DNA 中发生突变的热点。

研究证明,基因突变发生在 DNA 结构中的部位不同,其突变造成的影响也不同。Antonarakis^[19]在确定 β-球蛋白基因簇多态性的基础上,观察了患有 β-地贫的美国黑人中 β-地贫基因的突变位点,发现临床症状较轻的病人其点突变发生在接近于启动子的高度保守区-29 碱基处,出现了单个碱基取代 (A → G)。而另一种临床症状较严重的病人,是在 RNA 拼接部位发生了单个碱基取代 (A → G),突变结果阻

止了 RNA 的正常加工。从而使此病患者临床症状程度轻重不同得到解释。

关于 DNA 结构中变异位点的发生频率,其说不一,一般认为每 100—500 碱基对中有一个变异位点^[20]。总之,通过研究已经证明发生在 DNA 结构中的基因突变是一种普遍的现象,它反映出人类是在不断的变异中发展这一进化规律。

最近,Zarbl 等人^[21]利用化学致癌剂 NMU (N-nitroso-N-methylurea) 诱发大鼠乳腺癌,通过 DNA 多态性分析证明是 NMU 直接激活 Ha-ras-1 癌基因所致。已经证明, NMU 专一

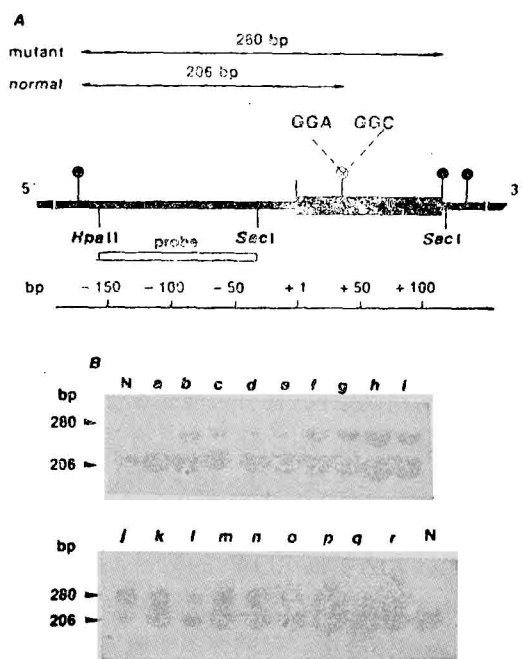


图3

性地引起 DNA 结构中 G → A 转变^[22]。Ha-ras-1 癌基因的激活是位于第 12 个密码子内第二个核苷酸发生了 G → A 转变,该突变破坏了限制酶 MnlI 的专一识别顺序 GAGG(图 3A 中 ⊗ 所示, GGA 和 GGC 分别为第 12 和第 13 个密码子。● 表示 MnlI 识别位点),结果产生了不同于正常型 (206 bp) 的突变型 DNA 多态性片段 (280bp)。图 3B 为用 MnlI RFLPs 检测 NMU 诱发乳腺癌的 DNA 酶切图谱。这些对于研究人类癌症的发病机理也有着重要意义。

3. 遗传病基因的检测

所有遗传病都涉及基因组 DNA 遗传特征的改变,所以原则上任何遗传病(或遗传表型)

都可以在基因水平上进行分析鉴定。DNA 重组技术提供了分离与染色体上任何基因位点相关的 DNA 片段的可能性^[23],因此重要的是找出与异常基因紧密连锁的特异性探针并确定 RFLPs 的位点。通过用 DNA 探针得到限制酶切图谱,然后将正常基因与异常基因加以比较,就能直接或间接分析异常基因中碱基顺序的改变,作出基因诊断。目前,关于利用 DNA 多态性检测遗传病基因已有不少报道(见表 1 和表 2)^[2]。其中报道最多并且最成功的例子是球蛋白基因突变的血液学诊断,临床上已经建立了一整套有效的产前诊断方法^[24]。在我国也已开展了这方面的应用研究^[25]。

表 1 利用 DNA 重组技术直接分析的部分遗传病

遗 传 病	基 因 探 针	作 者
抗凝血酶 III 缺乏症	抗凝血酶 III	Prochownick et al. (1983)
α-抗胰酶缺乏症	人工合成的寡核苷酸	Kidd et al. (1983)
糖尿病	胰岛素	Haneda, et al. (1983)
生长激素缺乏症	生长激素	Phillips, et al. (1981)
血友病 B	凝血因子 IX	Giannelli, et al. (1983)
成视网膜细胞瘤	染色体 13 中 DNA 片段	Cavence, et al. (1983)
成骨不全	pro α1(1) 胶原	Chu, et al. (1983)
次黄嘌呤鸟嘌呤转磷酸	HPRT	Wilson, et al. (1983)
核糖基酶 (HPRT) 缺乏症		
镰型细胞贫血	β-球蛋白	Geever, et al. (1981)
	人工合成寡核苷酸	Conner, et al. (1983)
地中海贫血	α-和 β-球蛋白	Little, et al. (1980)

表 2 利用 DNA 重组技术间接分析的部分遗传病

遗 传 病	基 因 探 针	作 者
糖尿病	胰岛素	Rotwein, et al. (1983)
成骨不全	pro α2(1) 胶原	Tsipouras, et al. (1983)
苯酮尿症	苯丙氨酸羟化酶	Woo, et al. (1983)
生长激素缺乏症(I 型)	生长激素	Phillips, et al. (1982)
地中海贫血	β-球蛋白	Boehm, et al. (1983)
镰型细胞贫血	β-球蛋白	Boehm, et al. (1983)
视网膜层分裂	λRC8*	Wieacker, et al. (1983)
萎缩性肌强直	补体 C ₃ 基因	Davies, et al. (1983)

* 克隆的 DNA 片段。

在利用 DNA 多态性分析进行产前诊断的过程中,使用多 RFLPs 位点检测可以提高风险估计的准确度。Dryia 等人^[26]利用特异探针检测了 13 号染色体上与成视网膜细胞瘤基因位点

连锁的 RFLPs。通过计算染色体 13q12—13q22 之间多态性位点与位于 13q14 的成视网膜细胞瘤基因位点之间的重组距离,他们认为利用任何存在于 13q12—13q22 之间的三个特异性探

针,预测家系内成视网膜细胞瘤基因携带者的准确度可达77%以上。

另外, Barker 等人^[27]还报道了存在于 onco 基因位点的 DNA 多态性。Dracopoli 等^[28]利用 DNA 多态性分析对恶性黑色素瘤基因进行了研究,认为黑色素瘤与多态性限制片段丢失有关。因此,肿瘤细胞内各种染色体异常(如姐妹染色体不等交换、畸形有丝分裂和重组等)均有可能通过 DNA 多态性分析进行检测,这无疑扩大了 DNA 多态性研究的应用范围。

四、结 语

DNA 重组和分子克隆技术的应用已使人类遗传学研究从间接地研究遗传表型转移到直接研究基因的结构、顺序、表达和遗传特征上来。DNA 多态性研究是一种基因分析的新方法,由于 RFLPs 普遍存在于整个人基因组中,使人们有可能利用这种新的遗传标记绘制出完整的人类基因组图谱。此外, DNA 多态性研究是鉴定 DNA 分子损伤、研究基因突变的有效手段。在临床应用中,研究与 RFLPs 紧密连锁的遗传病基因位点的突变,可为遗传病的早期诊断提供可靠的证据,这对于优生学有着重要意义。

目前, DNA 多态性研究虽然还受到某种程度上的限制,如 RFLPs 的数目和制备特异性探针等,但仍在不断取得进展^[29-32]。最近, Miller 等人^[33]还在酵母线粒体 DNA 中发现了位于 tRNA 合成基因位点的多态性。随着这方面研究的深入和对人类基因组结构的逐步了解,必将会对这些人类基因组中 DNA 多态性位点的起源、保持以及功能等问题作深入的探讨,从而有助于揭示生命现象本质。

参 考 文 献

- [1] Kan, Y. W. et al.: *Proc. Natl. Acad. Sci. USA*, 75, 5631, 1978.
- [2] 杜传书、刘祖洞:《医学遗传学》,人民卫生出版社, p. 153, 1983 年。
- [3] Studencki, A. B. et al.: *Am. J. Hum. Genet.*,

- 37, 42, 1985.
- [4] Skolnick, M. H. et al.: *Cytogenet. Cell Genet.*, 32, 58, 1982.
- [5] Cooper, D. N. et al.: *Hum. Genet.*, 66, 1, 1984.
- [6] Southern, E. M.: *J. Mol. Biol.*, 98, 503, 1975.
- [7] Rigby, P. W. J.: *J. Mol. Biol.*, 113, 237, 1977.
- [8] Wyman, A. R. et al.: *Proc. Natl. Acad. Sci. USA*, 77, 6754, 1980.
- [9] Skolnick, M. H. et al.: *Cytogenet. Cell Genet.*, 37, 210, 1984.
- [10] Clark, A. C. et al.: *Am. J. Hum. Genet.*, 37, 60, 1985.
- [11] David, C. et al.: *Am. J. Hum. Genet.*, 36, 565, 1984.
- [12] Varshney, U. et al.: *Mol. Biol. Med.*, 2, 193, 1984.
- [13] Fearon, E. R. et al.: *Am. J. Hum. Genet.*, 36, 329, 1984.
- [14] Botstein, D. et al.: *Am. J. Hum. Genet.*, 32, 314, 1980.
- [15] Naylor, S. L.: *Proc. Natl. Acad. Sci. USA*, 81, 2447, 1984.
- [16] Bell, G. I. et al.: "Banbury Report 14 recombinant DNA applications to human disease", USA, 319, 1983.
- [17] Gusella, J. F. et al.: *Proc. Natl. Acad. Sci. USA*, 79, 7804, 1982.
- [18] Barker, D. et al.: *Cell*, 36, 131, 1984.
- [19] Antonarakis, S. E. et al.: *Proc. Natl. Acad. Sci. USA*, 81, 1154-58, 1984.
- [20] Neel, J. V. *Am. J. Hum. Genet.*, 36, 1135, 1984.
- [21] Zarbl, H. et al.: *Nature*, 315(6018), 382, 1985.
- [22] Pegg, A. E.: *Cancer Invest.*, 2, 223, 1984.
- [23] Davies, K. E. et al.: *Hum. Genet.*, 58, 351, 1981.
- [24] Beutner, E.: *Clinical Laboratory Assays*, p. 231, 1980.
- [25] 吴冠芸等:《第五次全国生物化学学术会议论文摘要汇编》, 158, 1984 年。
- [26] Dryia, T. P. et al.: *Am. J. Hum. Genet.*, 65, 320-324, 1984.
- [27] Barker, D. et al.: *Am. J. Hum. Genet.*, 34, 67A, 1982.
- [27] Dracopoli, N. C. et al.: *Proc. Natl. Acad. Sci. USA*, 82, 1470, 1985.
- [29] Goodbourn, S. E. Y. et al.: *Mol. Biol. Med.*, 2, 223, 1984.
- [30] Bock, S. C. et al.: *Am. J. Hum. Genet.*, 37, 32, 1985.
- [31] Welker, D. L. et al.: *Mol. Cell Biol.*, 5, 273, 1985.
- [32] Murphy, P. D. et al.: *Nucleic Acids Res.*, 13, 3015, 1985.
- [33] Miller, D. L. et al.: *Nucleic Acids Res.*, 13, 859, 1985.

[本文于 1985 年 10 月 10 日收到]