*Mycobacterium leprae*の遺伝子について

中田 登*

国立感染症研究所ハンセン病研究センター

〔受付：2001年5月22日〕

キーワード：ゲノム、インテイン、偽遺伝子

*Mycobacterium leprae*ゲノムDNAの全配列が決定され公開された¹⁾。*M. leprae*ゲノムに関する研究は1980年代から行われており、guanine+cytosine含有率が約58%であること、ゲノムサイズが他のマイコバクテリアと比べて1/2~3/4と小さいことなどが知られていた。その後、多数の*M. leprae*遺伝子がクローニング、シーケンシングされたが、ゲノムシーケンシングに関する最初の報告は1993年に行われ、4つのcontig上に合計約2.8Mbのコスミド断片がマップされていた。最近公開されたゲノムシーケンスでは以前のものと比較して遺伝子マップが大きく修正されている。*M. leprae*ゲノムシーケンスから直ちにハンセン病治療に関して有益な情報が得られるわけではないが、*M. leprae*ゲノムの全体像が明らかになったことにより、ハンセン病治療、および*M. leprae*研究に新たなアプローチへの道が開かれたと言える。以下、筆者らの行った*M. leprae*の*katG*領域に関する分析を含めて、*M. leprae*のゲノムや個々の遺伝子に見られる特徴的な構造について述べる。

*trxB/A*遺伝子

*trxB*および*trxA*遺伝子はそれぞれチオレドキシンレダクターゼ(TR)、チオレドキシン(Trx)をコードする遺伝子であるが、*M. leprae*の*trxB/A*遺伝子はTRとTrxという2つの異なる機能を一つのタンパク質としてコードしている²⁾。TRとTrxは酵素と基質という関係があり、*Mycobacterium tuberculosis*を含めて一般に他の細菌では別々のタンパク質として発現するが、*M. leprae*ではこれら二つのタンパク質が一つのタンパク質として産生される。TRはNADPHの電子をTrxに渡して還元し、還元されたTrxは生体内の種々の重要な反応

で電子供与体として機能することが知られており、たとえば、活性酸素分子の処理などにも関わっていると考えられている。これら2つの遺伝子は*大腸菌*K-12株では染色体上の離れた位置に逆向きにコードされているが、*M. tuberculosis*を含む一部のマイコバクテリアでは二つの遺伝子がクラスターを形成しており、フレームを別にして1塩基のオーバーラップで隣接している。*M. leprae*ではこれら二つの機能が一つの遺伝子にコードされていて、言わば融合タンパクとして産生される。*M. leprae*のTR-Trx融合タンパクはタンパク濃度が低いときに効率よく働くことが示されており、また、*M. leprae*の*trxB/A*遺伝子を*Mycobacterium smegmatis*で発現させると細胞内での活性酸素中間体に対する抵抗性が增強されることが示されている³⁾。

*Corresponding author :

〒189-0002 東京都東村山市青葉町4-2-1

国立感染症研究所ハンセン病研究センター病原微生物部

Tel.042-391-8211 Fax.042-394-9092

E-mail:n-nakata@nih.go.jp

protein intron

通常DNAにコードされる遺伝子はmRNAに転写された後にタンパク質へと翻訳されるが、ある種の遺伝子ではタンパク質産物が翻訳された後に、その中央部が切り出されて前後の配列が結合し、成熟したタンパク質になることが観察されている。この反応はプロテインスプライシングと呼ばれ、切り出される中央部分をインティン、その両側のN末側をN-エクステイン、C末側をC-エクステインとそれぞれ呼ぶ。インティン部分をコードするDNAをprotein intronと呼び、これは多くの場合ある種のホーミングエンドヌクレアーゼと相同性を持ち、ゲノムと独立したある種の寄生的なジェネティックエレメントと考えられている。protein intronが見つかった *M. leprae* の遺伝子は *dnaB*、*gyrA*、*pps1*、*recA* であり、このうち *gyrA*、*recA* については実験的にインティン部分が切り出されることが確認されている。*recA* 遺伝子は、*M. leprae*、*M. tuberculosis* を含めた数種のマイコバクテリアでインティンの存在が示されているが、*M. leprae* と *M. tuberculosis* ではインティンの挿入部位やサイズが異なる⁴⁾。すなわち、*recA* 遺伝子におけるインティン遺伝子の挿入は *M. leprae* と *M. tuberculosis* とでは系統発生において別々に起こった事象であると考えられる。また、*recA* 遺伝子にインティンの存在が示されている他のマイコバクテリアでは全て *M. leprae* 型の挿入を持つことが知られている。細菌に見られるインティンの多くはマイコバクテリアに見つかったが、マイコバクテリアの遺伝子でもインティンが見られるのは *dnaB*、*recA*、*gyrA*、*pps1* の4種の遺伝子に限られており、*M. leprae* ではその全ての遺伝子にインティンが見られる。インティンの発見はゲノムシーケンシングが行われている菌種かどうかということも関係するが、*M. leprae* に多くのインティンが多く見つかったということについて、菌の増殖速度や病原性との関連が疑われている。

Pseudogene (偽遺伝子)

katG 遺伝子はカタラーゼとパーオキシダーゼの

両方の機能を持つ酵素をコードする遺伝子であるが、筆者らは *M. leprae* の *katG* 領域の塩基配列を明らかにして、これが pseudogene であることを報告した^{5,6)}。pseudogene は、機能している遺伝子の塩基配列と相同性の高い塩基配列を持つが、遺伝子としての機能を持たないDNAの配列のことを意味し、機能のある遺伝子から、塩基の置換、欠失、重複などによって生じると考えられている。*katG* 遺伝子領域のDNAシーケンスを *M. leprae* と *M. tuberculosis* でホモロジープロットによって比較すると、お互い高い相同性を持つことがわかるが(図1)、プロットに2カ所不連続な部分が見られることから、*M. leprae* の *katG* には約100bpの欠失領域が2カ所あることがわかる。また、*katG* 遺伝子のすぐ上流には鉄の取り込みに関係する *furA* 遺伝子があるが、これも *M. leprae* と *M. tuberculosis* である程度の相同性が見られる。*furA* の上流と *katG* の下流では相同性が消失しており、この両端で遺伝子組み換えが行われたことが示唆される。一方、*katG* 領域をオープンリーディングフレーム(ORF)分析で比較すると *M. tuberculosis* では約2kbの *katG* のORFが見られるが、*M. leprae* では全てのフレームに多くの終止コドンがみられるなど、*katG* 遺伝子としてのORFは見られないことがわかる(図2)。*furA* についても同様で、*M. tubercu-*

図1 *katG* 領域のホモロジープロット分析

losisにおいては約450bpのORFが見られるのに対して *M. leprae* ではそれが見られないことから、pseudogeneであることがわかる。ゲノムシーケンスの結果から *M. leprae* ゲノムには多くのpseudogeneが存在することがわかってきた¹⁾。それによると、*M. leprae* ゲノムの27%をpseudogeneが占めており、機能を持つ遺伝子が占める割合は49.5%であるという。*M. tuberculosis* ゲノムの90.8%が機能を持つ遺伝子で占められていることを考えると、*M. leprae* ゲノムの遺伝子密度は著しく低い。*katG* 遺伝子周辺領域の遺伝子マップを *M. leprae* と *M. tuberculosis* で比較しても、*M. leprae* ではこの領域に多くのpseudogeneが見られることがわかる(図3)。*M. leprae* のゲノムサイズが他のマイコバクテリアと比較して小さいこと、そして多くのpseudogeneを持つことは、*M. leprae* が系統発生において多数の遺伝子を失って進化してきたことを示している。

Repetitive element

M. leprae がそのゲノム中に特有のリピート配列

図2 *katG* 領域のオープンリーディング分析
各フレームの三角印は開始コドンを表し、縦の線は終止コドンを表す。黒色で示した部分はオープンリーディングフレームを表す。

を持つことは比較的早くから知られていた。このリピート配列は特に他の細菌で見つっているISなどとは相同性が見られず、これ自体が染色体上を移動するような性質はDNAシーケンスからは見られないが、進化の過程においてはそのような役割を持っていたのではないかと考えられている。このリピート配列は545bpの共通のコアを持ち、これが *M. leprae* ゲノムに散在している⁷⁾。例えば逆向きのリピートが隣接した場合にそれらがISのインバーテッドリピートIRのような役割をして組み換えを起こした可能性が考えられている。そのような部分的な組み換えが頻繁に起こった結果、*M. leprae* のゲノムが縮小されたと考えられる。

多型性

M. tuberculosis などではISの数や挿入部位などが株間によって異なることはよく知られているし、部分的な領域の欠落なども見つっているが、*M. leprae* は分離株間における多型性がほとんど見られないということが以前から知られていた。筆者の知識によれば、薬剤耐性に関わる変異を除いた場合、現在までにシーケンス上で多型性が知られている部分は4カ所しかない。筆者らの分析により、*groEL1* 遺伝子の近傍に78bpのリピート配列を持つ *M. leprae* 株が見つっている(未発表データ)。また、*rpoT* 遺伝子内の配列に分離株によって二つのタイプがあり、アスパラギン酸、イソロイシンに相当するGACATCという6

図3 *M. leprae* と *M. tuberculosis* の *katG* 領域の遺伝子地図¹⁾
白矢印は偽遺伝子を表し、黒矢印は機能を持つ遺伝子を表す。

塩基が4コピー見られる菌株と3コピーの菌株があることがわかっている⁸⁾。また、*polA*遺伝子の近傍の配列に多型性が見られるという報告がある⁹⁾。そしてゲノムマップ上の約2,785kbの位置に存在するTTC3塩基のリピートの数が分離株間で異なることが報告されている¹⁰⁾。このように*M. leprae*では菌株間における遺伝的多型性についてはごく僅かが知られているのみであるが、このことは*M. leprae*にpseudogeneが多いことが関係していると考えられる。例えばDNAポリメラーゼⅢのイブシロン鎖をコードする*dnaQ*遺伝子が*M. leprae*ではpseudogeneであるため、DNAに損傷が起きたときにそれをうまく修復できずに致死的になり、結果としてゲノムDNAの多型性が見られなくなるという可能性などが考えられている。

病原性に関連する遺伝子について

*M. leprae*の個々の遺伝子やゲノム全体のシーケンシングによって明らかになったことのうち、最も重要なことはやはり多数のpseudogeneの存在が明らかになったことであろう。DNAシーケンスによるORF分析の結果のみから個々の遺伝子がpseudogeneであると結論づけるのは安易であり、より詳細な分析が必要であると考えられるが、*M. leprae*において多くの遺伝子が不活化されているということは、*M. leprae*の増殖速度の遅さや、人工培地での培養が達成されていないこととの関連を強く示唆するものである。例えば*M. tuberculosis*では、グリオキシル酸回路で働くイソクエン酸リアーゼをコードする遺伝子として、*icl*と*aceA*という二つの遺伝子を持つが、*M. leprae*ではこのうち*icl*がpseudogeneとなっていることなどから、*M. leprae*は菌の生存にとって最小限かまたはそれに近い遺伝子のセットしか持っていないことが示唆される。

*M. leprae*は細胞内寄生菌であるという点では*M. tuberculosis*と共通しているため、宿主細胞内での生存に関して*M. tuberculosis*などと共通のメカニズムを持つことが考えられる。しかし、*M. tuberculosis*の宿主細胞内での生存に関わっているとされている遺伝子のうち多くのものは*M. leprae*には存在しないか、あるいは*katG*遺伝子のように

pseudogeneとなっており、*M. leprae*は、細胞内での生存を含めた病原性という点で*M. tuberculosis*とは異なったメカニズムを持つことが強く示唆される。

また、神経親和性という性質は*M. tuberculosis*に見られない*M. leprae*に特有の性質であるが、これを規定している遺伝子についても不明な点が多い。最近、*M. leprae*のシュワン細胞への接着に、21kDaタンパク質やPGL-1といった糖脂質が関わっていることが報告されたが^{11,12)}、一方でシュワン細胞への菌の取り込み率は*M. leprae*と他の菌で変わらないという報告もある¹³⁾。*M. leprae*と*M. tuberculosis*のゲノムシーケンスを比較した場合、*M. leprae*にのみに見られる遺伝子が165あるとされているが、それらの遺伝子が*M. leprae*に特有の性質を規定している可能性もあり、今後*M. leprae*の病原性を規定している遺伝子を明らかにすることが必要である。

文 献

- 1) Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR and Barrell BG. Massive gene decay in the leprosy bacillus. *Nature* 409: 1007-1011 (2001).
- 2) Wieles B, van Soolingen D, Holmgren A, Offringa R, Ottenhoff T and Thole J. Unique gene organization of thioredoxin and thioredoxin reductase in *Mycobacterium leprae*. *Mol Microbiol* 16:921-929 (1995).
- 3) Wieles B, Ottenhoff TH, Steenwijk TM, Franken KL, de Vries RR and Langermans JA. Increased intracellular survival of Myco-

- bacterium smegmatis containing the *Mycobacterium leprae* thioredoxin-thioredoxin reductase gene. *Infect Immun* 65: 2537-2541 (1997).
- 4) Davis EO, Thangaraj HS, Brooks PC and Colston MJ. Evidence of selection for protein introns in the *recAs* of pathogenic mycobacteria. *EMBO J* 13:699-703 (1994).
 - 5) Nakata N, Matsuoka M, Kashiwabara Y, Okada N and Sasakawa C. Nucleotide sequence of the *Mycobacterium leprae* *katG* region. *J Bacteriol* 179: 3053-3057 (1997).
 - 6) Nakata N. The absence of a catalase function in *Mycobacterium leprae*. *Indian J Lepr* 71: 87-92 (1999).
 - 7) Woods SA and Cole ST. A family of dispersed repeats in *Mycobacterium leprae*. *Mol Microbiol* 4:1745-1751 (1990).
 - 8) Matsuoka M, Maeda S, Kai M, Nakata N, Chae GT, Gillis TP, Kobayashi K, Izumi S and Kashiwabara Y. *Mycobacterium leprae* typing by genomic diversity and global distribution of genotypes. *Int J Lepr Other Mycobact Dis* 68: 121-128 (2000).
 - 9) Fsihi H and Cole ST. The *Mycobacterium leprae* genome: systematic sequence analysis identifies key catabolic enzymes, ATP-dependent transport systems and a novel *polA* locus associated with genomic variability. *Mol Microbiol* 16:909-919 (1995).
 - 10) Shin YC, Lee H, Lee H, Walsh GP, Kim JD and Cho SN. Variable numbers of TTC repeats in *Mycobacterium leprae* DNA from leprosy patients and use in strain differentiation. *J Clin Microbiol* 38:4535-4538 (2000).
 - 11) Shimoji Y, Ng V, Matsumura K, Fischetti VA and Rambukkana A. A 21-kDa surface protein of *Mycobacterium leprae* binds peripheral nerve laminin-2 and mediates Schwann cell invasion. *Proc Natl Acad Sci U S A* 96: 9857-9862 (1999).
 - 12) Ng V, Zanazzi G, Timpl R, Talts JF, Salzer JL, Brennan PJ and Rambukkana A. Role of the cell wall phenolic glycolipid-1 in the peripheral nerve predilection of *Mycobacterium leprae*. *Cell* 103:511-524 (2000).
 - 13) Band AH, Bhattacharya A and Talwar GP. Lack of *Mycobacterium leprae*-specific uptake in Schwann cells. *Int J Lepr Other Mycobact Dis* 54:71-8 (1986).

Genome features of *Mycobacterium leprae*

Noboru Nakata*

Department of Microbiology, Leprosy Research Center National Institute of Infectious Diseases

[Received: 22 May 2001]

Key words: genome, intein, pseudogene

Recent studies have revealed that the *Mycobacterium leprae* genome contains many pseudogenes. This short review summarizes the structural features of the *M. leprae* genome and genes.

*Corresponding author :

Department of Microbiology Leprosy Research Center
National Institute of Infectious Diseases
4-2-1 Aobacho, Higashimurayama-shi, Tokyo 189-0002, Japan
Tel.81-42-391-8211 Fax.81-42-394-9092
E-mail : n-nakata@nih.go.jp