

乳糖操纵子模型的建立与教学中若干问题的解析

邢万金

内蒙古大学生命科学学院, 呼和浩特 010070

摘要: 基因结构和表达调控机制是现代生命科学的研究热点和焦点。乳糖操纵子是大肠杆菌(*Escherichia coli*)分解代谢乳糖的一簇基因, 其基因组成与表达调控方式是最早被阐明的基因结构与调控机制, 因而成为微生物学、遗传学和分子生物学等多门专业课程讲解基因调控机制的经典教学案例和要求重点掌握的内容, 备受师生们的重视。该知识点虽然结论简单, 记忆容易, 但由于触及生命结构与功能的核心机制, 内涵丰富, 逻辑深奥, 理解困难。教师要充分发挥该教学案例的效果并非易事, 需要深入了解乳糖操纵子的基因结构和工作原理, 特别是科学家揭秘这些奥秘的科学背景和思维过程。本文通过回溯大肠杆菌乳糖操纵子发现和表达模式解析的历程, 追随 J. Monod 和 F. Jacob 等前辈名家的脚印, 聆听他们对实验结果的分析, 学习他们的科研思想和创新思维, 结合乳糖操纵子的 DNA 序列, 深入分析和理解乳糖操纵子表达的若干奇特现象的原因, 共同探讨如何充分发挥遗传学和分子生物学经典案例的教学价值。

关键词: 遗传学; 案例教学; 乳糖操纵子; 基因调控

The establishment process of lac operon model and the analysis of several teaching problems

Wanjin Xing

School of Life Sciences, Inner Mongolia University, Hohhot 010070, China

Abstract: Gene structure and expression regulation mechanism are the research hotspots and focus of modern life sciences. The lac operon is a cluster of genes through which *Escherichia coli* catabolizes lactose. It was first proposed by F. Jacob and J. Monod, who were also awarded the Nobel Prize in Physiology or Medicine in 1965 for their contributions. Thereafter, the lac operon has become the classic teaching case of the gene regulation mechanism in microbiology, genetics and molecular biology, and been highly valued by teachers and students alike. Although the conclusion is easy to follow and memorize, its rich connotation and esoteric reasoning has rendered it difficult to understand, neither is it easy for teachers to fully exploit the advantages of this teaching case. Therefore it is necessary to have an in-depth

收稿日期: 2019-03-27; 修回日期: 2019-05-13

基金项目: 遗传学课程群教学团队建设项目, 内蒙古大学精品资源共享课程建设项目, 混合式教学课程建设项目和研究生精品课程建设项目资助[Supported by the Teaching Team Construction Project of Genetics Course Group and the Curriculum Construction Projects of Inner Mongolia University]

作者简介: 邢万金, 博士, 教授, 研究方向: 细胞分子生物学。E-mail: xwanjin@imu.edu.cn

DOI: 10.16288/j.ycz.19-084

网络出版时间: 2019/5/24 7:27:00

URL: <http://kns.cnki.net/kcms/detail/11.1913.R.20190524.0726.002.html>

understanding of the genetic structure and working principle of the lac operon, especially the scientific background and thinking process through which scientists revealed these mysteries. In this paper, the historical discovery and analysis process of the *E. coli* lac operon was reviewed by following their footprints, listening to their analysis of experimental results. Based on the DNA sequences, the reasons for several unusual phenomena of lac operon expression were also discussed to exemplify the teaching value of the classic cases in genetics and molecular biology.

Keywords: genetics; case teaching; lac operon; gene regulation

从 Jacques Monod 观察到细菌在糖混合培养基中的两期生长模式, 至确认细菌产生 β -半乳糖苷酶并非原有的前体酶“适应”, 而是受诱导后的“从头合成”, 到提出乳糖操纵子模型, 法国巴斯德研究所的几位科学家长期专注于一个科学问题, 最终阐明了代谢乳糖的基因簇的结构和诱导表达机制。

在当今的高校本科教学中, 限于学时数和掌握的原始文献及了解的科学史, 教师在讲授乳糖操纵子模型时, 往往只是简单地介绍结论。教师的授课目标局限于给学生介绍一种原核基因表达调控机制, 学生的学习任务量化为记忆“调节基因”、“阻遏物”、“结构基因”和“操纵子”等几个名词, 以及乳糖、阻遏物与操纵子之间的简单相互作用, 而未能触及针对乳糖操纵子模型的创新性科研思维, 因而未能充分发挥乳糖操纵子案例的教学价值。

目前, 全社会和教育主管部门都在提倡培养学生的科研素质和创新能力, 这就要求教师更新教学理念, 改革授课方式。在课程教学过程中, 在讲解每个知识点的时候, 不仅要介绍原理和结论, 更要讲解科学家发现原理和推出结论的科学思维过程^[1,2], 而后者才是教学的重点。要了解创新知识点的来源、历程及其研究思路与方法, 只有系统梳理文献, 重现历史。本文继续秉承用科学史引导教学的理念^[3], 系统梳理乳糖操纵子的发现过程与科学家们的解析思路, 深入挖掘乳糖操纵子案例对于培养学生科研思维和分析能力的教学价值。

在讲解乳糖操纵子模型时, 通常会遇到几个逻辑悖论。例如, 乳糖操纵子模型认为细菌合成 β -半乳糖苷酶和透性酶需要先有乳糖进入细胞才能诱导编码这两个酶的基因表达, 但细菌细胞膜上先要有透性酶乳糖才能进入细胞; 进入细胞的乳糖分子本身并不能诱导操纵子表达, 必须先被 β -半乳糖苷酶

转化为异乳糖才能发挥诱导作用。即表达 β -半乳糖苷酶和透性酶需要乳糖诱导, 而乳糖反而需要先有这两种酶才能进入细胞并被转化成异乳糖, 由此就形成了逻辑上的怪圈, 到底是细菌先摄入乳糖, 后合成 β -半乳糖苷酶和透性酶? 还是先合成 β -半乳糖苷酶和透性酶, 后摄入乳糖以发挥其诱导能力呢? 此外, 乳糖操纵子模型认为该操纵子应该转录出一条完整的多顺反子 mRNA, 那么翻译出的 3 种酶应该几乎等量, 但实际上人们发现细胞内的转乙酰酶远比 β -半乳糖苷酶少, 这是为什么呢?

上述若干疑问是乳糖操纵子案例教学中容易被忽略但又必须直面的问题。其实这些问题不仅仅是教学问题, 也是深入研究基因调控分子机理的切入点, 因而早已被众多研究者所关注, 并设计实验进行了解答, 只是目前的教科书上鲜有对这些问题的提及和系统论述, 这就要求任课教师在备课时对此有所钻研。

深入挖掘教学案例涉及的科学问题和寻找问题的答案应该成为研究型教学改革为抓手, 惟其如此, 才能体现教师授课的价值, 也促使高校教师在教学过程中开展科研活动, 不断提升自身的学术修养, 在课堂上培养学生发现问题和解决问题的科研创新能力。本文以大肠杆菌 K12 基因组(LN832404.1)的乳糖操纵子基因序列为蓝本, 在 DNA 序列水平上分析了乳糖操纵子的工作原理和特点, 通过广泛查阅研究文献, 给出上述疑问的答案。

1 乳糖操纵子的发现

1.1 “酶适应”假说

1882 年, 德国真菌学家 Julius Wortmann^[4]报道

了细菌“*Bacterium termio*”只有生长在含淀粉的培养基中才产生淀粉酶(amylase)。尽管文中同时报道了无论培养基中是否有蔗糖,酵母都始终产生蔗糖酶(invertase),他发现的培养基成分会影响微生物产生某种酶的能力的现象还是引起了广泛的关注和跟进研究。如 1899 年,巴斯德研究所的 Emile Duclaux 发现灰绿曲霉(*Aspergillus glaucus*)和灰绿青霉(*Penicillium glaucum*)两种真菌能否产生蔗糖酶或酪蛋白酶(casease),取决于培养基中是否含有蔗糖或牛奶^[5]。赫尔辛基大学的 Henning Karstrom 观察到大肠杆菌能否分解乳糖也取决于培养基中是否含有乳糖^[6]。随着研究案例的积累,人们逐渐形成共识:微生物能否产生某种酶在很大程度上受培养基中是否存在此酶的底物的影响。Henning Karstrom 把这种现象称为酶适应(enzymatic adaptation),并把细菌分解糖的酶分为两类:一类为适应型(adaptive)酶,只有在底物存在时才能在细菌细胞中检测到该酶的活性;另一类为组成型(constitutive)酶,无论底物是否存在,在细菌细胞中都能检测到该酶的活性^[7]。

组成型酶可以理解细菌生命活动所必需的代谢反应催化剂。但适应型酶似乎只针对环境中的营养物质,这种现象的机理是什么呢?为了解释酶适应现象,剑桥大学的 John Yudkin^[8]提出了“mass action”假说,认为细菌细胞内的酶以两种形式存在:一种是“无活性的前体形式”(precursor, P),另一种是“有活性的酶形式”(enzyme, E),二者处于化学平衡状态。如果某种酶是适应型酶,在无底物(substrates, S)时平衡倾向于 P 形式;有底物时,酶与底物形成“酶-底物复合物”(E-S),使反应平衡转向 E 形式(图 1)。

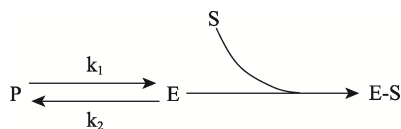


图 1 适应型酶活性改变的“mass action”假说

Fig. 1 The mass action hypothesis of “adaptive enzyme”

P: 无活性前体; E: 有活性的酶; S: 酶的底物; E-S: 酶-底物复合物; k_1 和 k_2 代表不同的反应速度常数。前体(P)与酶(E)处于化学平衡状态。加入的底物(S)会与酶(E)结合,打破前体-酶(P-E)平衡,导致更多的前体(P)转化为酶(S)。

1.2 Jacques Monod 的发现与思考

1941 年,巴斯德研究所的 Jacques Monod^[9]在博士毕业时报道,如果用两种糖,如葡萄糖和木糖(xylose)、葡萄糖和鼠李糖(rhamnose),作为碳源培养大肠杆菌,绘制的大肠杆菌生长曲线呈现出被一个明显的静止期分割开的两个指数生长期(图 2)。Monod 用“double growth”杜撰了一个名词“diauxie”来描述这种现象。当他就这个结果与导师 André Lwoff 讨论时,Lwoff 提醒他这种生长现象可能与酶适应有关。

Monod 当时未曾注意过酶适应。他随即查阅文献,认真研究了前人关于酶适应的论述,认为 Yudkin 的“mass action”假说虽然有道理,但并不能完全解释自己的实验结果,因为随着底物被适应型酶分解的越来越少,细菌的生长也应该越来越慢,而不应该是 S 型曲线(先快速增加,后停止,后再快速增加)^[10]。

另一方面,早在 1900 年,E. Duclaux 的博士生 Frédéric Diénert^[11]在其博士论文中报道,当在酵母培养基中加入两种糖(如葡萄糖和半乳糖)时,酵母不能同时产生分解这两种糖的酶,即一种糖的存在似乎会抑制酵母产生分解另一种糖的酶,Diénert 把这种现象称为葡萄糖效应(glucose effect)。葡萄糖效应提示细菌可能只产生一种酶,两种底物在细胞内竞争这种酶,这种酶分解不同的底物时就成为不同的酶。Monod 的细菌实验结果类似于这一现象,于是

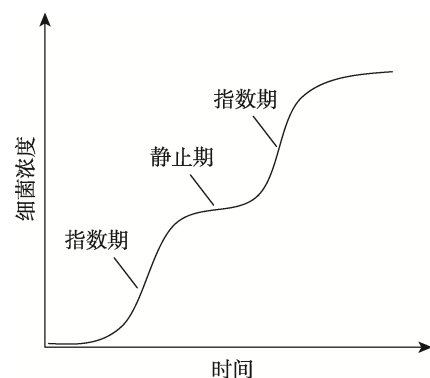


图 2 生长在葡萄糖和木糖混合培养基中的大肠杆菌的两期生长曲线

Fig. 2 Diauxic growth curve of *E. coli* in medium with glucose and xylose

生长曲线的两个指数期被一个静止期分割。

Monod^[10] 1942 年在 Diénert 的底物竞争假说和 Karstrom 的酶适应假说的基础上, 提出第一种糖是被细胞里已经存在的“组成型酶”分解, 第二种糖是被原先细胞内不存在的“适应型酶”分解。Diauxie 曲线的静止期可解读为细菌细胞内的某种酶在第一个指数生长期分解完第一种糖后, 需要一个缓慢的适应过程才能被第二种糖诱导去分解第二种糖。

但 Monod 的研究被德国入侵法国中断了, 他随即投笔从戎。1945 年秋天, 二战一结束, Monod 就被 Lwoff 招回巴斯德研究所继续解析大肠杆菌 Diauxie 生长曲线的生理生化机制^[12]。

酶适应现象可理解为微生物细胞为了利用环境中的新营养物质分裂增殖而做出的生理生化改变。但华盛顿大学的 Sol Spiegelman^[13]发现二倍体酵母细胞即使不增殖, 也有酶适应现象, 且酶活性以 S 型曲线模式增加。为了解释这种 S 型曲线模式, Spiegelman 在 Yudkin 的“mass action”假说的基础上提出了一个“自我复制”(self-duplication)假说, 认为细胞内总是存在一定量的蛋白质作为形成某种酶的前体(P), P 以速度常数 k 转化为有活性的某种酶(E), 但 E 很不稳定, 会以更高的速度常数 k' 转变为 P, 因而细胞内的 E 很少。底物 S 能增加 E 的稳定性, 有过量的 S 进入细胞, 就会形成 ES, 从而抑制 k' , 使 P 不断地转化为 E。 k 的数量级受基因和 E 共同控制, 无 S 时, k 主要受基因控制, 一旦有 E 生成 k 就主要受 E 控制, 加速把 P 转化为更多的 E, 结果好像 E(酶)能够自我复制。

按照“自我复制”假说推导出的酶活性增加曲线确实是 S 型, 与 Monod 观察到的 S 型曲线模式相符。但这一假说要求承认细胞质中的酶一旦产生就可以脱离细胞核内的基因控制而进行自我增殖, 即酵母细胞质内存在不受细胞核控制的“细胞质基因”(plasmagenes)。Monod 认为这违背了孟德尔遗传, 因此强烈反对。Monod^[14]认为细菌“适应”某种糖的现象可能并非由前体转化为活性酶的过程, 而是底物诱导酶蛋白从头合成的过程。因为有些细菌虽然丧失了“适应”某种糖的能力, 但细胞内仍然存在着分解这种糖的微弱酶活性。为了系统准确地分析细胞内的酶活性变化及其实质, Monod 首先需要一种能直观地测量酶活性的方法。

1.3 β -半乳糖苷酶活性的测定

在 1950 年以前, 人们用乳糖、麦芽糖、精氨酸和色氨酸等多种营养物质诱导细菌来研究“酶适应”, 但都缺乏测定酶活性的便利生化方法。其中测定 β -半乳糖苷酶活性的方法是测定乳糖溶液吸光度的变化或者测定乳糖被分解后释放的 CO_2 量。1950 年, 威斯康星大学的 Martin Seidman 和 Karl Paul Link^[15]应同校的 Joshua Lederberg 博士的请求, 合成了邻硝基苯- β -D-半乳糖苷(o-nitrophenyl- β -D-galactoside, ONPG)。 β -半乳糖苷酶能水解 ONPG, 生成黄色的邻硝基酚。Lederberg^[16]建立了用 ONPG 作底物测定 β -半乳糖苷酶活性的新方法, 并发现在用乳糖培养的大肠杆菌中存在强烈的 β -半乳糖苷酶活性, 而在用其他糖类培养的大肠杆菌中却只有非常微弱的 β -半乳糖苷酶活性, 说明乳糖会诱导 β -半乳糖苷酶的活性。那么, 在用乳糖培养的大肠杆菌中, β -半乳糖苷酶活性激增的原因到底是原有的 β -半乳糖苷酶前体被诱导产生了酶活性, 还是在细胞内重新合成了更多新的 β -半乳糖苷酶呢?

1.4 “酶适应”假说的否定, “酶诱导”假说的提出

为了回答上述问题, 1951 年, Monod 等^[17]分离纯化了大肠杆菌 β -半乳糖苷酶, 免疫家兔制备了兔抗 β -半乳糖苷酶的抗血清, 并结合 ONPG 法检测了用乳糖培养的大肠杆菌的 β -半乳糖苷酶含量和活性。他们发现如果培养基中的碳源只有乳糖, 大肠杆菌的 β -半乳糖苷酶活性确实非常高, 是以其他糖为碳源的 1000 多倍! 但同时用抗血清分析的结果显示, 细胞内的 β -半乳糖苷酶含量增加了, 并非原有的无活性 β -半乳糖苷酶前体适应乳糖而成为有活性的 β -半乳糖苷酶, 这与 Yudkin 的“mass action”假说矛盾, 说明细胞内 β -半乳糖苷酶活性增加的原因是因为乳糖诱导大肠杆菌合成了新的 β -半乳糖苷酶, 即早先的“酶适应”说法是错误的, 应该是“酶诱导”。于是, Monod 用“诱导”(induction)一词取代了“适应”(adaptation), 来描述“受诱导剂激活的酶蛋白合成”(即早先的“酶适应”)现象。

1.5 β -半乳糖苷酶与其诱导物的关系

既然细胞内并非事先准备了大量的前体等待着底物将其转变为有活性的酶，而是细胞本来就具备重新制造某种酶的能力，底物只是诱导细胞发挥这种制造能力，那么，必须是酶的底物才能诱导细胞产生这种酶吗？乳糖(图 3A)是 β -半乳糖苷酶的底物之一，确实能诱导大肠杆菌合成新的 β -半乳糖苷酶。其他类似乳糖的半乳糖苷衍生物呢？Monod 等^[17]尝试了手头能找到的所有半乳糖苷衍生物，发现并非所有的 β -半乳糖苷酶底物都能诱导该酶，能诱导 β -半乳糖苷酶的半乳糖苷衍生物也不都是该酶的底物。如苯酰半乳糖苷是 β -半乳糖苷酶的底物之一，但无诱导能力；甲酰半乳糖苷不是该酶的理想底物，却有诱导能力；而苯酰硫代半乳糖苷(thiophenyl-galactoside)不是该酶的底物，不仅无诱导能力，反而能抑制其他诱导剂的诱导作用。这些结果提示诱导剂并不与酶直接作用，可能是通过细胞内某种其他组分间接发挥诱导作用，这也与 Yudkin 的“mass action”假说矛盾。

尽管诱导剂并非直接作用于所诱导的酶，比较几种诱导剂的分子结构和诱导能力，仍然可以看出诱导剂的诱导能力与其自身的分子结构有关。如异丙基硫代 β -D-半乳糖苷(IPTG)(图 3D)和苯酰硫代半乳糖苷都不是 β -半乳糖苷酶的底物(因为糖苷键的 O

原子被 S 原子取代)(图 3E)，但 IPTG 是个强诱导剂，而苯酰硫代半乳糖苷不仅不能诱导，反而会抑制别的诱导剂的诱导作用。显然糖苷类分子的糖苷配基参与了诱导作用^[18]。尽管不同糖苷配基的空间构象不同，但诱导作用可能相同，也说明诱导剂并不需要与其所诱导的酶直接结合。那么，诱导剂的空间构象与谁嵌合？如何发挥作用呢？

1.6 β -半乳糖苷酶的遗传控制

早在 1903 年，Lucien Cuenot 就提出了“一个基因一个酶”假说^[19]，后来斯坦福大学的 George Beadle 和 Edward Tatum 用 X 射线诱发链孢霉(*Neurospora crassa*)突变，在鉴定突变体的生化性质时也发现基因突变是影响链孢霉营养物质代谢途径的某一步，于是在 1945 年也提出了“一个基因一个酶”假说^[20]。按照这一假说，大肠杆菌的 β -半乳糖苷酶也应该有一个对应的基因。实际上，1946 年 Monod 和 A. Audureau^[21]研究一株从 Lwoff 的肠道中分离出的大肠杆菌时就发现这株菌不能在只有乳糖为碳源的培养基上生长(lac^-)。但他们在培养过程中发现一种突变导致它恢复了利用乳糖的能力(lac^+)，而且这株 lac^+ 突变菌在葡萄糖和乳糖混合培养基中表现出典型的两期生长曲线。“酶适应”表型可以由突变产生，意味着这种现象也受基因控制。

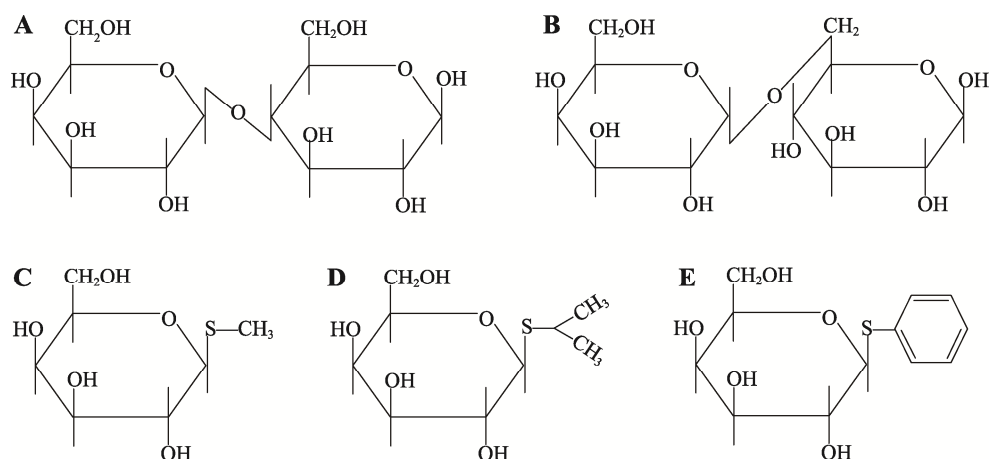


图 3 几种半乳糖苷衍生物分子结构

Fig. 3 Structural formula of several galactosidic derivatives

A: 乳糖(lactose); B: 异乳糖(allolactose); C: 硫代甲基- β -D-半乳糖苷(thiomethyl-beta-D-galactoside, TMG); D: 异丙基硫代 β -D-半乳糖苷(isopropyl-beta-D-thiogalactoside, IPTG); E: 苯酰硫代半乳糖苷(thiophenyl-galactoside)。

当时对酶和基因的结构、酶与基因的关系知之甚少。Monod^[14]认为酶的结构应该是个多聚体, 包含一个特异性的模块 B、一个非特异性模块 i、以及一个所有酶或蛋白质共有的组分, 而基因的作用就是操作这些模块, 用它们搭建出酶的活性部位, 即, 酶活性问题本质上应该是个遗传学问题, 应该用遗传学方法进行研究。但当时的遗传学研究方法主要是用二倍体有性生殖生物进行杂交, 然后分析杂交结果, 推断控制表型的基因和基因型, 尚无法用单倍体细菌进行杂交。

1.7 微生物遗传学的兴起

1946 年, 耶鲁大学 Lederberg 与 Tatum^[22]合作发现大肠杆菌能被处理产生营养缺陷型突变, 并能区分“雌雄性”, 通过接合(conjugation)产生重组型后代。这一发现暗示可以用类似真核生物杂交的方法研究大肠杆菌的基因和基因型。

1954 年, 巴斯德研究所 Lwoff 实验室的博士研究生 François Jacob 与 Élie Léo Wollman^[23]合作用高频重组(Hfr)雄性溶原菌与雌性非溶原菌进行杂交时发现将近一半的合子成为噬菌斑, 他们称之为合子诱导(zygotic induction), 提示溶原性大肠杆菌细胞内存在 λ 原噬菌体(prophage)与细菌的平衡机制, 而且这种平衡受溶原性细菌细胞中某种因子的维持, 所以原噬菌体不会裂解溶原性细菌, 但非溶原性细菌细胞内却没有这种调控因子, 因而会被裂解。那么, 溶原化调控因子的基因是什么? 在接合的过程中何时传到了雌性非溶原性细菌细胞中(从而导致非溶原性细菌裂解)呢?

为了揭示细菌接合过程中基因的传递细节, Wollman 用厨房的搅拌器定时搅动杂交菌液以中断它们的接合过程, 然后根据噬菌斑出现的时间推断控制溶原化的基因在什么时刻传到了雌性非溶原性大肠杆菌中。他们发现用这种中断杂交(interrupted mating)方法很容易分析大肠杆菌的遗传结构、基因组成和位置。特别是当受体菌突然出现一种新表型时, 就意味着它从供体菌中获得了一个新基因, 并能确定这个新基因相对于其他基因的位置。通过中断杂交实验, 他们发现在接合过程中, 雄性大肠杆菌会缓慢地把自己的染色体注射到雌性大肠杆菌细

胞内, 而且对于任何一株雄性菌, 注射的起点总是固定的。比较不同的 Hfr 菌株传递染色体基因的顺序, Wollman 等^[24]推断大肠杆菌染色体是环形的, 并证实了在染色体外还存在有独立的遗传物质——外质体(episomes), 如性因子(F 因子)。

1.8 其他乳糖代谢相关基因的发现

鉴于用生物化学手段再也无法进一步解析大肠杆菌酶诱导原因了, 而新兴的微生物遗传学研究已在巴斯德研究所展开, Monod 意识到该求助于微生物遗传学手段了。

Monod 团队的 G. Cohen 等^[25]发现两种不能利用乳糖的 *lac*⁻突变菌株, 一株不能产生 β -半乳糖苷酶, 因此不能代谢乳糖; 而另株虽能产生 β -半乳糖苷酶, 但却仍然不能代谢培养基中的乳糖, 他们称之为假突变体(cryptic mutant)。假突变体能产生 β -半乳糖苷酶, 为何不能利用培养基中的乳糖呢?

Monod 认为细菌要利用培养基中的乳糖, 首先需要把乳糖摄入细胞内部。而这个假突变体可能丧失了从培养基中摄取乳糖的能力。他们发现硫代甲基- β -D-半乳糖苷(thiomethyl-beta-D-galactoside, TMG) (图 3C)虽然不是 β -半乳糖苷酶的底物, 但是个强诱导剂。借助放射性同位素标记, 他们发现 TMG 能进入野生型大肠杆菌并在细胞内积累, 但不能进入假突变体大肠杆菌细胞内, 说明野生型大肠杆菌细胞膜上有一种特殊的蛋白能把 TMG(或乳糖)运入细胞内, 他们称之为透性酶(permease), 如果编码透性酶的基因(*Y*)突变了, 大肠杆菌就不能把糖苷类分子运入细胞内(*lacY*⁻)。同时发现 *Y* 基因自身的表达也受 β -半乳糖苷酶的诱导剂诱导。

Monod 团队的 I. Zabin 等^[26]随后在分析硫代半乳糖苷诱导前后大肠杆菌内含物成份变化时发现少量的乙酰硫代半乳糖苷(6-0-acetyl-thiogalactoside), 说明细胞内出现了转乙酰酶。同时发现半乳糖苷转乙酰酶(Galactoside transacetylase)也受 β -半乳糖苷酶的诱导剂诱导表达。半乳糖苷转乙酰酶可把乙酰辅酶 A 的乙酰基转到半乳糖苷、葡萄糖苷和乳糖苷、IPTG 及 TMG 等分子上。它们被乙酰化后排出细胞, 无法再被摄入细胞, 从而达到解毒目的^[27]。

β -半乳糖苷酶、透性酶和转乙酰酶共同受乳糖

诱导,提示大肠杆菌乳糖代谢相关基因的表达调控是个“模块化”的开关模式,也进一步证明诱导剂与所诱导的酶并无直接关系,而是通过某种类似开关的第三者发挥了诱导作用。因此解析 β -半乳糖苷酶、透性酶和转乙酰酶模块的诱导机制,实质上就是解析这些酶基因表达的调控机制。首要的工作是了解编码这些酶的基因是以什么形式的“模块”存在于细菌染色体上。

1.9 乳糖操纵子的组成

Monod 与 Jacob 等^[28]通过人工突变鉴定了大肠杆菌 β -半乳糖苷酶、透性酶和转乙酰酶基因及其响应乳糖诱导的突变位点,并用中断杂交法进行了基因作图。他们根据染色体作图的结果和对突变体表型分析的结论于 1960 年提出了操纵子(operon)概念,把一组协同表达的基因称为一个操纵子。

大肠杆菌的乳糖操纵子由编码 3 个酶的基因 Z ($lacZ$)、 Y ($lacY$)和 A ($lacA$)组成。操纵子的上游有 1 个基因 $lacI$ 和介于 $lacI$ 与 $lacZ$ 之间的操纵区 $lacO$,共同构成乳糖诱导模块(图 4)。 $lacI$ 决定 3 种酶能否被诱导,称为调节基因(regulatory gene)。 $lacZ$ 、 $lacY$ 和 $lacA$ 分别决定能否产生有活性的 3 种酶,称为结构基因(structural genes)。 $lacI$ 突变体不需要诱导剂也能产生大量的 β -半乳糖苷酶、转乙酰酶和透性酶,即结构基因组成型表达(constitutive expression)。 $lacZ$ 突变体不能被诱导产生有活性的 β -半乳糖苷酶,但能被诱导产生正常的转乙酰酶。 $lacA$ 突变体不能被诱导产生有活性的转乙酰酶,却能被诱导产生正常的 β -半乳糖苷酶。但 $lacY$ 突变体类似于 $lacI$ 突变体,不仅不能被诱导产生透性酶,而且也不能被诱导产生 β -半乳糖苷酶,其中绝大部分突变体(并非全

部)也不能被诱导产生转乙酰酶!由此可见,调节基因 $lacI$ 的突变同时影响 3 个结构基因的表达,而结构基因突变则一般只影响其自身表达($lacY$ 突变体是个例外,导致诱导剂不能进入细胞内)。显然,调节基因 $lacI$ 并非分别直接控制 3 个结构基因,而是通过一个“诱导机关”间接发挥控制作用,同时开关 3 个结构基因。而且这个“诱导机关”显然也不是 3 个结构基因本身的组成部分。那么调节基因 $lacI$ 与“诱导机关”如何协作控制结构基因表达呢?

1.10 调节基因控制乳糖操纵子的机制

基因 $lacI$ 在理论上可能通过两种模式控制结构基因表达:一种是“正控制”(positive control)模式,即野生型 $lacI^+$ 是结构基因表达必需的激活元件之一,但需要诱导剂的辅助才能激活结构基因表达,而突变型 $lacI$ 则不需要诱导剂辅助,独自即可激活结构基因表达。另一种是“负控制”(negative control)模式,即野生型 $lacI^+$ 抑制结构基因表达,而诱导剂的作用是解除 $lacI^+$ 的抑制作用,使结构基因表达;突变型 $lacI$ 则不能抑制结构基因表达,所以结构基因不需要诱导剂即可表达。大肠杆菌采用了那种模式呢?在一次学术演讲后,核物理学家 Leo Szilard 建议 Monod 考虑“负控制”机制。

其实 Monod 还想过 $lacI^+$ 的另一种作用方式,即诱导剂才是激活结构基因表达的主角, $lacI$ 通过诱导剂辅助控制结构基因表达:野生型 $lacI^+$ 抑制大肠杆菌细胞内产生“内源性诱导剂”,因而需要外源诱导剂进入细胞才能开启结构基因表达,而突变型 $lacI$ 的大肠杆菌细胞能源源不断地产生“内源性诱导剂”,因而不需要外源诱导剂即可开启结构基因表达。

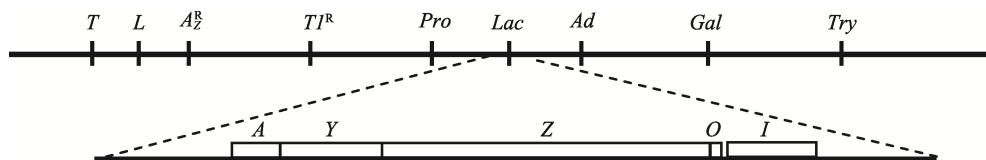


图 4 大肠杆菌 K12 的乳糖位点图谱及乳糖操纵子的基因组成

Fig. 4 The lac locus of *E. coli* K12 and structure of the lac operon

上:包含乳糖位点的大肠杆菌 K12 部分染色体图谱。下:乳糖操纵子的 3 个结构基因: Z ($lacZ$, 编码 β -半乳糖苷酶)、 Y ($lacY$, 编码透性酶)和 A ($lacA$, 编码硫代半乳糖苷转乙酰酶),以及位于紧邻结构基因上游的操纵区 O ($lacO$)和调节基因 I ($lacI$, 编码阻遏物)。根据文献^[29]绘制。

上述 3 种模式均难以用生物化学方法鉴定。不过, 深受孟德尔遗传学理论影响的 Monod 意识到, 如果是二倍体生物的杂合体($lacI^+ lacI^-$), 在“正控制”模式下, 等位基因 $lacI^-$ 对 $lacI^+$ 显性, “负控制”模式下, 则 $lacI^+$ 对 $lacI^-$ 显性。无论谁显性, 都将说明“内源性诱导剂”的想法是错误的, 即不存在假想中的“内源性诱导剂”。可是, 大肠杆菌是单倍体生物, 不存在天然的杂合体, 怎么才能证明 $lacI^+$ 和 $lacI^-$ 谁是显性基因呢?

加州大学伯克利分校的 Edward Adelberg^[30]曾到巴斯德研究所访学, 1959 年报道了在分离 Hfr 大肠杆菌菌株的过程中发现一种新形式的 F 因子, 带有染色体上的部分基因, 称为 *sfa* (sex factor attachment), 后被称为 F' 。Jacob 和 Adelberg^[31]发现在接合过程中 F' 能携带包含 *lac* 操纵子在的几个供体染色体的基因转移到受体大肠杆菌细胞。由于 F' 能独立于染色体稳定地存在于细菌细胞中, 导致其所携带的供体染色体基因在受体细胞内与受体细胞原有的相同基因同时存在, 类似二倍体真核生物的等位基因, 被称为部分二倍体(merozygote) (图 5)。Jacob 等^[32]把这种借助 F' 因子传递若干染色体基因的过程称为 F-duction 或 sex-duction (性导)。

性导形成的部分二倍体类似于高等生物的同源染色体, 其上携带的基因可形成杂合体(如图 5 的 $F' lacI^+ lacZ^- / lacI^- lacZ^+$), 可用于鉴定大肠杆菌野生型基因与突变型基因的显隐性关系并研究其功能。于是 Mond 和 Jacob 合作, 与来巴斯德研究所访学的美国学者 Arthur Pardee 一起, 用性导方法研究了 β -半乳糖苷酶基因及其诱导表达模式, 史称“PaJaMo (Pardee, Jacob, Monod)”实验^[33]。他们用性导和部分二倍体实验鉴定了 $lacI^+$ 和 $lacI^-$ 的显隐性。带有

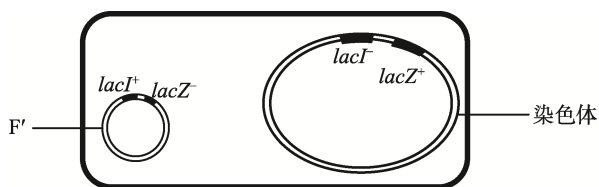


图 5 $lacI$ 和 $lacZ$ 部分二倍体

Fig. 5 Merozygote of $lacI$ and $lacZ$

带有 $lacI^+ lacZ^-$ 基因的 F' 进入基因型为 $lacI^- lacZ^+$ 基因大肠杆菌细胞中, 导致 $lacI$ 和 $lacZ$ 基因成为二倍体。

$lacI^+ lacZ^-$ 和 $lacI^- lacZ^+$ 部分二倍体的大肠杆菌乳糖操纵子(图 5)表现诱导型表达, 说明 $lacI^+$ 对 $lacI^-$ 显性。即, 一条染色体上的 $lacI^+$ 基因能抑制另一条染色体上的 $lacI^-$ 基因的作用。提示在野生型($lacI^+$)大肠杆菌细胞内, 结构基因确实是被某种物质抑制着(负调控), 需要外源诱导剂进入细胞解除抑制; 而在突变型($lacI^-$)细胞内, 抑制剂(repressor)消失, 抑制解除, 不需要外源诱导剂, 结构基因得以组成型表达。

那么, 这种抑制剂是 $lacI^+$ 基因本身呢, 还是 $lacI^+$ 基因的表达产物呢? 如果是 $lacI^+$ 基因本身, 即 $lacI^+$ DNA 直接抑制其相邻的结构基因表达(顺式作用), 就不应该抑制位于另一条染色体上与 $lacI^-$ 基因相邻的结构基因的表达, 即不会表现显性性状。如果 $lacI^+$ 基因是通过其产物抑制结构基因表达(反式作用), 由于基因产物游离在细胞质中, 既可以抑制同一条染色体上的结构基因, 也可以抑制另一条染色体上的结构基因表达, 就能表现显性性状。

“PaJaMo”实验结果显示基因 $lacI^+$ 对 $lacI^-$ 显性, 说明基因 $lacI^+$ 起反式作用, 是借助某种游离在细胞质内的阻遏物(suppressor)来抑制结构基因的表达。因而, 对 $lacI^-$ 菌株中结构基因组成型表达的最简单解释就是 $lacI^-$ 突变导致不能产生阻遏物或者阻遏物丧失了抑制功能。即“负控制”模式是正确的, 不存在“内源性”诱导剂, 外源诱导剂的作用实质上是进入细胞后解除阻遏物对结构基因表达的抑制。

$lacI^+$ 基因产物在何处同时关闭 3 个结构基因表达呢? 除了 $lacI^-$ 隐性突变, Jacob 和 Monod^[34]还分离到一种组成型表达的显性突变 $lacO^c$ (operator, 操纵区), 即使部分二倍体有野生型基因 $lacI^+$ 和野生型基因 $lacO$, $lacO^c$ 染色体上的 $lacZ$ 基因和 $lacA$ 基因仍然是组成型表达, 表明 $lacO$ 是顺式作用元件, 即 DNA 本身, 只能控制与其在同一条 DNA 上的相邻结构基因。这说明基因 $lacO$ 与 $lacI$ 组成了一个抑制 β -半乳糖苷酶表达的二元系统, 二者缺一不可, 其中一个基因突变($lacI^-$ 或 $lacO^c$)都将破坏抑制效应, 导致结构基因组成型表达。 $lacO^c$ 突变能同时抑制 β -半乳糖苷酶和转乙酰酶两个基因表达, 说明 $lacO$ 元件也不是结构基因本身, 而是结构基因之外的独立区域。

根据 Francis Crick^[35]1958 年提出的中心法则, 储存在基因里的遗传信息需要传到 RNA 中, 然后再传到蛋白质里才能发挥作用。那么 *lacI*⁺ 基因是通过 RNA 还是通过蛋白质发挥作用呢?

1.11 调节基因产物的化学性质

为了回答上述问题, Monod 和 Jacob^[34]根据早先人们发现的序列像 DNA 的 RNA (DNA-like RNA), 提出了“信使”RNA (messenger RNA, mRNA) 概念, 认为 DNA 模板像复制自身一样按照碱基互补原则在细胞核内合成一条 mRNA, 藉此把 DNA 中的遗传信息转录到 mRNA 中。mRNA 到细胞质中与核糖体结合, 用所携带的遗传信息指导合成蛋白质。即 DNA 先转录出 mRNA, mRNA 再指导合成蛋白质。

因此 *lacI*⁺ 基因应该是通过 mRNA 或蛋白质抑制结构基因表达。进一步实验发现在细菌交配形成部分二倍体之前先给受体菌培养基中加入 5-甲基色氨酸或氯霉素抑制受体菌合成蛋白质, 形成的部分二倍体仍能被诱导产生 β -半乳糖苷酶。因此 Jacob 和 Monod^[34]1961 年提出操纵子模型的时候认为 *lacI*⁺ 基因所控制的阻遏物 (aporepressor) 应该是它所转录的 mRNA。他们^[36]坚信阻遏物不会是小分子化合物, 虽然没有证据表明是核酸, 但也不太可能是蛋白质。

然而, 俄勒冈大学的 Tadao Horiuchi 和 Aaron Novick^[37]随后发现了温度敏感型 *lacI* 突变。以往的研究表明, 对温度敏感的大分子一般是蛋白质, 而不是核酸, 说明 *lacI* 基因产物更有可能是蛋白质。Jacob 和 Monod^[38]也意识到如果阻遏物是 mRNA, 难以解释它与小分子诱导剂的相互作用。后来哈佛大学的 Walter Gilbert 和 Benno Muller-Hill^[39]用 IPTG 分离了阻遏物并鉴定了其化学性质, 确认与 IPTG 结合的阻遏物是蛋白质而不是核酸。

1.12 调节基因产物的别构效应

lacI⁺ 基因产生的阻遏蛋白与 *lacO* 元件如何同时抑制 3 个结构基因表达呢? 阻遏物与 *lacO* 元件同时抑制 3 个紧密连锁的结构基因, 且这种抑制效应能被糖苷类小分子诱导剂解除, 意味着阻遏物既能与 *lacO* 元件相互作用, 又能与糖苷类诱导剂相互作用,

且两种相互作用存在竞争性抑制。这就要求阻遏物本身具有特殊的空间构象。什么类型的蛋白质具有这样特殊的空间结构呢?

巴斯德研究所的 Jean-Pierre Changeux^[40]发现苏氨酸脱氨酶 (threonine-deaminase) 在酶活性不受影响的情况下可以失去抑制剂的结合位点, 说明抑制剂的结合位点可以不是酶的活性位点, Monod 和 Jacob^[36]称之为别构抑制 (allosteric inhibition), 并把这类酶称为别构酶 (allosteric enzyme)。这类酶应该有两个位点, 一个是酶活性位点, 另一个是与抑制剂结合的位点。受此启发, Monod 和 Jacob^[41]提出乳糖代谢的阻遏物应该至少具有两个不重叠的结合位点: 一个是与 DNA 结合的活性位点 (active site); 另一个是与小分子别构效应物 (allosteric effector) 结合的别构位点 (allosteric site)。与 DNA 结合的活性位点会因别构位点与诱导剂结合而发生变构 (allosteric transition), Monod 和 Jacob 称之为别构蛋白 (allosteric protein)。

1.13 乳糖操纵子的调控模型

Monod 和 Jacob 根据别构蛋白理论, 提出了乳糖操纵子的调控机制 (图 6): *lacI* 组成型表达, 产生一种可变构的阻遏物, 与 *lacO* DNA 结合, 阻止 RNA 聚合酶与启动子 *lacP* 结合, 从而抑制操纵子的 3 个结构基因转录。当培养基中有乳糖 (且没有葡萄糖) 时, 乳糖被细胞膜上的透性酶 (*lacY* 编码) 运入细胞内, 与阻遏物结合, 引起阻遏物变构, 不能再与 *lacO* DNA 结合, 从 *lacO* 上脱落, RNA 聚合酶得以与 *lacP* 结合, 转录操纵子的 3 个结构基因。其中 *lacZ* 编码的 β -半乳糖苷酶可不断分解进入细胞内的乳糖, 生成葡萄糖, 供细胞代谢。当培养基和细胞内的乳糖被分解完后, 阻遏物又与 *lacO* DNA 结合, 阻止 RNA 聚合酶与 *lacP* 结合, 关闭操纵子。

乳糖操纵子模型第一次阐明了环境中的营养物质调控细菌基因表达的分子机制, 与 DNA 双螺旋结构模型一起引导遗传学乃至生物学整体研究进入了分子水平, 成为生命科学发展史上的一个里程碑, 并作为经典案例被载入遗传学、微生物学、细胞生物学和分子生物学等多门学科的教科书中, 是遗传学课程必讲的案例。然而, 粗浅地介绍 Jacob 和

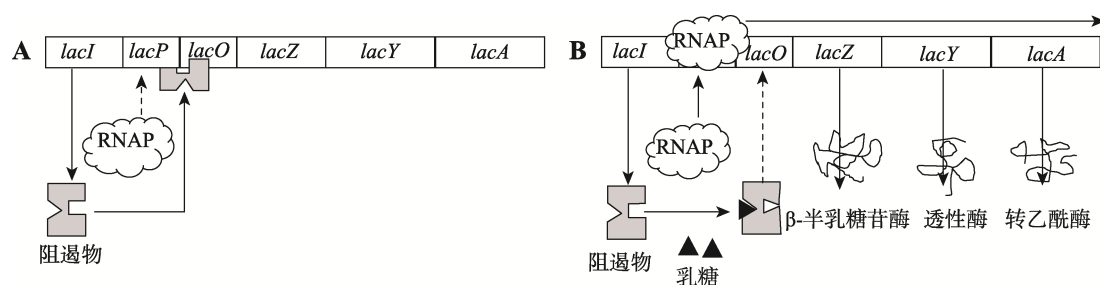


图 6 大肠杆菌乳糖操纵子的组成结构及其调控原理

Fig. 6 The structure and regulation mechanism of the *E. coli* lac operon

大肠杆菌乳糖操纵子由紧密连锁的 3 个结构基因 *lacZ* (编码β-半乳糖苷酶)、*lacY* (编码透性酶)和 *lacA* (编码硫代半乳糖苷转乙酰酶), 以及位于紧邻结构基因上游的 2 个调控元件, *lacP* (启动子)和 *lacO* (操纵区)组成。它们共同受临近上游的一个调节基因 *lacI* (编码阻遏物)的控制。A: 当培养基中无乳糖时, 阻遏物与 *lacO* 结合, 阻止 RNA 聚合酶(RNAP)与 *lacP* 结合。RNAP 不能转录结构基因。B: 当培养基中有乳糖时, 进入细胞内的乳糖(实际上转变为异乳糖)与阻遏物结合, 使阻遏物不能再与 *lacO* 结合, RNAP 可以与 *lacP* 结合, 转录结构基因。

Monod 模型容易, 结合后来对大肠杆菌乳糖代谢各个环节的精细研究结果精准讲解该模型并非易事, 师生都必须面对若干难以解释的现象甚至悖论。下文将分享回答这些问题的思路。

2 乳糖操纵子教学中的若干问题解析

2.1 乳糖操纵子调控模型的悖论

上述调控模型揭示, 大肠杆菌进化出了一种巧妙的诱导机制, 用代谢的靶标分子作为启动代谢通路的钥匙, 形成了一个自动启闭回路。环境中有乳糖时自然启动表达分解乳糖的基因, 当乳糖被分解完后, 分解乳糖的基因也随即被关闭, 既能利用靶标, 又能最大限度地节约自身的资源。

但大肠杆菌这一自动启闭回路似乎并不完美, 因为 Monod 等从测定大肠杆菌的β-半乳糖苷酶活性开始就发现即使培养基中没有乳糖等糖苷类化合物, 大肠杆菌也并非完全不表达β-半乳糖苷酶, 只是表达水平很低而已。当培养基中有乳糖等诱导剂时, 表达量会上千倍地增加。

既然是诱导型调控, 当培养基中没有乳糖时, 大肠杆菌应该彻底关闭乳糖操纵子才是最节省的模式, 为什么它允许操纵子低水平表达呢? 原来, 水溶性的乳糖并不能自由扩散穿过细胞膜的磷脂双分子层, 而是需要膜上的透性酶蛋白帮助才能进入大肠杆菌细胞内(IPTG 进入细胞不依赖于透性酶)。由

于编码透性酶的基因(*lacY*)自身也是乳糖操纵子的成员, 所以无论培养基中是否有乳糖, 大肠杆菌都得先合成少量的透性酶装配到细胞膜上, 随时准备把外界的乳糖运入细胞内。

其次, 尽管乳糖能诱导野生型大肠杆菌的乳糖操纵子表达, Monod 等^[42]1965 年发现乳糖并不能诱导 *lacZ* 突变株大肠杆菌产生转乙酰酶, 说明乳糖自身并不能诱导乳糖操纵子, 需要β-半乳糖苷酶的帮助才能成为该操纵子的诱导剂。那么β-半乳糖苷酶到底对乳糖进行了何种改造使其成为诱导剂?

印第安纳大学的 Müller-Hill 等^[43]1964 年报道乳糖的异构体异乳糖(allolactose) (图 3B)也是乳糖操纵子的有效诱导剂之一。乳糖与异乳糖的区别是糖苷键的位置不同, 乳糖是β-1,4 糖苷键, 而异乳糖是β-1,6 糖苷键。索尔克研究所的 Alan Jobe 和 Suzanne Bourgeois^[44]意识到乳糖在细胞内可能被β-半乳糖苷酶催化变成异乳糖才能发挥作用, 于 1972 年证实乳糖操纵子的天然诱导剂其实是异乳糖, 且诱导效率优于 IPTG。即β-半乳糖苷酶有 3 种功能: (1)把一个乳糖分子直接水解为一个半乳糖和一个葡萄糖分子; (2)催化乳糖发生转糖基作用(transglycosylation)形成异乳糖^[45]; (3)把异乳糖也进一步分解为半乳糖和葡萄糖^[46]。被透性酶运入大肠杆菌细胞内的一部分乳糖还要被β-半乳糖苷酶催化变成异乳糖才能诱导乳糖操纵子大量表达, 因此无论培养基中是否有乳糖, 大肠杆菌都得先合成少量的β-半乳糖苷酶, 在细胞质中随时准备把透性酶运入的乳糖转变为异乳糖。

上述两个事件要求至少两个乳糖操纵子的结构基因必须是低水平的组成型表达,当被迫以乳糖为生时才能诱导产生更多代谢乳糖的酶。因此,乳糖操纵子实际上是组成型表达基础上的诱导型表达,尽管能在很大程度上避免资源浪费,但还无法做到在不利用乳糖时彻底关闭乳糖操纵子。假如大肠杆菌细胞表面有个乳糖受体,接收到培养基中的乳糖信号后才能激活细胞内的乳糖操纵子表达,可能更为经济。

大肠杆菌如何实现组成型和诱导型的统一呢?理解乳糖操纵子表达特点的关键在于该操纵子的调控序列结构。

2.2 乳糖操纵子调控区的结构特点

在大肠杆菌 K-12 基因组(LN832404)中,乳糖代谢基因簇位于基因 *mhpR* 与 *cynX* 之间(图 7)。从 *lacI* 编码区的最后 17 bp 至 *lacZ* 基因起始密码子前的 141 bp DNA (图 7, 左下)中包含了 3 个启动子^[47] (*lacP2*、*lacP3*、*lacP1*)和一个“隐蔽启动子”(cryptic promoter) (*lacPc*)^[48]; 2 个操纵区(*lacO1*、*lacO3*)和一

个分解代谢激活蛋白(CAP-cAMP)结合位点。*lacP3*、*lacP2* 和 *lacP1* 均位于 CAP-cAMP 结合位点与 *lacO1* 之间,将受制于葡萄糖和阻遏物双重抑制。其中 *lacP2*^[49] 和 *lacP3* 的活性很差^[47,50], *lacP1* 与大肠杆菌启动子的一致序列(consensus sequence)同源性最高(图 7, 右上),是最强的启动子,是教科书里列举的 *lacP*。

“隐蔽启动子”*lacPc* 位于 CAP-cAMP 结合位点上游,不受葡萄糖影响。其-35 box 位于上游 *lacI* 基因的 3'端编码区内,-10 box 虽然与 *lacO3* 重叠,但 *lacO3* 与阻遏物结合非常弱^[51],因此 *lacPc* 也几乎不受阻遏物影响,可以驱动乳糖操纵子组成型表达,但核心序列与一致序列相差较大(图 7, 右上),因而启动活性并不强,这是大肠杆菌乳糖操纵子低水平组成型表达的原因之一。

其次,*lacI* 与 *lacZ* 之间的 123 bp DNA 内应该有个转录终止子(*lacIt*),以结束 *lacI* 转录,但实际上这段序列中并无典型的不依赖于 rho 终止子(rho-independent terminator)序列。耶鲁大学的 Heidi Horowitz 和 Terry Platt^[52]在这段 DNA 中鉴定出一个

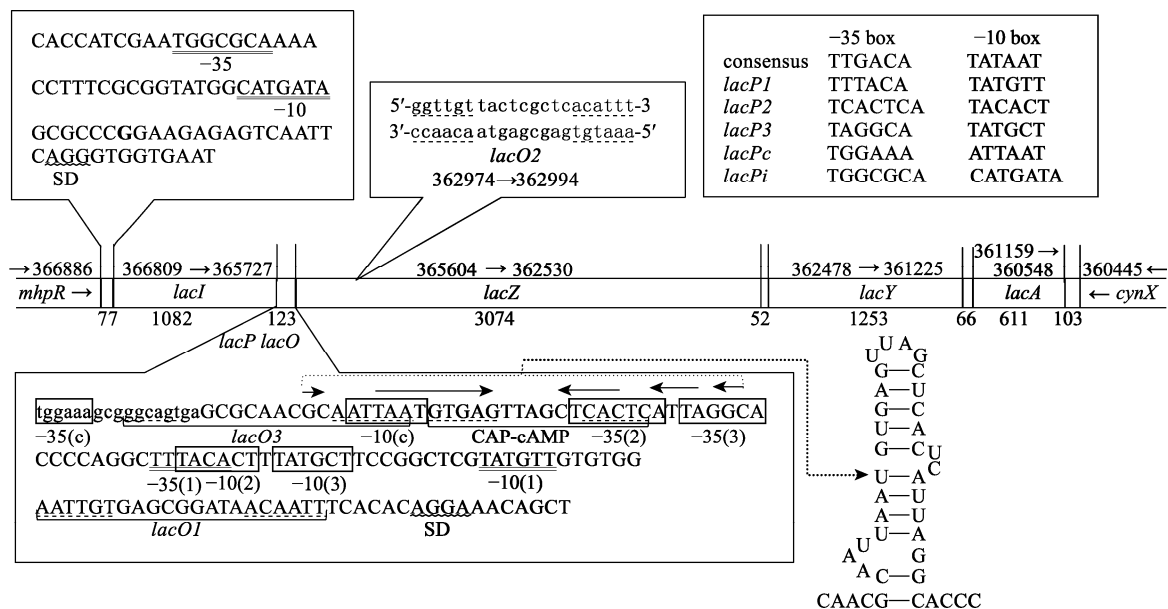


图 7 大肠杆菌乳糖操纵子的调控区序列特点

Fig. 7 Characteristics of the regulatory sequences of the *E. coli lac* operon

序列来自 GenBank LN832404。基因上方的数字是各基因起始密码到终止密码的核苷酸序号,下面的数字是对应区段的长度(核苷酸数)。小写字母表示编码区的碱基,大写字母表示上一个基因的终止密码至下一个基因的起始密码之间的碱基。4 个启动子的核心元件分别用-35(c)-10(c)、-35(2)-10(2)、-35(3)-10(3)和-35(1)-10(1)表示。双下划线和方框内字母表示启动子的-35 box 和-10 box 序列,虚下划线表示 *lacO1*、*lacO2* 和 *lacO3* 及 CAP-cAMP 结合位点的回文序列,波浪下划线表示 SD 序列。箭头下面的序列是 *lacI* 基因的终止子(*lacIt*),其转录物能形成右侧的茎环结构。

转录终止子(图7, 左下箭头下的序列), 所转录的RNA能形成茎环结构, 但其后不带一串U(图7, 右下)。体外实验显示这个终止子只有20%~90%的转录终止效率, 但不受rho或nusA蛋白的影响。由此可见*lacI_t*的终止效率不强, 可能导致RNA聚合酶在转录完*lacI*基因后偶尔继续转录, 把其后的*lacZ*等结构基因也转录出来, 这可能是乳糖操纵子低水平组成型表达的另一个原因。

2.3 *lacI* 基因自身的表达调控

在大肠杆菌染色体上只有一个乳糖操纵子。*lacI*基因只需遏制住这一个操纵子转录即可, 因此*lacI*基因既要组成型表达, 又不需要表达很多。大肠杆菌如何做到恰如其分地表达*lacI*基因呢?

野生型*lacI*的启动子(*lacP_i*)虽然是个组成型启动子, 但核心序列中GC含量太高, 且与一致序列差异较大(图7, 右上), 因此它的启动效率确实非常弱, 在每一代细胞中只转录1~2次, 每个细胞只有约10个阻遏蛋白分子^[39]。

那么, 如此少的阻遏蛋白分子如何高效地抑制了操纵子转录呢?

2.4 阻遏蛋白高效抑制操纵子转录的机制

阻遏蛋白分子与操纵区*lacO₁*的亲合力很高。*lacO₁*位于*lacP_i*的-10 box下游, 与转录起始位点重合(图7, 左下)^[53], 因此, 如果阻遏物结合到*lacO₁*上, 则很可能阻止RNA聚合酶与启动子结合或者阻止已经结合的RNA聚合酶向下转录操纵子。

不仅如此, 乳糖操纵子DNA中另有2个能与阻遏物结合的位点, 分别称为*lacO₂*^[54]和*lacO₃*^[55]。*lacO₂*位于*lacO₁*下游, 在*lacZ*基因内部的互补链上, 距*lacZ*的起始密码子370 bp(图7, 上中), 距*lacO₁*约140 nm, 与*lacO₁*方向相反; *lacO₃*位于*lacO₁*的上游, 在CAP-cAMP结合位点之前。*lacO₁*的序列特点是回文序列(AATTGTgagcggataACAATT), 因此每个*lacO₁*DNA可以结合两个阻遏蛋白分子。*lacO₂*序列(AATTGT agcgagtaACAACC)与*lacO₁*同源性较高, 与阻遏物的亲合力约相当于*lacO₁*的50%。由于阻遏蛋白能互相结合成4聚体, 结合在*lacO₂*上的两个阻遏蛋白分子与结合在*lacO₁*上的两个阻

遏蛋白分子^[56]能互相结合成4聚体, 使乳糖操纵子DNA弯曲成为一个小环^[57]更有效地阻止环内的RNA聚合酶前进^[58]。此外, 借助阻遏蛋白的4聚化作用, 结合到*lacO₂*上的两个阻遏蛋白分子将大大提高邻近的*lacO₁*吸引阻遏蛋白分子的机会。因此, 10个阻遏蛋白分子即可有效地与*lacO₁*和*lacO₂*结合, 抑制操纵子转录。*lacO₃*序列(GGCAGTgagcg-caacGCAATT)与*lacO₁*同源性较差, 与阻遏物的亲合力很低, 对乳糖操纵子的调控作用不大。

用尽可能少量的阻遏物去关闭操纵子也有利于诱导剂快速移除阻遏物, 开启转录。一旦结合在*lacO₁*和*lacO₂*上的阻遏蛋白被诱导剂移除, RNA聚合酶就会按顺序转录乳糖操纵子的3个结构基因^[59]。

在最后一个结构基因*lacA*的下游有一段DNA, 其转录物会形成两个茎环结构^[60], UGUUU紧随第2个茎环(图8), 构成了典型的不依赖rho的转录终止子, 能有效地终止结构基因的转录, 转录出一条带有3个开放阅读框(ORF)的多顺反子(polycistronic)mRNA。

2.5 乳糖操纵子的转录后调控

由于*lacZ*、*lacY*和*lacA*的起始密码子前均有SD序列, 因此核糖体可以分别独立翻译多顺反子mRNA上的3个ORF, 产生3条独立的多肽, 且理论上产量相等。但人们发现乳糖操纵子的表达实际上具有自然极性(natural polarity)^[61], 即第一个结构基因*lacZ*的产物β-半乳糖苷酶的分子数是最后一个结构基因*lacA*的产物转乙酰酶分子数的3~5倍。华盛顿大学的David Kennell和Howard Riezman^[62]测定和计算了乳糖操纵子的表达量, 也得出结论类似的结果。理论上同时转录和翻译的乳糖操纵子为何不能产生等量的β-半乳糖苷酶和转乙酰酶呢?

β-半乳糖苷酶分子数远多于转乙酰酶分子数的原因在理论上可能有3方面: (1) RNA聚合酶可能并非总能把操纵子完整地转录出来, 而是在转录到*lacZ*基因后的某处提前终止了, 导致包含*lacZ*的mRNA比包含*lacA*的mRNA多; (2) 核糖体翻译*lacZ*mRNA的起始效率比翻译*lacA*mRNA高; (3) *lacA*mRNA比*lacZ*mRNA降解的更快。究竟是哪种原因呢?

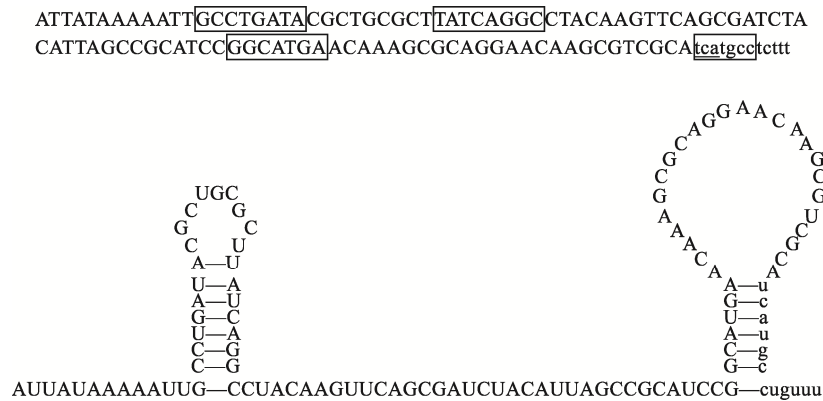


图 8 *lacA* 基因终止密码与下游 *CynX* 基因(位于互补链上)之间的序列及其转录产物

Fig. 8 The transcripts and DNA sequences between the termination codon of *lacA* and initiation codon of downstream *CynX* genes

小写字母为 *CynX* 基因的 3'端, 下划线为 *CynX* 基因终止密码的互补序列。方框内的序列转录的 RNA 可形成茎环结构(下), 右面的茎环结构为典型的不依赖 rho 转录终止子。

首先想到的原因就是转录提前终止, 而乳糖操纵子内部的结构基因间隔区内确实有回文序列, 转录出的 mRNA 可能形成转录终止子样的茎环结构(图 9)。美国国立癌症研究所的 Benoit De Crombrugghe 等^[63]的体外转录实验发现, 添加高浓度 rho 因子会导致乳糖操纵子转录提前终止, 暗示乳糖操纵子内部有依赖于 rho 因子的转录终止子。但华盛顿大学的 Louis Lim 和 Kennell^[64]发现 RNA 聚合酶在细胞内总能把乳糖操纵子完整地转录出来。加州大学旧金山分校的 George Murakawa 等^[65]分析了细胞内乳糖操纵子转录出来的 mRNA, 也发现 90% 以上的 *lac* mRNA 长度完整。他们还鉴定了 *lacZ* 与 *lacY* 基因之间的回文序列(图 9B)在细胞内终止转录的能

力。尽管这个回文序列转录出的 RNA 茎环结构之后紧随 4 个 U, 非常像不依赖 rho 的转录终止子, 但在细胞内却不能有效地终止转录。因此, *lacZ-Y* mRNA 和 *lacY-A* mRNA 形成的茎环结构(图 9)并非转录终止子, 可能是防止 3'外切核酸酶攻击 mRNA 的 3'端^[65]。因此转乙酰酶产量低的原因并非乳糖操纵子转录提前终止。

另一种原因可能是 *lac* mRNA 的稳定性, 即包含 *lacA* 密码的 mRNA 部分被降解。mRNA 的失活机制可能有 3 种: (1)游离末端被外切核酸酶进攻, 如色氨酸操纵子 mRNA 的 5'外切核酸降解(5' exonucleolytic breakdown)^[66,67]; (2)mRNA 被从内切酶从中间切断; (3)mRNA 内部的特定位点被切断, 如

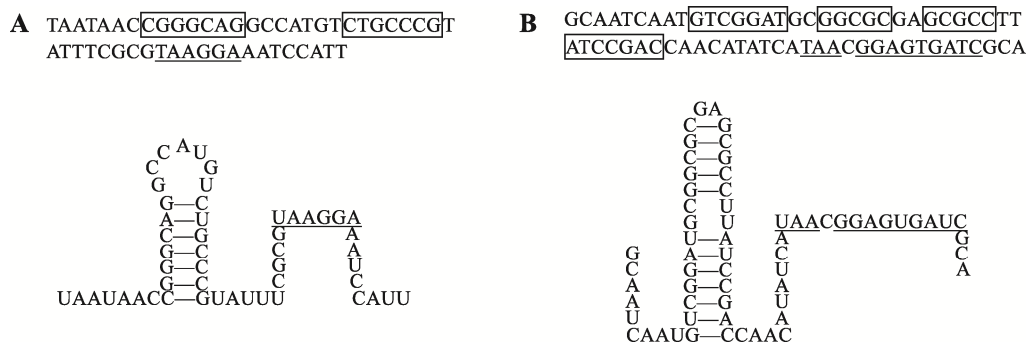


图 9 乳糖操纵子内部的基因间隔区序列及转录出的 RNA 可能形成的茎环结构

Fig. 9 The intergenic sequences of the *lac* operon and the stem-loop structures of the corresponding RNA

A: *lacZ* 终止密码与 *lacY* 起始密码之间的 DNA 序列(上)和转录出的 RNA 可能形成的茎环结构(下); B: *lacY* 终止密码与 *lacA* 起始密码之间的 DNA 序列(上)和转录出的 RNA 可能形成的茎环结构(下)。方框表示回文序列, 下划线表示 SD 序列。

多顺反子 mRNA 上每个基因的核糖体结合位点^[68]。巴斯德研究所的 A. Kepes^[59]分析了乳糖操纵子的转录、翻译和 mRNA 降解(mRNA decay)顺序,并测定了 lac mRNA 的降解速率,发现编码转乙酰酶的 mRNA 的半衰期比编码 β -半乳糖苷酶的短,认为这可能是转乙酰酶产量低的原因。后来华盛顿大学的 Martin Blundell 等^[69]也研究了 lac mRNA 降解的机制,发现 lacA mRNA 的半衰期比 lacZ mRNA 短约 2 倍。

Lim 和 Kennell^[64]、Blundell 和 Kennell^[70]在分析乳糖操纵子的转录物时还发现 lac 多顺反子 mRNA 在细胞内被切割成大约单个基因长度的 mRNA 片段,然后各自以不同的速率按 5'→3'方向降解。结合核糖体少的 mRNA 片段(lacA mRNA)比结合核糖体多的 mRNA 片段(lacZ mRNA)更易被降解^[71]。lac mRNA 的降解模型说明 lacA mRNA 降解更快是转乙酰酶产量低的原因之一。

lacA mRNA 片段的翻译效率也应该是影响转乙酰酶蛋白产量的原因之一。Kennell 和 Riezman^[62]分析了乳糖操纵子 mRNA 上的核糖体密度,发现 lacA mRNA 上的核糖体只有 lacZ mRNA 上的约 1/5,说明转乙酰酶的翻译效率低。此外 lacA 基因的起始密码子是 UUG 而不是 AUG,可能也会导致翻译起始效率差。总之,转乙酰酶产量远低于 β -半乳糖苷酶可能是由 lac mRNA 降解模式和效率及 mRNA 翻译效率的差异共同引起的。

3 结语

本文梳理了乳糖操纵子的发现过程,总结了乳糖操纵子表达研究的部分结果,结合乳糖操纵子的 DNA 序列结构,回答了乳糖操纵子教学过程中可能遇到的若干问题。限于篇幅,本文未涉及葡萄糖影响乳糖操纵子表达的分子机制。全文沿科学史脉络,用原始文献的数据,以乳糖操纵子为案例,力图真实地再现科学家们发现问题和解决问题的科研思想和历程,让这个遗传学和分子生物学的经典教学案例在培养本科生的科学素养和创新能力中发挥更大的作用。作者也希望通过乳糖操纵子科研背景的

深入挖掘引导读者自觉挖掘其他案例的科研背景和教学价值。

参考文献(References):

- [1] Xing WJ, Morigen. Cultivating the scientific research ability of undergraduate students in teaching of genetics. *Hereditas (Beijing)*, 2016, 38(11): 1030–1038.
邢万金, 莫日根. 在遗传学课堂教学中培养本科生科研素质. *遗传*, 2016, 38(11): 1030–1038.
- [2] Xing WJ, Morigen. Su HM. An exploration for research-oriented teaching model in biology teaching. *Hereditas (Beijing)*, 2014, 36(7): 732–738.
邢万金, 莫日根, 苏慧敏. 生物学教学中研究型教学方法与内容的探索. *遗传*, 2014, 36(7): 732–738.
- [3] Morigen, Xing WJ, Fan LF, Hada, Hasi A. Teaching practice of molecular biology in a scientific story-directed. *J Biol*, 2016, 33(5): 112–116.
莫日根, 邢万金, 范丽菲, 哈达, 哈斯阿古拉. 科学史引导的分子生物学教学架构和实践. *生物学杂志*, 2016, 33(5): 112–116.
- [4] Wortmann J. Untersuchungen über das diastatische ferment der bacterien. *Hoppe-Seyler's Z. Physiol. Chem*, 1882, 6(4–5): 287–329.
- [5] Duclaux E. Traité de Microbiologie: Diastases, Toxines et venins, Vol. II. *Masson et Cie*, 1899.
- [6] Karström H. Über die enzyymbildung in bakterien und Über einige physiologische Eigenschaften der untersuchten Bakterien. *Ann Acad Sci Fen A*, 1930, 33: 2.
- [7] Karström H. Enzymatische adaptation bei mikroorganismen. *Ergebnisse der Enzymforschung*, 1938, 7: 350–376.
- [8] Yudkin J. Enzyme variation in micro-organisms. *Biol Rev*, 1938, 13(1): 93–106.
- [9] Monod J. Sur un phénomène nouveau de croissance complexe dans les cultures bactériennes. *CR Acad Sci(Paris)*, 1941, 212: 934–936.
- [10] Monod J. Recherches sur la croissance des cultures bacteriennes. *Paris: Hermann & Cie*, 1942.
- [11] Diénert MF. Sur la fermentation du galactose et sur l'accoutumance des levures à ce sucre. *E. Charaire*, 1900.
- [12] Freeland JH. The eighth day of creation: the makers of the revolution in biology 1996. New-York: Cold Spring Harbor Laboratory Press.
- [13] Spiegelman S. The physiology and genetic significance of enzymatic adaptation. *Annals of the Missouri Botanical Garden*, 1945, 32(2): 139–163.

- [14] Monod J. The phenomenon of enzymatic adaptation-and its bearings on problems of genetics and cellular differentiation. *Growth*, 1947, 11(4): 223–289.
- [15] Seidman M, Link K. New Compounds. substituted salicylaldehydes and derivatives. *J Am Chem Soc*, 1950, 72(9): 4324–4324.
- [16] Lederberg J. The beta-d-galactosidase of *Escherichia coli*, strain K-12. *J Bacteriol*, 1950, 60(4): 381–392.
- [17] Monod J, Cohen-Bazire G, Cohn M. Sur la biosynthese de la β -galactosidase (lactase) chez *Escherichia coli*. La specificite de l'induction. *Biochim Biophys Acta*, 1951, 7: 585–599.
- [18] Herzenberg LA. Studies on the induction of beta-galactosidase in a cryptic strain of *Escherichia coli*. *Biochim Biophys Acta*, 1959, 31(2): 525–538.
- [19] Hickman M, Cairns J. The centenary of the one-gene one-enzyme hypothesis. *Genetics*, 2003, 163(3): 839–841.
- [20] Horowitz NH. One-gene-one-enzyme: remembering biochemical genetics. *Protein Sci*, 1995, 4(5): 1017–1019.
- [21] Monod J, Audureau A. Mutation and enzymatic adaptation in bact. coli-mutabile. *Ann Inst Pasteur*, 1946, 72(11/12): 868–878.
- [22] Lederberg J, Tatum EL. Gene recombination in *Escherichia coli*. *Nature*, 1946, 158(4016): 558.
- [23] Jacob F, Wollman E. Induction spontanée du developpement du bacteriophage lambda au cours de la recombinaison génétique chez *Escherichia coli*. *C R Hebd Seances Acad Sci*, 1954, 239(3): 317–319.
- [24] Wollman EL, Jacob F. Mechanism of the transfer of genetic material during recombination in *Escherichia coli* K12. *C R Hebd Seances Acad Sci*, 1955, 240(25): 2449–2451.
- [25] Cohen GN, Rickenberg HV. La galactosidepermease d'*Escherichia coli*. *Ann Inst Pasteur*, 1956, 91: 693–720.
- [26] Zabin I, Kepes A, Monod J. On the enzymic acetylation of isopropyl- β -d-thiogalactoside and its association with galactoside-permease. *Biochem Biophys Res Commun*, 1959, 1(6): 289–292.
- [27] Andrews KJ, Lin EC. Thiogalactoside transacetylase of the lactose operon as an enzyme for detoxification. *J Bacteriol*, 1976, 128(1): 510–513.
- [28] Jacob F, Perrin D, Sánchez C, Monod J. L'opérongroupe de gènes à expression coordonnée par un opérateur. *C R Hebd Seances Acad Sci*, 1960, 250(9): 1727–1729.
- [29] Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, 1961, 3: 318–356.
- [30] Adelberg E, Burns SN. Genetic variation in the sex factor of *Escherichia coli*. *J Bacteriol, Genetics*, 1960, 79(3): 321–330.
- [31] Jacob F, Adelberg E. Transference of genetic properties by incorporation with the sexual factor of *Escherichia coli*. *C R Hebd Seances Acad Sci*, 1959, 249(1): 189–191.
- [32] Jacob F, Schaeffer P, Wollman EL. Episomic elements in bacteria. *Symp Soc Gen Microbiol*, 1960, 10: 67–91.
- [33] Pardee AB, Jacob F, Monod J. The genetic control and cytoplasmic expression of “Inducibility” in the synthesis of β -galactosidase by *E. coli*. *J Mol Biol*, 1959, 1(2): 165–178.
- [34] Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, 1961, 3(3): 318–356.
- [35] Crick FH. On protein synthesis. *Symp Soc Exp Biol*, 1958, 12: 138–163.
- [36] Monod J, Jacob F. Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol*, 1961, 26: 389–401.
- [37] Horiuchi T, Novick A. A thermolabile repression system. *Cold Spring Harb Symp Quant Biol*, 1961, 26: 247–248.
- [38] Jacob F, Sussman R, Monod J. On the nature of the repressor ensuring the immunity of lysogenic bacteria. *C R Hebd Seances Acad Sci*, 1962, 254: 4214–4216.
- [39] Gilbert W, Müller-Hill B. Isolation of the lac repressor. *Proc Natl Acad Sci USA*, 1966, 56(6): 1891–1898.
- [40] Changeux JP. The feedback control mechanisms of biosynthetic L-threonine deaminase by L-isoleucine. *Cold Spring Harb Symp Quant Biol*, 1961, 26: 313–318.
- [41] Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. *J Mol Biol*, 1963, 6: 306–329.
- [42] Burstein C, Cohn M, Kepes A, Monod J. Rôle du lactose et de ses produits métaboliques dans l'induction de l'opéron lactose chez *Escherichia coli*. *Biochimica Et Biophysica Acta (Bba)-Nucleic Acids and Protein Synthesis*, 1965, 95(4): 634–639.
- [43] Mueller-Hill B, Rickenberg HV, Wallenfels K. Specificity of the induction of the enzymes of the lac operon in *Escherichia coli*. *J Mol Biol*, 1964, 10: 303–318.
- [44] Jobe A, Bourgeois S. Lac Repressor-operator interaction. VI. The natural inducer of the lac operon. *J Mol Biol*, 1972, 69(3): 397–408.
- [45] Wallenfels K, Malhotra OP. Galactosidases. *Adv Carbohydr Chem*, 1961, 16: 239–298.
- [46] Huber RE, Kurz G, Wallenfels K. A quantitation of the factors which affect the hydrolase and transgalactosylase activities of beta-galactosidase (*E. coli*) on lactose. *Biochemistry*, 1976, 15(9): 1994–2001.

- [47] Xiong XF, de la Cruz N, Reznikoff WS. Downstream deletion analysis of the lac promoter. *J Bacteriol*, 1991, 173(15): 4570–4577.
- [48] Krebber A, Burmester J, Plückthun A. Inclusion of an upstream transcriptional terminator in phage display vectors abolishes background expression of toxic fusions with coat protein g3p. *Gene*, 1996, 178(1–2): 71–74.
- [49] Malan TP, McClure WR. Dual promoter control of the *Escherichia coli* lactose operon. *Cell*, 1984, 39(1): 173–180.
- [50] Donnelly CE, Reznikoff WS. Mutations in the lac P2 promoter. *J Bacteriol*, 1987, 169(5): 1812–1817.
- [51] Fried M, Crothers DM. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res*, 1981, 9(23): 6505–6525.
- [52] Horowitz H, Platt T. A termination site for LacI transcription is between the CAP site and the lac promoter. *J Biol Chem*, 1982, 257(19): 11740–11746.
- [53] Gilbert W, Maxam A. The nucleotide sequence of the lac operator. *Proc Natl Acad Sci USA*, 1973, 70(12): 3581–3584.
- [54] Reznikoff WS, Winter RB, Hurley CK. The location of the repressor binding sites in the lac operon. *Proc Natl Acad Sci USA*, 1974, 71(6): 2314–2318.
- [55] Oehler S, Eismann ER, Krämer H, Müller-Hill B. The three operators of the lac operon cooperate in repression. *Embo J*, 1990, 9(4): 973–979.
- [56] Kania J, Müller-Hill B. Construction, isolation and implications of repressor-galactosidase-beta-galactosidase hybrid molecules. *Eur J Biochem*, 1977, 79(2): 381–386.
- [57] Krämer H, Niemöller M, Amouyal M, Revet B, von Wilcken-Bergmann B, Müller-Hill B. Lac repressor forms loops with linear DNA carrying two suitably spaced lac operators. *Embo J*, 1987, 6(5): 1481–1491.
- [58] Becker NA, Peters JP, Maher LJ 3rd, Lionberger TA. Mechanism of promoter repression by Lac repressor-DNA loops. *Nucleic Acids Res*, 2013, 41(1): 156–166.
- [59] Kepes A. Sequential transcription and translation in the lactose operon of *Escherichia coli*. *Biochim Biophys Acta*, 1967, 138(1): 107–123.
- [60] Hediger MA, Johnson DF, Nierlich DP, Zabin I. DNA sequence of the lactose operon: the *lacA* gene and the transcriptional termination region. *Proc Natl Acad Sci USA*, 1985, 82(19): 6414–6418.
- [61] Zabin I, Fowler AV. Chapter III: β -galactosidase and thiogalactoside transacetylase. *CSH Monograph Archive* 1970, 1: 27–47.
- [62] Kennell D, Riezman H. Transcription and translation initiation frequencies of the *Escherichia coli* lac operon. *J Mol Biol*, 1977, 114(1): 1–21.
- [63] de Crombrughe B, Adhya S, Gottesman M, Pastan I. Effect of Rho on transcription of bacterial operons. *Nat New Biol*, 1973, 241(113): 260–264.
- [64] Lim LW, Kennel D. Evidence against transcription termination within the *E. coli* lac operon. *Mol Gen Genet*, 1974, 133(4): 367–371.
- [65] Murakawa GJ, Kwan C, Yamashita J, Nierlich DP. Transcription and decay of the lac messenger: role of an intergenic terminator. *J Bacteriol*, 1991, 173(1): 28–36.
- [66] Morse DE, Mosteller R, Baker RF, Yanofsky C. Direction of *in vivo* degradation of tryptophan messenger RNA—a correction. *Nature*, 1969, 223(5201): 40–43.
- [67] Morikawa N, Imamoto F. Degradation of tryptophan messenger. On the degradation of messenger RNA for the tryptophan operon in *Escherichia coli*. *Nature*, 1969, 223(5201): 37–40.
- [68] Morse DE, Mosteller RD, Yanofsky C. Dynamics of synthesis, translation, and degradation of trp operon messenger RNA in *E. coli*. *Cold Spring Harb Symp Quant Biol*, 1969, 34: 725–740.
- [69] Blundell M, Craig E, Kennell D. Decay rates of different mRNA in *E. coli* and models of decay. *Nat New Biol*, 1972, 238(80): 46–49.
- [70] Blundell M, Kennell D. Evidence for endonucleolytic attack in decay of lac messenger RNA in *Escherichia coli*. *J Mol Biol*, 1974, 83(2): 143–161.
- [71] Lim LW, Kennell D. Models for decay of *Escherichia coli* lac messenger RNA and evidence for inactivating cleavages between its messages. *J Mol Biol*, 1979, 135(2): 369–390.

(责任编辑: 谢建平)