

分枝杆菌噬菌体的分子生物学研究进展及其应用

刘 文* 胡 巍¹ 王洪海

(复旦大学遗传学研究所 上海 200433)

(山东淄博师专生物系¹⁾ 山东 255013)

关键词 分枝杆菌噬菌体 L5, 超感染免疫性, 重组 BCG 疫苗, 萤光素酶报告噬菌体

分类号 Q939.94 **文献标识码** C **文章编号** ISSN-0253-2654(1999)-01-58-62

世界卫生组织估计,世界上有 1/3 多的人感染结核分枝杆菌 (*Mycobacterium tuberculosis*, TB), 并导致每年有 300 万人死亡。最近几年, 由于抗药性和多重抗药性结核杆菌 (MDR-TB) 的出现以及人免疫缺损病毒 (HIV) 侵袭, TB 在世界范围内复苏, 现行的医疗手段越来越无能为力, 急需一种新的、有效的方法来预防、诊断和治疗结核^[1]。

病原性分枝杆菌生长缓慢, *M. tuberculosis* 繁殖一

代大约需 24h, 在固体培养基上形成一个菌落至少需要 3~6 周时间; *M. leprae* 不能体外培养, 必须在鼠脚垫或犏狨中培养; 快速生长的非病原菌 *M. smegmatis* 倍增的时间为 2~4h, 需 3~4 周形成菌落, 这就为研究分枝杆菌提供一些方便。但快速和慢速生长的菌株都趋于成

* 山东淄博师专在职研究生

1997-11-11 收稿, 1998-05-11 修回

簇生长,所以菌落经常由一簇而不是由一个细胞形成。

这些特点使分枝杆菌的遗传分析比较困难和复杂^[2]。

分枝杆菌噬菌体的分子生物学研究为分枝杆菌提供了一个新的相对简单的遗传分析系统。其中 L5 噬菌体是迄今为止研究比较详细的一个。

1 分枝杆菌噬菌体 L5 的一般特性

分枝杆菌噬菌体 L5 是 1960 年从 *M. smegmatis* 的培养物和混浊噬菌斑中分离的,能使 *M. smegmatis* 形成稳定溶源菌,并有广泛的宿主范围,包括 *M. tuberculosis*、BCG 和一些 *M. avium* 亚菌株^[3,4],但也有资料报道,L5 在 *M. tuberculosis* 和 BCG 中不能有效地形成噬菌斑,而且在缓慢生长的分枝杆菌中尚未发现温和噬菌体^[5,6]。与 L5 亲缘关系较近的包括烈性噬菌体 D29 和 FRA1 均能感染快速和慢速生长分枝杆菌,D29 还能吸附 *M. leprae* 细胞。含有一个拷贝 L5 的溶源菌对 L5 和 D29 的超感染具免疫性。L5 具有其它溶源性噬菌体不具备的特征,例如:L5 不仅合成自己的 DNA 聚合酶和三个 tRNA,而且在早期裂解生长时关闭宿主的合成系统。这些特征提示:L5 的裂解和溶源生长过程的遗传调节不同于其它温和噬菌体。

已发现 L5 和 *E. coli* λ 噬菌体在总体上有相似性,但在序列水平几乎没有亲缘关系。L5 的大小和形状与 λ 噬菌体相似,其 20 面体的头部含有 DNA,长而灵活的尾部不能收缩,没有侧面的纤毛^[2]。gp17 是主要的头部外壳亚单位蛋白,形成精致的共价交联,相似的蛋白结构在许多其它分枝杆菌噬菌体中也观察到,可能头部亚单位交联是这些病毒的一个共同的特征。从头部和尾部已分别纯化出五种蛋白成份,但对 L5 的组装过程了解甚少^[3]。

2 L5 噬菌体的基因组结构及表达

L5 基因组为一条线形双链 DNA 分子,两端有短的单链粘性末端(COS 位点),其 DNA 序列已经测定,含 52297bp。整合位点 attP 位于近基因组中间处,把基因组分成左臂和右臂。用计算机方法推断有 88 个编码蛋白的基因,近 COS 位点处有约 2kb 非编码区,三个 tRNA 基因位于基因组左端 4kb 处。L5 基因的常规密码子表也已得到^[2,3,7]。

基因组左臂上紧密连接的 1-32 基因从 COS 位点到 attP 向右转录,在 L5 感染宿主的后期表达,它们的表达产物参加新的病毒颗粒的形态构建;右臂的 33-88 基因

从 COS 位点到 attP 向左转录,在感染早期表达(也有资料表明基因 33,即噬菌体整合酶基因是向右转录的^[2]),负责 DNA 复制和噬菌体的扩增^[5]。大多数 L5 基因用 AUG 或 GUC 作为翻译起始,但至少 5 个基因用 UUG 作为翻译起始(分别是基因 27、33、44、50、53)。虽然对噬菌体启动子的精确位置了解甚少,但 L5 基因的表达对利福平敏感,表明宿主 RNA 聚合酶起作用。L5 基因 44 编码它自己的 DNA 聚合酶,而其它的温和噬菌体都用宿主的聚合酶进行 DNA 复制,翻译起始位点在 UUG 密码子的 gp44 含 3'-5' 外切酶和 DNA 聚合酶活性,与 DNA 聚合酶 I 的 klenow 区域相似,计算机数据库比较表明 gp44 和细菌 DNA 聚合酶之间有明显的氨基酸同源性^[3]。

L5 与一般的温和噬菌体不同,在裂解生长时关闭宿主合成系统(包括应激蛋白),溶源菌被诱导后,宿主菌的蛋白合成停止并持续整个诱导过程。诱导后 10min 内可看到 L5 噬菌体蛋白的合成并分成两类:早期和晚期蛋白合成。至少有 12 种不同早期蛋白在诱导后 10min 内出现,45min 进入裂解生长时消失。诱导后 20~25min 晚期蛋白合成并持续整个裂解生长期直到细胞裂解(接近 2.5h)^[3]。

3 gp71 与溶源性维持和超感染免疫性

L5 溶源菌或含基因 71 的菌对 L5 或其亲缘的烈性噬菌体 D29 的超感染具有免疫性。溶源性维持和超感染免疫性需要 L5 基因 71 编码的一个类似阻遏物的蛋白(gp71)。gp71 含 183 个氨基酸残基,有 DNA 结合蛋白(包括噬菌体阻遏物)相似的螺旋-转角-螺旋的 DNA 结合区,且末端富含酸性氨基酸(66 个 C 端氨基酸中有 19 个为谷氨酸和天冬氨酸)。基因 71 的启动子可能位于上游 71 和 72 基因之间,在其它基因不存在时仍有活性,基因 71 下游可能有中止子阻止下游基因的表达^[8]。

L5 在 *M. smegmatis* 中形成模糊噬菌斑,形成透明噬菌斑的 L5 突变体根据基因型和表型特征分成三类:第一类形成透明噬菌斑的 L5 突变体完全裂解细菌细胞,不能感染含 L5 基因 71(一个拷贝或染色体外多个质粒拷贝)的 *M. smegmatis* 菌株,该类突变体可能是由于基因 71 部分或全部缺失或基因突变造成。第二类 L5 突变体形成的噬菌斑比第一类混浊,在这类噬菌体的感染区可能得到处于溶源性、对超感染有免疫性的菌落,

在培养物上清液中可得到与原来相同的突变噬菌体,说明溶源菌的形成不是回复突变或污染造成,该类突变体影响溶源性的建立,而不影响溶源性维持功能^[8]。Sarkis等实验说明:基因72~82区域与溶源性建立有关,是第二类突变体的靶位点^[5]。第三类突变体能感染含单拷贝基因71的*M. smegmatis*,但不能感染含多拷贝基因71或L5溶源的菌株,从而推断基因71有突变,突变体表达的突变gp71足以影响含单拷贝基因71的菌株,使之失去超感染免疫性,但不能使多拷贝质粒表达的大量gp71失活。另外,基因71内的点突变可以形成热诱导突变体:30℃时形成模糊噬菌斑,42℃时形成透明噬菌斑。第一类和热诱导突变体显然由基因71的隐性突变产生。透明噬菌斑突变体由于基因71缺失或突变导致裂解生长,热诱导突变体由于温度的改变使gp71失活导致从低温到高温时裂解生长被诱导,从而证明gp71对超感染免疫性和溶源状态的维持都是必需的。基因71是透明噬菌斑表型突变的主要靶位点。另外,gp71在溶源菌中的表达可能被未知的L5基因产物激活或从一个远距离的上游启动子转录^[8]。Nesbit等提出基因71上游71~72区域有三个启动子,基因组右臂末端还有一个向左转录的启动子均表达基因71^[10]。

由此可见,基因71在噬菌体生活周期的调节中扮演了一个重要的角色。实验表明在L5生活周期调节中溶源性有效建立需一个或更多的基因。基因71~88的大部分区域对噬菌体生长是非必需的^[2]。

4 整合与切除

L5溶源性的建立伴随着L5基因组位点专一地插入细菌染色体中。这个过程被作用于噬菌体attP位点和细菌attB位点的噬菌体整合酶所催化。attP和attB位点有相同的43bp的核心区域,重组发生在这个区域。

L5的整合位点attP紧邻于int基因的5'侧^[11]。构建含L5 int基因和T7启动子的质粒并转化*E. coli*,观察整合酶高表达和在体外催化整合重组的能力,发现L5整合酶的翻译起始位点是UUG,分子量为41.8ku,在25℃和37℃均能催化重组,42℃时则催化效率很低,有效的重组不需要Mg²⁺或ATP,无亚精胺或加整合酶之前先加拓扑异构酶I时催化效率很低,整合酶的N端片段对重组功能非常重要。还发现*M. smegmatis*的抽提物能刺L5整合酶介导的重组,但*E. coli*抽提物和纯化的IHf、HU因子则不能,说明可能存在一种新的分枝

杆菌整合因子^[12,13]。

int基因的下游和43bp核心区域的上游是潜在的RNA二级结构区,前者可能在int基因表达调控中起作用,后者起转录终止子作用,当L5溶源性整合后,它阻止转录从细菌基因组进入原噬菌体基因组^[11]。

*M. smegmatis*的attB位点含两个tRNA基因(tRNA^{Gly}和tRNA^{Phe}),attB的43bp核心区域与tRNA^{Gly}基因的3'末端重叠,因此整合后tRNA基因有效的重建。attB位点是高度保守的,43bp核心区域在BCG、*M. smegmatis*和L5基因组中只有一个碱基差异,差异位于tRNA^{Gly}可变环中,通过转化率看出对转化无明显影响^[11]。

通过把L5 attP-int区域插入到含卡那霉素抗性(来自Tn903)的pUC119载体中构建的整合质粒(integration-proficient plasmid)如:PMH94可以有效地转化*M. smegmatis*、BCG和*M. tuberculosis* H37Ra,而且形成稳定溶源状态^[11,14]。由温和噬菌体构建的载体整合到宿主染色体中只需噬菌体编码的整合酶,而原噬菌体的切除,除了需要整合酶外,还需要噬菌体编码的切除酶,因此带有attP和int基因的载体可以有效的转化分枝杆菌并保持稳定^[15]。

5 L5在分枝杆菌遗传学方面的应用

5.1 构建重组BCG疫苗

用BCG构建活的重组疫苗即把其它病原菌的抗原基因插入到BCG中,扩大它的保护能力,目的是构建一个可以在婴儿出生时用的抗多种疾病的多价活疫苗。这个方法的可行性和最终完成需要满足以下两个特殊条件。

首先,新导入的抗原基因的稳定性是很重要的,否则将在疫苗制备或接种时丢失,或者效能被严重损害。染色体外质粒载体在实验室里可用抗生素选择来维持,但对于人类疫苗则不是一个理想选择,因为存在着把抗生素抗性传到其它病原菌的危险。另一种办法是用遗传重组把质粒插入细菌染色体,它将随染色体DNA的复制而复制。衍生于L5的位点专一整合载体是非常简单而有效的。因为int基因紧邻attP位点,含这两个位点的DNA插入不能自主复制的分枝杆菌质粒内,构成的载体能有效地转化分枝杆菌,转化子的attB位点已整合一个拷贝的质粒。由于L5切除酶基因(位于attP右侧)不在载体中,即使在不选择的情况下插入序列也是非常稳定的。这个系统在*M. smegmatis*

和BCG中都运转很好,它可能被应用于分枝杆菌的其它种。由于噬菌体的整合通常包括52kb病毒基因组的插入,所以载体具有插入大量外源抗原基因的潜力^[2,14]。

第二是选用合适的选择标志。由于常用的抗生素基因不适用于活疫苗,转化子的选择标志可用L5基因71。把基因71插入整合载体,转化后用裂解分枝杆菌的L5透明噬菌斑突变体或D29选择。由于转化子具有对L5(及其突变体)和D29超感染免疫性而非转化子不具有此特征,故L5突变体和D29均能杀死非转化细胞,保留转化子。用透明噬菌斑突变体选择时,最好用基因71的缺失突变体,尽量避免用突变形成的突变体,防止回复突变而造成的假阳性。用突变体或D29选择染色体外多拷贝质粒转化子也有效。可见基因71作为选择标志对于染色体外或整合质粒、快速或慢速生长菌种都适用^[8]。

除L5基因71外,还可用*E.coli* β -半乳糖苷酶基因或萤光素酶报告噬菌体作为选择标志。

5.2 构建萤光素酶报告噬菌体 (Luciferase reporter Mycobacteriophage, LRMs) L5的另外一个潜在应用是构建重组萤光素酶报告噬菌体。日益流行的抗药性结核分枝杆菌菌株是有关公共健康的重要问题。由于细菌的缓慢生长,临床样品的药敏试验需要几周或几个月,这期间病人不能得到有效治疗,加速了病人的死亡,诱使新的抗性的发展,危害与病人接触的人。萤光素酶报告分枝杆菌噬菌体(LRMs)提供了另外一个快速确定临床分离物药敏状况的方法。这个新方法靠噬菌体把萤光素酶基因导入宿主细胞并使细菌发光。光的产生依赖于菌体感染,萤光素酶基因的表达和ATP水平,在报告噬菌体感染*M.tuberculosis*后几分钟内就能检测到光。死细胞不能发光,故普通菌株在含有抗结核药物条件下被LRMs感染后不产生光信号,而抗药性菌株能继续产生光信号。因为少量的光就能检测到,相对来说只需要少量的细菌,避免了生长期过长的要求,LRMs把药敏检测时间从几周缩短到几天^[9]。

初步实验揭示用L5和其亲缘种构建的萤光素酶报告噬菌体可以有效地控制萤光素基因表达的时间和水平,提高检测灵敏度。L5基因组左臂上有许多开放阅读框架,大多对噬菌体的生长是必须的。左端tRNA基因之间即基因8和9(tRNA^{Trp}和tRNA^{Gln})之间有NheI限

制位点(bp4441),把FFlux插入该位点似乎不影响噬菌体的必需功能。由于L5基因组含有几个NheI位点,故采用两步构建重组噬菌体:首先构建PGS24质粒(含L5 4kb片段6~12基因及在NheI位点插入FFlux),然后把L5噬菌体导入含pGS24的*M.smegmatis*中,筛选出两种重组噬菌体:phGS15和phGS18,都可作为萤光素酶报告噬菌体,其中phGS15含4.3kb缺失(基因72~77),phGS18含1.3kb缺失(基因83~88)。克隆过程中偶然形成的分枝杆菌启动子使插入的原噬菌体高水平表达萤光素酶。FFlux基因启动子位于FFlux DNA片段和L5基因连接点30bp处,如果与其它细菌的启动子结构相似的话,其-10区位于FFlux片段内,-30区位于邻近的噬菌体序列中。出人意料的是,FFlux在裂解生长期的表达比溶源生长时低,原理尚不清楚^[5]。phGS18在适当的时间内通过简单实验检测宿主存在的灵敏度接近PCR法,比PCR优越的是它只检测活菌,而且可检测对抗生素的表型反应^[5]。

LRMs的另一个重要的应用是寻找新的抗结核药物,传统的药物筛选方法由于*M.tuberculosis*生长缓慢而很困难,利用LRMs可自动、快速、简便地检测和发现新的抗结核药物。

5.3 建立分枝杆菌的遗传分析系统 *M.tuberculosis*和其它病原性分枝杆菌的遗传分析缺少成熟的遗传系统,L5等温和性噬菌体的分子生物学研究提供了一个有用的病毒系统。除上面提到的应用以外,L5在分枝杆菌的转化、抗药性机理、保护性和致病性的抗原决定簇、细胞壁的结构和生物合成、分枝杆菌基因调控机制等方面的研究都有广泛的应用潜力。为结核病的快速诊断、防治和治疗提供新的途径^[2,15]。

参 考 文 献

- [1] Bloom B R, Murray C T L. *Science*, 1992,257: 1055~1064.
- [2] Hatfull G F. *Features*, 1994,60(5):255~260.
- [3] Hatfull G F, Sarkis G J. *Mol Microbiol*,1993,7(3): 395~405.
- [4] Snapper S B, Lugosi L, Jekkel A, *et al.* *Proc Natl Acad Sci USA*, 1988,88:6987~6991.
- [5] Sarkis G J, Jacobs W R, Hatfull G F. *Mol Microbiol*, 1995,15(6):1055~1067.
- [6] Hatfull G F, Jacobs W R. *American Society for Microbiology*,1994,165~183.

- [7] Oyaski M, Hatfull G F. Nucl Acids Res, 1992, 20: 3251.
- [8] Mary K D W, Jacobs W R, Hatfull G F. Mol Microbiol, 1993, 7(3): 407~417.
- [9] Jacobs W R, Barletta R G, Rupa Udani, *et al.* Science, 1993, 260: 819~822.
- [10] Nesbit C E, Levin M E, Mary K, *et al.* Mol Microbiol, 1995, 7(6): 1045~1056.
- [11] Lee M H, Pascopella L, Jacobs W R, *et al.* Proc Natl Acad Sci USA, 1991, 88: 3111~3115.
- [12] Lee M H, Hatfull G F. J Bacteriol, 1993, 175(21): 6836~6841.
- [13] England P M, Wall S, Mcfadden J. Mol Microbiol, 1991, 5: 2047~2052.
- [14] Stover C K, Delacruz V F, Feurst T R, *et al.* Nature, 1991, 351: 456~460.
- [15] Hatfull G F. Trends in Microbiology, 1993, 1(8): 310~314.