

噬菌体 MS2 RNA 中的基因重叠与裂解基因

单链 RNA 噬菌体共分为 4 个血清类群。第一类 RNA 噬菌体基因组 (包括 MS2、R 17 和 f2 等) 含有 4 个蛋白质编码的信息。1976 年 Fiers 等发表 MS2RNA 全序列 (3569 个核苷酸) 后, 提出其中三个基因——A 蛋白基因、(130—1308)、外壳基因、(1335—1724) 和合成酶基因 (1761—3491)。最近, Fiers 小组又确定了 MS2 RNA 中的第 4 个基因——裂解基因 (Lysis gene, 简称 L 基因) 的位置并研究了此基因的表达方式。L 基因是 MS2RNA 中 1678—1902 的一段核苷酸, 与外壳基因和合成酶基因重叠, 其蛋白产物与寄主细胞的裂解有关。L 基因的翻译取决于外壳基因的翻译, 它不能单独与核糖体结合形成翻译起始复合物。在核糖体沿外壳基因从 5'→3' 端前进, 到达 L 基因之前的几个密码时, 由于遇到两个因移码而产生的终止密码, 导致核糖体移框, 从而使外壳蛋白的合成终止, 裂解基因随之起译。

Fiers 等在 L 基因表达的研究中使用了 DNA 重组法。这比直接使用 MS2 RNA 的方法具有两个优点: (1) 可以得到任何位点的点突变体; (2) 可以研究单一基因的表达。

用限制性内切酶把通过反转录得到的 MS2 DNA 切割成各种大小的亚基因组, 用质粒 pla 831 作载体构建成各种 pMS (plasmid-MS2) 重组质粒。在重组质粒 pMS20 中, 虽然 L 基因缺失了 58 个核苷酸片段 (1764—1822), 但仍含有完整的外壳基因和一部分合成酶基因, 实验证明 pMS20 的克隆不发生细胞裂解。在 pMS 16 中, 除去了大部分 A 蛋白基因, 在 pMS 17 中, 完全除去了 A 蛋白基因, 这两个质粒的 L 基因和外壳基因是完整的, 并带有部分合成酶基因。它们的克隆都能使细胞裂解。由以上三个质粒的实验说明: (1) 所确定的 L 基因位置是正确的; (2) A 蛋白基因与细胞裂解无关。

再切割 pMS17 建成 pMS1 (只含 1628—2057 一段核苷酸)。当 pMS1 的外壳基因大部分被除去后, 就丧失了细胞裂解的能力。其可能的解释是 (1) L 基因受外壳基因的极性控制; (2) 外壳基因本身是裂解细胞所需要的。为了证实上述的哪一种解释是正确的, 他们再改建 pMS1 即把外壳基因插到 pMS1 的 L 基因, 构建成 pMS1.9。结果发现带 pMS1.9 的克隆不发生裂解作用, 但凝胶电泳分析证实此克隆有外壳蛋白产生。说明 L 基因表达需要有天然基因排列顺序; 外壳蛋白本身与细胞裂解无关, 支持了“外壳基因对 L 基因的表

达起极性控制作用”的解释。

除去质粒 pMS25 (带有 869—2057 一段序列, 但其中 1013—1365 有 352 个核苷酸缺失) 中外壳基因的起始区, 它的克隆就不发生裂解。即使在 pMS25 的 L 基因后再连接上完整外壳基因, 细胞裂解也不再发生。这排除了 pMS 25 裂解细胞能力的丧失是由于外壳基因起始区缺失的可能性, 并说明 L 基因的表达需要 L 基因与外壳基因部分重叠, 进一步支持“L 基因表达受外壳基因极性控制”的观点。

根据 MS2 RNA 中无意义密码子存在和 MS2 RNA 二级结构的证据, 曾提出极性控制分子基础的主要特征是在合成酶起始区和在外壳蛋白编码区的核苷酸序列之间存在着碱基配对。当核糖体通过这个区域时, 暴露了合成酶的核糖体结合位点, 使合成酶的翻译起始。在与外壳基因相重叠的 L 基因的表达中却具有与众不同的机制。在体外构建了一系列含有 MS2 DNA 移框突变的质粒。凡是含有 +1 移框顺序并由此造成核糖体在 1672 位核苷酸处遇到 UAA 终止码的质粒, 都有诱导细胞裂解的能力; 凡是含有 -1 移框顺序并由此造成在 1672 位核苷酸处无 UAA 终止码出现的质粒, 都无裂解细胞的能力。此外还观察到, 外壳基因的质粒, 由于整个密码子的缺失不产生移框, 使细胞不能裂解。以上实验结果说明, L 基因没有独立的与核糖体结合的起译区。核糖体只在外壳基因起始区起始, 然后在 L 基因之前的 UAA 处终止外壳蛋白翻译, 发生移框, L 基因随之起译。这样, 一个核糖体在不离开模板的情况下发生两次起译。

L 基因的表达水平也受起始码 AUG 与终止码 UAA 之间距离的控制。距离越短, 核糖体再度起译的几率就越大。

MS2 RNA 的移框模型是否适用于其它重叠基因尚属未知, 但已发现 MS2 RNA 与 $\phi \times 174$ 和 G4DNA 重叠基因的类似之处是它们都发生 +1 移框, 并由此造成 UAA 终止码的出现。

基因重叠的意义一方面是在有限的 RNA 或 DNA 中储存更多的遗传信息, 另一方面是代表了一种新的基因表达调控方式—移框调节机制。使 MS2 感染的细胞中只有累积了足够量的外壳蛋白时才开始合成裂解蛋白, 从而保证噬菌体在寄主菌中得以大量增殖。

[Nature, Vol. 295, 35, 1982.]

(张继仁 摘译)