

## 高速ゲノム解析システムの開発

### Development of New System for High Throughput Genome Analysis

(研究プロジェクト番号：JSPS-RFTF 97L00102)

#### プロジェクトリーダー

服部 正平 東京大学医科学研究所・助教授  
(平成9年度～平成10年度プロジェクトリーダー、平成11年度コアメンバー)  
柴 忠義 北里大学理学部・教授  
(平成11年度プロジェクトリーダー、平成9年度～平成10年度コアメンバー)



(服部正平)

地球上の生物は、様々な形態や能力を持っており、それらの設計図となるものがゲノムです。ゲノムは、DNAと言う物質でできており、G、A、T、Cと呼ばれるたった4文字(塩基)が並んだ構造をしています。生物の複雑さとゲノムDNAの文字数との間には、かなりの相関があることが分かっています。最近ゲノムの塩基配列決定がほぼ終了したヒトの場合、ゲノムDNAの文字数が約30億塩基であるのに対して、単細胞生物の大腸菌では、約470万塩基です。

微生物ゲノムDNAの塩基配列の決定には、ホールゲノムショットガン法と呼ばれる方法が用いられます。ホールゲノムショットガン法は、自動DNAシーケンサー(塩基配列読みとり装置)の登場により短期間に大量の塩基配列データの産生が可能になったことから、現在では微生物のゲノム配列決定に欠くことのできない方法となっています。

ホールゲノムショットガン法の各ステップを簡単に表すと、以下の通りになります(図を参照)。1)微生物から、ゲノムDNAを抽出する。2)ゲノムDNAをランダムに1000

～2000塩基に断片化する。3)各断片の塩基配列決定を行う(ひとつの断片で500塩基くらいのデータが得られます)。4)コンピュータを用いて、各DNA断片の塩基配列をつなぎ合わせる(アセンブル)。アセンブルとは、ホールゲノムショットガン法の核心部分であり、得られた塩基配列データ同士を総当たりで相同性を比較し、相同性の高い配列同士を次々と連結していくことを言います。

ホールゲノムショットガン法を用いて大腸菌のゲノム解析を行った場合、10万個のDNA断片からそれぞれ500塩基の塩基配列データを得て、アセンブルを行うこととなります。ゲノムDNAはたった4文字で書かれているため、似かよった文字の配列がたくさん存在します。したがって、アセンブルは単純で簡単な作業ではなく、最後は人間の手作業によって連結します。

ゲノムDNAの塩基配列が決定した後、生体内の主要な部品であるタンパク質の設計図(遺伝子)を見つけだし、名前や特徴付けを行います。こうして最終的にその微生物の特徴を知ることができるようになります。

#### 生物のゲノムの大きさ

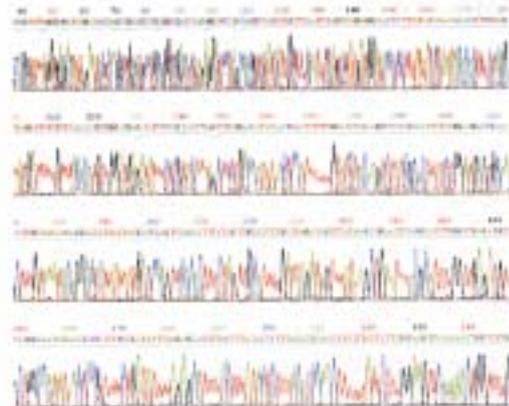
生物は固有のゲノムDNAをもっていますが、その大きさと生物の複雑さとの間には、かなりの相関があります。

ゲノムDNAの文字数	
ヒト	30 億
マウス	30 億
イネ	4.5 億
ショウジョウバエ	2 億
酵母	1340万
大腸菌	470万
マイコプラズマ	58万

#### 自動DNAシーケンサーの結果

自動DNAシーケンサーからは下の4色の波形データが得られます。

G—黒、A—緑、T—赤、C—青となっています。



## ホールゲノムショットガン法の流れ



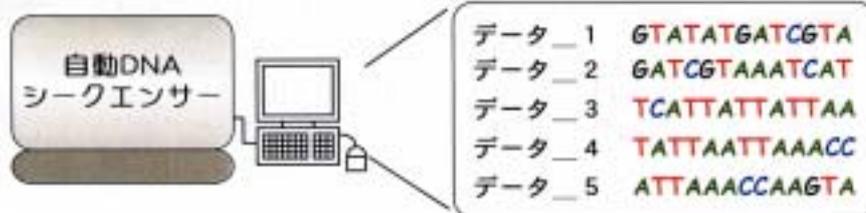
↓  
微生物からゲノムDNAを取り出す



↓  
小さな断片にする



↓  
小さな断片ひとつひとつのDNA配列を  
自動DNAシーケンサーで読み取る



↓  
読み取られた配列データをコンピューターを  
用いてつなぎ合わせる (アセンブル)

GTATATGATCGTA  
GATCGTAAATCAT  
TCATTATTATTAA  
TATTAATTAACC  
ATTAACCAAGTA

↓  
間違っている部分の読み直しや読み足  
りない部分をつけ足して、完成させる

GTATATGATCGTAAATCATTATTATTAATTAACCAAGTA

完成！！