

载脂蛋白基因克隆技术的研究进展

吕新跃 暴晓梅

(武警医学院, 天津 300162)

主题词 载脂蛋白; 基因; 克隆; 方法; 动脉粥样硬化

摘要 动脉粥样硬化的形成受遗传和环境多因素的影响, 某些载脂蛋白的基因或基因变异与动脉粥样硬化的发病有关。近年来有关基因克隆技术及其在克隆载脂蛋白基因方面的应用进展迅速。这些技术基本可分为三大类, 即功能克隆、定位克隆和候选克隆。

动脉粥样硬化(atherosclerosis, As)作为一种多基因病(即一切遗传因素的影响不为零的疾病), 其发病受遗传与环境多因素的相互作用。病理改变的一个显著特征是血浆脂质代谢失调。其中个别患者的发病与某一个或少数基因的遗传变异直接相关, 如低密度脂蛋白(low density lipoprotein, LDL)受体基因突变所致的家族性高胆固醇血症, 但多数患者的致病效应还是由多个基因及不良环境因素共同作用累积叠加而逐步形成的。外因是通过内因而起作用, 任何环境因素都是通过一个或多个靶基因的结构改变或改变靶基因的表达调控, 或是改变数个相关靶基因之间的相互关系而起作用。因此, 克隆及研究载脂蛋白基因、或其突变基因及影响动脉粥样硬化发生的其它相关基因的研究工作已经成为脂蛋白及动脉粥样硬化分子生物学研究的核心内容。本文就近年来有关基因克隆的新技术策略及应用作一简介。

迄今, 已成功地克隆分离出人类约一千多个基因, 其中200多个基因与人类疾病有关。已克隆了几十个载脂蛋白、脂蛋白受体、影响脂代谢酶类的基因。所采用的基因克隆技术可大致分为三类: 功能克隆(functional cloning)、定位克隆(positional cloning)与候选克隆(candidate cloning)^[1]。

1 功能克隆

人的各种性状(含疾病)都与蛋白质的功能有关, 功能克隆是从基因的蛋白质表达与功能着手来分离基因。这又包括:

1.1 抗体筛选

将纯化的蛋白制备出抗体, 通过免疫反应把在多聚糖体

上正在合成的相应肽链连同其模板 mRNA(或 cDNA)一起免疫沉淀下来, 这可获得该基因编码序列(或转录文本)。多数的人类载脂蛋白基因基本上是通过这一方法克隆的。最近, 我们也利用这一技术, 从构建的表达型 cDNA 文库中克隆到不易感 As 动物北京鸭、树鼩载脂蛋白 AI 及树鼩载脂蛋白 CI 三个 cDNA 的完整序列^[2,3], 三个新基因均被 GenBank 接收。

1.2 通过部分氨基酸测序

将获得的氨基酸顺序反推出一个可能的兼并核苷酸顺序为探针, 来筛选 cDNA 及基因组文库 DNA, 以获得相应的基因。如人载脂蛋白 AI cDNA 的克隆。

1.3 比较 mRNA 的克隆策略

有人把此类也归属于功能克隆。其根据是蛋白质功能的差异, 首先是反映在组织中 mRNA 表达的差异。这一克隆技术发展较快, 已形成了多种克隆方法, 并被广泛应用。

1.3.1 消减杂交(subtractive hybridization)筛选^[4,5] 这一方法又可分为三种类型: 即 DNA-DNA 杂交、cDNA-mRNA 杂交和 cDNA-cDNA 杂交。后两种杂交类型可以获得在某一组织中缺失或特异表达的消减 cDNA 文库(subtractive cDNA library), 它常用于克隆两种组织或同一组织在不同生理和病理状态表达有差异的基因。由于减数文库富集了仅在一种组织或在一种状态下表达的 cDNA 克隆, 经筛选的克隆总数减少了约数十倍。这一方法要求一个细胞系(cDNA 文库)具有特异的目的 DNA 区域(检测 DNA), 而另一细胞系(另一 cDNA 文库)拥有除特异目的 DNA 区域以外的上一细胞系其他所有的 DNA(驱赶 DNA)。将含有目的 DNA 的细胞标记制备 cDNA 单链与另一细胞系标有生物素的 DNA 变性

后杂交,使检测 DNA 中与驱赶 DNA 相同的大部分 DNA 与驱赶 DNA 复性成双链,这些标有生物素的杂交链通过卵清蛋白结合,或通过磁珠减数等技术去除杂交的双链,从而检测 DNA 中特异的目的 DNA 被游离出来,游离的 cDNA 再次重复上一次筛选后,可测序或再以此为探针,筛选 cDNA 文库,以获得完整的 cDNA 序列。cDNA 消减杂交对检测有转录活性的基因很灵敏,利用该法可很快同时对具有表型不正常或缺失的基因片段进行检测。

1.3.2 代表性差别筛选分析(representational difference analysis, RDA)^[6] 这是在 cDNA 消减杂交筛选的基础上发展起来的一种技术。该法首先用一种限制性内切酶切割 cDNA,将他们与接头连接后,PCR 进行扩增。因较小的 DNA 片段易于扩增,故 PCR 较易获得进行杂交的较小 DNA 片段(即扩增子)。在扩增了“检测”(tester)与“驱动”(driver)扩增子后,酶切除接头,只在检测扩增子 5' 端接上新的接头,再将检测扩增子与过量驱动子混合杂交复性,这样只有检测扩增子中特有的 DNA 片段才会复性形成两端均有接头的双链,而其他 DNA 片段或只一端有接头、或两端均无接头。用单链杂交酶消化掉未复性的单链,然后依接头设计引物进行 PCR,即可获得仅在检测扩增子中存在的特异 DNA 片段。

RDA 省去了 cDNA 消减杂交筛选中物理分离等步骤,同时它仅需一部分代表性小片段 DNA 分子为复性模板,降低了复性模板的复杂度。必要时还可用数种酶处理 DNA 的一系列扩增子,可弥补可能出现的信息丢失。

1.3.3 mRNA 差异显示(mRNA differential display, mRNA DD)^[7,8] 高等生物约有 10 万个不同的基因,但仅有 15% 的基因在任何个体细胞中均表达(看家基因)。因此,在不同类型组织细胞、不同分化时期及疾病状态时,特异表达的基因的克隆亦显得十分重要。譬如,不同载脂蛋白基因的表达具有明显的组织特异性和种属特异性,同时也具有胚胎及不同发育时期表达的时空性,这一技术的出现为研究基因表达的阶段性、及某一特定环境或疾病状态下基因差异表达调控及分离提供了重要的克隆手段。其原理是提取不同个体、不同组织或细胞来源的 mRNA,或提取相同类型的组织细胞在发育的不同时期或经过不同处理后的 mRNA,经反转录合成 3' 端 cDNA,以此为模板经 PCR 扩增,得到能在测序胶上分辨的差异条带,将这些特异的条带再经过克隆及 RNA 印迹验证后,测序并以此为探针,筛选 cDNA 文库,最终获得完整的差异表达序列,并进行 GenBank 检索分析。差异显示自建立后,方法上也做了许多改进,由最初的均由十个碱基组成的 3' 端 12 条锚定引物和 5' 端 20 条引物,减化到 1996 年 3' 端仅需 3 条引物,5' 端带上酶切位点(HIND Ⅲ)的 8 条引物,这 24 种组合即可覆盖几乎所有的 mRNA。同时 PCR 扩增条件、核苷酸的标记及分辨率均得到了改善和提高,特别是优化了引物设计,提高了分离差异表达新基因的可能性。

功能克隆的特点是不依赖基因组信息。其缺点是必须对相应的蛋白质有一定的了解。但功能克隆技术成熟、方法直接,费用相对较低,迄今还是基因克隆的首选策略。

2 定位克隆^[9]

随着人类基因组计划的实施,直接从染色体的某一特定位置点或区域克隆某一基因,称为“定位克隆”(positional cloning)。基因的定位克隆并不需要事先获得已纯化的蛋白质,但必须明确待克隆的基因在染色体上的位置及相应的 DNA 标志。因而,定位克隆的策略分为基因定位、基因组分析和 cDNA 鉴定等三个步骤。自 1986 年首次用定位克隆技术分离第一个基因——X 连锁慢性肉芽肿病基因到最近,近百种遗传病基因已被克隆,获得了极大的成功。

2.1 基因定位

主要根据 连锁分析,④易确定的某一(疾病)性状可能与某一染色体特定位置的缺失或易位的关系来定位。并利用 RFLP、微卫星重复序列等 DNA 标志将基因的位置压缩至 2 000 kb 以内。

2.2 基因组分析

对选出可能存在待分离基因的染色体区段的 DNA 克隆(可向 GDB 等购买或索取),如酵母人工染色体(YAC)、P1(P1 人工染色体)或 Cosmid,将这些克隆按顺序排列,进一步分离测序。因为即便基因定位在 0.5 cm 以内(约 500 kb),或 1 个 YAC 克隆,仍可能面对 10 个以上的基因。

2.3 编码序列(cDNA)的筛选

2.3.1 外显子捕获(exon trapping) 含有内含子的基因在转录后必需剪接掉内含子才会得到成熟的 mRNA,依据剪切特有的外显子与内含子的边界识别序列,Duky 等构建了专用载体,在基因组 DNA 中(YAC 等)若含有外显子,则会被特定的克隆出来,将其转录到真核细胞中进行转录并剪接,产生载体衍生的 mRNA,再通过 RT-PCR,检测被捕获的外显子。当然该策略不适用于不含内含子的基因。

2.3.2 种属间序列同源性(Zoo-blot)^[10] 基于不同种属的编码序列同源性大大高于非编码区,用另一种属的保守编码序列标记后与待查种属的染色体克隆杂交,筛选相应的基因。最近,Allan 等人用小鼠载脂蛋白 E/C1 相关基因与含有人 19 号染色体载脂蛋白 E/C1/C2 区段的 P1 克隆杂交获得了一个人的新基因——载脂蛋白 C 基因^[16]。

2.3.3 CpG 岛筛选法 CpG 双核苷酸广泛存在于人类基因组 DNA 中,并常与表达顺序(所有的看家基因及 40% 的组织特异性基因)5' 端相连,可通过稀切酶(Eag Ⅳ、Sac Ⅲ、BssH Ⅲ)酶切来检测^[11]。将连接 CpG 岛附近的 DNA 片段克隆测序进行分析,可分离出基因的编码区。缺点是并非所有的基因都与 CpG 岛连锁。

2.3.4 cDNA 文库的筛选 用 YAC 或 Cosmid 筛选 cDNA 是目前在染色体小区段内寻找基因的重要方法之一。包括两种常用方法:(a)噬菌斑杂交法:YAC DNA 经脉冲电泳纯化后,以 Alu-PCR 方法进行标记,对含有 cDNA 文库的尼龙膜杂交。(b)直接筛选法(Direct selection):YAC DNA 纯化后固定于小片尼龙膜上,与 cDNA 杂交,特异的 cDNA 被洗膜后,进行 PCR 扩增,再重新克隆。

2.3.5 编码序列富集法(magnetic bead capture of expressed sequences)^[12] 其中具有代表性的是表达序列的磁珠捕捉

法,即纯化的 YAC 或 Cosmid 补平接上接头,用生物素标记的引物 PCR 扩增;cDNA 文库的插入片段也用载体引物扩增。然后,两者严格杂交,杂交的基因组 DNA-cDNA 复合物用链霉抗生物素蛋白胶连的磁珠捕捉,去掉非特异的 cDNA,捕捉到的 cDNA 洗脱后,再进行 PCR 扩增以富集编码序列。

2.3.6 不依赖 DNA 片段的策略 上述方法均利用了克隆的染色体 DNA,如 YAC、Cosmid 等,但在缺乏某一区域染色体 DNA 克隆片段(YAC、PI)的情况下,对该区域的 DNA 感兴趣时,可以采用以下克隆策略。

利用杂种细胞建立区域特异性 cDNA 文库^[13]:该方法可快速从含有人染色体或染色体片段的杂种细胞(人-鼠杂种细胞)中分离到转录顺序,并可用 PCR 扩增。因为核不均一 RNA(hnRNA)含有种属特异性的 Alu 重复序列,便发展形成了以 hnRNA 为模板,选择合成人特异性的 cDNA。一种是应用一致性 5' 剪接顺序或随机六聚体寡核苷酸为引物,对杂种细胞的 hnRNA 反转录合成 cDNA,随机用人类基因组 DNA 重复序列作杂交筛选,获得人特异性 cDNA。另一种是采用人特异性 Alu 重复顺序为引物,hnRNA 为模板,合成 cDNA,所得到的克隆不必再进行筛选。Alu 顺序为人特异性重复序列,由于鼠 DNA 中不存在这种 Alu 顺序,故人-鼠杂交细胞中人的 cDNA 总是优势合成。

④染色体显微切割和微克隆法(microdissection and microcloning)^[14]:此法不需要建立杂种细胞系,但要知道目的基因的染色体定位和需要显微切割仪。利用该技术可任意切割整条染色体、染色体区域或单一条带,所得的片段接上载体后进行 PCR 扩增。由克隆的片段可以获得大量的单拷贝探针,通过 DNA 序列分析和合成引物转化为 STS(sequence tagged site),进而构建人类染色体 STS 图谱和 YAC、phage 或 cosmid 重叠群;也可构建染色体特异区域探针;还可筛出多肽 DNA 标志,进行基因连锁分析;也可用来筛选 cDNA 文库,获得染色体特异性或区带特异性的 cDNA。

2.4 致病突变位点的筛选或确认疾病相关基因

致病突变位点的筛选就是对病人及先证者致病相关基因的筛选。可用部分或全长 DNA 来进行检测,以确认该病所特有的缺失或点突变。这些筛查技术包括 酶切杂交技术;④PCR 扩增,产物测序可发现单碱基替换及缺失;④FISH 技术发现染色体易位及缺失; 异源双链错配分析; 单链构象多肽性(SSCP)等。突变检测要注意区别正常的多肽性,因此,要选择多个无亲源关系的个体与患者的基因进行比较分析。

3 候选克隆

候选克隆是人类基因组计划的发展的产物,是在已定位克隆的基因越来越多的情况下衍生出的克隆策略。它又分为定位候选克隆(positional candidate cloning)和功能候选克隆(functional candidate cloning)。

定位候选克隆:根据(疾病)基因的连锁分析或染色体分析基本定位,即从基因组数据库(GDB)检索该标志附近与该区域内所有已知的基因,已知的 EST,用它们直接进行致病

突变的筛选。

功能候选克隆:根据待分离基因的可能功能,检测 GDB 和 GenBank 中基因功能区域,将可能含有功能区域的基因进行(致病)突变检测。

候选克隆的另一功能是确定功能已知但与疾病尚无关联的基因在基因组中的位置,再在定位区域附近检索寻找已定位的疾病,一个成功的例子是载脂蛋白 E4 等位基因与老年性痴呆发病关系的确定。

目前,基因克隆技术已得到很大的发展。预计到 2005 年,将完成人类基因组全部 DNA 的测序工作,并完成分辨率很高的转录图谱^[15]。那时,定位候选策略将取代其他克隆策略,成为首选的克隆策略。然而,我们对动脉粥样硬化等多基因病的认识还不深,揭开这些多基因病的奥妙将是我们面临的极其重要的课题。基因克隆技术将在这严峻挑战中得到发展,并发挥决定性作用。

参考文献

- Kahn P. Gene hunters close in on elusive prey. *Science*, 1996, **271**: 1 352~ 354
- Lu XY, Chen BS, Wang KQ. Study on cDNA sequences of tree apolipoprotein AI and CI. *Atherosclerosis*, 1997, **134** (1-2): 19
- Lu XY, Chen BS, Wang KQ. Cloning and sequencing of Beijing duck apolipoprotein AI cDNA. *Atherosclerosis*, 1997, **134** (1-2): 38
- Jones KW, Shaper MH, Cheverte M, et al. Subtractive hybridization cloning of a tissue specific estigsher: TSE1 encodes a regulatory subunit of protein kinase A. *Cell*, 1991, **66**: 861~ 872
- Bonaldo MF, Lennon G, Soares MB. Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res*, 1996, **6**: 791 ~ 806
- Rosok O, Odeberge J, Rode M, et al. Solid- phase method for differential display of genes expressed in Hematopoietic stem cells. *Biotechniques*, 1996, **2** (1): 114~ 121
- Liang P, Pardee AB. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, 1992, **257**: 967~ 971
- Lisitsyn N, Lisitsva NI, Meigler M, et al. Cloning the differences between two complex genomes. *Science*, 1993, **259**: 946~ 949
- Gardiner K, Mural RJ. Getting the message identifying transcribed sequences. *Trends Genet*, 1995, **11**: 77~ 79
- Call KM, Glaser T, Ito CY, et al. Isolation and characterization of zine finger polypeptide gene at the human chromosome 11 Wilm's tumor locus. *Cell*, 1990, **60**: 509~ 520
- Allkmets R, Kashuba, Bertil P, et al. NotI linking clones as a tool for joining physical and genetic maps of human genome. *Genomics*, 1994, **19**: 303~ 309
- Tagle DA, Swaroop M, Lovett M, et al. Magnetic bead capture of expressed sequences encoded with large genomic segments. *Nature*, 1993, **361**: 751~ 753
- Corbo L, Maaley JY, Nelson DL, et al. Direct cloning of human transcripts with HnRNA from hybrid cell lines. *Science*, 1990, **249**: 652 ~ 655
- Guan XY, Meltzer PS, Cao J, et al. Rapid generation of region-

specific genomic clones by chromosome micodissection: isolation of DNA from a region frequently deleted in melanoma. *Genomics*, 1992, **14** : 680~ 684

- 15 Collinge FS. Positional cloning moves from preditional to traditional. *Nat Genet*, 1995, **9** : 347~ 350

- 16 Allan CM, Walker D, Segrest JP, et al. Identification and characterization of a new human gene (apo C4) in the apolipoprotein E, C iv, and C ϵ gene locus. *Genomics*, 1995, **28** : 291~ 300

(1999-04-19 收到)

(此文编辑 朱雯霞)