

LIM 类同源异形框基因研究进展*

REVIEW OF STUDY ON LIM CLASS HOMEBOX GENE

王 勇 刘宗柱 张培军

(中国科学院海洋研究所 青岛 266071)

* 同源异形框基因(Homobox gene)是 McGinnis 等^[11]与 Scott 等^[14]在研究果蝇同源异形基因(Homoeotic gene)时发现的。当时他们发现果蝇同源异形基因中都有一个异常保守的核苷酸序列,该序列共有 180 个核苷酸,编码 60 个氨基酸,他们就称该序列为同源(异形)框(Homobox)。同源框所编码的氨基酸序列叫同源异形结构域(Homodomain),同源异形结构域的作用是与靶基因的 DNA 序列结合,从而调控其表达。后来发现,同源框基因分布很广,在动物、植物和真菌中都有。其中在动物界,从腔肠动物到人的各大分类单元中的后生动物都有同源框基因^[7]。同源框基因数量很多,种类也不少,其中 LIM 类同源框基因是最近几年发现的一类同源框基因,与动物胚胎发育中多种细胞谱系的建立有关。

1 LIM 类同源框基因的发现与命名

早在 1988 年,Way 和 Chalfie^[21]在研究线虫(*Caenorhabditis elegans*)的触觉神经元时发现,线虫 *mec-3* 基因是线虫的 6 个触觉神经元正常分化所必需的基因。他们用转座子标记法(Transposon tagging)克隆了这个基因并测序。当时他们只注意到 *mec-3* 基因序列中的同源框特点,而没有注意在同源异形结构域上游的富含半胱氨酸的区域。1990 年 Freyd 等^[6]和 Karlsson 等^[9]在同期发表了各自的研究结果。Freyd 等^[6]克隆了线虫(*C. elegans*)的一个基因 *lin-11*,并测序。该基因在线虫胚胎发育中控制阴门前体细胞的不对称分裂。在该基因编码的蛋白质分子中,在同源异形结构域之前有两个串联重复的序列,该序列是一个富含半胱氨酸的模色(motif)。而 Karlsson 等^[9]则克隆了大鼠胰岛素基因增强子结合蛋白的基因 *isl-1*,并测序。他们在 *isl-1* 基因的蛋白产物中也发现,在同源异形结构域之前有一个富含半胱氨酸的模色,与 *lin-11* 基因的基本一致。这两个小组又分别把自己的研究结果与对方的以及 Way 和 Chalfie 的研究结果相比较,发现这种模色在 *lin-11*, *isl-1* 和 *mec-3* 中的几乎一样,因此,他们分别取这三个基因的第一个字母,组成了这种模色的名字,即 LIM。LIM 模色所编码的氨基酸序

列称为 LIM 结构域(LIM domain),含有 LIM 模色的同源框基因则称为 LIM 类同源框基因(LIM class homobox gene)。以后,又陆续从黑腹果蝇(*Drosophila melanogaster*)中克隆了 *apterous*(ap)基因^[3],从鸡中克隆了 *islet-1*, *islet-2*, *Lim-1* 和 *Lim-3* 基因^[18],从鱼中克隆了鱼 *isl-1* 类基因^[8],从海鞘中克隆了 *Hrlim* 基因^[20]等等,至今已克隆了几十个 LIM 类同源框基因。此外,还发现了不含同源框的 LIM 基因,如只含 LIM 结构的基因、C 末端为 LIM 结构而 N 末端为其他结构的基因等^[4]。最近,有人根据 Scott 于 1992 年提出的同源框基因命名的建议,提出将 LIM 类同源框基因更名为 LHX 基因^[4]。为避免叙述混乱,本文暂不采用这一命名法。

2 LIM 结构域的结构与功能

在 LIM 类同源框基因的研究中,除了进行基因克隆以外,还对 LIM 结构域的结构和功能作了系统研究。LIM 结构域在组氨酸和半胱氨酸位点上非常保守,但无论在不同的 LIM 基因之间还是在同一 LIM 基因的两个 LIM 结构域之间,其他的氨基酸保守性很差,因此, LIM 结构域的保守性远没有位于其下游的同源异形结构域的保守性好。LIM 结构域的空间结构很像一个双锌指结构^[4],尽管有人曾推测 LIM 结构域既可以结合 Fe^{2+} 也可以结合 Zn^{2+} ^[10],但最近的研究证明, LIM 结构域只与 Zn^{2+} 结合^[12]。由于锌指蛋白能与 DNA 结合并调节基因活性,因此推测 LIM 结构域可能也有类似功能,只是尚未证实。现在已有证据表明, LIM 结构域确实能与蛋白质相互结合。例如,有一种只有 LIM 结构域的蛋白 CRP,能形成二聚体,而二聚体的形成就依靠 LIM 结构域^[5]。在 *Mec-3* 和 *Isl-1* 蛋白的体外实验中发现,蛋白质中的同源异形结构域与 DNA 的结合能被 LIM 结构域抑制,而删除 LIM 结构域则可以恢复同源异形结构域与 DNA 结合的活性^[13, 22]。体内实验的结果与此很相似,如果向爪蟾胚

* 中国科学院海洋研究所调查研究报告第 3037 号。实验海洋生物学开放研究实验室研究报告第 149 号。
收稿日期:1996 年 8 月 5 日

胎中注射完整的 *Xlim1* 基因的 mRNA, 则并不表现出 *Xlim1* 基因的活性; 而注射去除 LIM 结构域或者 LIM 结构域已突变的 *Xlim1* 的 mRNA, 则表现出 *Xlim1* 基因的活性^[16]。体内和体外实验的结果均提示, LIM 结构域可能通过分子内或分子间的相互作用调节蛋白质的功能。

3 LIM 类同源框基因的生物学功能

3.1 无脊椎动物 LIM 类同源框基因的生物学功能

不同种动物的 LIM 类同源框基因的生物学功能有较大差别。在线虫中, 如果缺失 *mec-3* 基因, 这 6 个神经元就不会表现出其独特的分化特点而转化成其他类型的神经元, 因此, *mec-3* 基因决定线虫的 6 个触觉神经元的正常分化^[21]。由于 *mec-3* 基因是同源框基因, 提示 *mec-3* 基因的蛋白产物通过与靶基因的 DNA 序列结合并调节这些决定神经元分化特征的基因的转录, 来决定神经元的分化。与 *mec-3* 基因相类似, *lin-11* 基因在线虫阴门前体细胞发育命运的决定中发挥作用^[6]。

在果蝇 (*Drosophila melanogaster*) 中, *apterous* (*ap*) 基因有多种功能, 这一点与大多数脊椎动物的 LIM 类同源框基因相类似。已发现, *ap* 基因参与果蝇翅的图式形成^[3], 并在肌肉前体细胞中表达, 其在肌肉细胞中的表达明显地与肌肉分化有关^[2]。另外, 在果蝇神经系统中也有 *ap* 基因的广泛表达^[2]。

Wada 等^[20]克隆了海鞘中的 LIM 类同源框基因 *Hrlim*, 并研究了该基因在海鞘胚胎各发育阶段的时空

表达。研究发现, *Hrlim* 在海鞘胚胎中的表达以原肠形成界, 可分为两个阶段。第一个阶段, 在 32-细胞期 *Hrlim* 的转录最早在内胚层谱系的细胞中被特异激活并持续到原肠形成期; 在原肠形成之前, *Hrlim* 也在将来形成脊索的分裂球中短暂表达。第二个阶段, 在原肠形成之后, 表达图式完全变了, *Hrlim* 在中胚层和内胚层的表达消失, 而在外胚层的预定神经系统的一组细胞中表达。这些表达图式提示, 在海鞘的胚胎发育中 *Hrlim* 起了几种不同的作用: 在原肠形成之前, 尽管 *Hrlim* 的基因产物并不构成内胚层决定子, 但 *Hrlim* 基因可能受内胚层决定子的调控, 因而 *Hrlim* 可能与内胚层和脊索的细胞谱系的决定有关; 而在原肠形成之后, 可能与神经组织中某些细胞的特化有关。因此, 综合以上资料可以看出, 在无脊椎动物中, LIM 类同源框基因与细胞谱系的决定以及细胞分化过程的启动与维持有关。

3.2 脊椎动物 LIM 类同源框基因的生物学功能

现在已发现的脊椎动物 LIM 类同源框基因数量很多, 这里只讨论经过充分鉴定的基因。在用功能缺失 (Loss of function) 和功能获得 (Gain of function) 方法研究基因的生物学功能之前, 往往只能从基因的表达图式来推测其生物学功能, 因此, 全面地了解某个基因的表达图式就显得十分重要。表 1 列出了几类经过充分鉴定的脊椎动物 LIM 类同源框基因的表达图式, 据此, 可以初步推测这些基因在胚胎发育中的可能的生物学功能。

表 1 脊椎动物 LIM 类同源框基因的表达图式^[4]

基因	表达区域
<i>Isl-1</i>	中枢神经系统和外周神经系统中的大部分, 垂体、甲状腺的神经分泌细胞, 胰岛细胞, 角质细胞和成纤维细胞
<i>Lim-1</i>	蛙原肠胚的组织者、脊索和脊索前方的中胚层, 前肾、中肾和后肾, 中枢神经系统大部
<i>LH-2</i>	B-淋巴细胞、T-淋巴细胞, 脑的大部, 垂体的促性腺和促甲状腺部分的细胞谱系
<i>Lim-3</i>	主要局限于神经细胞和神经内分泌细胞
<i>Gsh-4</i>	脊髓和后脑的腹部两侧区域
<i>Lim-5</i>	原肠胚时在外胚层, 胚胎发生后期则主要在间脑

首先, 部分 LIM 类同源框基因参与图式形成。图式形成, 即指胚胎细胞形成空间上有序的分化组织的活动。例如, 爪蟾的 *Xlim1* 基因在原肠胚时期的胚胎组织者区表达。通过向爪蟾 2-细胞胚胎中注射 *Xlim-1* 基因 mRNA 的实验, 发现在两个 LIM 结构域处都发生

突变的 *Xlim-1* 基因(已丧失 LIM 结构域活性)具有 5 种活性: (1) 神经诱导; (2) 粘液腺形成; (3) 激活 *goosecoid* 基因; (4) 在表达 *Xbra* 基因的细胞中的肌肉诱导; (5) 在脊索形成中与 *goosecoid* 基因联合作用。而以上这些活性均为 *Spemann* 组织者的特点, 提示

Xlim-1 基因可能在爪蟾胚胎发生的图式形成中起作用^[16]。

其次,大多数 LIM 基因与细胞分化有关。爪蟾的 Xlim 5 和斑马鱼的 Lim 5 基因在神经胚期的表达局限于发育中的神经板的前部区域;在 2 d 龄爪蟾和 24 h 的斑马鱼胚胎中, Xlim 5 和 lim 5 基因强烈表达,表达区位于间脑。这一结果提示, Lim 5 基因在神经细胞的分化,尤其是间脑的早期特化中发挥作用^[17]。另外,鸡 Lmx1 (C-Lmx1) 基因在鸡胚胎发育早期于预定附肢中胚层的背侧表达,以后就局限于附肢芽中胚层的背侧。C-Lmx1 基因的表达可能受外胚层中的 Wnt7a 基因调控,因为 Wnt7a 基因能诱导 C-Lmx1 在中胚层腹侧异位表达。在附肢向外生长的过程中, C-Lmx1 在中胚层腹侧的异位表达可引起附肢中胚层的腹侧转化为背侧。因此可以推测, C-Lmx1 基因在鸡附肢发育中特化背侧细胞的发育命运^[19]。而鸡的另外几个 LIM 基因, is1-1, is1-2, lim-1 和 lim-3 则控制运动神经元的分化,并且有证据表明,这几个基因使得运动神经元的轴突能有选择地沿不同路径在不同的位置到达肌肉靶细胞^[18]。

4 LIM 类同源框基因的作用特点

4.1 LIM 同源异形结构域蛋白质充当发育特异的转录因子

已知含有同源异形结构域的蛋白质都可充当转录因子^[1],而且,从 LIM 类同源框基因的生物学作用来看,它们都充当调节基因。因此毫无疑问,其编码蛋白质-LIM 同源异形结构域蛋白必定充当发育特异的转录因子。

4.2 脊椎动物 LIM 类同源框基因的协同作用

与无脊椎动物的 LIM 类同源框基因不同,脊椎动物的 LIM 类同源框基因具有协同作用,即几个基因可以联合起来决定某一性状,而且不同的基因组合可以决定不同的性状。在这方面,对鸡的胚胎运动神经元的发育调控有较为详细的研究。Tsuchida 等^[18]克隆了鸡的一个 LIM 类同源框基因家族,包括 is1-1, is1-2, lim-1 和 lim-3,共 4 个基因。这些基因的联合表达确定了几类运动神经元的发育命运,这些神经元在脊髓中分成几个柱,并选择了不同的轴突路径。这 4 个基因就在这几类运动神经元轴突路径形成之前和脊髓中运动神经

元柱出现之前表达。Tsuchida 等^[18]发现,这 4 个 LIM 类同源框基因基于它们选择轴突路径的能力又使得躯体运动神经元继续分化。其中 is1-1, is1-2 和 lim-3 的共同表达决定了中间运动神经元柱的中间分支(The medial subdivision of the median motor column, MMCm),这些神经元对来自生皮肤节的信号作出反应; is1-1 和 is1-2 决定了中间运动神经元柱的两侧分支(The lateral subdivision of the median motor column, MMCl),和两侧运动神经元柱的中间分支(The medial subdivision of the lateral motor column, LMcm),这些神经元将轴突延伸到侧壁中胚层的腹侧; lim-1 和 is1-2 则决定两侧运动神经元柱的两侧分支(The lateral subdivision of the lateral motor column, LMCl),此处的神经元将轴突延伸到侧壁中胚层的背侧。总之, LIM 类同源框基因使得运动神经元的轴突能有选择地沿不同路径在不同的位置到达肌肉靶细胞。

5 结语

综上所述, LIM 类同源框基因保守性很强,在动物界广泛存在,因而作为一个基因家族,不同的 LIM 类同源框基因有着不同的表达图式以及不尽相同的生物学功能都是预料之中的。但作为一个基因家族,它们的共同点是显而易见的,即在结构上, LIM 类同源框基因的编码蛋白都有同源异形结构域和 LIM 结构域;在生物学功能方面,都与细胞谱系的决定和/或细胞分化有关。因此,对于 LIM 类同源框基因的深入研究,将有助于从分子水平上理解细胞谱系的决定和细胞分化等胚胎发育中的基本现象,从而为在分子水平上彻底阐明发育的基本机制奠定基础。

参考文献

- [1] 王 勇、张培军、郎刚华,1996。海洋科学 3: 33~ 35。
- [2] Bourgoin, C., S. E. Lundgren *et al.*, 1992. *Neuron* 9: 549-561.
- [3] Cohen, B., M. E. McGuffin *et al.*, 1992. *Genes Dev.* 6: 715-729.
- [4] Dawid, I. B., R. Toyama *et al.*, 1995. *Life Sci.* 318: 295-306.

(下转 60 页)

(上接 27 页)

- [5] Feuerstein, R., X. Wang *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91 : 10 655-10 659.
- [6] Freyd, G., S. K. Kim *et al.*, 1990. *Nature* 344: 876-879.
- [7] Gehring, W. J., M. Affolter *et al.*, 1994. *Annu. Rev. Biochem.* 63 : 487-526.
- [8] Gong, Z., C. Hui *et al.*, 1995. *J. Biol. Chem.* 270: 3 335-3 345.
- [9] Karlsson, O., S. Thor *et al.*, 1990. *Nature* 344: 879-882.
- [10] Li, P. M., J. Teichert *et al.*, 1991. *Proc. Natl. Acad. Sci. USA* 88 : 9 210-9 213.
- [11] McGinnis, W., R. L. Garber *et al.*, 1984. *Cell* 37 : 403-408.
- [12] Michelsen, J. W., K. L. Schmeichel *et al.*, 1993. *Proc. Natl. Acad. Sci. USA* 90 : 4 404-4 408.
- [13] Sanchez-Gacia, I., H. Osada *et al.*, 1993. *EMBO J.* 12: 4 243-4 250.
- [14] Scott, M. P. and B. J. Weiner, 1984. *Proc. Natl. Acad. Sci. USA* 81: 4 115-4 119.
- [15] Taira, M., M. Jamrich *et al.*, 1992. *Genes Dev.* 6 : 356-366.
- [16] Taira, M., H. Ohtani *et al.*, 1994. *Nature* 372: 677-679.
- [17] Toyama, R., P. E. Curtiss *et al.*, 1995. *Dev. Biol.* 170: 583-593.
- [18] Tsuchida, T., M. Ensini *et al.*, 1994. *Cell* 79 : 957-970.
- [19] Vogel, A., C. Rodriguez *et al.*, 1995. *Nature* 378 : 716-720.
- [20] Wada, S., Y. Katsuyama *et al.*, 1995. *Mech. Dev.* 51 : 115-126.
- [21] Way, J. C. and M. Chalfie, 1988. *Cell* 54: 5-16.
- [22] Xue, D., Y. Tu and M. Chalfie, 1993. *Science* 261: 1 324-1 328.