

# 雌激素受体的结构与功能

靳嘉瑞 李载平

(中国科学院上海生命科学联合开放实验室)

(中国科学院上海生物化学研究所)

甾体激素是一类具有专一生理作用的重要活性物质。大约在 20—30 年前人们就已注意到甾体激素与蛋白质结合后可直接与 DNA 相互作用，而且还提出调节基因表达的设想。随后不少实验室分离、纯化了受体蛋白。但是有关受体蛋白在细胞浆和细胞核内存在的形式，它的性质和结构特点，激素受体蛋白复合物如何参与基因表达调控等问题仍然不十分清楚。由于近年来遗传工程技术和分子生物学研究的进步，才为这一领域的研究开拓了新的前景。

当前分子生物学研究的总目标之一是要阐明真核细胞在生长、发育和分化等重要生命过程中基因表达调控的程序是如何实现的。已知

在转录水平上这一调控过程是通过蛋白因子 (Trans-acting factors) 和 DNA 顺序上的专一调控元件 (Cis-acting elements) 相互作用来实现。甾体激素与受体蛋白结合所形成的复合物可以看作是一类蛋白质因子。它通过调节靶基因的转录来实现某些生理功能。因此有关甾体激素受体以及其他核受体家族成员的结构与功能的研究成为一个十分活跃的研究领域。

就雌激素受体 (ER) 而言，近几年在 Chambon 实验室继组建了它的 cDNA 克隆之后，进一步采用定向变异等手段，在研究分析功能区的定位，探讨激素-受体复合物是如何激活靶基因转录等方面取得了引人注目的进展。现将主要结果作一简要综述。

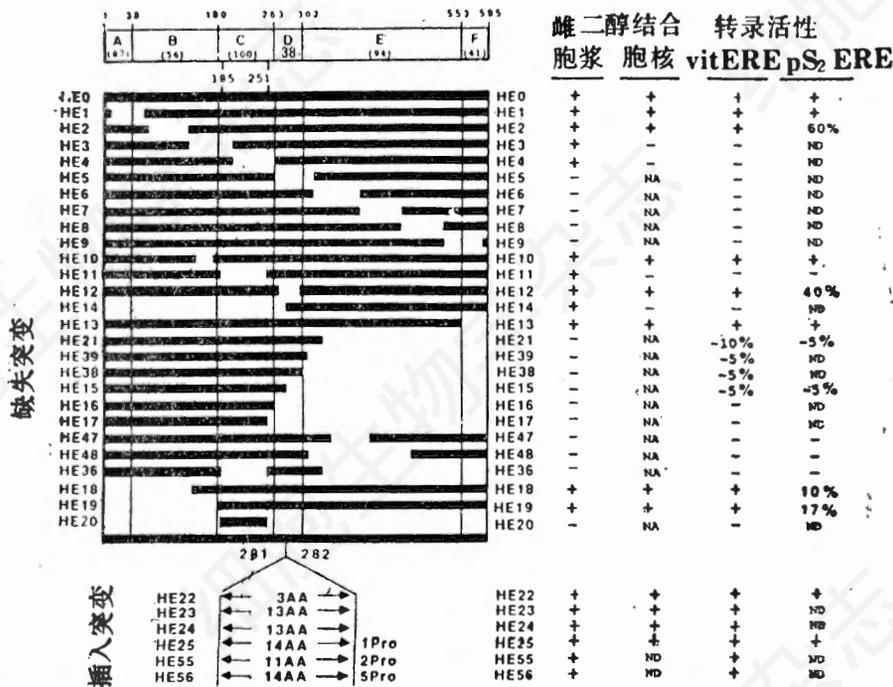


图 1 人和鸡雌激素受体氨基酸顺序的比较以及人雌激素受体基因插入和缺失突变对其功能的影响

## 一、mRNA 和蛋白的结构

从人乳腺癌 MCF-7 细胞株 cDNA 库中筛选出的 ER cDNA 克隆, 其 mRNA 所具有的阅读框架为 1785 个核苷酸, 能编码 595 个氨基酸。利用 SI 酶图谱法和引物延伸法可以测知人的 ER mRNA 的侧翼顺序以及转录起始点。同时发现在 -27 位置和 -103 位置处有类似 TATA 和 CAAT 的转录元件顺序<sup>[1,2]</sup>。

这个 mRNA 的翻译产物, 为 66 kD 的蛋白质。根据人和鸡的 cDNA 结构的比较, 可以区分为 6 个具有不同保守性的区域(图 1)<sup>[3]</sup>。A, C, E 三个区在人与鸡之间分别有 87%, 100%, 和 94% 的氨基酸是相同的; 而 B(56%), D(38%) 和 F(41%) 的保守性明显低一些。

人们根据保守性, 推测 C 和 E 可能是具有重要功能的区段。从结构上看, C 区是一个含有 Lys, Cys, Arg 相当高的亲水区。它可能是 ER 识别靶基因的部位, 是 DNA 结合区。E 区具有疏水区的特点, 能与甾体激素相结合。

雌激素受体与其他甾体激素受体、甲状腺激素受体和 erb A 等的顺序比较, 发现 ER 的 C 区和 E 区与 GR(糖皮质激素受体)、PR(孕酮受体) 和 erba 等大约有 45—91% 和 15—55% 的相似性。D 区则不同, 长度和顺序在所有各类受体间没有保守性。虽然 ER 的 A 区在人和大鼠, 人和鸡之间的相似性分别为 95% 和 87%, 但在不同的受体间并不具有多大的保守性。往往由于 A 和 B 区的长度不同而决定了某些受体的大小。

总之, 由于比较了各不同受体间的保守性, 人们得出的一个结论是, 这一类受体蛋白属于一个核受体家族, 能分别与配体结合后作用于靶基因, 调节基因的转录。目前发现属于这一家族的不仅包括了 C-erbA 基因, 还包括能结合维生素 A 酸的受体<sup>[4]</sup>, 以及最近发现的在人癌肝细胞内与 HBV 基因整合位点相邻的

## 维生素 A 酸的新受体<sup>[5,6]</sup>

## 二、人 ER cDNA 的表达

将含有完整编码区的人 ER cDNA 插入真核细胞表达载体 pKCR<sub>2</sub> 组建成的质粒 HEO 能在 HeLa 细胞内表达, 用蛋白印迹法和免疫沉淀法测知其表达产物分子量大小与 MCF-7(人乳腺癌细胞)细胞内 ER 相同, 而且与雌激素结合能力也相似。如果把 ER cDNA 插入表达载体 Bluescribe 的 T<sub>7</sub> 启动子下游, 在体外合成的 mRNA 可以在兔网织红细胞系统中翻译成 66 kD 蛋白, 它同样也能与雌二醇结合<sup>[7]</sup>。

此外, 人的 ER 基因在酵母细胞中的表达也获得成功<sup>[8]</sup>, 这为进一步研究 ER 的结构功能提供了较大量的 ER, 为研究真核细胞基因调控因子提供了一个很好的系统。

## 三、E 区是与激素结合的部位

E 区的非保守部分对于与配体的专一性结合是十分重要的。E 区的缺失突变如 HE<sub>15</sub>, HE<sub>6</sub>, HE<sub>7</sub>, HE<sub>8</sub> 和 HE<sub>9</sub> 失去了与 E<sub>2</sub>(雌二醇)的结合能力(图 1)。相反, HE<sub>1</sub>—HE<sub>4</sub> 或是 HE<sub>12</sub>—HE<sub>13</sub> 的缺失突变并不影响与 E<sub>2</sub> 的结合能力。此外, 保留完整 E 区的 HE<sub>14</sub> 基因在 HeLa 细胞中的表达产物具有与 E<sub>2</sub> 结合的能力。这就进一步说明了 E 区具有与激素结合的功能, 同时也证明了这一功能并不依赖于 ER 的 N 端。总之, 上述实验结果告诉我们, ER 的第 301 至 552 这一区段的氨基酸是能与 E<sub>2</sub> 结合的疏水区, 只是在内部缺失的情况下才失去与 E<sub>2</sub> 的结合能力<sup>[9]</sup>。

## 四、C 区能与 DNA 结合

完整的 ER 能与 DNA 结合, 但只有与 E<sub>2</sub> 结合后所形成的激素受体复合物才与 DNA 形成较紧密的结合。例如, 在没有 E<sub>2</sub> 存在时, 发现经 ER 基因转染的 HeLa 细胞中, 至少 80% 的 ER 是在胞浆内, 而在 E<sub>2</sub> 存在时, 大

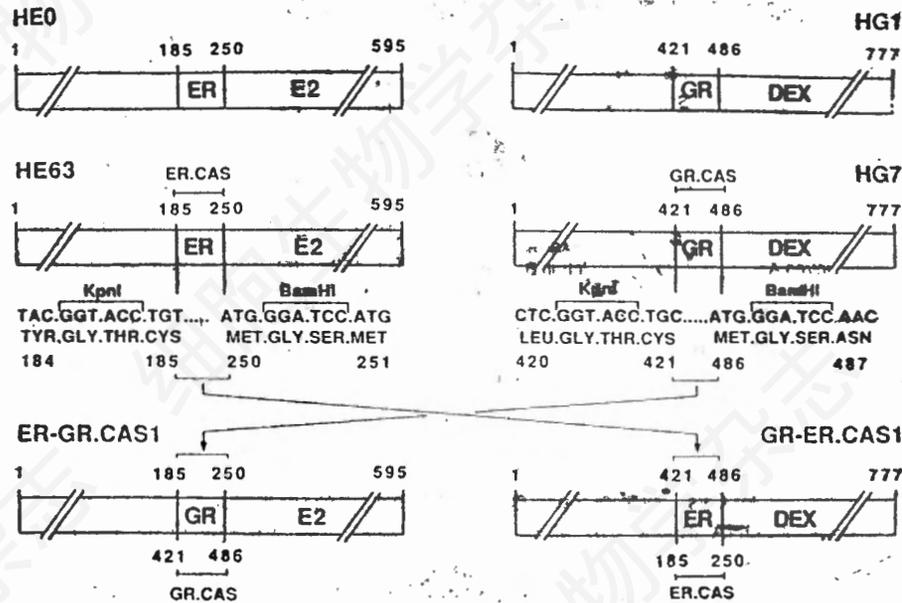


图2 嵌合体基因组示意图

HEO 和 HG1 分别为人的 ER 和 GR 的野生型。HE<sub>63</sub> 和 HG<sub>7</sub> 分别在相应的顺序中引入新的酶切点 KpnI 和 BamHI, ER-GR.CAS1 和 GR-ER.CAS1 为交换 ER 和 GR 的 DNA 结合区而形成的新嵌合体基因

约 90% 受体 E<sub>2</sub> 复合物能紧密与细胞核结合在一起。用缺失突变的实验也可说明 C 区是与 DNA 结合的部位, 如 HE<sub>3</sub> 和 HE<sub>4</sub> 的 C 区部分缺失后就不能与核 DNA 结合, 也不能激活靶基因的转录。比较 HE<sub>10</sub>、HE<sub>11</sub> 和 HE<sub>12</sub>, 只有 HE<sub>11</sub> 缺失 C 区核心部位第 185 至 251 共 66 个氨基酸就不能激活靶基因的转录。这表明保守的富 Cys 区在功能上是十分重要的。一系列有关 C 区突变的结果说明它是 ER 与 ERE (ER 结合的专一 DNA 顺序, Estrogen Response Element) 结合的重要部位 (图 1)。

利用嵌合体基因的研究结果, 为上述的推断提供了更有力的证据<sup>[9,10]</sup>。实验是按如下设计进行的, 把位于 ER 的 C 区富含 Cys 的第 185—250 氨基酸用 GR 的 C 区第 421—486 氨基酸所取代, 仍保留 ER 其他部位的完整性。在转染的 Hela 细胞中, 该嵌合体受体能激活带有 GRE (与 GR 结合的 DNA 顺序) 的 CAT (氯霉素转乙酰基酶) 指示基因的表达, 不激活带有 ERE 的 CAT 指示基因的表达。因此, 这

一段被交换的 C 区是决定激素受体复合物识别靶基因的专一元件 (图 2)。

进一步分析 C 区的结构, 可以推想, 其富含 Cys 的特征可能会形成稳定的指形结构<sup>[11]</sup>。这样一个 Cys—X<sub>2</sub>—Cys—X<sub>16</sub>—Cys—X<sub>2</sub>—Cys 图形与某些真核细胞转录因子的指形结构的特征氨基酸的顺序相似, 如爪蟾 5 S RNA 基因的转录因子 TF IIIA, 以及大家所熟知的 SP<sub>1</sub> 因子等也都具有这种所谓的“锌指结构” (Zinc Finger) (图 3)。

最近还提出 ER C 区 66 个氨基酸又可分为两个部分, 即 CI 和 CII<sup>[11]</sup>。CI 的指形结构能识别专一 DNA 顺序, 其中近 C 端的谷氨酰胺、甘氨酸、丙氨酸又是决定识别专一性的关键部位<sup>[12]</sup>。而 CII 是第二个指形结构, 它是非专一性的 DNA 结合区。由凝胶滞留法 (gel retardation) 证明结合在 ERE 上的 ER 是二聚体<sup>[13]</sup>, CII 很可能是 ER 分子间相互作用所需要的<sup>[10]</sup>。对稳定 ER 与 ERE 的结合起重要作用。

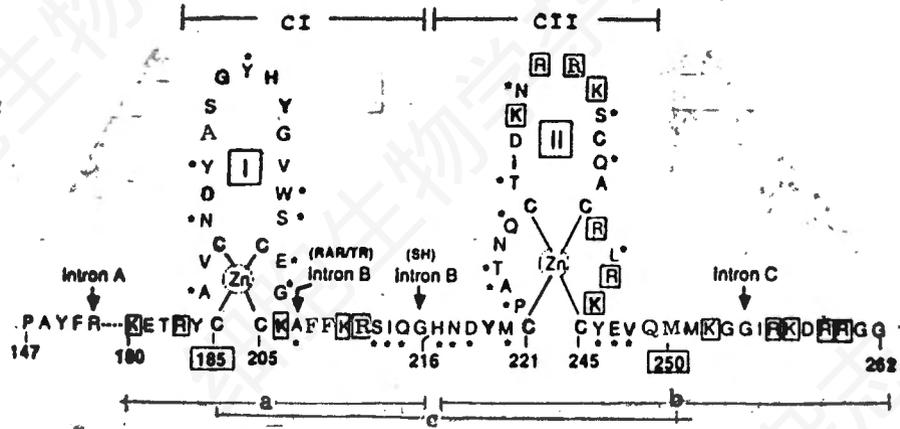


图3 根据ER的C区氨基酸组成区分的CI和CII区以及可能存在于该二区的Cys.残基与Zn<sup>2+</sup>四面体配位所形成的“锌指结构”

a. 专一的DNA结合区 b. 非专一的DNA结合区 c. ER的C区66个氨基酸核心区

五、C区和E区都是激活转录所必需的基本元件

C区虽然可以专一的识别靶基因的ERE, 但是仅有这一保守的66个氨基酸, 如HE<sub>20</sub>, 并不能激活指示基因的表达(图1)。HE<sub>15</sub>包含ER的A、B、C及部分D区虽然也同样保留了识别靶基因的功能, 但只有相当于HEO的5%的激活转录的能力<sup>[7]</sup>。用蛋白印迹法或免疫沉淀法, 以及与HEO竞争的实验都说明: HE<sub>15</sub>基因能在细胞内表达并进入细胞核内, 但激活靶基因表达的能力却明显地降低了, 这是由于失去了E区所致<sup>[9]</sup>。

酵母GAL<sub>4</sub>是一个已经了解比较透彻的转录因子, 它可以在Hela细胞中表达并识别它的DNA专一顺序。利用这个系统可以进一步分析ER分子E区激活转录的功能。包括E区在内的ER C端与GAL<sub>4</sub>的DNA结合区组建而成的嵌合体基因(图4), 在有E<sub>2</sub>存在下可以激活指示基因的转录, 从而说明在ER的E区具有激活基因转录的功能。但是这一区域的多肽结构并不含有像其他转录因子所具有的酸性氨基酸集团(Acidic Blob)或是两性螺旋(amphipathic helix)的特点。它的结构特点和

激活转录的机制仍需进一步的研究<sup>[14,15]</sup>。

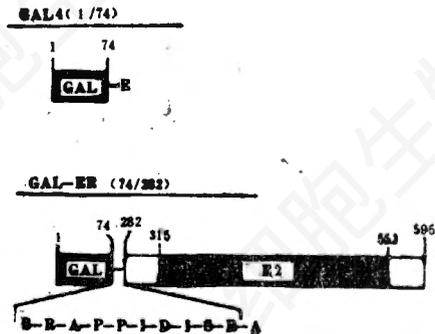


图4 酵母GAL<sub>4</sub>蛋白的DNA结合区第1—74个氨基酸及GAL DNA结合区与ER的C端(包括部分D区及E、F区)所组建的嵌合体受体基因

总之, ER的N端的DNA结合区(DBD)和C端的激素结合区(HBD)是两个重要的功能区。后者不仅是与激素的结合部位, 而且参与ER二聚体的形成。可见这两个重要的功能区段只有相互协同作用, 才能使ER紧密结合ERE并激活指示基因的转录活动。但是, 它们又是分别独立的两个元件。如果把含有DBD与HBD的ER分子的这两个部分反向连接而重组成的受体基因HD, 如图5所示, 同样具有激活指示基因转录的活性。同时, 这一结果也说明两个重要的功能区的相对位置对维持受体

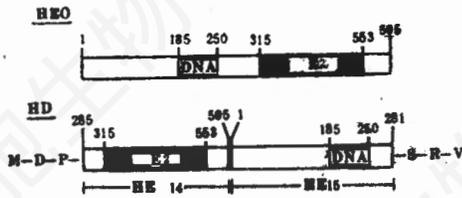


图 5 人 ER 的 N 端区与 C 端区反向连接组成的“反向”ER 受体 HD、HEO 为野生型 ER

基因实现其功能并不十分重要。

### 六、ER 的 N 端 AB 区在激活转录中的作用

ER 的 N 端 AB 区占有 1—179 个氨基酸。无论是部分缺失还是 AB 区全部缺失的突变都对 ER 激活转录的功能没有很大的影响。如 HE<sub>2</sub> 及 HE<sub>18</sub> 和 HE<sub>10</sub> (图 1) 在 HeLa 细胞中的表达产物仍能与激素和核 DNA 紧密结合, 并激活转录。可是如果用 PS<sub>2</sub>-CAT 作为指示基因则其效率大受影响, HE<sub>2</sub> 则只有相当于 HEO 的 60% 作用。可见 AB 区是对某些启动子如 PS<sub>2</sub> 基因启动子的激活是重要的<sup>[1]</sup>。在鸡胚成纤维细胞的实验中还进一步证实 AB 区是有细胞专一的激活某些启动子转录的功能。在 ER 的 E 区和 AB 区分别具有激活转录的作用, 但在氨基酸组成性质上是与 GR 不同的二个功能区<sup>[10]</sup>。

### 七、D 区的作用

D 区包括 39 个氨基酸, 连接着重要的 C 和 E 区。利用插入或缺失突变来研究 D 区的作用, 发现插入 3 个至 14 个氨基酸的突变株, 并不影响 ER 的功能, HE<sub>12</sub> 是在 D 区缺失 30 个氨基酸, 仍能与 E<sub>2</sub> 结合产生 ER-E<sub>2</sub> 复合物, 并与核 DNA 结合而激活基因的转录。由此看来通过插入或缺失改变 D 区的长度都不会影响到 ER 的功能。这说明 HBD 和 DBD 之间的距离对 ER 行使功能没有什么影响<sup>[9]</sup> (图 1)。

### 八、雌激素受体的作用机制

ER 作用在靶基因的 ERE, 而启动转录活动是需要先与 E<sub>2</sub> 结合。目前对它的解释是, 没有与 E<sub>2</sub> 结合的 ER, 它的激素结合区起着“掩盖”(masking) DNA 结合区的作用, 从而影响 ER 与 DNA 的结合。所谓“掩盖”可能是本身构象上的原因或是由于某些蛋白的作用, 如 90 kD 的热休克蛋白若与 ER 的激素结合区相互作用, 则可能阻止 ER 与 ERE 的结合或是形成二聚体。在加入 E<sub>2</sub> 后则使激素结合区发生构象的改变, 90 kD 蛋白的释放使 DNA 结合区去“掩盖”, 同时也会导致 ER 的二聚作用。至于 E<sub>2</sub> 本身可能在激活基因转录活动中的作用也值得重视。

ER 复合物作用在靶基因的 ERE 调节基因转录活动, 可以作为一个增强子结合因子来探讨转录的调控机制<sup>[17]</sup>。该因子通过与其他转录因子之间相互作用形成转录起始复合物, 在 RNA 聚合酶 II 的参与下与相应的专一顺序相互作用起始基因转录。根据目前的有关研究结果说明, 转录机制具有一定的保守性, 如酵母的 GAL<sub>4</sub> 因子在哺乳动物细胞中能同样的激活 TATA 元件的转录, 而人的 ER 也可在酵母中激活相关基因的转录起始。所以说在酵母和哺乳动物细胞之间存在共同的调控因子和增强转录的机制。

对于 ER 作为一个受激素诱导的增强子因子的研究都是在暂时性表达系统中进行的, 究竟能在多大程度上反映 ER 基因在染色质水平上的作用机制, 目前仍是一个值得探讨的问题。

### 摘 要

人的雌激素受体根据其氨基酸顺序的保守性可以分成 6 个区, 其中最富于保守性的是 DNA 结合区和激素结合区。DNA 结合区负责与专一的 DNA 顺序结合。激素结合区不仅能与配体(雌激素)结合, 参与形成二聚体, 而且具有激活靶基因转录活动等重要功能。雌激素

受体蛋白-激素复合物可以被认为是一个受配体诱导的转录因子,它与具有增强子功能的DNA顺序结合后调节靶基因的转录活动。

### 参 考 文 献

- [1] Watter, P. et al., 1985, *Proc. Natl. Acad. Sci.*, 82:7889—7893.
- [2] Green, S. et al., 1986, *Nature*, 320:134—139.
- [3] Krust, A. et al., 1986, *EMBO. J.*, 5: 891—897.
- [4] Petkovic, M. et al., 1988, *Nature*, 330: 444—450.
- [5] de The, H. et al., 1987, *Nature*, 330: 667—670.
- [6] Brand, N. et al., 1988, *Nature*, 332:850.
- [7] Kumar, V. et al., 1986, *EMBO. J.*, 5: 2231—2236.
- [8] Metzger, D. et al., 1988, *Nature*, 334: 31—36.
- [9] Kumar, V. et al., 1987, *Cell*, 51:941—951.
- [10] Green, S. and Chambon, P., 1987, *Nature*, 325:75—78.
- [11] Green, S. et al., 1988, *EMBO. J.*, 7: 3037—3044.
- [12] Mader, S. et al., 1989, *Nature*, 338: 271—274.
- [13] Kumar, V. and Chambon, P., 1988, 55: 145—156.
- [14] Webster, N. et al., 1988, *Cell*, 52:169—178.
- [15] Webster, N. et al., 1988, *Cell*, 54:199—207.
- [16] Tora, L. et al., 1989, *EMBO. J.*, 8: 1981—1986.
- [17] Green, S. and Chambon, P., 1988, *Trends Genet*, 4:309—314.

## 果蝇前后图式基因调控的层次性\*(上)

赵 德 标

(中国科学院上海细胞生物学研究所)

### 一、引 言

由于分子遗传学的进步,对于基因如何代代相传以及个别基因的表达调控,我们已有相当的了解。然而对发育过程中,基因如何按一定的时空秩序,依次表达,并导致性状的发育,则所知甚少。这方面的研究还刚开始深入。已有人提出不同的理论模型,以解释在发育的时间和空间,不同基因的转录如何被启动和关闭,使胚胎发育得以不断进行<sup>[1-3]</sup>。但所假设的调节分子及其作用机理并未得到实验的证实。只是近年来,主要由于对果蝇早期胚胎的发育遗传学研究的长足进展,才使我们对胚胎发育中基因表达的调控有了比较深入和系统的了解。

70年代初,通过观察致死胚胎的表皮等结构图式,开始系统地寻找影响果蝇胚胎早期

发育的突变<sup>[4-7]</sup>。至今发现的突变中,许多可使胚胎极性和体节图式发生不同改变,从而使研究发育过程中,基因表达的时空调节规律成为可能。近三年来,由于几项重要技术的应用,使果蝇发育遗传学研究发展极为迅速。这些技术包括多线染色体上一定区域DNA的微克隆(DNA-microcloning),生殖细胞转化,多线染色体DNA和胚胎RNA原位杂交,胞质移植,胚胎整体标记,以及微注射技术等。这些技术的结合,有可能揭示不同突变基因在胚胎发育中转录、表达和功能,以及它们之间相互作用<sup>[8,9,19,24,58]</sup>。

本文主要综述影响果蝇胚胎前后极性和体节图式的突变基因,着重介绍这些基因在胚胎

\* 庄孝儒教授对本文的整体构思和主要观点的阐明提出了宝贵意见,并对本文作了多次修改,特表谢意。