

イントロンの起源と進化：ゲノム進化における新たな機能の解明

● 剣持 直哉

宮崎大学フロンティア科学実験総合センター

＜研究の目的と進め方＞

真核生物の核ゲノムの遺伝子には、多数のイントロン（スプラ イセオソーム型イントロン）が存在する。これらのイントロンは どこから来たのか？ なぜ必要なのか？ いまだ多くの謎が残されて いる。私達は、これまでのリボソームタンパク質（RP）遺伝 子を用いた解析により、イントロンはミトコンドリアが共生した 後に核ゲノムに出現し、その後いっきに拡散した可能性があるこ とを示してきた。そして、核小体に存在する低分子RNA（snoRNA）がイントロン内にコードされていることから、イントロンは機能性ノンコーディングRNAのキャリアとして機能し ており、これを介して種分化に大きな影響を与えたのではないかと考えるにいたった。本研究は、ゲノム進化におけるイントロンの獲得と消失のパターンを推定し、その出現時期を明らかにするとともに、その動態と種分化との関連を明らかにすることを目的とする。

そのためのアプローチとして、①オルガネラのリボソームタンパク質遺伝子のイントロンを解析し、その出現時期を推定する。②真菌類のゲノムを比較解析することにより、新規イントロンの獲得機構を推定する。③イントロンにコードされたノンコーディングRNAを解析することにより、イントロンの新たな役割（機能性RNAのキャリア）を提示する。④以上の解析を行うための情報基盤として、リボソームタンパク質遺伝子およびsnoRNAのデータベースを整備する。

＜2008年度の研究の当初計画＞

1. RPGデータベースを用いたイントロンの獲得と消失のパターン推定：リボソームタンパク質（RP）遺伝子は、全ての生物が持ち、保存性が高く、種類も多いことから、イントロンの進化を解析する材料として有利なデータセットを与える。まず、基盤情報となるデータベース（RPG）を拡充する。これを用いて、イントロンの位置を最尤法で解析し、生物種ワイドなイントロンの獲得と消失のパターンを推定する。
2. オルガネラRP遺伝子を用いたイントロン出現時期の推定：ミトコンドリアのリボソームは細胞内共生したバクテリア由来とし、その遺伝子の多くはミトコンドリアゲノムから核ゲノムへ移行している。ミトコンドリアRP（MRP）遺伝子は多数のイントロンをもつが、これらが細胞内共生後に獲得されたものである可能性が高いことが、これまでの解析で明らかになっている。そこで、MRP遺伝子と細胞質のRPの獲得と消失のパターンを比較し、細胞質のRP遺伝子のイントロンが出現した時期の推定を試みる。同様に、クロロプラストのRP遺伝子のイントロンの位置を、ミトコンドリアと細胞質のRP遺伝子のもものと比較する。これらの解析のために、MRPとクロロプラストのRP遺伝子の情報を中心に収集する。
3. 非翻訳型RNA遺伝子のイントロンにコードされたsnoRNAの

解析：後生動物のsnoRNAは他の遺伝子のイントロンにコードされる場合が多く、その宿主遺伝子には、mRNAに十分な長さのORFがなく、タンパク質をコードしていない（非翻訳型）ものがある。私達は、一部の真菌類で非翻訳型と思われる宿主遺伝子を発見した。これらの遺伝子には、7つのエキソンが推測され、6種類のsnoRNAが各イントロンに1個ずつコードされていた（イントロン内在型）。別の真菌類では、オーソログのsnoRNA遺伝子が、イントロン内在型と同じ順序でゲノム上に並んでいたが、ポリシストロニックな発現をしていた（ポリシストロニック型）。そこで、イントロン内在型snoRNAの発現様式を解析することにより、これらsnoRNA遺伝子クラスターの祖先型を推定し、snoRNA遺伝子の挿入とイントロン進化との関連を明らかにする。

4. snoRNAデータベースの整備：snoRNA遺伝子を種間で比較することにより、snoRNA遺伝子がイントロンに入り込んだ時期を推定する。そのために、snoRNAデータベースを拡充し、解析の基盤とする。snoRNAのオーソログ、遺伝子の構成（モノシストロニック型、ポリシストロニック型、イントロン内在型）について情報を整備する。

＜2008年度の成果＞

1. RPGデータベースを拡充し、45生物種、7,000カ所のイントロンの獲得と消失のパターンを推測した。その結果、原生生物、植物、後生動物に共通の祖先生物には、すでに多数のイントロンが存在していた可能性が示された。また、共通の祖先生物から現存する生物にいたる進化の過程で、イントロンの獲得と消失は活発に起きており、生物種によりその頻度・パターンは異なることが明らかになった。
2. イントロンの出現時期を推定するためにオルガネラRP遺伝子の解析を行った。17生物種のミトコンドリアRP遺伝子のイントロン（1,332カ所）と、同じく17生物種の細胞質のRP遺伝子のイントロン（2,706カ所）について、獲得と消失の頻度を推定し、そのパターンを比較した。その結果、ミトコンドリアRP遺伝子においても、後生動物の祖先生物には多数のイントロンが存在していたこと、最近分岐した生物においてはイントロンの獲得と消失はあまり活発ではないことが示され、細胞質のRP遺伝子と同様な結果が得られた。これまでの解析で、ミトコンドリアRP遺伝子のイントロンはミトコンドリアが細胞内共生した後に出現した可能性が高いことが明らかになっている（Yoshihama et al. *PLoS Genet*, 2006）。今回、ミトコンドリアと細胞質のRP遺伝子のイントロンの獲得と消失のパターンが似ていることが示されたことより、細胞質RP遺伝子のイントロンも細胞内共生後に獲得されたものである可能性が考えられた。これは、イントロンは真核生物が誕生した後に出現し

たという、「イントロン後生説」を支持するものである。

3. 情報解析により、真菌類の一種であるイネいもち病菌で非翻訳型RNA遺伝子のイントロンにコードされている6種のsnoRNAを発見した。そこで、これらのRNAが実際に発現しているかどうかを確認するために、イネいもち病菌のトータルRNAを回収しRT-PCRを行った。その結果、宿主遺伝子のmRNAおよび内在するsnoRNAの発現が確認された。これらのsnoRNAは各イントロンに1個ずつコードされており、クラスターを形成していた。

4. snoRNAに関する情報の整備を行った。snoRNA遺伝子はヒトで約600種類、出芽酵母で約80種類が公共データベースに登録されており、他の生物でも多数の報告があるが、これらのオーソログの情報は十分に体系化されていなかった。snoRNA遺伝子の挿入とイントロンの進化を解析するために、snoRNAのオーソログの情報を整備する必要があると考え、多種生物のsnoRNAの情報を網羅したデータベースsnOPYを構築した。予備的なバージョンをこれまでも公開していたが、今回はデータを大幅に増やして、全体の構成も全面的に変更した。現在までに、35生物種、約17,000種のsnoRNA遺伝子を収集しているが、そのうち6生物種、1,842遺伝子に関する情報を公開している。

上記以外に、東京大学の菅野純夫教授と鈴木稔准教授との共同研究により、次世代シーケンサーから得られた配列をもとに決定したイントロン内のプロモーターの情報を得た。これらはイントロン内で発現している機能性RNAに関係している可能性がある。得られた11,506箇所のプロモーターの宿主遺伝子を解析したところ、RP遺伝子のイントロン内に89箇所、MRP遺伝子のイントロンに43箇所のプロモーターがマップされた。

<国内外での成果の位置づけ>

イントロンの進化を研究するための材料として様々な遺伝子が使われているが、RP遺伝子は、すべての生物が持っており種類も豊富なことから優れたデータセットであると考えられる。また、オルガネラの遺伝子もあり、起源や進化の解析にきわめて有利である。私達は長年、RP遺伝子の解析に従事し、遺伝子の構造解析、転写開始点の決定、データベースの構築など基盤情報の拡充を図ってきた。RPGでは、すべてのデータはマニュアルでチェックされており、非常に精度の高いデータセットとなっている。これを用いたイントロン進化の解析はきわめて優位性が高いといえる。

一方、snoRNAに関するデータベースに関しては、ヒトと出芽酵母に限定したもの（それぞれLBME、UMass-Amherst）が既に公開されており、これらは、標的RNA上の修飾塩基、snoRNA遺伝子のゲノム上の位置など、詳細な情報を提示している。しかしながら、オーソログに関する情報は不十分である。snOPYは、snoRNA遺伝子のオーソログの関係を主として扱うデータベースである。Ensemblよりアノテーションを参考に、これまでにない規模（17,000種）のsnoRNA遺伝子の情報を収集した。これらの情報の信頼性やオーソログ関係はキュレーターによる確認作業を通して決定し、一部を公開している。snOPYのような多種類の生物のsnoRNA遺伝子を横断的に網羅しているデータベースは他になく、今後、イントロンの意義を解析する重要なツールになると期待している。

<達成できなかったこと、予想外の困難、その理由>

イントロンの出現時期を推測するために、MRP遺伝子だけでなく、クロロプラストのRP遺伝子のイントロンの情報も用いて解析を行う予定であったが、十分なデータの収集ができなかった。これは技術的な問題ではなく、他の解析（snoRNAのデータ収集）に予想以上に手間と時間がかかってしまったためである。また、イネいもち病菌でsnoRNAの発現機構を解析するための実験系の構築に遅れが生じている。これは、イネいもち病菌の取扱が初めてであったことで、培養の条件等の設定に手間取ってしまったためと考えている。今後は順調な進展が期待される。

<今後の課題>

イントロンの進化を解析するためには、データベースの充実が欠かせない。RPGに関しては、クロロプラストRP遺伝子のデータを収集する必要がある。snOPYに関しては、真菌類のデータを充実させる。現在、慶應義塾大学の榊原康文教授と共同で「真菌類ゲノムにおけるsnoRNA遺伝子の同定」に取り組んでおり、自動解析により多数のsnoRNAの情報を得るためのアルゴリズムを開発中である。いずれにしても、マニュアルでキュレートしたデータの整備が重要であると考えている。

<成果公表リスト>

1) 論文／プロシーディング

1. 0812191600

Uechi, T., Nakajima, Y., Chakraborty, A., Torihara, H., Higa, S., Kenmochi, N.: Deficiency of ribosomal protein S19 during early embryogenesis leads to reduction of erythrocytes in a zebrafish model of Diamond-Blackfan anemia, Hum Mol Genet, 17, 3204-3211 (2008) .

2. 0901141814

Chakraborty, A., Uechi, T., Higa, S., Torihara, H., Kenmochi, N.: Loss of ribosomal protein L11 affects zebrafish embryonic development through a p53-Dependent apoptotic response, PLoS ONE, 4, e4152 (2009) .

2) データベース／ソフトウェア

・リボソームタンパク質遺伝子データベース (RPG)

<http://ribosome.med.miyazaki-u.ac.jp/>

・snoRNA データベース (snOPY)

<http://snoopy.med.miyazaki-u.ac.jp/>

3) 共同研究

・真菌類ゲノムにおけるsnoRNA遺伝子の同定
榊原康文 (比較ゲノム)

・snoRNAによるRNA修飾の同定
鈴木勉 (比較ゲノム)

・イントロン内にコードされた機能性RNAの探索
菅野純夫、鈴木稔 (基盤ゲノム)