

## 核リボソーム RNA の二次構造と分子系統学への利用

傳法 隆, Dian Hendrayanti, 市村輝直

リボソーム RNA (rRNA) は、生命活動に不可欠なタンパク質の合成工場であるリボソームの主要な構成要素であり、その遺伝子は核ゲノム以外にミトコンドリアや葉緑体のようなオルガレラ中にも存在する。真核生物の核 rRNA は、一般に、RNA 合成酵素である RNA ポリメラーゼ I によって 18S rRNA (Small subunit: SSU) - 5.8S rRNA - 25 ~ 28S rRNA (Large subunit: LSU) が内部介在配列 Internal transcribed spacer (ITS) を介して繋がった一つの転写単位として合成された後、酵素によって ITS が切り出され、それぞれ成熟 RNA 分子となる。これらの転写単位はプロモーターやターミネーターなどを含むスペーサー領域 Intergenic spacer (IGS) によって区切られつつ、DNA 上をタンデムに複数個連続して存在している。一方、5S rRNA の存在様式は生物によって異なり、下等な生物では上記の転写単位の後方にリンクしている場合が多く、反対に高等な生物ではこれらとは別な場所に単独で反復して存在するが多い。また、後者の 5S rRNA は、別種の酵素である RNA ポリメラーゼ III によって合成されている。

これら 4 種類の成熟 rRNA 分子は多数のリボソームタンパク質と複合して、巨大なリボソーム分子を構成している。真核生物のリボソームは、全体で 80S の沈降係数をもつが、さらに 60S と 40S からなる二つの粒子に分かれる。60S 粒子には、上記の rRNA の内 LSU, 5.8S rRNA 及び 5S rRNA が、また 40S 粒子には SSU が各 1 分子ずつ含まれている。チェックらによるリボザイムの発見 (Kruger et al. 1982) 以前は、酵素活性はタンパク質に固有の性質であるという考えが支配的で、rRNA は単にリボソームの構造を保持するためだけに必要なものと考えられていた。しかしながら、今日では、リボソーム上でのタンパク質合成反応の主要な部分に rRNA が深く関与していることが分かっている。例えば、原核生物の大腸菌を用いた研究で、ペプチド結合の伸長反応に関与する酵素、ペプチジルトランスフェラーゼの活性中心が LSU (大腸菌の場合、23S rRNA) のドメイン V に局在することが示されている。

このような酵素活性を示す rRNA はリボソーム内でどのような形で存在しているのだろうか。

RNA 分子の中で、最も早くから構造の解明が進んでいたのは全塩基長が約 70 塩基と短いトランスファー RNA (tRNA) であった。この分子は、部分的に二重らせん (ヘリックス) 構造を形成したクローバー葉型の二次構造をとり、さらに高次に折り畳まれ、L 字形の三次構造がとることが既に解明されている。しかしながら、これよりも塩基長の長い rRNA 分子では、まだこのように完全な三次元モデルは構築できていない。なぜならば、その前段階である二次構造モデルが未だ完全ではないからである。原核生物の SSU と LSU の構造では、今までのところ、全体のわずか 3% と 2% のヌクレオチドが三次構造の相互作用に関与しており、また 60% と 58% のヌクレオチドが二次構造の相互作用に関与していることが分かっている。

RNA の二次構造は、一本鎖 RNA が鎖内で水素結合による塩基対を形成し、部分的に A 形に近い二重らせん構造 (DNA のワトソン・クリックモデルである B 形と異なり、右巻きのらせんは 1 回転あたり 11 ヌクレオチド残基含む。A 形は、B 形よりも主溝が狭くかつ深く、副溝がかなり浅いという特徴がある) をとることにより形成される。RNA の二次構造モデルの予測では、ワトソン・クリック型である A-U 対と G-C 対の他に、G-U 対も考慮されている。実際、G-U 対は塩基間に 2 つの水素結合を形成して安定に存在することができ、tRNA がリボソーム上でメッセンジャー RNA (mRNA) と結合するときにもこの塩基対が形成されていることが知られている。

RNA の二次構造を表す基本的な学術用語には、以下のものがある。まず、二重らせん部分の構造をステムと呼ぶ。ステム内で二重らせんの両側に存在するヌクレオチドが塩基対を形成できず、一本鎖のままである場合その構造をインターナルループと言い、片側だけ塩基対を形成できないヌクレオチドを持つ構造をバルジループと呼ぶ。ヘアピンループはステムの先端にあり、RNA 鎖の方向が逆転する構造である。また、ジャンクションは、3 つ以上のステムが閉鎖系内で接続されるときに形成される。

RNA の二次構造を予測する手段として、ここでは相互に関連のある二つの方法を取り上げる。一

つは生物学では伝統的に取り入れられている比較分析法で、もう一つは mfold を用いる方法である。

比較分析法では(Pace et al. 1999) まずできるだけ多くの生物から、高次構造をとりかつ相同な部分を持った RNA の塩基配列を集め、アライメント後、補償的な塩基置換 compensating base pair changes (CBCs) が起こっている箇所(アライメントの列に相当)を探索する(Mai and Coleman 1997)。例えば、アライメント中の 5' 側のある列で形質状態 A が、また少し離れた 3' 側の列で形質状態 U がそれぞれ優占していた場合、もし 5' 側の列である生物(アライメントの行に相当)の形質状態が A から U に変化したとき、3' 側の列でも U と塩基対が形成できるような塩基置換、すなわち形質状態 A または G に変化していたならば、CBCs が起こったと考える。このような塩基置換が起こっている箇所では、構造的にも何らかの関連があることが予想され、もし複数の列に連続して CBCs が起こっているならば、これらの配列でステム構造を形成している可能性が高い。

mfold(熱力学的な実験データから得られたパラメーターに基づき、個々のステム構造やループ構造の安定性の評価値を計算し、その合計したエネルギー関数を最小化するアルゴリズムを組み入れたソフトウェア)は、単一の塩基配列から最も低い自由エネルギーを持った構造と、ある一定量の自由エネルギーを増加させたときに現れるいくつかの準最適構造を予測する。また、準最適構造の構築に関するすべての塩基対をグラフィックで表示できるエネルギードットプロット energy dot plot (図 1 参照)も提供する。なお、原理や操作方法に関する詳細は、Mathews et al. (1998)や Zuker et al. (1999)を参照して頂きたい。ここでは、以下に簡単な操作方法を述べる。まず、mfold を利用するためには、以下のウェブサイトインターネットでアクセスしなければならない(<http://bioinfo.math.rpi.edu/mfold/ma/form1.cgi>)。現在、このソフトウェアはバージョン 3.1 で、無料でダウンロードできる。ただし、そのプログラムやスクリプトは Unix 環境または Intel ハードウェアでの Windows 98/NT に適応している。

始めに、二次構造を推定したい塩基配列名を登録後、tab キーで次の横長のウィンドウに移行し、前もってコピーしておいた GenBank, EMBL, FASTA 及び IntelliGenetics のいずれかのフォーマットで書かれた塩基配列をここにペーストする。次の縦長のウィンドウは、二次構造の構築に制限情報 constraint informations を与えるときに使用する。全部で

以下に列挙するように 5 種類ある。ただし、この制限情報の使用は、試験的にある条件下でどのような構造変化が起こるかを確かめたいときや、比較分析法などによって分子系統学的に支持される制限を与えたいときに限るべきである。制限情報が不適切な場合は二次構造は得られないので気を付けなければならない。

制限 1: 相手を指定せずに連続した配列に塩基対を形成させたいとき、 $F_{i0k}$  と表記する。例えば、F2305 では 5' 側から数えて 23 番目の塩基から 27 番目の塩基までに 5 塩基対を形成させる。相手はどの塩基でもかまわない。

制限 2: 相手を指定して連続した配列に塩基対を形成させたいとき、 $F_{ijk}$  と表記し、その基部側の塩基対を  $i$  と  $j$  で指定する。例えば、F21103 では 2 番目と 110 番目の塩基間、3 番目と 109 番目の塩基間、及び 4 番目と 108 番目の塩基間で 3 塩基対を形成させる。

制限 3: 相手を指定せずに連続した配列に塩基対を形成させたくないとき、 $P_{i0k}$  と表記する。例えば、P2305 では 23 番目の塩基から 27 番目の塩基まで一本鎖になる。

制限 4: 相手を指定して連続した配列に塩基対を形成させたくないとき、 $P_{ijk}$  と表記する。例えば、P21103 では 2 番目と 110 番目の塩基間、3 番目と 109 番目の塩基間、及び 4 番目と 108 番目の塩基間での塩基対形成を禁止する。ただし、どちらも他の塩基との間で塩基対を形成することは制限されない。

制限 5: 区域ごとに塩基対形成を禁止させたいときの表記法は  $P_{i-jk-l}$  である ( $i, j, k, l$ )。例えば、P2-110 2-110 と表記された場合、2 番目の塩基から 110 番目の塩基まで内部での塩基対形成が禁止される。ただし、この区域外との塩基対形成は制限されない。

次の項目で、RNA の塩基配列が直鎖状であるのか、または環状であるのかを選択する。このバージョンでは、その下のフォールディングの温度は 37 に固定されていて、他の条件は選択できない。希望するフォールディング数の上限は upper bound の欄内(デフォルト値は 50)に直接入力できるが、出力数は以下の percent suboptimality(デフォルト値は 5%)と window のパラメーターによっても影響される。Percent suboptimality の欄には、実際に計算される最小自由エネルギーの百分率で与えられる自由エネルギーの増加分を設定できる。デフォルト設定では、最小自由エネルギーからプラス 5% の範囲内で、エネルギードットプロットの作図や

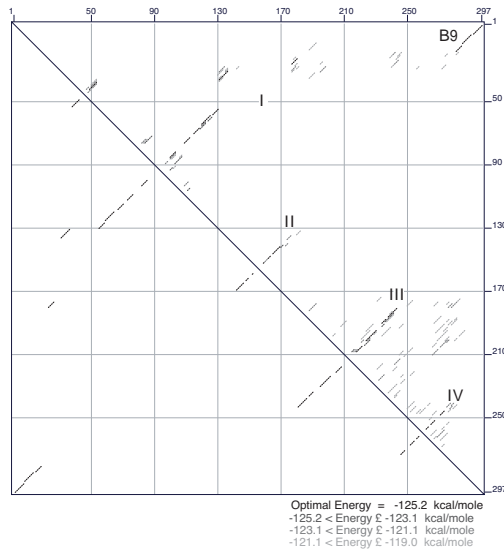


図1 *Closterium ehrenbergii*の単為生殖株 99-15-2 (与那国産)において、5.8S rRNAとLSUから構成されるB9ヘリックスとITS-2を含む領域で二次構造を推定したとき得られたエネルギードットプロット(制限条件なし)。上側のトライアングル内には、最適エネルギー(-125.2 kcal/mole)からプラス5% (6.2 kcal/mole)の範囲内で現れる準最適構造の構築に関わる全ての塩基対のエネルギードットプロットを示している。また、黒いドットは、上下とも最適エネルギーでのプロットを示している。図から、二次構造モデルでヘリックスIIIとIVに相当する区域では、準最適条件下で形成可能な塩基対が多数存在することが分かる。

準最適構造の構築が行われる(図1参照)。また、Windowのパラメーターは、直接入力しなければ塩基配列の長さに依存した数値が自動的に選択される(Zuker et al. 1999)。このパラメーターの値は小さいときは互いにほとんど変わらない構造が多数与えられ、反対に大きいときはかなり異なった構造が少数与えられる傾向にある。したがって、短い配列のとき大きい値を与えると、推定されるフォールディング数は減少する。Maximum distance between paired basesの項目のパラメーターは、デフォルトでは無限大に設定されているので全く制約はない。しかし、例えば1000塩基からなる直鎖状の配列のときに50と入力した場合、塩基間の距離が50塩基を越えるものどうしでは塩基対の形成ができず、塩基対の形成が短い範囲内に限定された構造が与えられる。

最後に、いくつかの出力データのフォーマットを順次設定後、長さが500塩基以下の場合にはAn immediate jobを選択し、your\_\_nameの部分に名前を入力する。また、長さが500塩基を越える場合にはA batch jobを選択し、正確なメールアドレスを入力して、Fold RNAボタンをクリックする。デー

タは、An immediate jobではすぐに取得できる。また、A batch jobでも多少時間はかかるものの、フォールディングが完了するとe-mailで知らせてくる。ただし、データは二日以内に取得しなければ消失するので注意が必要である。

このようにして得られたRNAの二次構造は、分子系統学の分野ではアライメントのときに強力な武器となる。なぜならば、RNA遺伝子はタンパク質をコードしている遺伝子とは異なり、三つの塩基が一組で一つのアミノ酸を指定しているという規則性がないため、一次配列を眺めただけではアライメントの基準となる部位の推定が困難だからである。最も効率的なアライメント方法は、clustal Xのようなアライメントソフトウェアを使って機械的にアライメントを行った後、mfoldで推定した個々のRNAの二次構造を参照しながら、共通の基準となる構造を探して、マニュアルで調整するやり方であろう。このようにして得られたアライメントは、フィードバックして、比較分析法による特定の生物群に共通に保存されている二次構造の推定に役立つものと思われる。

RNAの二次構造を分子系統学の研究に利用した事例を一つ紹介したい。我々の研究室では、接合

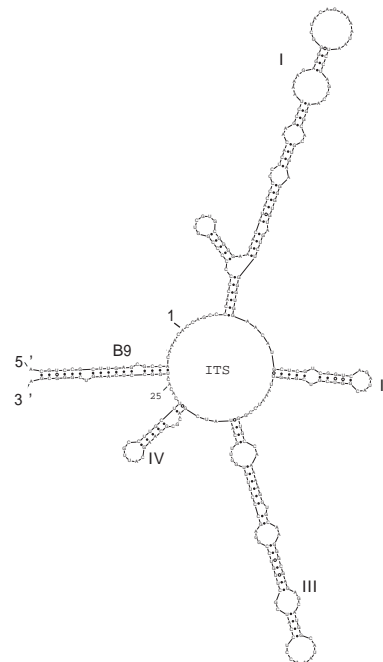


図2 *Closterium ehrenbergii*の単為生殖株 99-15-2 (与那国産)における、B9ヘリックスとITS-2の二次構造モデル(最適エネルギー = -111.4 kcal/mole)。系統学的に支持される制限条件を付与している。この図のヘリックスIVは制限条件なしでは現れない。

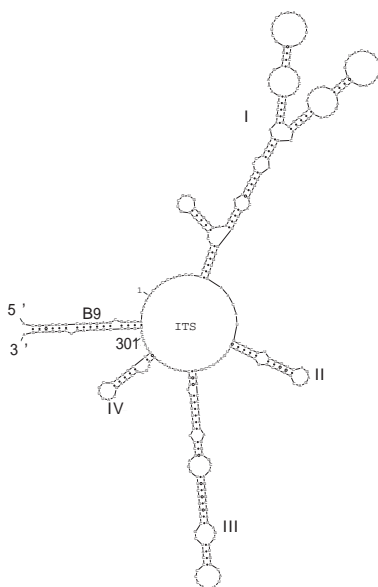


図3 *Closterium ehrenbergii* の単為生殖株 98-34-13 (ベトナム産) における, B9ヘリックスとITS-2の二次構造モデル(最適エネルギー = -118.8 kcal/mole)。ヘリックスIの先端部分で重複が起こっていることが分かる。系統的に支持される制限条件の他に, 重複構造が分かりやすいように制限条件を付与している。

藻類に属するミカヅキモの仲間を材料として, 種分化の研究を行っている。これらの種の多くは単一クローンの培養で接合胞子を形成するホモタリズムの交配様式を示すが, 一部の種では相補的な交配型であるプラスとマイナスの株を混合することによって接合胞子を形成するヘテロタリズムの交配様式を示す。また, 接合することなく, 単為生殖によってのみ厚い細胞壁を持った休眠胞子を形成する集団もある。この属の *Closterium moniliferum-ehrenbergii* 種複合体では, 核 rDNA の SSU から構築した系統樹において, *C. moniliferum* Erenberg ex Ralfs v. *moniliferum* のホモタリックな系統が, 他の交配様式をもつ系統よりも早く分岐していることが報告されている (Denboh et al. 2001)。また, ITS-2 では, *C. ehrenbergii* Meneghini ex Ralfs の与那国島産の単為生殖株は石垣島産の交配群 K (生物学的種に相当し, *C. ehrenbergii* では現時点で A から S まで) と 100% の相同性 (長さは 258 nt) を示すが, ベトナム産の単為生殖株 (長さは 301 nt) とは 83.4% の相同性しか示さない。この種複合体全体のアライメント (62 分類群) では, このベトナム産の単為生殖株だけが, ある部分で 42 ヌクレオチドからなる長い配列が纏まって挿入されて

いることが確認されている。さらに, この ITS-2 から構築された系統樹では, ベトナム産の単為生殖株と沖縄産の株が姉妹群になることも分かっている。

緑色植物の ITS-2 の二次構造モデルは, 一般に, 4本のヘリックス構造をとり, ヘリックスIIにはピリミジン-ピリミジンミスマッチ構造が広く保存されている (Mai and Coleman 1997)。これらの構造は, 世界中から集められた *C. moniliferum-ehrenbergii* 種複合体のヘテロタリックな系統でも完全に保存されている (図2, 3参照)。図2には与那国産の単為生殖株 99-15-2, 図3にはベトナム産の単為生殖株 98-34-13 における B9ヘリックスとITS-2 からなる二次構造モデルを示している。これらの図の比較から, ベトナム産の単為生殖株では, ITS-2 のヘリックスI で 42 nt からなる先端の二次構造が重複していることが明らかである。この結果から, 次の二つのことが類推できる。

- (1) ベトナム産の単為生殖株 98-34-13 の祖先は, 沖縄にルーツがあり, そこから移動してきた。
- (2) *C. ehrenbergii* の単為生殖する 2 つの系統 99-15-2 と 98-34-13 は, ヘテロタリックに交配する系統である交配群 K から進化した。

このような手法を用いることによって, 古生代から出現し種分化してきた形態種や現在もおお種分化を続けていると考えられる交配群について, 進化の道筋と交配様式の推移を検証することが可能となる。

#### 文献

- Denboh, T., Hendrayanti, D. and Ichimura, T. 2001. *J. Phycol.* 37:1063-1072.
- Kruger, K., Grabowski, P. J., Zaug A. J., Sands, J., Gottschling, D. E. and Ceck, T. R. 1982. *Cell* 31:147-157.
- Mai, J. C. and Coleman, A. W. 1997. *J. Mol. Evol.* 44:258-271.
- Mathews, D. H., Andre, T. C., Kim, J., Turner, D. H. and Zuker, M. 1998. p. 246-257. In: Leontis, N. B. and SantaLucia, J. J. (eds.) *Molecular modeling of nucleic acids*. American Chemical Society, Washington, DC.
- Pace, N. R., Thomas, B. C. and Woese, C. R. 1999. p. 113-141. In: Gesteland, R. F., Ceck, T. R. and Atkins, J. F. (eds.) *The RNA world*, Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Zuker, M., Mathews, D. H. and Turner, D. H. 1999. p. 11-43. In: Barciszewski, J. and Clark, B. F. C. (eds.) *RNA biochemistry and biotechnology*. Kluwer Academic Publishers, Dordrecht/Boston/London.

(北海道大学北方生物圏フィールド科学センター)