

8月1日に新班が立ち上がりました

## 「比較ゲノム研究班」の発足にあたって

後藤 修

比較ゲノム研究班 班長

近年数多くの生物種のゲノム塩基配列が解読されています。それらのゲノム塩基配列はそれぞれの生物種の遺伝情報全体を担っていますが、その内容を理解し有効に活用するには、過去から現在に至る時間軸を考慮に入れることが不可欠です。リチャード・ドーキンス風に言うと、今現在地球上に生存する生物のすべては、30億年以上にわたり連続として続く遺伝子の流れの一断面を具現化したものに他ならないからです。新しく発足する「比較ゲノム研究班」ではこの視点に重点を置き、様々な生物種のゲノム配列を相互に比較しながらより高次の知見へと高めることを目指しています。CBRCにはすでにゲノム情報解析を行うチームが二つ存在し、内容的に重なる部分も少なくありませんが、協力関係を保ちながらも以下のような独自の研究課題に取り組みたいと考えています。

### 1. 遺伝子ファミリーごとのゲノム横断的遺伝子予測

ゲノム配列解析の最初かつ最も重要な課題は、ゲノム配列がコードする遺伝子を同定することです。従来いくつかの異なる原理に基づく計算手法が開発されており、それらを用いて多くのゲノム配列上の遺伝子が予測されています。しかし、ヒトなどごく一部を除いて、それらの予測が十分に信頼できるものであることは検証されていません。私たちは、同じファミリーに属する遺伝子(同族遺伝子)を一括して相互比較することにより遺伝子予測精度を検証する方法を開発しています。さらにそれを発展させ、ファミリー単位で予測精度を向上させる手法の開発も行っています。その際、同一生物種由来であるか異種生物種由来であるかに関わりなく同族遺伝子として取り扱います。現在、代表的な多重遺伝子族であるチトクロームP450超遺伝子族を対象に、この方法の有効性を確認する計画を進めています。

### 2. 転写産物のゲノム配列への効率的マッピング手法

cDNA, EST, タンパク質アミノ酸配列などの転写産物に対応するゲノム上の位置を同定する作業をゲノムマッピングと呼んでいます。前項とも関連しますが、マッピングの結果は遺伝子の内部構造を正確に知る有力な手掛かりとなります。私たちは、独自の戦略に基づくマッピング手法を創案し、イントロンの存在を意識した「スプライスアライメント法」と組み合わせることにより、効率的で高精度に遺伝子の内部構造を推定できる方法を開発しています(図1)。現在、次世代シーケンサーからのデータにも対応できるようにこの手法の拡張・改良を試みています。

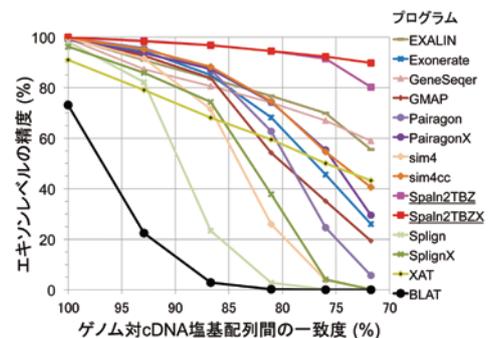


図 1. スプライスアライメントの精度比較

### 3. 全長ゲノム配列間比較法

異なる生物種のゲノム塩基配列を直接比較することは比較ゲノム研究の最も基本となる手段です。しかし、哺乳動物ゲノムなど、30億文字に及ぶ配列の相互比較は近年のコンピュータの性能向上をもってしても難しい課題です。私たちは「配列の粗視化」という技法を用いて計算の効率化を図る試みを行っています(図2)。私たちの方法は効率的であるとともに、転座や逆位などゲノム配列の大規模な再編成を伴う難しい状況にも比較的容易に対応できるという特徴があります。今後、複数のゲノム配列の同時比較への拡張などを目指す予定です。

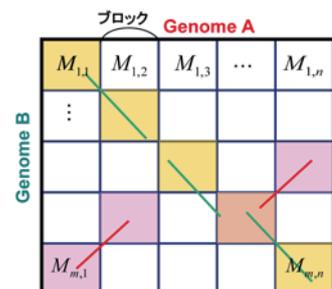


図 2. 粗視化アライメント法