D 17 日本DNAデータバンクの紹介

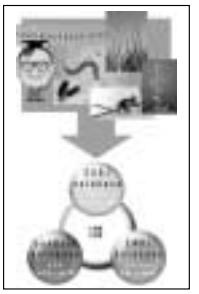
ゲノム情報を解析する

展示責任者 DDBJ

展示責任者所属
国立遺伝学研究所生命情報DDBJ研究センター

生命科学のめざましい発展の基盤として、DNA塩基配列から得られる知識は欠かすことのできないものとなっています。現在では、DNAの塩基配列情報はインターネットを通じて世界中の研究者が無償で入手できるシステムが確立されています。そのシステムの担い手の1つがDDBJ(図1)です。DDBJは、欧州のEBI/EMBLおよび米国のNCBI/GenBankとの密接な連携のもと、『DDBJ/EMBL/GenBank 国際塩基配列データベース』を構築しています(図2)。3つのDNAデータバンクによる国際協調の結果、日本の研究者は自分が明らかにしたDNA塩基配列情報を、DDBJを通して国際塩基配列データベースに登録することが出来ます。「ゲノムひろば」で紹介される数多くのゲノム研究で明らかにされた大量かつ有益なDNAの塩基配列情報も、そのほとんどがDDBJを通して登録され、全世界に公開されました。本ポスターでは、ゲノムの塩基配列情報がどのようなプロセスを経てデータベースに組み込まれ研究者に提供されるか、これまでどのようなゲノムDNAがDDBJを通して全世界に公開されたか、などについて紹介します。

(図1) DDBJのホームページ
(http://www.ddbj.nig.ac.jp/Welcome-j.html)

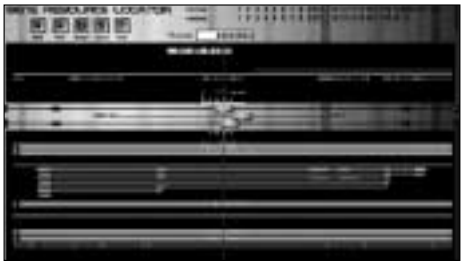
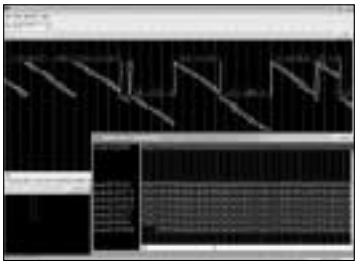


(図2) 様々な生物種のゲノム配列は国際塩基配列データベースを通して全世界に公開されます。
(画像出典:国立遺伝学研究所 遺伝学電子博物館 http://www.nig.ac.jp/museum/index.html、
同角谷研究室 http://www.nig.ac.jp/labs/AgrGen/epigenetics.html、
水田の生き物 イネの生活 http://www2.saganet.ne.jp/mono_k/ikimono/moine06.html、
国立科学博物館 微小藻の世界 http://www.kahaku.go.jp/special/past/bisyoso/ipix/mo/3/3_5.html)

D 18 ゲノムを解読するには、どんな計算が必要なのでしょう?

ゲノム情報を解析する

展示責任者 森下真一／土井晃一郎／瀬々　潤

展示責任者所属
東京大学大学院情報生命科学専攻

ゲノムの解読や遺伝子の解析にはコンピュータソフトウエアが必要だといわれます。どのような場面で「計算」が必要なのでしょうか?　たとえばヒトゲノムは約30億個の文字から成り立っていますが、文字をすべて解読するには、ゲノムを細かい数千万個の断片にして、それから繋ぎ合わせてゆきました。実はこの繋ぎ合わせる作業は結構大変で、自動化するため、いろいろと計算方法を工夫します。たとえばグラフ構造という数学の手法が巧みに使われます。それから、長さが数千の遺伝子配列を、非常に長いゲノム配列の中を探すのも時間がかかる作業です。高速で見落としがない方法を考えようとすると、確率論とアルゴリズムを組み合わせた設計方法が必要になります。さらにゲノムから読み取られる遺伝子の振る舞いを一度に観測できる装置の出現で、数万個の遺伝子の様子も追跡できるようになりました。数万個パラメータ(遺伝子)を処理できるデータマイニングという統計学の拡張が使われます。このようにゲノム解読には、さまざまな数学や計算の考え方が使われますが、基本的となる考え方をやさしく紹介しようと思います。