

好熱菌ゲノムデータベースの作成

Construction of Thermophile Genome Database

小山芳典

Yoshinori Koyama

(産総研 生物機能工学)

(Institute for Biological Resources and Functions, AIST.)

e-mail: y-koyama@aist.go.jp

塩基配列決定技術の進歩により数多くの微生物ゲノム塩基配列が決定されてきたが、対象を好熱菌のみに絞った容易に使えるゲノムデータベースは作成されていなかった。そこで全塩基配列が決定されている好熱菌についてパソコン上で容易にその遺伝子の表示、検索および BLAST サーチができるシステムを構築した。

全塩基配列が決定されて GenBank に登録されている微生物ゲノムのデータから、生育至適温度が摂氏 60 度以上であると判断される好熱菌 46 種を選び、NCBI より GenBank 形式、塩基配列 Fasta、タンパク配列 Fasta をダウンロードした。GenBank 形式のデータは BioPerl の bp_genbank2gff.pl コマンドにより GFF3 形式のデータに変換した。これらを連結後、bp_bulk_load_gff.pl コマンドで MySQL プログラムにデータを導入した。配列 Fasta データは formatdb コマンドで Blast Data に変換した。

GMOD(the Generic Model Organism Database project)で開発されたゲノムブラウザー Gbrowse を Macintosh Mac OSX 上にインストールした。これによりマウスだけで 46 種の好熱菌から表示したいものを選択し、希望する領域の遺伝子構造を表示可能になった。また NCBI で開発されている Web タイプの wwwblast をインストールすることにより、46 種の好熱菌の塩基配列およびタンパク配列の高速 Blast 検索 (Blastp なら 1 秒) が可能になった。導入したプログラムはともに Apache Web サーバー上の CGI プログラムとして動くので、ネット接続された他のパソコンからも利用可能である。



図 1. 46 種の好熱菌から HB8 株を選ぶ

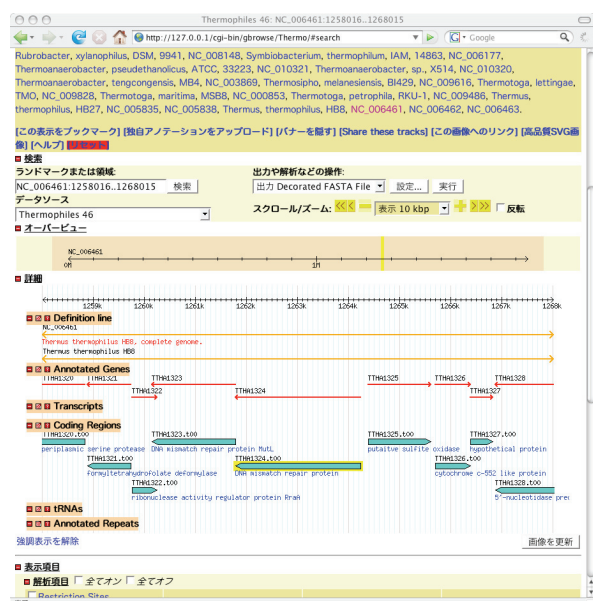


図 2. HB8 株 mutS 遺伝子付近の構造の表示