ヒトゲノム情報解読に向けた新技術

マーティン・フリス
配列解析チーム
研究員

私の研究の目指すところは、ヒトや他の生物を形作る遺伝子コードを解明することです。特に、胚の発生において遺伝子発現を調節するDNAコードの解読に関心を寄せています。多くの場合、DNAコードは種の壁を越えて（ヒトと魚類など）高度に保存されているゲノム領域に存在することから、当該領域にある、極めて出現頻度の高いモチーフ（over-represented motifs）を見つけ出す研究を配列解析チームの大里直樹氏と共同で行っています。発見したモチーフの一部が、発生に関わる既知の転写因子（Hox遺伝子群など）のターゲットであるといった有望な結果が得られています。

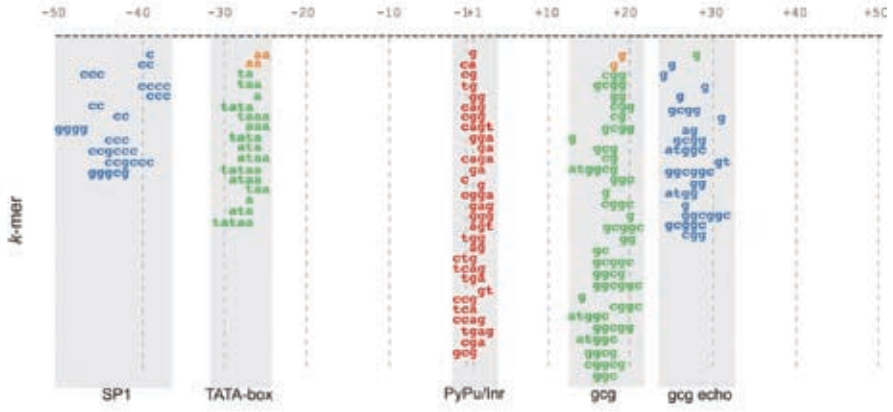
遺伝子発現の最初のステップである転写開始機構も研究テーマです。転写開始頻度は染色体上の階層的クラスタリングで説明できることを示しました。さらに、小クラスター内での転写開始部位の正確な選択に関与するDNAコードを突き止めました。現在、遺伝子発現が変化する際にどのクラスター層が調節されるのかを解明しようとしています。

その他に、配列の比較と整合を行うための新たな手法を開発中です。このような手法は、計算生物学（前述のプロジェクトなど）にはたいてい欠かすことのできないものですが、BLASTなどの既存ソフトウェアでは、新しいDNA配列決定技術によって生成される大規模なデータセットを処理するのは困難です。チーム長のポール・ホートン氏との共同研究では、LASTという新しいソフトウェアを開発しました。LASTを使えば、標準的なデスクトップコンピュータで、ヒトゲノムサイズの2つのデータセットをおよそ1日で比較することができます。LASTの高効率の決め手となっているのは、サフィックスアレイと呼ばれる興味深いデータ構造です。また、配列アライメントの不正確な部分をいかに検出し、回避するという重要ながら見逃されがちな問題の研究を進めています。

要するに、ゲノム配列中の生物情報を解読する手法を開発し、その手法を応用することが私の研究内容です。

Reference

http://www.cbrc.jp/~martin/



A code for transcription initiation