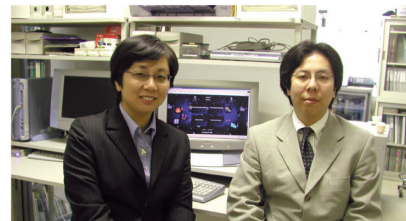


ユーザー訪問 数理システム

染色体の解釈に統計解析技術を利用 新しい研究手法をもたらしたS-PLUS

鹿児島大学理学部地球環境科学科 助教授
鹿児島大学理学部数理情報科学科 助教授
鹿児島純心女子短期大学 講師

宮本 旬子 Junko Miyamoto
宿久 洋 Hiroshi Yadobisa
末永 勝征 Katsuyuki Suenaga



鹿児島大学理学部助教授 宮本旬子氏(左)と宿久洋氏(右)

統計解析を手軽に行える 実用的なプログラム

染色体とは、生物の遺伝情報を記録している物質であるDNAが、ヒストンというタンパク質に巻き付いた、非常に小さな複合体である。染色体は生物の種類によって本数や細胞分裂中期の時の形などが決まっているため、それらを比較することにより、進化に関する情報が得られると考えられてきた。染色体研究では、初めに顕微鏡写真から染色体の大きさや縞模様の様子を単純化した模式図を作る。図の作成には、染色体写真を拡大し、糸やマップメーターで1つ1つ計測したり、切り取った写真の並べ替えなど手間のかかる煩雑な作業が多い。しかしこれらは顕微鏡写真のデジタル化、自動計測プログラム、描画ソフトなどの発達により効率化が進んできた。

しかしこれまで、模式図同士を比較する有効なツールは開発されておらず、長年の経験により蓄積された知識に頼るのみという状況だった。そのうえ、完成した模式図についての検討も、他の種と似ている、似ていないという記述的表現にとどまり、数値情報を統計的に処理してグルーピングするという試みは一般化されていなかったのである。

そのような染色体研究の状況を一変すべく、鹿児島大学理学部助教授の宮本旬子氏、

宿久洋氏、そして鹿児島純心女子短期大学講師の末永勝征氏らにより開発されたのが、統計解析ツールを組み込んだ染色体核型解析プログラムだ。このプログラムはS-PLUSを用い、web上で各研究者が作成した染色体の模式図から数値データを抽出し、統計処理を行うというものである。研究者はインターネットでアクセスし、自分の作成した染色体の模式図を画像データとしてプログラムに読み込ませ、認識されているかチェックするだけで、染色体の本数や、染色体の腕部分の長さ、染色体上に見られる縞模様の間隔といった、各種の特徴を示す数値データが自動的に算出される。さらに複数の染色体や複数の種類の関係をグラフや樹状図で簡単に作成することもできる。この開発により、これまで模式図を描くまでが主だった染色体の研究に、統計解析を活用し、複数の種の染色体の類似性を具体的に示すことができるようになった。

DNAだけでは分からない 生物進化の様子を解析

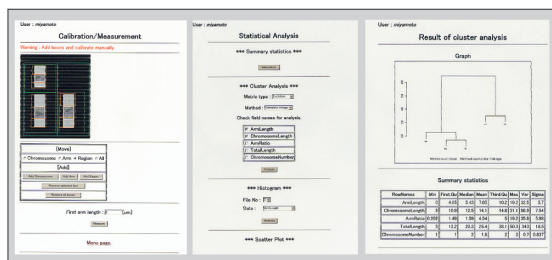
宮本氏の専門は野生植物の系統学である。生物進化研究の進展、特にDNAの塩基配列を比較することで、種の分化の系譜が分かるようになってきた。しかしDNAの調査だけで、実際の生物進化の様子を全て捉えられるわけではない。進化の様子を表す際、1本の系統が次々と二股に枝分かれしていく系統図が描かれるが、実際の生物では枝分かれした種同士の雑種が生まれる網目状の関係や、染色体数が倍に増える倍数化が見られる。これまでのDNA研究では、枝分かれまでは確認できるが、網

目状の系統関係や倍数化の様子まではよく分からない。宮本氏は染色体研究こそが網目状の進化や倍数化の理解に直結するのではないかと考え、染色体の形や塩基の配置を反映した模様の違いを客観的に比較し、数値解析できるツールを必要としていたのだ。

そして具体的なプログラム構成を考え、この開発をサポートしたのが宿久氏と末永氏。「データを入力し、解析結果をその場で確認できるサービスの実現には、データベースとは別に解析エンジン専用のサーバーを持つ必要がありました。そこでサーバーサイドでの解析サービスに耐えられる統計ソフトとして実績のあった、S-PLUSの導入を決めました。簡単なカスタマイズで様々な統計解析の機能を加えることができる拡張性も魅力でした」(宿久氏)。宿久氏の専門は計算機統計学であり、統計ソフトの開発も研究領域に入る。ちょうど研究者の間で個々のコンピューターからサーバーにアクセスし、解析サービスを受けるシステムが話題になっており、導入検討時、一般に販売されている統計ソフトの中でクライアント・サーバーシステムにいち早く対応していたのもS-PLUSだった。

「このプログラムは今年の春に完成したのですが、その前に卒業研究をまとめた学生たちが『早く作って欲しかった』と残念がっていました。『染色体の論文は記載的で面白くない』という評価もあったので、このプログラムが統計解析的な論文を書くための手助けになればと思います」(宮本氏)。

DNA研究はコンピューターの方で塩基配列解読、遺伝子発現プロファイル、相同解析、多型解析など様々な場面で多くの成果を上げてきた。同様に染色体研究でも、今回のプログラム開発によるコンピューターでの統計解析手法が確立されたことで、生物学全体に新たな知見をもたらすことが期待されている。



染色体核型解析プログラム。染色体領域を自動的に認識し、解析法や表示法を選択することで、目的のグラフや図を作成できる。