

## 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（スプライソ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。

### <研究開始時の研究計画>

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細菌由来を由来とし、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. 新規に獲得されたイントロンの推定：菌類は、比較的ゲノムサイズが小さく、相同遺伝子が多く存在し、ゲノム解析の進んだ生物種が多いことから、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となりうる。イントロンが獲得、あるいは消失する機構は未だ不明であるが、最近になって変動したイントロンには、その痕跡が残されている可能性が高いと考えられる。私達は、菌類の近縁種に着目したイントロンの獲得と消失のパターンを、最尤法を用いて推定し、これまでに、新規に獲得された可能性のあるイントロン14個を見出した。本研究では、ホスト遺伝子の発現と14個の新規イントロンのスプライシングが正常に行なわれていることを確認する。同時に、基盤情報となるデータベース（RPG）を拡充し、上記と同様の手法で新規に獲得、あるいは消失したイントロンを見出す。

4. 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、そのホスト遺伝子には、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もの（非翻訳型）がある。私達は、一部の菌類で非翻訳型と思われるホスト遺伝子を発見した。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）。別の菌類では、オルソログの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。

5. snoRNAデータの充実化：snoRNAによるターゲットRNAの部位特異的な修飾は、ターゲットRNAと相補鎖を形成する約10ntの配列により可能となる。このため、異種生物間におけるsnoRNAのオルソログを推測することは、snoRNA全長に渡る相同性が低い場合でも、ターゲットRNAの修飾塩基の位置と周辺配列を比較することにより可能である。本研究では10生物種の800個のsnoRNA遺伝子のオルソログを決定する。これらのデータを用いて、進化の過程におけるsnoRNA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）の変遷を推測し、ホスト遺伝子のイントロンの獲得と消失のパターンと比較する。これにより、イントロン内在型snoRNA遺伝子が出現した時期を推定する。

### <研究期間の成果>

1. イントロンの獲得と消失のパターンを推定：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. オルガネラRP遺伝子を用いてイントロン出現時期を推定：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. 菌類で新規に獲得されたイントロンを解析：イントロンが獲得、あるいは消失する機構は未だ不明であるが、最近になって変動したイントロンには、その痕跡が残されている可能性が高いと考えられる。私達は、菌類の近縁種に着目したイントロンの獲得と消失のパターンを、最尤法を用いて推定し、フザリウムのゲノム中に新規に獲得された可能性のあるイントロン14個を見出した。これらイントロンを獲得したタンパク質遺伝子とイントロンの存在を確認するために、フザリウムからゲノムDNAとトータルRNAを回収しPCR法で解析した。その結果、14カ所のうち10カ所については、タンパク質遺伝子が発現していること、および、イントロンが成熟mRNAから除去されていることを確認した。残り4カ所は、タンパク質遺伝子の発現がないこと、さらに、私達が予測したイントロン配列がゲノム中に存在しないことが明らかになった。つまり、インフォマティクスと実験に用いた株のゲノム配列が異なっていることが明らかになった。これら4カ所のイントロンの配列は、既知のトランスポゾンと相同性が高かったことから、フザリウムのゲノムではトランスポゾンの転移が活発であることが示唆された。また、タンパク質遺伝子については、ゲノム配列のアノテーションにエラーがあった可能性があげられるが、別の可能性として、インフォマティクスで用いた株では、トランスポゾンが挿入したことによりタンパク質遺伝子の発現が活性化された、あるいは、実験に用いた株では、トランスポゾンが消失したことによりタンパク質遺伝子の発現が抑制された可能性があると考えた。これを確認するためには、今後、インフォマティクスで用いた株のゲノムDNAとトータルRNAで調べる必要がある。これらは、真核生物におけるイントロンの獲得機構を解明するための有力な手がかりとなると考えている。

4. 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 (イントロン内在型)、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、インフォマティクスにより菌類の一種であるイネいもち病菌においてsnoRNA遺伝子クラスターを見出し、それらがイントロン内在型であること、また宿主遺伝子は非翻訳型であると予測した。これらを確認するために、イネいもち病菌のゲノムDNAとトータルRNAを回収しPCR法で解析した。その結果、宿主遺伝子のmRNAと内在するsnoRNAの発現が確認され、また、宿主遺伝子は7つのエキソンから構成されること、および、6種類のsnoRNAは各イントロンに1個ずつコードされていることが明らかになった。出芽酵母では、オーソログのsnoRNA遺伝子が同様なクラスターを形成して

いるが、ポリシストロニックな転写を経て発現することが既に報告されている (ポリシストロニック型)。そこで、snoRNA遺伝子クラスターの進化の過程を推定するために、菌類21種、後生動物5種、植物2種およびマラリア原虫のゲノム配列を用いて、インフォマティクスによりオーソログのsnoRNAについて遺伝子構成を予測した。その結果、菌類にはsnoRNA遺伝子クラスターが存在し、そのうち、イネいもち病菌を含む9種ではイントロン内在型、出芽酵母を含む10種ではポリシストロニック型であることが明らかになった。また、菌類以外の生物7種では、オーソログsnoRNAの遺伝子はゲノム上にクラスターを形成していなかったことから、10億年前の菌類の祖先生物では既にクラスターが形成されていたと推測した。次に、イントロンの獲得と消失のパターンを推測し、snoRNA遺伝子の構成とイントロン進化との関連を調べた。その結果、snoRNA遺伝子の構成がイントロン内在型であった9生物種では、進化の過程で獲得したイントロンが多く、snoRNA遺伝子の構成がポリシストロニック型であった11生物種では、進化の過程で消失したイントロンが多いことが推測された。これらの結果は、snoRNAが進化の過程でイントロンを利用して発現する機構を獲得したことを示唆していると考えている。

5. snoRNA遺伝子データベースを構築：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snoOPYを構築した。予備的なバージョンをこれまで公開していたが、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、これらの情報を公開している。

6. 当初計画にはなかったその他の事項：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>

イントロンの起源・進化・機能に関してはいまだ不明な点が多く残っている。本研究では、RP遺伝子とsnoRNAに着目した独自の解析によりこの問題にチャレンジし新たな知見を得た。

1. イントロンの起源：イントロンの出現時期に関して、二つの立場に分かれて活発な議論が展開されている。Gilbertらは、イントロンの前生説 (introns early) を唱え、イントロンの起源はとても古いものであるが、真正細菌や古細菌のゲノムからはすでに失われており、一方、ヒトなど真核生物のゲノムにはまだ残っていると主張する。これに対して、Cavalier-Smithらは、イントロンの後生説 (introns late) を唱え、イントロンは真核生物において出現し徐々に蓄積されてきたと主張する。これまで、様々なデータセットを用いて解析が行われてきたが、いまだこの議論に決着はついていない。本研究では、ミトコンドリア (真正細菌由来) と細胞質 (古細菌由来) タイプのRP遺伝子を比較することにより、イントロンはミトコンドリアが細胞内共生をした後に出現したとの結論を得た。この結果はイントロンの後生説を支持しており、これまで混沌としていた論争に大きなインパクトを与え

た。

2. **イントロンの進化**：これまで種々の遺伝子を用いて解析が行われてきたが、RP遺伝子を用いた本研究は、これまでになく広い生物種をカバーする精度の高い結果を提示した。前述の通り、RP遺伝子はすべての生物が有し、種類も豊富で保存性が高いため、イントロン位置の比較が容易であった。そのため、進化の過程におけるイントロンの獲得と消失のパターンを正確に推定することが可能となった。過去の他の解析では、遺伝子数は多いものの、生物種は限られており、詳細なイントロン進化の解析は困難であった。RP遺伝子は、この点を打破する強力なデータセットとして認知されるにいたった。

3. **イントロンの機能**：イントロンの機能のひとつとして機能性RNAのキャリアーとしての役割を提唱し、snoRNAに着目した解析を行った。特に、非翻訳型RNA遺伝子のイントロンにコードされたsnoRNAの解析は、イントロンの進化と機能を調べるための有効なモデルであることが明らかになった。さらに、本解析は近年注目されている機能性RNAの起源と進化を理解するうえでも重要な意義がある。

4. **データベースの整備**：本研究では2つのデータベース（RPGとsnOPY）を作製し公開している。RPGはRP遺伝子に関する情報を整備したもので、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を提供している。すでに、国際的な標準データベースとして広く認知されており、連日1,000件を超えるアクセスがある。一方、snoRNAに関しては、snOPYの他に、ヒトと出芽酵母に限定したもの（それぞれLBME、UMass-Amberst）が公開されている。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に関して詳細な情報を提示しているが、オーソログに関する情報は不十分である。一方snOPYは、snoRNA遺伝子のオーソログを広範に扱う点で特徴的なデータベースである。既存のデータベースにはない規模（17,000種）でsnoRNA遺伝子の情報を提供している。こちらも多くの研究者に利用されており、2009年8月以降は連日数千件のアクセスがある。

#### <達成できなかったこと、予想外の困難、その理由>

新規にイントロンを獲得あるいは消失する際にはたらく機構を明らかにするために、菌類のゲノム配列を用いて新規に獲得されたイントロンを検出する予定であったが、必要な菌類の情報を十分に集めることができなかった。また、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、RPG データベースにクロロプラスト RP 遺伝子の情報を収めることができなかった。これらの問題は、集めたデータの確認作業に予想以上の時間がかかったために生じたもので、今後も継続的に作業を進めることで解決できると考えている。

また、イネいもち病菌でsnoRNAの発現機構を解析するための適切な実験系を作成することができなかった。これはひとつには、イネいもち病菌の遺伝子組換え株の作成が予想より難しく、また慣れていないこともあり、試行錯誤が多く手間がかかっているためと考えられる。他は、技術的な問題ではなく、snOPY データベースに一時的に収められたsnoRNAの情報について、キュレーターによるオーソログの決定、ターゲットとなるRNA配列の確認などに予想以上の時間が取られていることがあげられる。特に、菌類、植物に関して正確な情報が必要であるため作業に手間がかかっている。

#### <今後の課題、展望>

ゲノム配列が明らかになった菌類の情報を収集し、新規に獲得されたイントロンの予測に用いる。同時に、フザリウムで見出された新規イントロン14カ所のうち、トランスポゾンと相同性のある4カ所について、遺伝子の発現とスプライシングを確認する。そのために、インフォマティクスに用いた株のゲノムDNAとトータルRNAを入手する。また、イネいもち病菌で見出されたイントロン内在型snoRNAについて、発現様式を解析するための実験系を確立する。これにより、菌類の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の関連を明らかにしたい。加えて、snOPYデータベースに納められたsnoRNA遺伝子の情報について、キュレーターによるオルソログの決定を完了させる。その後、これらのデータを用いて、進化の過程におけるsnoRNA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）の変遷を推測する。同時に、クロロプラストRPのデータを中心にRPGデータベースを拡充し、イントロンの獲得と消失のパターンとsnoRNAの遺伝子構成の進化の過程を比較する。これにより、イントロン内在型snoRNA遺伝子が出現した時期を推定する。

#### <研究期間の全成果公表リスト>

##### 1) 論文

- 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(1), e4152 (2009).
- 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(20), 3204-3211 (2008).
- 0805071917  
Marygold, S., Roote, J., Reuter, G., Lambertsson, A., Ashburner, M., Millburn, M., Harrison, P., Yu, Z., Kenmochi, N., Kaufman, T.C., Leever, S.J., Cook, K.R.: The ribosomal protein genes and *Minute* loci of *Drosophila melanogaster*. *Genome Biol*, 8, R216 (2007).
- 0705081907  
Nguyen, H.D., Yoshihama, M., Kenmochi, N.: The evolution of spliceosomal introns in alveolates. *Mol Biol Evol*, 24(5), 1093-1096 (2007).
- 0702021545  
Yoshihama, M., Nguyen, H.D., Kenmochi, N.: Intron dynamics in ribosomal protein genes. *PLoS ONE*, 2(1), e141 (2007).
- 0701041543  
Nguyen, H.D., Yoshihama, M., Kenmochi, N.: Phase distribution of spliceosomal introns: implications for intron origin. *BMC Evol Biol*, 6, 69 (2006).
- 0701041554  
Uechi, T., Nakajima, Y., Nakao, A., Torihara, H., Chakraborty, A., Inoue, K., Kenmochi, N.: Ribosomal protein gene knockdown causes developmental defects in zebrafish. *PLoS ONE*, 1(1), e37 (2006).
- 0601241052  
Yoshihama, M., Nguyen, H.D., Kenmochi, N.: Analysis of ribosomal protein gene structures: implications for intron

evolution. *PLoS Genet*, 2(3), e25 (2006).

9. 0601241031

Nguyen, H.D., Yoshihama, M., Kenmochi, N.: New Maximum Likelihood Estimators for Eukaryotic Intron Evolution. *PLoS Comput Biol*, 1(7), e79 (2005).

2) 学会発表

1. 吉浜 麻生, 岩切 淳一, 有江 力, 剣持 直哉: イントロンと共進化する snoRNA の遺伝子構成, 第 11 回日本進化学会大会, 札幌, 2009.9.4
2. Naoko Hirano, Sayomi Higa, Tamayo Uechi, Naoya Kenmochi : Functional analysis of rRNA modification enzyme dyskerin in zebrafish: A vertebrate model for Dyskeratosis Congenita, Ribosome Synthesis, Regensburg, Germany, 2009.8.27
3. 吉浜 麻生, 岩切 淳一, 有江 力, 剣持 直哉: イントロン内在型 snoRNA はいつ誕生したのか? 第 10 回日本 RNA 学会年会, 新潟, 2009.7.28
4. Naoya Kenmochi, Hung Dinh Nguyen, Maki Yoshihama : Evidence for “introns-late” from the Evolution of Intron Phase Distribution, The 8th International Workshop on Advanced Genomics, 東京, 2009.6.16
5. 岩切 淳一, 吉浜 麻生, 比嘉 三代美, 立石 大樹, 中尾 彰宏, 剣持 直哉: snoRNA 遺伝子データベース snOPY を用いた遺伝子複製様式の解析, 日本分子生物学会第 9 回春季シンポジウム, 宮崎, 2009.5.11
6. 吉浜 麻生, 岩切 淳一, 有江 力, 剣持 直哉: イントロン内在型 snoRNA の進化, 日本分子生物学会第 9 回春季シンポジウム, 宮崎, 2009.5.11
7. 岩切 淳一, 吉浜 麻生, 比嘉 三代美, 立石 大樹, 中尾 彰宏, 剣持 直哉: snoRNA データベース (snOPY) の構築: イントロンにコードされる機能性 RNA 遺伝子の進化, BMB2008, 神戸, 2008.12.9
8. 吉浜 麻生, 岩切 淳一, 有江 力, 剣持 直哉: イントロン内在型 snoRNA の進化から見る真核生物におけるイントロンの役割, BMB2008, 神戸, 2008.12.9
9. 吉浜 麻生, 岩切 淳一, 剣持 直哉: snoRNA 遺伝子とイントロンの進化, 第 7 回新しい RNA/RNP を見つける会, 宮崎, 2008.9.9
10. 吉浜 麻生, 岩切 淳一, 有江 力, 比嘉 三代美, 剣持 直哉: スプライセオソーム型イントロンと機能性 RNA の進化, 第 10 回日本進化学会大会, 東大駒場キャンパス, 2008.8.25
11. 吉浜 麻生: イントロンの機能と進化: snoRNA 遺伝子クラスターの解析, 第 10 回 RNA ミーティング, 札幌, 2008.7.23

3) 図書

1. 剣持直哉, クバプロ, 真核生物の small RNA: 新しい RNA の分子生物学 (印刷中)
2. 0812191606  
剣持 直哉, 生体の科学: 特集 現代医学・生物学の仮説・学説 2009: イントロン内在型機能性 RNA 遺伝子の進化 (2009)

4) データベース/ソフトウェア

1. 0507070232 (データベース)  
RPG, リボソームタンパク質遺伝子データベース  
<http://ribosome.med.miyazaki-u.ac.jp/>
2. 0507070233 (データベース)  
snOPY, snoRNA データベース  
<http://snoopy.med.miyazaki-u.ac.jp/>