

## ヒト腸内フローラのメタゲノム解析

### Metagenomics of microbial community in human intestine

服部正平 (Masahira HATTORI)

北里大・北里生命研 / 理研・GSC

**【背景】**メタゲノム解析は、自然界の複雑多様な細菌集団(叢)のゲノムシーケンス情報を(培養を経ないで)大量に獲得し、そのインフォマティクス解析によって、構成菌種や遺伝子組成等の細菌叢の全体像を解析する方法である。今日では、メタゲノム解析は技術的に応用可能であり、既に、多数の難培養細菌から構成される海洋や土壌等の環境細菌叢のプロジェクトが米国を中心に進められている。

ヒト腸内フローラに関しては、今日までに、細菌の分離培養と 16S rRNA 配列等をベースにした解析手法による多くの成果が蓄積されている。最近の大規模な 16S rRNA 解析では、約 400 種の細菌が同定されている。一方で、ヒト腸内フローラ構成細菌のゲノム解析の多くは、ビフィズス菌、大腸菌、ラクトバチラス、バクテロイデスなど、プロバイオティクスに属する乳酸菌類や日和見感染症の起原菌に限られ、その数は論文発表と進行中あわせて 30 種にも満たない(今日までにゲノム解析されている細菌種の総数は約 1,300 種)。ヒト腸内フローラ構成細菌のゲノム解析数が少ない理由として、それがきわめて多くの細菌種から構成され、その 80%以上が嫌気性で難培養性または培養条件の確立がきわめて困難な菌種であること等が挙げられる。

**【目的】**難培養細菌も含めた全細菌を対象にした 16S rRNA 解析においても、注意すべき点または問題点がいくつかある。たとえば、(1) 共通プライマーで PCR 増幅しない rRNA 配列をもつ菌種の存在を否定できない、(2) 用いる PCR 条件において、すべての 16S rRNA 配列をバイアスなくまた組成比を変えずに増幅している保証がない、(3) 菌種数の全体をカバーする大量の 16S rRNA 配列を解析する必要がある、(4) 1つの菌種がもつ 16S rRNA のコピー数は菌種ごとに異なるので、16S rRNA 数の比は必ずしも細菌の組成比を直接表わしていない、(5) 機能と直結した遺伝子情報が得られない、等を挙げることができる。このような腸内フローラを解析する上での諸問題に打ち勝つ方法として、また、圧倒的多数を占める難培養菌も含めた構成細菌種の遺伝子情報を獲得する目的で、メタゲノム解析を腸内フローラ解析に応用した。本シンポジウムでは、これまでに得られた結果の一部を報告する。

**【方法】**伊藤喜久治博士(東大農)との共同研究によって、約 30 人のふん便由来腸内フローラを単離し、そのゲノム DNA を調製した。そのうち、8 サンプルについて、ショットガンライブラリー作製、鋳型 DNA 調製、シーケンス反応、配列データの取得を行った。トータルで 400Mb 以上の配列データを得た。配列データに対する Blast 等の情報学的解析によって、フローラの菌種組成、COG 解析、個人フローラ間およびほか環境細菌叢との比較解析等を行った。

【結果】結果の一部を以下に示す。

- (1) Blast 検索およびメタゲノム解析用に開発した新しい遺伝子予測プログラムによって、1つのフローラにつき少なくとも4万個の遺伝子を同定した。
- (2) UniProt データベースの蛋白質に対して、予測された遺伝子のうちの約70%が有意なアミノ酸配列相同性をもっていた。
- (3) アミノ酸配列相同性は45-100%の範囲であったが、約70%および95%以上の配列類似度付近に、腸内細菌遺伝子の多くが相同性を示した。
- (4) COG データベースに対する解析から、2,000種類以上のCOGを特定した。
- (5) 600種以上の菌種が特定され、ACE および Chao1 estimators を用いた外挿法から、少なくとも800菌種の存在が示唆された。
- (6) 成人フローラでは、*Bacteroides*、*Clostridium*、*Bacillus* が、幼児では、*Escherichia* およびほかの *Enterobacteria* が、おもな優先菌種となった。

【考察】今日の公的データバンクに登録されている細菌ゲノムデータの大部分が、病原菌由来やまったく異なった環境に由来した菌であるため、常在菌メタゲノムデータからの菌種の特定は推定の域を出ない。一方、腸内フローラから得られた配列データのうち、その約70%に遺伝子機能を特定できた。さらに、個人間のCOGや菌種組成の違い、ほかの環境メタゲノムとの比較等の解析を行うとともに、構成菌種の分離とその個別シーケンス等も進めている。

最近、ヒト常在菌の大規模なゲノム解析をめざした国際コンソーシアムによる「ヒトメタゲノム計画」が提案されている。その第1回会議が2005年10月末にパリで開催された。この会議には、著者らも含めて日米欧中から約70名が参加し、ゲノム解析、機能解析、リソース/データ共有など、幅広い課題について議論が行われた。この計画では、大きく2つのチャレンジがある。ひとつは構成菌種の網羅的な個別ゲノム解析、もうひとつは個人フローラのメタゲノム解析である。前者は多くの未知難培養菌の分離とそれらの培養方法の確立が含まれ、後者は健康および病態フローラの大規模シーケンシングとインフォマティクス解析が必要となる。これら2つは相補完するもので、メタゲノム解析によって、フローラを構成する菌種数や組成比の見積もり、未知菌種の存在の確認ができ、個別解析が進めば、メタゲノムデータをより高い精度で解析できる。著者らはこの国際計画に参加、貢献するとともに、将来的に数千人から数万人の規模での個人フローラのメタゲノム解析を現実的なスピード及びコスト面で実行できる技術や体制を我が国に確立する必要があると考えている。これらが達成できれば、腸内フローラと健康や病気の関係をより精密に解析することができ、得られる成果は、学術分野はもとより我が国の医療や食品分野の産業にも計り知れない波及効果をもつと期待している。本研究は、おもに平成17年度開始の文科省特定領域研究「基盤ゲノム」科学研究費で行われている。おもな参加グループは、北里大学、奈良先端科学技術大学院大学、東大新領域、東大農、三菱総研、海洋開発研究機構、宮崎大学、麻布大学、徳島大学、理研GSCである。