

- Cell, 59:881-892.
- [23] Tautz D. 1988, *Nature*, 332:281-284.
- [24] Manseau L. and Schupbach T., 1989, *TIG.*, 5:400-405.
- [25] Strecker T. R. et al., 1989, *Science*, 243: 1062-1066.
- [26] Sprenger F. et al., 1989, *Nature*, 338: 478-483.
- [27] Bonner T. L. et al., 1986, *Nucleic Acid Research*, 14:1009-1015.
- [28] Lohs-Schardin M. 1982, *Aoux Arch Dev. Biol.*, 191:28-36.

## 真核细胞重复顺序 DNA 研究的一些进展

张 玉 砚

(中国科学院上海细胞生物学研究所)

自 1968 年 Britten 等<sup>[1]</sup>发现真核细胞基因组普遍分布有重复顺序以来,对重复顺序的重复频率、在基因组和染色体上的分布、其有否转录产物以及有否功能等方面开展了一系列的研究。综合现已有的资料,对重复顺序 DNA 的起源、演变及功能等有如下几种推论:(1) 重复顺序 DNA 中相当部分是无意义的,即所谓的“自私”DNA;(2) 在基因组上,重复顺序和单一顺序呈间隔排列,因而推测一些散布的重复顺序可能作为基因间的间隔顺序;(3) 卫星 DNA 在染色体上的分布特征,提示其具维持染色体结构的功能;(4) 部分重复顺序的转录呈时、空间的动态变化,可能对其邻近基因的表达起调控作用;(5) 在演化过程中,重复顺序不断地变异,为新基因的形成打下基础;(6) 有证据指出,大多数哺乳类重复顺序家族是通过回位机制(Retroposition mechanism),由功能基因演化而来<sup>[2]</sup>。

### 基因多态性的分子基础

在遗传病中,除去几个已知与单基因缺损有关外,大多数遗传特征尚难以基因定位。近十年来,限制片段长度多态性(Restriction fragment length polymorphism, RFLP)研究提示了一条探索遗传病分子基础的途径。Kan 和 Dozy<sup>[3]</sup>最先在人的 $\beta$ -珠蛋白基因中发现了 DNA 多态性位点,至今已发现有多种 RFLP。

RFLP 在基因组上的分布各异,有的位于编码顺序内,有的则在编码顺序外。这种多态性变化可遗传,因此有可能与有缺陷的基因紧密连锁。

在人的基因多态性的研究中指出,由 14—65 个核苷酸构成的重复单位(Repeat unit)的重复串联是人类基因多态性的重要来源之一,称可变的串联重复(Variable tandem repeats, VTRs),又称为小卫星(Minisatellite)。人基因组分布有大量的 VTRs。由于 VTRs 重复单位数目的变化,或者由于 VTRs 区的高度不稳定性而导致重排等,使人基因组显示出限制性片段长度的变异<sup>[4]</sup>。

人类胰岛素基因的高度多态性是由于富 G+C 的串联重复顺序 ACAGGGGTGTGGGG 的丰度发生变化,重复单位数目的变化范围是 26—209 个<sup>[5]</sup>。引起 Ha-ras 基因多态性的分子基础是与 Ha-ras 编码顺序紧密连接的,由 CACTCCCCCTTCTCTCCAGGGGAGCCA 核心顺序重复串联组成的重复顺序 DNA<sup>[6]</sup>。Simmler 等<sup>[7]</sup>发现人的假常染色体区是减数分裂期发生重组的热点区,与分布于该区的小卫星 DNA 有关。

由于串联重复顺序区的易变性是可以检测的,因而作为探针,用于家系分析和亲代身份的鉴定,以及探讨肿瘤易感人群的遗传基础。Krontiris 等<sup>[8]</sup>, Hayward 等<sup>[9]</sup>, Corell 等<sup>[10]</sup>

的工作认为,由 Ha-ras 连接的 VTR 产生的等位基因与肿瘤易感性有关。但也有文献报道,认为两者间无明显相关性。Jeffreys 等<sup>[11]</sup>工作肯定地认为,人 DNA 的串联重复超易变区的自发突变率高,因而易产生新的长度等位基因,在家系追踪中可以检出。他们在人的肌红蛋白基因的内含子区筛得一由 4 个 33 bp 顺序串联组成的小卫星 DNA。33 bp 顺序中有共同核心顺序 GGAGGTGGGCAGGAAGG, 可用于家系分析。

### 调节控制基因的表达

以鸟类和哺乳类的前病毒基因组两端的长末端重复顺序(Long terminal repeat, LTR)的调控作用研究得较多。LTR 的 U 3 区顺序有 TATA, CAAT 和 AATAAA 转录调节信号。前病毒 5'/LTR 起启动子作用, 3'/LTR 则起多聚腺苷酸化和终止作用。当转录从上游阅读至 LTR 时, 则此 LTR 显示终止子作用。当从 LTR 下游起始转录时, 此 LTR 则呈现强的启动子作用。另外, 在 LTR 还有与 SV<sub>40</sub> 增强子核心顺序同源的顺序, 主要作用于邻近基因的活化。

Cheng 等<sup>[12]</sup>将 LTR 与胸腺苷激酶基因连接后构建质粒, 转染 LTK<sup>-</sup> 细胞。LTR 能促使胸腺苷激酶基因的转录和胸腺苷激酶的产生。Cullen 等比较了几个鸡逆转录病毒 LTR 的促转录活性, 由于 LTR 增强子顺序活性各异, 所以其转录活性不同。Gowda 等<sup>[13]</sup>构建了含有 LTR-neo-LTR 和 LTR-CAT-LTR 质粒, 并在上游 LTR 的 -299 和 -140 部位造成顺序缺损。转染细胞后测其转录水平。结果指出, 在 LTR 的 U 3 区内 -208--201 NT 和 -141--119 NT 顺序是转录必需的。

人的 Alu 短重复顺序也具有调控作用。Chung 等<sup>[14]</sup>从胰岛素刺激肝癌细胞生长后的 mRNA 构建的 cDNA 库中, 筛选到一个与 Alu 核心顺序有 83% 同源性的 cDNA。胰岛素短期处理或处理后早期, 均能诱导此 Alu 转录本的出现和增加。Perelygina 等<sup>[15]</sup>从 HeLa 细胞

得到的与 Alu 重复顺序结合的蛋白, 能与 SV<sub>40</sub> 病毒基因组的调控区相互作用。Podgornaya 等<sup>[16]</sup>的实验表明, Alu 重复顺序 DNA 与人核蛋白有多位点结合。

小鼠 B<sub>2</sub> 重复顺序的转录产物 B<sub>2</sub>RNA, 在转化细胞、胚胎细胞以及血清刺激下的小鼠成纤维细胞中的积累, 推测它也具有一定的调控作用。Clemens<sup>[17]</sup>比较了 B<sub>2</sub>RNA 共同顺序(Consensus)与几种可诱导的 mRNA 的顺序, 发现 B<sub>2</sub>RNA 与几种 mRNA 3' 端顺序间有一定的顺序互补性。有证据指出, 真核细胞 mRNAs 对核酸酶的敏感性与其 3' 端结构有关。B<sub>2</sub>RNA 3' 端顺序的互补, 使 mRNA 能抵抗核酸酶的作用, 增强了 mRNA 的稳定性。换句话说, B<sub>2</sub>RNA 起到调节细胞质中 mRNA 浓度的作用。

以 AATGGA 6 个 bp 为顺序单位的重复顺序可介导病毒诱导的人干扰素 β 基因的转录活性。Fujita 等<sup>[18]</sup>用化学合成的 6 bp 寡聚核苷酸串联重复顺序为核心顺序, 能起到增强子作用。

### 在细胞生长分化和胚胎发育中的作用

Kuroiwa 等从小鼠中度重复顺序 DNA 库, 筛选在细胞核中表达的中度重复顺序 DNA, 分析其表达特征可归纳为三类: (1) 优先在一类器官中表达; (2) 在所有器官内都表达; (3) 在部分器官中表达。又已知几种重复顺序转录本与细胞生长分化和发育有关, 如(1)从相关的结构基因得来的 mRNA, 如组蛋白, 肌动蛋白 mRNA; (2) 小核 RNA; (3) hnRNA 和一些胞质 polyA RNA 中含有重复顺序转录本; (4) rRNA 等。但有关这方面的研究, 大多是分析重复顺序 DNA 与细胞生长分化的相关性。

Vasseur 等<sup>[19]</sup>在小鼠胚胎发生早期, 分析 B<sub>2</sub> 重复顺序的转录产物。B<sub>2</sub>RNA 的积累与卵裂有关。在胚泡中, 内部细胞比滋养外胚层含有更多的 B<sub>2</sub>RNA, 而在 7.5 天胚胎中, B<sub>2</sub>RNA

仅见于外胚层和中胚层。Davidson等<sup>[20]</sup>以爪蟾为材料,指出,卵细胞营养细胞期重复顺序 mRNA 表达很微弱,后随细胞分裂逐步增加,至原肠胚期达最高水平,成年时又下降。Young等<sup>[21]</sup>以果蝇为材料,其 *Copia* 和 412 重复顺序的表达与果蝇发育阶段有关。近二、三年来, U1 小分子 RNA 与胚胎发育阶段的相关性,引起人们的注意。各种重复顺序家族在胚胎发育不同时期转录水平各异的结果提示,重复顺序元件在胚胎发育过程中可能作为协调调节顺序起作用。

Brulet等<sup>[22]</sup>从小鼠筛选到一具有转座子结构的适度重复顺序 DNA, 命为 E. T 家族, 在未分化的胚胎癌细胞系中有转录, 但诱导细胞分化后, 检测不出其转录本。Chou等(1984)用二甲基亚砜(DMSO)处理 HL<sub>60</sub> 细胞使达到终末分化。从终末分化细胞的 mRNA 逆转录得 cDNA, 50% 含有重复顺序。含重复顺序的百分比, 比相应的对照组高 20 倍; 其中有的含有 *Alu* 顺序。Herget等<sup>[23]</sup>体外培养诱导大鼠成肌细胞 L<sub>6</sub> 细胞分化形成肌管。从分化肌细胞 mRNA 的 cDNA 库中筛选得到一些具阶段专一性的 cDNA (Stage-specific cDNA), 含有重复顺序组分。进一步以阶段专一的 cDNA 中之一 pL 6-411 为探针, 与成肌细胞和肌管形成过程的胞质 RNA 进行分子杂交反应, 分析重复顺序元件在细胞分化中的可能作用。结果表明, 与 pL 6-411 相关的转录本在 L<sub>6</sub> 细胞肌管形成过程中逐步表达和积累。在 pL 6-411 上有 TGGCA 结合蛋白识别部位。已知 TGGCA 结合蛋白的结合部位多见于转录的增强子, 腺病毒复制起点, 以及一些基因 DNA 的超敏感点。所以, 推测 TGGCA 结合蛋白与靶部位的相互作用, 调节了阶段和组织专一的基因表达。

Glaichenhaus等<sup>[24]</sup>从多瘤病毒转化的、以及正常大鼠成纤维细胞的 cDNA 基因库中筛选出 4 个 cDNA 克隆。进一步分析结果指出, 在这 4 个 cDNA 中均含有一段共有的重复顺序,

有 110 bp 长, 其产物与细胞生长状态相关。另外, 在 Ha-ras 和 myc 基因转化的细胞中, 其转录水平也可提高 10—50 倍。Kirschmeier等(1982, 1983)在体外用化学致癌剂、放射等方法诱发啮齿类细胞转化过程中, 伴随一系列含有类 LTR 顺序的表达。Housey等(1985)在体内用化学致癌剂诱发皮肤细胞癌变后, 检测到一系列含有 LTR 顺序的表达。推测在细胞转化过程, 一些调节基因转录的 DNA 顺序的结构和功能可能发生紊乱, 或者伴有内源性逆转录病毒的表达。我们实验室在 DENA 诱发大鼠肝癌发生过程中观察到, 与适度重复顺序 DNA 互补的肝核 RNA 逐步减少<sup>[25]</sup>。从大鼠重复顺序 DNA 库筛选得一约 5 kb 重复顺序 DNA 片段, L 5 B-4。其转录本在大鼠肝和肝癌细胞核、质中的分布有差异。与正常肝细胞相对比, 其转录本在肝癌细胞核中减少; 而在肝癌细胞多聚核蛋白体 RNA 中, 不但相对量增加、而且还呈现有 3 条与 L 5 B-4 DNA 片段可杂交的 RNA 带谱, 正常肝细胞多聚核蛋白体 RNA 中则无。与处于不同生长时态和病态的肝组织 RNA 对比分析结果指出, 肝癌细胞出现的 3 条 RNA 带谱具有一定的专一性。已知 L 5 B-4 DNA 的一段亚片段对转录起负调控作用, 推测其在肝细胞癌变中可能起一定作用<sup>[26]</sup>。

Flores等<sup>[27]</sup>在小鼠心脏细胞观察到, 位于染色体外的、由不同重复顺序组成的共价闭合 DNA 量的变化, 与小鼠年龄有一定相关性。

#### 用于探讨种属关系及其演化意义

对已知高度重复顺序与适度重复顺序 DNA 的分析, 及其与某些功能基因同源性的比较, 提供证据指出, 大多数哺乳类散布的重复顺序家族是起源于功能基因。又由于所有真核细胞重复顺序结构间极少相似, 所以, 长的散布重复顺序(LINES)和短的散布重复顺序(SINES)的产生机制各异。一般它们可通过转

座、复制、重排或整合等方式演变，而回位过程对 LINES 和 SINES 的形成是共同的。

从 SINES 顺序分析结果中发现，SINES 中有的与 tRNA 有部分顺序同源性，而且结构上，甚至次级结构上也很相似；有的则与小分子 RNA 基因有部分顺序同源性。如啮齿类 2 型 Alu 顺序家族、大鼠 ID 核心顺序、小鼠 B<sub>2</sub> 家族、兔 C 家族、牛或绵羊 73 bp 重复顺序和小鼠 B<sub>1</sub> 家族均分别与不同的 tRNA 顺序有同源性；人的 Alu 顺序家族和小鼠 B<sub>1</sub> 家族则与 7 SRNA 有同源性。7 SRNA 基因从果蝇到人类是高度保守的，而与之有部分同源性的 Alu 顺序和 B<sub>1</sub> 家族则分别在一定范围内分布，少保守。也就是说，啮齿类的 SINES 在灵长类中没有发现，反之也同。可能它们的亲代起源是同一的，在演化过程中，发生变化后不断扩增而形成不同的家族<sup>[28-31]</sup>。

对于 LINES，除去长短上与 SINES 有不同外，仍具有开放阅读框码 (ORF)。如大鼠 2.4 kb 重复顺序，在一条链上有两个 ORF，在另一条链上有一个 ORF，有多肽顺序的起始信号和终止信号。根据核苷酸顺序推测出的氨基酸顺序，其中有一个膜多肽顺序<sup>[32]</sup>。小鼠 LIMd 元件的顺序分析，具有两个 ORF，推导出氨基酸顺序与逆转录酶的一些区段有同源<sup>[33]</sup>。人 TBG 41 是 Kpn 家族中的成员之一，可代表全长；其一些区段分别与人脑啡呔原 3' 旁侧顺序、人 U 3snRNA 假基因旁侧顺序、人免疫球蛋白加工基因 3' 旁侧顺序、小鼠激肽释放酶和大鼠促乳素的内含子区有同源性外，从其 ORF 核苷酸顺序推导出的氨基酸顺序，与人的转铁蛋白和乳转铁蛋白有高度同源性<sup>[34]</sup>。

从上述有关重复顺序 DNA 起源及其演化过程的资料，人们设想利用经筛选得具有一定专一性的重复顺序 DNA 片段，可进行有关种属关系及其在生物演化中作用等方面的研究。

## 摘 要

本文从如下几方面介绍了真核细胞重复顺序 DNA 结构功能研究的一些进展：(1) 串联重复顺序是构成基因多态性的分子基础，基因多态性分析已用于家系分析；(2) 一些短的重复顺序如 LTR 上具有转录调节信号，对基因表达起调控作用；(3) 一些重复顺序的表达，与细胞生长分化有一定的相关性；(4) 重复顺序在种属演化中具有一定的意义。

## 参 考 文 献

- [1] Britten, RJ, et al., 1968, *Science*, 161: 592-540.
- [2] Deininger, PL, et al., 1986, *TIG.*, 2: 76-80.
- [3] Kan, YW, et al., 1978, *PNASUSA*, 75: 5631.
- [4] Colb, M, et al., 1986, *NAR*, 14: 7929-7938.
- [5] Bell, GI, et al., 1982, *Nature*, 295: 31-35.
- [6] Capon, DJ, et al., 1983, *Nature*, 302: 33-37.
- [7] Simmler, MC, et al., 1987, *EMBO J.*, 6: 963-969.
- [8] Krontiris, TG, et al., 1985, *Nature*, 313: 369-374.
- [9] Hayward, NK, et al., 1988, *Hum Genet*, 78: 115-120.
- [10] Corell, B. et al., 1988, *Hum Genet*, 79: 255-259.
- [11] Jeffreys, AJ, et al., 1985, *Nature*, 314: 67-73. 1988, *Nature*, 332: 278-281.
- [12] Cheng, SM, et al., 1984, in *Gene transfer and cancer*, Pearson, ML, et al., (eds). Raven press, NY, p 331-335.
- [13] Gowda, S, et al., 1988, *Virology*, 162: 243-247.
- [14] Chung, LM, et al., 1988, *Biochem Biophys Res Commun.*, 156: 1287-1292.
- [15] Perelegina, LM, et al., 1987, *Mol Biol Reports*, 12: 111-116.
- [16] Podgornaya, OI, et al., 1988, *FEBS Letters*, 232: 99-102.
- [17] Clemens, MJ, 1987, *Cell*, 49: 157-158.
- [18] Fujita, T. 1987, *Cell*, 49: 357-367.
- [19] Vasseur, M, et al., 1985, *EMBO J.*, 4:

- 1749—1753.
- [20] Davidson, EH, et al., 1982, *Nature*, 297: 633—635.
- [21] Young, MW, et al., 1981, in Cold Spring Harbor Symposia on Quantitative Biology vol. 45 pt 2 p 629—640.
- [22] Brulet, P, et al., 1983, *PNASUSA*, 80: 5641—5645.
- [23] Herget, T, et al., 1986, *EMBO J.*, 5: 659—664.
- [24] Glaichenhaus, N, et al., 1987, *Cell*, 50: 1081—1089.
- [25] 张玉砚等, 1982, 实验生物学报, 15: 137—141.
- [26] Xu YN. et al., 1989, *Chinese J of Cancer Research*, 1(5):7—14.
- [27] Flores, SC, et al., 1988, *NAR* 16: 3889—3905.
- [28] Sakamoto, K, et al., 1985., *J Mol Evol.*, 22: 134—140.
- [29] Lawrence, CR, et al., 1985, *NAR.*, 13: 4239—4252.
- [30] Daniels, GR, et al., 1985, *Nature*, 317: 819—822.
- [31] Deininger, PL, et al., 1986, *TIG.*, 2: 76—80.
- [32] Achten, S, et al., 1986, *J Mol Biol.*, 192: 489—502.
- [33] Loeb, DD, et al., 1986, *Mol cell Biol.*, 6: 168—182.
- [34] Hattori, M, et al., 1985, *NAR.*, 13: 7813—7827.

## 研究工作

# 小鼠胃粘膜细胞和人胃癌细胞间隙连接的超微结构研究与促癌变剂的影响\*

高燕 林仲翔 吕桂芝 周立新 韩亚玲

(北京市肿瘤防治研究所细胞生物学研究室)

鲁崎唔 吴莲英

(中国科学院生物物理研究所电镜室)

间隙连接介导的细胞连接通讯在调控细胞增殖和分化中有重要作用<sup>[1-4]</sup>。早期的研究<sup>[5,6]</sup>发现肿瘤细胞缺乏连接通讯功能。进一步的研究<sup>[7-10]</sup>表明肿瘤细胞连接通讯抑制的程度与细胞的恶性程度,特别是与细胞的锚着不依赖性增殖能力和转移能力呈正相关;许多促癌变剂有抑制细胞连接通讯的作用。由此提出的假说认为细胞连接通讯的抑制是癌变机理的一部分,很可能是癌变促进(Tumor Promotion)的关键因素,细胞连接通讯的抑制使癌变启动细胞(Initiated Cells)得以逃避周围细胞的调控而继续恶变和扩增最后形成癌肿<sup>[11,12]</sup>。日益增多的关于细胞连接通讯研究资料支持这一假说,指出细胞连接通讯功能在癌变促进的过

程中逐渐丧失<sup>[11-13]</sup>。

细胞连接通讯依赖于相邻细胞膜上间隙连接的数目和/或其通透性的大小。冷冻蚀刻电镜技术能够提供间隙连接结构的基础资料,对于了解与细胞连接通讯变化有关的癌变机理是十分必要的,然而这样的研究报道还不多<sup>[14,15]</sup>。我们曾观察到人胃癌 MGC-803 细胞缺乏间隙连接通讯功能<sup>[16]</sup>,提示细胞连接通讯抑制也是胃癌癌变的机理之一。为了解细胞连接通讯抑制的超微结构基础以及与促癌变可能的连系,本工作应用冷冻蚀刻电镜技术比较观察胃癌 MGC-803 细胞与正常小鼠胃粘膜上

\* 本工作由国家自然科学基金项目资助,