

脊椎动物血浆蛋白质的进化

郝 军 山

(山西大学分子科学研究所,太原)

提 要

分子生物学的迅猛发展,使我们有可能从分子水平来探讨生物进化这一经典命题。分子进化遂成为进化生物学中的一个新的研究领域。脊椎动物血浆中有数百种蛋白质,依据其一级结构的类似性可划分为仅仅几个家族。蛋白质的进化主要涉及到基因复制(包括加倍)和外显子改组。对不同种属同一种蛋白质的氨基酸排列顺序的比较,可以给出一些有关蛋白质进化的知识。这方面业已取得的成就极大地推进了进化生物学的发展。

关键词 血浆蛋白质,进化,脊椎动物

一、历史回顾

Pauling 等(1962)根据生物进化过程中一些分子(如血红蛋白和细胞色素 C)每年每氨基酸替代率均为特定的演化常数,遂提出了分子进化钟(molecular evolutionary clock)假说来描述生物的演化。血浆蛋白质的进化研究是分子进化研究中的一个重要研究领域。创始人美国加州大学 Doolittle 教授。他早在哈佛大学当研究生时主攻人血液蛋白质研究,1962年从事比较生理学研究后,很自然开始了人与脊椎动物血浆蛋白质的比较研究^[1]。在过去的26年间,他领导的研究小组对脊椎动物血浆蛋白质及其进化进行了广泛而又深入的研究,由此不仅发现了一些特征性蛋白质,也给出了有关血浆蛋白质进化的渊源^[2-7]。继 Doolittle 之后,国际上有关研究工作日渐深入^[8-19]。国内关于分子进化研究工作较少,目前仅有若干综述^[8-11]。

血浆蛋白质进化研究主要包括诸血浆蛋白质的氨基酸序列比较及进化机制研究。1983年美国乔治敦大学全国生物医学研究基金会创建了“蛋白质顺序库”。随着电子计算机应用于这

方面的研究工作,业已取得了较大进展。由此而重建的系统树,不仅印证了以往主要依靠化石证据构建的系统树,而且能以氨基酸或核苷酸替代率为演化常数来刻画尚无化石证据证明的物种间或物种内的分野时期,并能估计基因簇中不同基因的分化情况,再现物种进化史和基因进化史。

二、同源性分析

典型的脊椎动物血浆含有多达 600 多种蛋白质成分,应用高分辨率的二维电泳可把它们分开^[25]。这些血浆蛋白质可以按功能分类,主要有血清白蛋白,铁传递蛋白,激动素,珠蛋白,血纤维蛋白原,血浆铜蓝蛋白,免疫球蛋白和脂蛋白,另外还有些多肽激素及生长因子。血液中还有几个独特的级联系统(cascade system,包括凝血系统和补体系统)。正是因为有这些物质的存在,血浆才能行使其输送营养物及废物、抵御微生物的侵染等重要功能。然而,问题在于所有这些蛋白质从何而来?在无脊椎动物中有与之对应的物质存在吗?如果有,它们之间的区别和联系又是怎样的呢?

有关脊椎动物血浆蛋白质如何进化的早期

研究发现,珠蛋白(一种选择性结合细胞间血液蛋白质的蛋白质)和胰凝乳蛋白酶类同源^[12]。以后又相继发现许多其他血浆蛋白质之间存在或多或少的同源序列^[6,7,13,14]。抗凝血酶 III (一种依赖肝素的蛋白酶抑制剂)其 30% 的序列和 α_1 抗胰蛋白酶同源。令人惊奇的是卵白蛋白和这两种蛋白酶抑制剂具有类似的同源序列。后来应用计算机来比较蛋白质一级结构又发现血管紧张肽的前体也属于这一家族。

人们很早就认识到 C3, C4, C5 这三种补体是彼此相关的。因为它们都通过蛋白质水解作用释放出类似的生物活性肽。但是后来发现 C3, C4 同 α_2 巨球蛋白有不易区分的化学敏感性。这三种蛋白质都含有特殊的环硫酯,可被亲核试剂如甲酰胺分解,然而,当时由于科学家们对环硫酯的化学方面的兴趣而忽略了把其同源序列深入至进化论去加以研究。后来对 α_2 巨球蛋白和三种补体 (C3, C4, C5) 的一级结构相似程度分析之后,遂提出 α_2 巨球蛋白是其他三

种蛋白质赖以衍生的亲本祖先^[5]。最近又发现了这一家族的另一个成员——妊娠带蛋白 (pregnancy zone protein) 其 60% 的序列和 α_2 巨球蛋白同源^[15]。

脊椎动物的血液凝集是一个非常复杂且巧妙的平衡体系。其中涉及到几种蛋白酶酶原和几个非酶蛋白因子。后者包括血纤维蛋白原 (构成凝血块) 和蛋白因子 V 和 VIII。如果人体缺失蛋白因子 VIII 就会导致血友病。这三种非酶蛋白质均由凝血酶活化。凝血酶是与胰蛋白酶家族有关的另一种丝氨酸蛋白酶。长期以来科学家们就猜想这三种非酶蛋白质是相关的。现在我们知道蛋白因子 V 和 VIII 二者确有同源序列,但它们并非和血纤维蛋白原有关,令人惊奇的是和血浆铜蓝蛋白有关^[5]。

表 1 列举了具有同源序列的某些脊椎动物血浆蛋白质。我们可以预见,随着研究范围的拓宽并向纵深发展,以后将会收集到更多有关血浆蛋白质序列的资料。

表 1 具有同源序列的某些脊椎动物血浆蛋白质

EGF 型:	β-2 型:
表皮生长因子前体	β_2 -糖蛋白
肿瘤生长因子	补体因子 B
低密度脂蛋白受体	补体因子 H
凝集因子 IX	因子 XIII b 链
凝集因子 X	α-2 型:
蛋白质 C	α_2 -巨球蛋白
组织血纤维蛋白原激活剂	妊娠带蛋白
尿激酶	补体 C3
补体 C9	补体 C4
激动素 "Finger"	补体 C5
激动素	α-1 型:
组织血纤维蛋白原激活剂	α_1 -抗胰蛋白酶
蛋白酶原 "Kringel"	α_1 -抗胰凝乳蛋白酶
血纤维蛋白原	抗凝血酶 III
组织血纤维蛋白原激活剂	血管紧张肽原
尿激酶	血浆铜蓝蛋白:
凝血酶原	血浆铜蓝蛋白
C9 型:	蛋白因子 V
补体	蛋白因子 VIII
低密度脂蛋白受体	血清白蛋白
(激动素?)	血清白蛋白
Notch (果蝇)	α -胎儿蛋白
Lin-12 (线虫)	

三、进化机制

现在认为蛋白质的进化机制有下述几种:

1. 基因复制

主要指基因加倍。如血清白蛋白有一种从 190 个残基进化而来的三倍结构。铁传递蛋白显然是一种双倍结构: 其两个含有 340 个残基的序列中 50% 的序列是相同的,且每一半都能结合一个铁原子。人血纤维蛋白原的 α 链有一段 13 个残基的片段出现好几次。许多脊椎动物血浆蛋白质的进化涉及典型的真核基因 (10—15kb) 的全复制。在有些情况下,用内含子的分布来确证两种蛋白质具有共同祖先与否的推断。如,血管紧张肽原的序列虽然只有约 20% 的序列和 α_1 抗胰蛋白酶同源,但其内含子的分布却是一样的,这就证实了这两种蛋白质属于同一家族。

2. 外显子改组 (Exon Shuffling)

外显子改组指基因中已有的外显子的不同组合和装配从而产生新的功能性蛋白质的过程。问题在于外显子改组是真核生物进化的主

要因素吗？现在流行的看法认为大多数蛋白质是通过重要的外显子的装配而进化的。但是也有人发现被交换的单位是小而完整的，通常包含几对二硫键。在产生新的脊椎动物蛋白质中涉及到的基因加倍其范围从较短序列的重复到 15kb 的重复，显然外显子改组仅是扩展真核生物蛋白质进化的几条途径之一。然而，正象科学家们刚发现内含子时就提出的那样，有些研究者们仍然看重内含子所包含遗传信息的重要的选择价值。从生物学功能考虑，内含子阻止有丝分裂和减数分裂期间复制基因中的错配对，这对真核生物来说是极为重要的。

3. 融合蛋白质

有些蛋白质显然是基因融合（或涉及加工剪接过程）的结果。例如血浆补体蛋白质 C1q 其亚基的氨基末端部分和胶原类似，暗示胶原基因的一部分业已和目前还尚未鉴定的基因之某一部分发生融合。C1q 的类胶原部分构成了可以和触发补体级联系统的抗原-抗体复合物相互作用的主干^[5]。

4. 镶嵌蛋白质

最近有几项研究表明业已确定的外显子单位的大规模重排产生一组镶嵌蛋白质。例如，组织血纤维蛋白原激活剂（TPA）显然具有至少三种蛋白质的片段：血纤维蛋白原本身，表皮生长因子（EGF）及其前体（EGFP），以及激动素（FN）。如此镶嵌得到的单位精确落在内含子的位置。有趣的是 EGF 型单位也出现于血液凝集和纤维蛋白溶解中所涉及到的几种其他蛋白酶原的氨基末端。这几种蛋白质包括凝集因子 IX 和 X，蛋白质 C 和尿激酶。这种联系甚至更为广泛。如在低密度脂蛋白受体（LDLR）的序列中也发现了 EGF 型单位。事实上 EGFP 的 270 个残基部分（包括一段含游离半胱氨酸的较长区域，以“y”表示，和三个类 EGF 单位）与其他几种蛋白质有 40% 的同源序列（见图 1）。说明它们有共同的祖先。另一个有趣的发现是补体 C9 的序列中含有一 39 个残基的片段很象 LDLR 氨基末端部分中的重复单位，这些单位也装配于 TPA 中。涉及镶嵌

单位的蛋白质至少有 12 种^[16-19]。

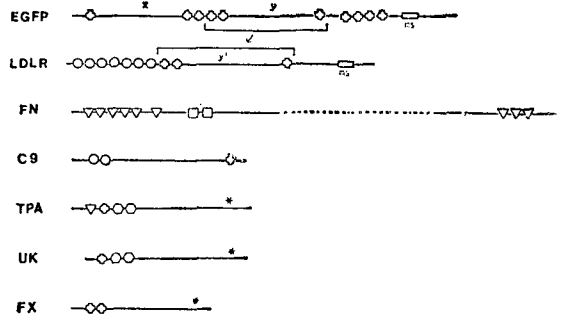


图 1 由外显子穿梭引起的几种脊椎动物血浆蛋白质含有同源“模块”的图解比较

EGFP: 表皮生长因子前体; LDLR: 低密度脂蛋白受体; FN: 激动素(粘连蛋白); C9: 补体成分 9; TPA: 组织血纤维蛋白原激活剂; UK: 尿激酶; FX: 血凝因子 X; ms: 越膜片段; *: 丝氨酸蛋白酶的活性位点

四、重建系统树

我们如何去了解脊椎动物血浆蛋白质的祖先呢？对有特定且灵敏试验的蛋白质（如铁传递蛋白）而言，直接研究蛋白质当然是合情合理的。Lee 等 1978 年就直接研究了铁结合蛋白^[20]。但其他蛋白质（如血清白蛋白）就难于驾驭了。而对那些具有相同功能的蛋白质——脂蛋白类，铁或铜传递体类等，问题在于要对同源蛋白质中的类似的东西加以区别，这就意味着要对很多材料进行测序。不难想象用传统的氨基酸测序法必然会陷入迷津。显而易见我们可以不去研究蛋白质而去研究其基因。重组 DNA 技术的新进展使得这方面的探索加重了取胜的筹码。这一战略取决于鉴定脊椎动物蛋白质所保留的（因此很有可能在其无脊椎动物亲族中也存在）诸特点。应用合成适当的寡核苷酸探针可以扫描原索动物和无脊椎动物的基因文库。由此而重建系统进化树。

如果所有被试哺乳动物都有某种蛋白质（如乳清蛋白），而其他脊椎动物没有这种蛋白质时，我们就可以设定乳清蛋白是在哺乳动物出现时（或出现前）‘新生’的。对五个不同生物类群做过有关其血红蛋白、血纤维蛋白原、补体蛋白质的氨基酸序列比较之后，绘制的进化树

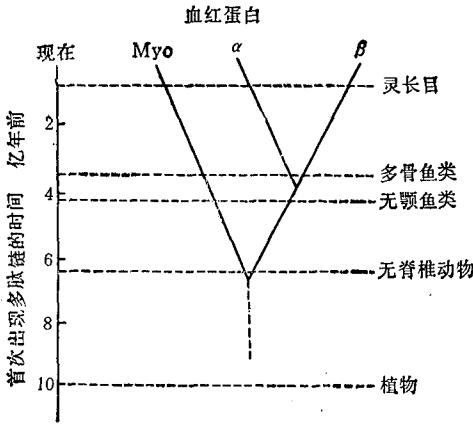


图2 血红蛋白及肌红蛋白 (Myo) 的进化树
图中三条多肽链的相似度一致

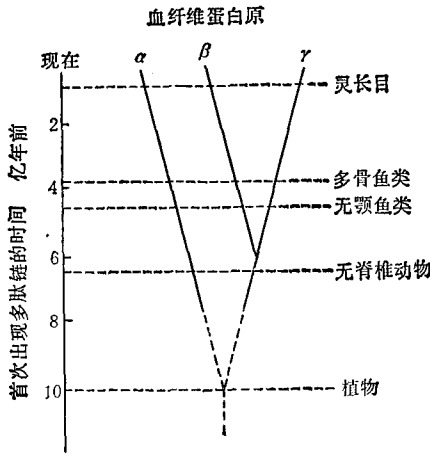


图3 基于人血纤维蛋白原 α β 和 γ 链 aa 序列
差异绘制人血纤维蛋白原的进化树

表2 一些主要脊椎动物类群的分野时期¹⁾

脊椎动物类群	亿年前 $\bar{x} \pm SD$
人类	0.28 \pm 0.02
灵长目类	0.80 \pm 0.10
啮齿目类	0.80 \pm 0.10
鸟类	2.00 \pm 0.20
爬行类	2.00 \pm 0.20
两栖类	3.20 \pm 0.30
多骨鱼类	3.75 \pm 0.35
圆口类	4.50 \pm 0.50

1) 主要依据化石分析

分别见图2、图3和图4。以血红蛋白为例，显然，4亿年前出现的那些生物（如鱼类）只有含一条多肽链的血红蛋白（没有 β 链！）。如此绘制的进化树印证了由化石分析得到的有关物种

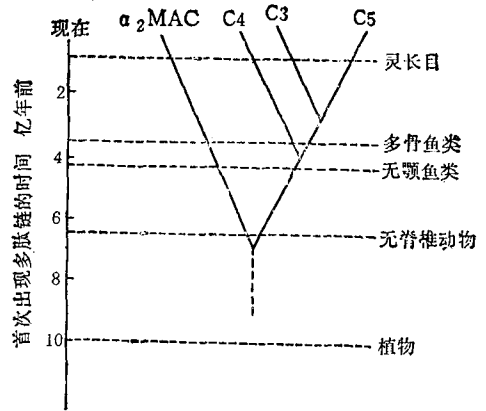


图4 基于有限的序列数据给出 α_2 巨球蛋白 (α_2 MAC)
同补体 C3, C4, C5 假设关系

分野时期划分的结果(见表2)。

五、结 论

脊椎动物血浆蛋白质的进化是由基因复制和外显子改组等引起的。对血浆中数百种蛋白质进行氨基酸序列比较可以找出几个具有同源序列的家族。依据序列分析绘制的物种分野的进化树和由化石分析得到的结果相符。计算机研究的引入无疑会拓宽分子进化研究的范围。有关脊椎动物血浆蛋白质进化的研究结果可以作为达尔文生物进化论的生化佐证。

本文蒙袁静明教授审阅,并提出修改意见,在此致谢。

参 考 文 献

- 1 Doolittle R F, Oncley J L, Surgenor D M. *J Biol Chem*, 1962; 237: 3123
- 2 Fuller G M, Doolittle R F. *Biochemistry*, 1971; 10: 1311
- 3 Doolittle R F, Fuller G M. *Biochim Biophys Acta*, 1972; 263: 805
- 4 Doolittle R F. In: *The Proteins*. 2nd ed, Vol IV, New York: Academic Press, 1979: 1—118
- 5 Doolittle R F. In: *The Plasma Proteins*. 2nd ed, Vol IV, New York: Academic Press, 1984: 317—360
- 6 Doolittle R F. *TIBS*, 1985a; 10: 233
- 7 Doolittle R F. *Biol Bull*, 1987; 172: 26!
- 8 陈建华. 遗传, 1985; 3: 41
- 9 吕宝忠. 自然杂志, 1987; 10: 846
- 10 吕宝忠. 自然杂志, 1988; 12: 904
- 11 王克夷, 冯佑民. 生物化学与生物物理进展, 1989; 15 (3): 176

蝎毒中的昆虫神经毒素的生化研究

吉永华

(中国科学院上海生理研究所)

提 要

开展对蝎毒中昆虫神经毒素的研究是近几年国内外同行较关注的课题之一。本文扼要地介绍了蝎昆虫神经毒素在一般理化特性,结构与功能的关系及作用的专一性和机理等基础理论研究方面上的一些进展概况。

关键词 蝎昆虫神经毒素,收缩或松弛麻痹,二硫键变位,钠离子通道

蝎属蛛形纲,节肢动物门,以苍蝇一类的小昆虫为食。一般地认为,在蝎毒(scorpion venom)中应以昆虫毒素为主,但事实上,迄今在被研究的各种蝎毒中均发现其主要的毒性成分却是对哺乳动物具有很强致死效应的神经毒素(MTx),其在蝎粗毒中的含量高达10%—50%,相反,昆虫神经毒素(ITx)的含量却低于1%。因此,人们对蝎MTx的了解,无论是在其理化特性,结构与功能的关系,药理与免疫性质以及作用机制等方面都远比对蝎ITx的要广泛、深入的多。有关对蝎MTx的生化研究概况已在前文作了一般的综述^[1]。本文将简要地介绍近年来对蝎ITx的一些研究进展及动态情况。

一、定义及生物鉴定

对蝎哺乳动物,昆虫及甲壳虫神经毒素(CTx)的划分是根据把一定剂量的来自蝎毒中的多肽物质注射进不同实验动物的体内,即

哺乳动物为小鼠,昆虫为苍蝇的幼虫和一种陆地生长的小甲壳虫(isopod)或给药在不同的离体动物标本中后,引起其中某一动物或离体标本的明显麻痹中毒反应,相应地将其物质定称为哺乳动物,昆虫或甲壳虫神经毒素。可是现有的研究资料表明,对蝎MTx和CTx的定义似乎并不严格分明,即这两类毒素对上述三种动物或离体的动物标本均有不同程度的毒性作用,只不过对哺乳动物或甲壳虫的麻痹作用更敏感些而已。而蝎ITx则比较专一性地作用于昆虫的神经系统^[2]。因此,相对而言,对蝎ITx的定义较为合理些。

蝎ITx的生物鉴定方法最早是由Zlotkin等人建立并用此法首次从非洲蝎*Androctonus australis* Hector毒中筛选到了一个昆虫神经毒素(AaH IT)^[3]。鉴定的材料为一种食肉的苍蝇幼虫,体重约100—200mg。此外,蟋蟀,蟑螂和蝗虫等活体昆虫也可用来鉴定蝎ITx。不

12 Banyal L, Varadi A, Patthy L. *FEBS Lett*, 1983; 163: 37
 13 Muller-Esterl W, Fritz H, Kellerman J *et al. FEBS Lett*, 1985; 191: 221
 14 Ripoche J A, Day J, Willis A C *et al. Biosci Rep*, 1986; 6: 65
 15 Sottrup-Jensen L, Folkersen J, Kristensen T *et al. Proc Natl Acad Sci USA*, 1984; 81: 7353
 16 DeScipio R G, Gehring M R, Podack E R *et al. Proc Natl Acad Sci USA*, 1984; 81: 7298

17 Stanley K K, Kocher H P, Luzio J P *et al. EMBO J*, 1985; 4: 375
 18 Derynck R, Roberts A B, Winkler M E *et al. Cell*, 1984; 38: 287
 19 Brown J P, Twardzik D R, Marquardt M *et al. Nature*, 1985; 313: 491
 20 Lee M Y, Huebers H, Martin A W *et al. J Comp Physiol*, 1978; 127B: 349

[本文于1989年1月3日收到]