

フジテレビジョン賞

tRNA遺伝子からひも解く遺伝情報制御の
新たなルール

—新規バイオインフォマティクス技術「SPLITS」の開発と応用—

慶應義塾大学大学院 政策・メディア研究科 政策・メディア専攻
後期博士課程 2年

菅原 潤一

1 緒 言

本研究は新規バイオインフォマティクス技術「SPLITS」を開発し、Transfer RNA (tRNA) 遺伝子の解析を進めたことで、遺伝情報制御に関する新たなルールを見出したものである。

tRNA は細胞において「タンパク質」を合成する役割を持つ分子である。より詳しくは、タンパク質の構成要素となる各種アミノ酸を運搬し、遺伝暗号(コドン)の指示に従いながらタンパク質の鎖を正しく伸長してゆくのが tRNA である(図 1a)。地球上の全ての生物のゲノムには、全遺伝暗号(61 種類)をアミノ酸に翻訳するための各種 tRNA の遺伝子セットが必ずコードされていなければならない(図 1b)、その役割上、tRNA は最も重要な遺伝子ファミリーの一つとして認識されている。

近年 DNA シーケンサーの技術進歩に伴い、急速な勢いで多種生物の完全長ゲノムが決定されている。決定されたゲノム配列から tRNA 遺伝子を予測し、その特徴を解析することは、その生物の進化的位置づけや、その生物の基礎となるタンパク質合成系の作りを議論する上で欠かすことができない。さらに、完全長ゲノム中に必須 tRNA 遺伝子の全セットが揃っているかどうかは、決定されたゲノム配列の確からしさを検証する上で一つの基準となるものであり、その確認は必ずおこなわれる。こうした背景から、tRNA 遺伝子を DNA 配列中から予測するための情報学的な技術は、これまで数多く開発されてきた。実際、tRNA 遺伝子の予測ソフトウェアは、分子生物学の分野において最も使用されるソフトウェアの一つに挙げられるだろう。

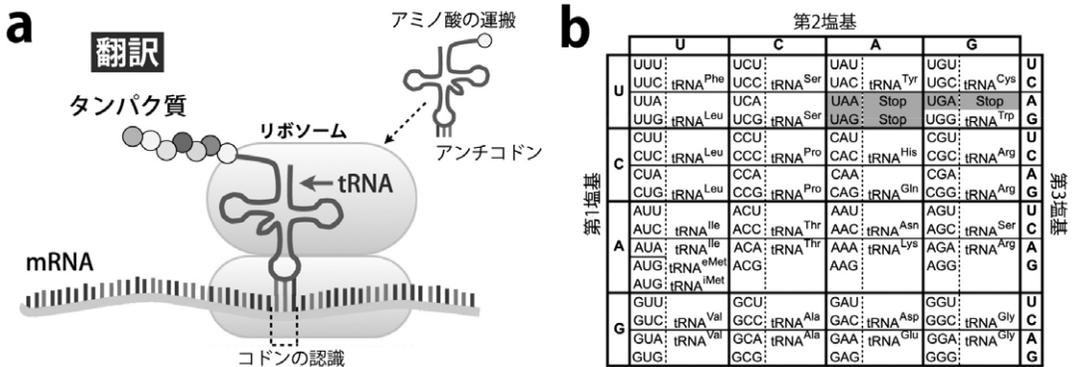


図1. tRNAの役割

(a) tRNA は決まった一つのアミノ酸をチャージし、タンパク質合成工場であるリボソームまで運搬する。tRNA のアンチコドンとよばれる3塩基の並びが、mRNA のコドンと結合することで、コドンとアミノ酸の対応をとる。(b) 各コドンに対応する tRNA の最小遺伝子セット。3塩基の並びで64通りのコドンが表現される(内3つはストップコドンとよばれ、アミノ酸には翻訳されない)。表に示す通り、塩基のG-U結合を考慮しても最小33のtRNA遺伝子が必須となる。

一方、近年急速な勢いで多種生物のゲノムが解読される中、特にアーキアとよばれる原始的な生物ドメインや、一部の真核生物種において、必須 tRNA 遺伝子の幾つかが発見されないケースが多発していた。こうした「見つからない tRNA 遺伝子」の存在は、新しい情報学的手法確立の必要性を訴えかけていたと共に、従来の手法では検出することのできない「未知

のルール」に制御された遺伝子が隠されている可能性を示唆するものであった。

そこで私は、こうした「見つからない tRNA 遺伝子」を発見するための新規ソフトウェア、SPLITS の開発に取り組んだ。本論文では、従来手法とは全く異なる視点から考案された SPLITS の tRNA 探索アルゴリズムと、それにより発見された「奇想天外な遺伝子」を紹介したい。

2 新規 tRNA 遺伝子予測ソフトウェア「SPLITS」の開発

本章では、既存の tRNA 遺伝子予測ソフトウェアが抱える課題を明確にした上で、今回開発に取り組んだ SPLITS の新規性を説明し、その予測精度を検証する。

2.1 既存ソフトウェアが抱える課題

クローバーリーフを探せ

DNA 配列から tRNA 遺伝子を探索(予測)するための情報学的手法は過去に多数開発されてきた。それらは主に、入力された DNA の文字列を左(5'側)から右(3'側)にスキャンして、tRNA 最大の特徴であるクローバーリーフ型の二次構造(図2)を形成可能な領域を見つけ出すというものである(図3)。最も代表的なソフトウェアとして「tRNAscan-SE」が挙げられ、その精度は極めて高く、試験に用いられた tRNA 遺伝子の 99% 以上を検出し、擬陽性候補は 15,000,000,000 長の塩基配列中からたったの一つ程度と報告されている [1]。

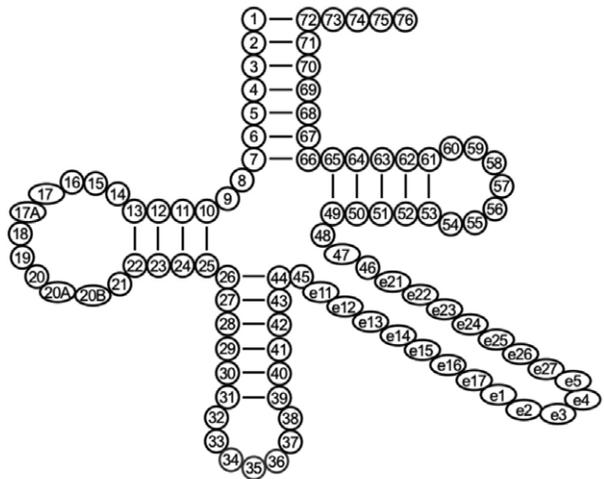


図2. tRNAのクローバーリーフ構造モデル

tRNA の各塩基を円と楕円で示し、5'側から番号を割り当てた。円は普遍的に観察される塩基を、楕円は一部の tRNA で例外的に観察される塩基を表す。mRNA のコドンの特異的に認識する部位(アンチコドン)を赤線で示す。塩基間の結合を実線で示す。ほぼ全ての tRNA はこのモデルに従う。

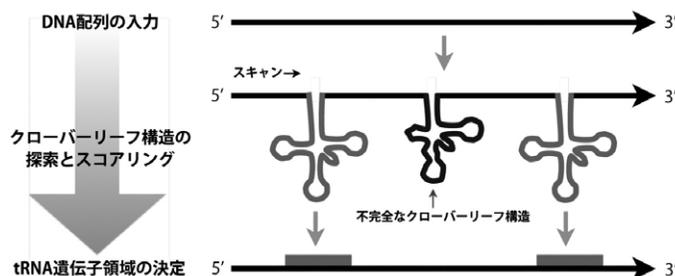


図3. tRNA 遺伝子予測の戦略

tRNA 遺伝子領域を任意の DNA 配列中から予測する際は、入力配列を 5'側から 3'側にスキャンして、tRNA の特徴であるクローバーリーフ構造を形成可能な領域を探索する。モデルに基づき tRNA らしさが定量化され、基準値を満たす領域が tRNA 遺伝子領域として決定される。

クローバーリーフが見つからない

一方、様々なゲノムが解読されるにしたがいで、クローバーリーフ構造が形成されない tRNA 遺伝子の存在が明らかになってきた。一つは「イントロンを介在する tRNA 遺伝子」であり(図4)[2]、イントロンがクローバーリーフ構造の形成を邪魔するため、従来手法で遺伝子を検出できなくなる(図5)。もう一つは、これは非常に特殊な例ではあるが、2005年に発見された「Split tRNA」が挙げられる(図4)[3]。この遺伝子は寄生性のアーキア *Nanoarchaeum equitans* (*N. equitans*)のゲノムから発見されたものであり、tRNA の左半分と右半分がゲノム上で分かれてコードされ、そこから別々に発現した RNA が細胞内でハイブリダイゼーションした後、エンドヌクレアーゼによるスプライシングを受けて合体することで、一つの機能性分子が合成されるというものだ。クローバーリーフが右半分と左半分で別々のゲノム領域にコードされる Split tRNA 遺伝子も、当然従来手法で検出することはできない(図5)。

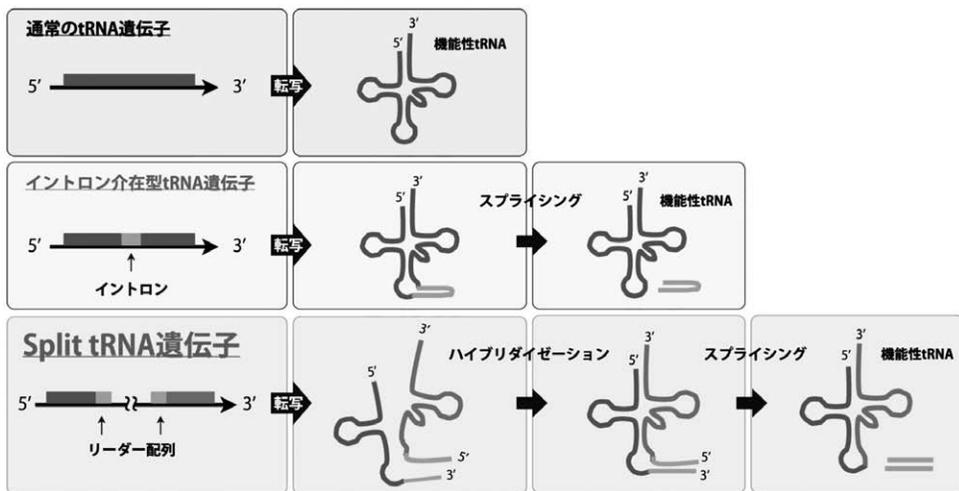


図4. クローバーリーフ構造を形成しないtRNA.

比較のため通常の tRNA 遺伝子も示す(上段)。中段はイントロン介在型 tRNA の発現を示しており、イントロンを橙色で表す。アーキアと真核生物では、tRNA のイントロンはスプライシングエンドヌクレアーゼにより除去される。下段に Split tRNA の発現を示す。Split tRNA の RNA 断片はリーダー配列(橙色)とよばれる 10 ~ 20 塩基の相補的な配列を有しており、それらが塩基対結合することで断片同士がハイブリダイゼーションする。この中間体が、イントロン介在型 tRNA と同様にエンドヌクレアーゼによりスプライシングされることで、断片同士が合体し機能性 tRNA となる。

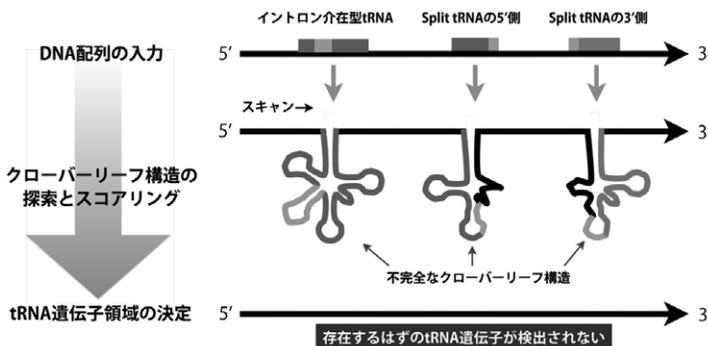


図5. 従来のtRNA 遺伝子予測の課題

従来の tRNA 遺伝子予測方法でイントロン介在型 tRNA 遺伝子や Split tRNA 遺伝子を予測できない理由を図示する。イントロンの介在や遺伝子の分断に伴い、ゲノム上でモデルの基準値を満たすクローバーリーフ構造が形成されないことが主たる要因となる。

2.2 SPLITS の新規性—仮想スプライシング法の開発

上述の通り既存手法で検出できない tRNA の共通点は、スプライシングという過程を挟むがために、最終的にはクローバーリーフ構造が形成されるものの、ゲノム上でその構造が見出せないことであった。ならば、細胞の中で起こるスプライシングをコンピューター上で再現することができれば、上述のような tRNA にもクローバーリーフ構造を形成させることが可能になるのではないかと考えた。

そこで私は、tRNA がスプライシングを受ける際に形成する普遍的な構造モチーフ、Bulge-Helix-Bulge (BHB) モチーフ (図 6) [2] に着目した。BHB モチーフは、Split tRNA のリーダー配列、及びイントロン介在型 tRNA のイントロンで形成される立体構造であり、エンドヌクレアーゼが同モチーフを認識することでスプライシングが進行する。この構造モチーフを情報学的に予測することができれば、(1) Split tRNA を形成する RNA 断片の組合せを決定できることに加え、(2) 予めリーダー配列領域やイントロン領域を決定し、クローバーリーフ構造の形成を阻害するそれらの領域を対象 DNA 配列中から取り除く (スプライシングする) ことができる。そうすれば、従来のソフトウェアを用いて、Split tRNA やイントロン介在型 tRNA 等のクローバーリーフ構造を、正當に評価させることが可能になると考えた。

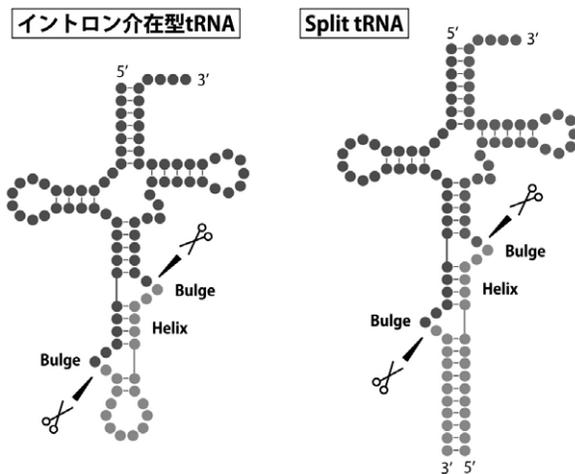


図 6. tRNA が形成する BHB 構造モチーフ

BHB モチーフの概要図。このモチーフはイントロン介在型 tRNA と Split tRNA で共通している。イントロン及びリーダー配列を橙色の円で示す。図に示す通り、BHB モチーフは 3 塩基のバルジに挟まれた 4 塩基対のヘリックスから成る。スプライシングエンドヌクレアーゼによって切断される箇所を三角で示した。イントロン及びリーダー配列は、中心のヘリックスからそれぞれ 3' 側に 2 塩基目の箇所 (いずれもバルジ内部) で切断される。ここでは、イントロンが tRNA の 37 番目と 38 番目の塩基 (図 2 参照) の間に介在される例を示したが、その他の様々な箇所にも介在することが知られている。

こうした設計思想に基づき、新規 tRNA 遺伝子予測ソフトウェア「SPLITS」を開発した (図 7) [4]。SPLITS には、アイデアのコアとなる「BHB モチーフの予測」及び BHB モチーフの位置情報に基づく *in silico* での「仮想スプライシング」という 2 つの処理と、イントロン介在型 tRNA 遺伝子を予測するための「シス・モード」、Split tRNA 遺伝子を予測するための「トランス・モード」の二つのモードが実装されている (図 8)。

BHB モチーフの予測は、(1) 位置特異的スコア行列を用いた一次構造の評価、(2) 二次構造予測による BHB モチーフ形成能の評価、(3) 及び予測された BHB 構造の自由エネルギー計算に基づく安定性評価の三点からおこなう。各パラメーターは、過去に報告された既知の BHB モチーフを全て検出できるものにチューニングした。

一方で、既知の BHB モチーフを全て検出しようとする、当然擬陽性候補が多くなるという問題に直面する。そこで私は、「もしそのモチーフが本物ならば、スプライシング後、最も奇麗なクローバーリーフ構造を形成するようになるはずだ」という仮定に基づき、「仮

想スプライシング」の処理において、図9に示す工夫を施した。後述するが、これにより SPLITS は極めて高い精度で、tRNA の BHB モチーフを正確に検出することが可能になっている。

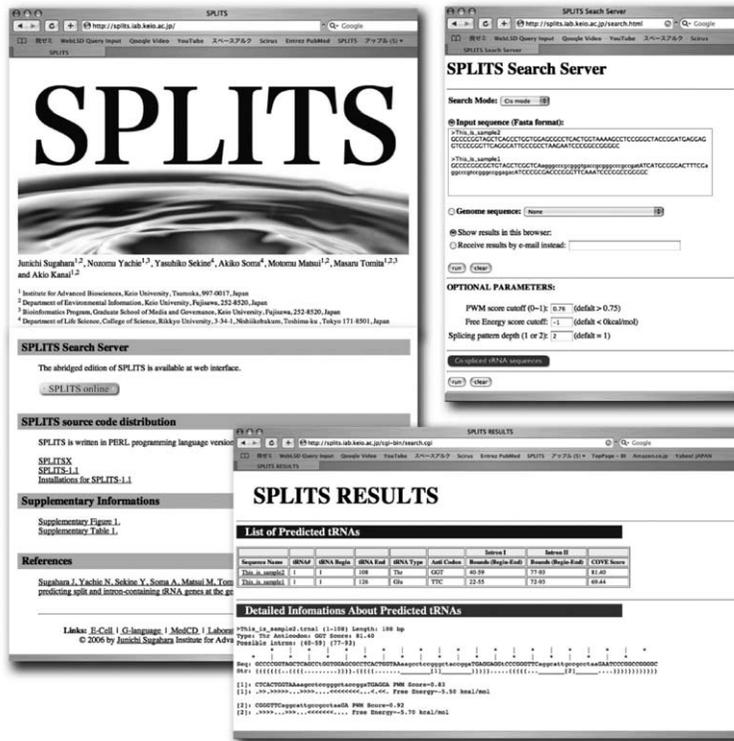


図7. SPLITSのウェブインターフェース

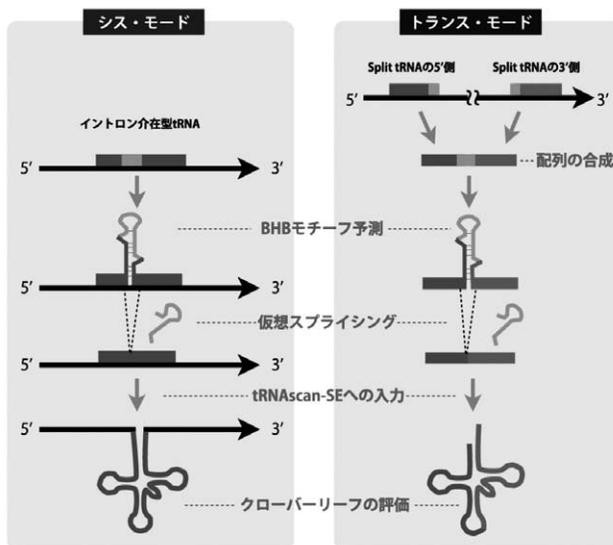


図8. SPLITS の仮想スプライシング

SPLITS におけるシス・モードとトランス・モードそれぞれの処理の概略を示す。図5と対比させながら見て頂きたい。BHB モチーフの予測により、Split tRNA の組み合わせの決定や、イントロンもしくはリーダー配列領域の除去をおこなない、tRNAscan-SE 等によるクローリーフ構造の評価が可能となる。シス・モードでは、最初に tRNA 様領域を Blast 等のアライメントで網羅的に検出し(当然その段階では多くの擬陽性 tRNA 候補を含む)、同領域に対して BHB モチーフ予測の処理をおこなう。トランス・モードではまず、パターンマッチングにより既知 tRNA の 5' 側及び 3' 側末端配列と相同な領域を検出し、それぞれ適当な長さに伸張したものを *in silico* で合体させ、BHB モチーフ予測の処理に回す。それらのペアが BHB モチーフをうまく形成できれば、Split tRNA の断片である可能性が高くなる。

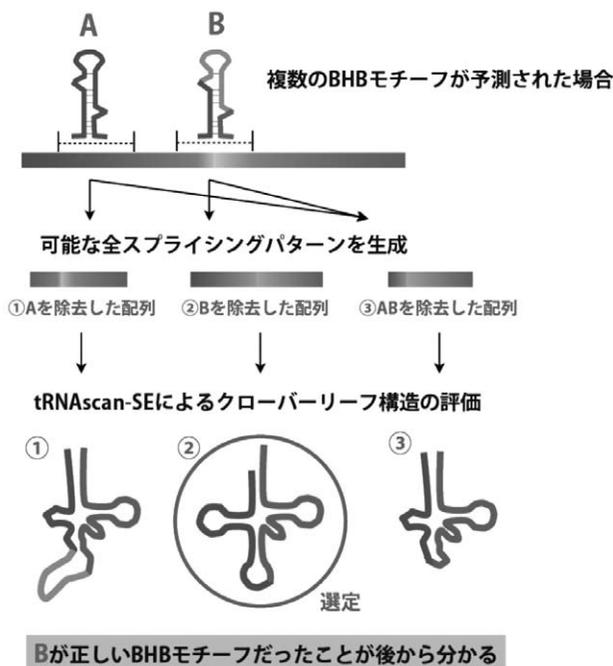


図9. BHB モチーフを予測する際の工夫

BHB モチーフを予測するには、多くの擬陽性候補が検出される。それらの中から本物を選ぶための工夫を、左図で説明する。

例えば一つの配列の2カ所からBHBモチーフが予測された場合(仮にA, Bとする), Aのみをスプライシングした配列, Bのみをスプライシングした配列, AとB両方をスプライシングした配列というように、全ての可能なスプライシングパターンを生成する。In silicoでスプライシングした全ての配列をtRNAscan-SEに入力し、クローバーリーフ構造等のtRNAらしさを評価する。

その結果、最もスコアが高くなるスプライシングパターンを一つ選定し、最終的な候補とする。即ち、左図で言えばBというモチーフを除去したパターンを最終候補とするため、間接的ではあるが、Aというモチーフは擬陽性候補であったという判断が下る。

2.3 SPLITSの精度検証

SPLITSの精度を検証する目的で2つの実験をおこなった。

15種のアーキアゲノムにおいて、イントロンの介在によりクローバーリーフ構造の形成が阻害されるtRNA候補31例[2]をSPLITS(シス・モード)とtRNAscan-SEで検出させ、その精度を比較した。この結果、同コントロールのtRNAscan-SEによる検出精度が38.7%(12/31)だったのに対して、SPLITSは100%(31/31)を正確に検出した。またその際、報告されていたイントロンの介在位置をSPLITSが全て正確に予測したことから、上述したBHBモチーフの予測と仮想スプライシングのアルゴリズムが正常に動作していることが確認された。

次に、*N. equitans*で報告された計6例のSplit tRNA遺伝子[3]の検出をSPLITS(トランス・モード)で試みた。この結果も同様に、報告されたリーダー配列部位(すなわちtRNAの5'側半分と3'側半分が結合するジャンクション)の正確な位置情報とともに、全てのSplit tRNA遺伝子を検出することが確認された。

3 SPLITSの応用—新規に発見された奇想天外なtRNA遺伝子—

アーキアや一部の真核生物ゲノムでは、本来コードされているはずの幾つかの必須tRNA遺伝子が見つからないケースが報告されていた。本章では、そうした「見つからないtRNA遺伝子」の探索を目的としたSPLITSによる網羅的な解析結果を報告する。

3.1 アーキアにおける tRNA 遺伝子の予測

隠された第三のイントロン

完全長ゲノムが解読された 47 種のアーキアゲノムを観察すると、61 種のコドンに対応する必須 tRNA の内、各生物種において平均約 10% が未同定であった。私はまず SPLITS のシス・モードをベースとした解析により、アーキアから未知 tRNA 遺伝子を網羅的に検出することを試みた。この結果、上記 47 種から計 2,143 の tRNA が予測され、各種において平均 99% 以上の必須 tRNA 遺伝子を検出することに成功した。その一例として *Thermoflum pendens* (*T. pendens*) での tRNA 遺伝子予測結果を図 10 に示す。上述の tRNA 遺伝子候補はいずれも、(1) tRNA が機能する上で必要な分子的特徴を有すること、(2) 未同定であった必須 tRNA 遺伝子の穴をピンポイントで埋める形で検出されたことから、信頼できる遺伝子候補として評価されている [5]。

この解析の過程で、単一分子内に最大 3 つものイントロンを介在する新規 tRNA 分子種が検出された(一例を図 11a, b に示す)[6, 7]。中でも特に、*T. pendens* で見つかった tRNA^{Pro}(UGG) は、特筆すべき特徴を有していた。なんと本 tRNA の両サイドの 2 つのイントロンは、中心のイントロンの BHB モチーフとオーバーラップしていたのである。即ち、中心のイントロンは両サイドのイントロンがスプライシングされないと、出現しないのである(図 11a, c)[6]。こんなに小さい分子の中に 3 つものイントロンが存在する、という面白さに加え、スプライシングが段階的に進行し、ゲノム上では観察されないイントロンが RNA 上で出現するという現象は大変興味深いものであった。なお、本 tRNA の発現と段階的なスプライシングは、慶應義塾大学の喜久田薫氏による RT-PCR とシーケンサーを用いた解析により実証されている。

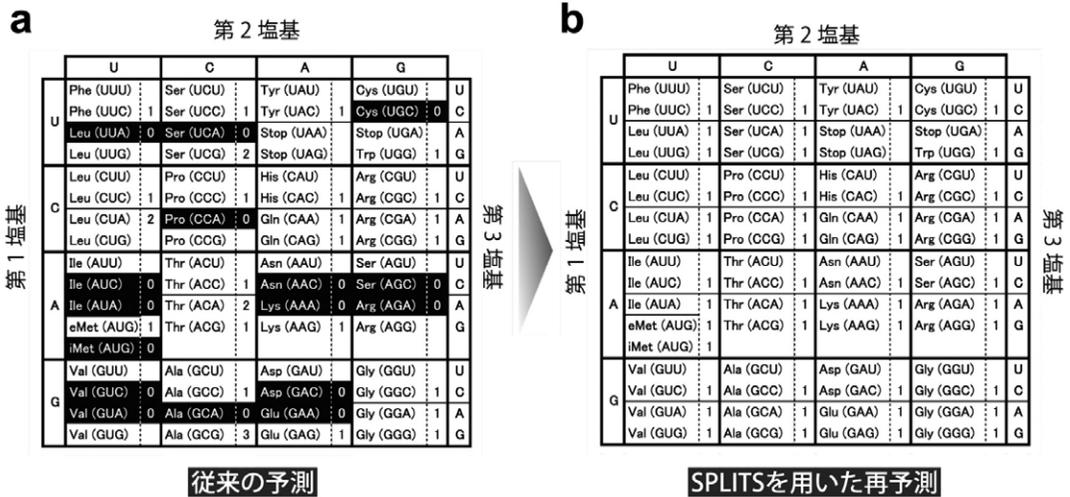


図10. *T. pendens*における各コドンとそれに対応するtRNA遺伝子数

アーキアの一種 *T. pendens* におけるコドン表と、各コドンに対応する tRNA 遺伝子の数をまとめた。(a) tRNAscanSE を用いた従来の予測結果では、幾つかのコドンに対応する必須 tRNA 遺伝子が見つからない(黒塗りつぶしで示す)。(b) SPLITS を用いた再予測の結果、見つからない必須 tRNA 遺伝子を全て検出することに成功した。

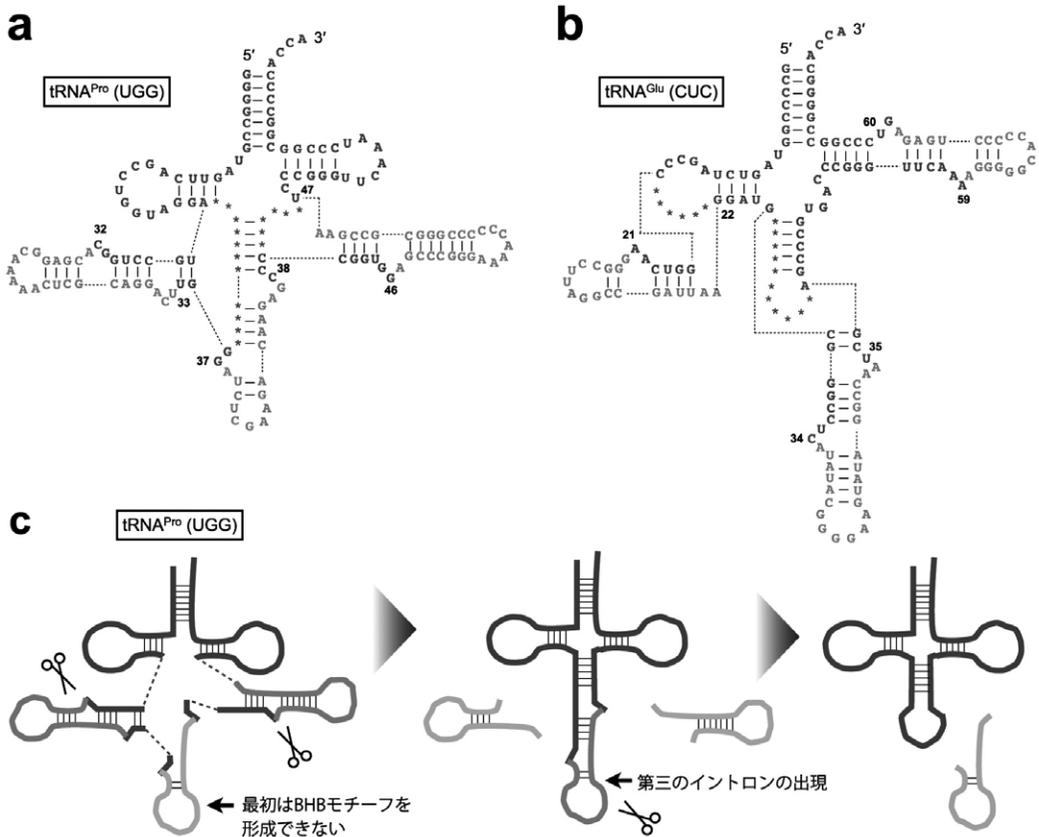


図11. 3つのイントロンを介在する新規tRNA分子種

(a) *T. pendens* で発見された3つのイントロンを介在する新規 tRNA の二次構造図。(b) *Pyrobaculum calidifontis* で発見された3つのイントロンを介在する新規 tRNA の二次構造図。二次構造はいずれも SPLITS が予測したものである。(c) *T. pendens* の tRNA^{Pro}(UGG) から明らかとなった3つのイントロンを介在する tRNA の段階的スプライシング。

Split tRNA の新種発見

上述の解析でもまだ、幾つかのコドンに対応する必須 tRNA が見つからない種が残された。そうした種を対象に今度は SPLITS のトランス・モードを適応し、Split tRNA 遺伝子の探索を試みた。この結果、超好熱好酸性アーキアである *Caldivirga maquilingsensis* (*C. maquilingsensis*; 図 12) のゲノムにおいて、これまで tRNA が見つかっていなかった GGG, GCG, GCA, GAA の4つのコドンに対応する tRNA が、Split tRNA 候補として予測された(図 13)。これら新規 Split tRNA 候補は、慶應義塾大学の藤島皓介博士による RT-PCR とシーケンサーを用いた解析により、各 RNA 断片の発現と成熟体 tRNA の形成が確認され、その成果は米科学紀要(PNAS)に掲載された [8]。

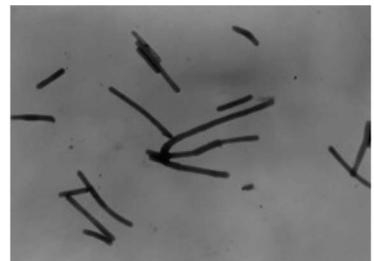


図12. *C. maquilingsensis* の細胞

C. maquilingsensis はフィリピン温泉から単離された高温(約85℃)、酸性(pH 3.7~4.2)の環境を好む好熱好酸性アーキアである。ゲノムサイズは約200万塩基であり、好熱性アーキアとしては標準的である。写真は理化学研究所バイオリソースセンターの伊藤隆氏による提供。

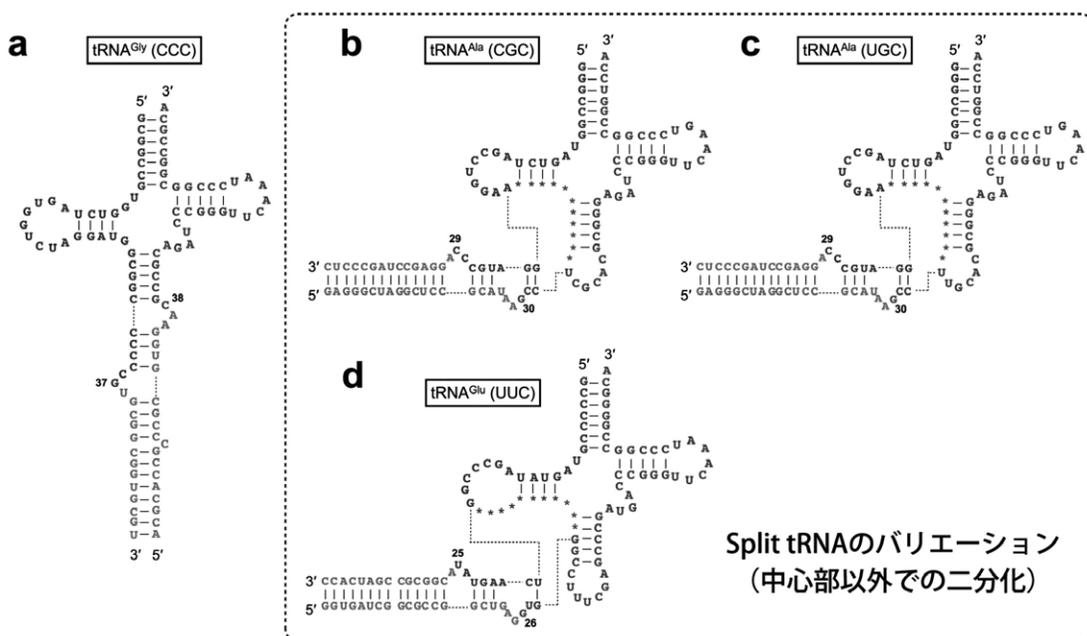


図13. *C. maquilingensis*で発見された4つの新規Split tRNA

SPLITSにより予測された新規 Split tRNA の二次構造図を示す (a-d)。(a)は *N. equitans* の Split tRNA と同様, tRNA の中心部(37, 38 番目の塩基の間)で二分化されるものである。一方, (b), (c) は 29, 30 番目の塩基, (d) は 25, 26 番目の塩基で二分化される。(b), (c), (d) のように, 中心部以外で二分化されるケースは今回初めて発見された。

本 Split tRNA 遺伝子発見の新規性は二つある。一つは, Split tRNA 遺伝子は 2005 年に寄生性のアーキア *N. equitans* で報告されて以来, これまで一つも他種では発見されてこなかった。*N. equitans* が極めて小さなゲノムを持つ細胞であったことから, Split tRNA 遺伝子がゲノム縮小に伴い発生した偶然的な産物であるのか, 一般的な種においても観察される現象であるのかは, 明らかになっていなかった。今回 *C. maquilingensis* で発見された Split tRNA 遺伝子は, 自由生活性の生物種における初の Split tRNA 遺伝子発見となり, 遺伝子の Split 化という現象の一般性の議論に新たな知見を加えた。第二に, *N. equitans* で発見された Split tRNA はいずれも tRNA が中心部分(37 番目と 38 番目の塩基の間)で二分化されるものであったが, 今回 *C. maquilingensis* では, 29 番目と 30 番目の塩基(図 13b, c), 25 番目と 26 番目の塩基(図 13d)など, 多様な箇所でも二分化される Split tRNA が初めて観察された。

3.2 遺伝情報の逆転現象—Permuted tRNA 遺伝子の発見—

SPLITS による奇妙な遺伝子構造の発見

次に, SPLITS の真核生物ゲノムへの応用例を紹介する。

完全長ゲノムが決定されている真核生物の中で, 必須 tRNA の幾つかが見つかっていなかった生物種, 原始紅藻 *Cyanidioschyzon merolae* (以下, シズン; 図 14)のゲノムを対象に, SPLITS による解析を試みた。

SPLITS を用いて「見つからない tRNA 遺伝子」を探索した結果, シズンのゲノム上で興味

深い構造を有する tRNA 遺伝子候補が検出された。本 tRNA 候補は、tRNA を形成する 5' 側半分と 3' 側半分が約 10 ~ 70 塩基離れた距離にコードされていたのだが、その際に、5' 側と 3' 側の配列が「逆さま」に配置されるという、不思議な構造を有したものであった(図 15a)。

上述の配置をとる遺伝子候補がシゾンゲノムから計 8 つ(tRNA^{Ala} [UGC], tRNA^{Gln} [CUG], tRNA^{Gly} [CCC], tRNA^{Leu} [UAA], tRNA^{Lys} [UUU], tRNA^{Met} [CAU], tRNA^{Thr} [AGU], tRNA^{Thr} [CGU]) 同定された。いずれも、これまで見つからなかった必須 tRNA と一致していたことから、本候補が発現し機能している可能性が期待された。また、詳細な観察の結果、3' 側半分がコードされる領域の上流にのみ、転写プロモーターと思われる TA リッチな配列が存在していたことから(図 15a)、これらの 3' 側配列と 5' 側配列は、一続きに転写されることが予想された(図 15b)。

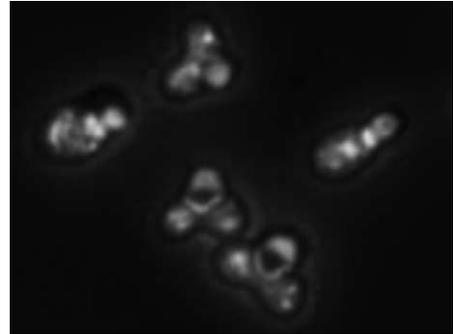


図14. シゾンの蛍光顕微鏡写真

原始紅藻シゾンはイタリアの温泉に生育する単細胞性の紅藻であり、42℃、pH2.5 を至適生育条件とすることから極限環境微生物とも見られる。細胞のサイズは直径約 1.5μm であり、ゲノムサイズは約 1,600 万塩基と、真核生物の中では最も小さい。写真は立教大学理学部 黒岩常祥教授による提供。

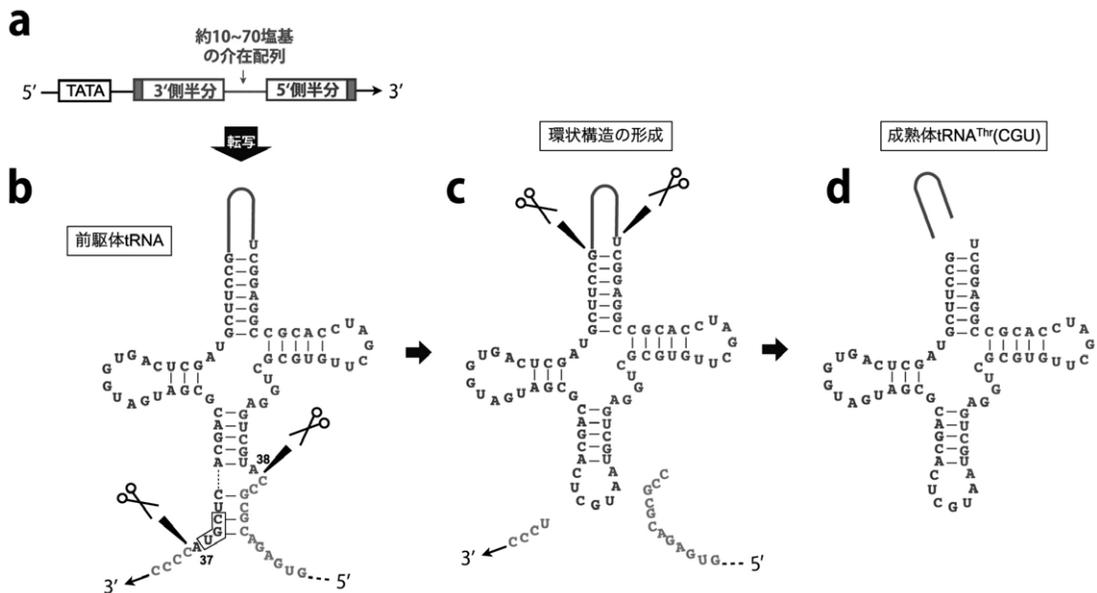


図15. Permuted tRNA遺伝子の構造と発現制御。

(a) Permuted tRNA 遺伝子のゲノム上における構造と、(b) 転写された RNA が形成すると予測される二次構造、(c, d) 及び実験的に示唆された転写後のプロセッシング過程を模式的に示す。スプライシングにより RNA が中間体として環状構造を形成し(c)、結合した末端とは逆側が切断されることで(d)、5' 側と 3' 側の遺伝情報が逆転することを特徴とする。

遺伝子発現と機能性の検証

上述の情報学的な解析結果と考察を、シゾンの tRNA 解析チームであった立教大学の関根靖彦准教授、相馬亜希子博士に報告し、その発現とプロセッシング過程、及び機能性を Northern hybridization や RT-PCR 等のウェットな解析により検証して頂いた。最終的に上記 8 遺伝子候補が全て発現していること、また予想通りそれらは一続きの RNA として転写されていることが確認された。さらに本 tRNA は、末端同士が結合してスプライシングを受けることで一度環状構造を形成し(図 15c)、次に 5' 側半分と 3' 側半分の繋いでいた介在配列部分が切断されることによって、成熟体 tRNA が形成されることが明らかとなった(図 15d)。驚くべきことにこの過程で、ゲノム上で逆さまにコードされていた情報が、RNA 上で逆転することになる。

少なくともシゾンにおいては、最終的に Split tRNA 遺伝子は発見されなかった。一方、Split tRNA に次ぐ新たな遺伝情報制御のルールが見出された。我々はシゾンで発見した上記の tRNA を「Permuted tRNA」と名付け、その成果を米科学誌 Science に発表した [9]。

Permuted tRNA で見られた「遺伝情報の逆転」という現象をより分かりやすく説明するためのモデルを図 16 に示す。

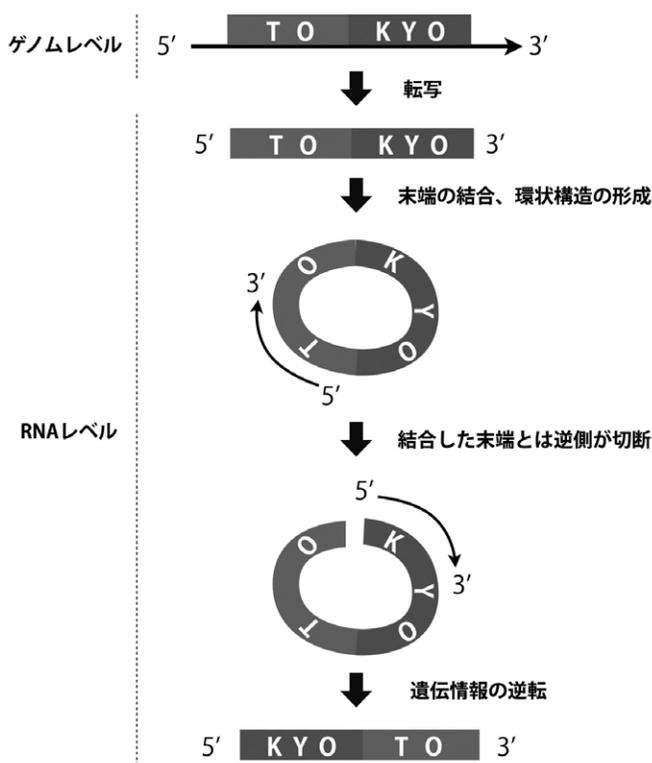


図16. 遺伝情報の逆転を説明するためのモデル

Permuted tRNA 遺伝子で観察された遺伝情報の逆転という現象を、よりわかりやすく説明するためのモデルを示す。例えば、ゲノム上で TOKYO という遺伝情報をコードする遺伝子があったとする。これが RNA に転写され後、その 5' 末端と 3' 末端が結合されることで、一度環状構造が形成される。今度は、最初に結合された末端とは対照に位置する領域が切断されることにより、再び線上の RNA に戻る。このとき、ゲノム上でコードされていた TOKYO という情報の、"TO" と "KYO" の位置が逆転し、RNA 上で KYOTO という情報が形成されることになる。

4 結 言

本研究では、(1)コンピューター内での仮想スプライシング法を考案し、tRNA 遺伝子予測ソフトウェア SPLITS を新規に開発した。(2)同ソフトウェアを用いて多種生物ゲノムを網羅的に解析した結果、3つのイントロンが段階的にスプライシングされる tRNA や、(3)自由生活性の生物種では初となる Split tRNA 遺伝子を発見することに成功した。(4)さらに、原始紅藻シズンのゲノム解析からは、tRNA の新規分子種「Permuted tRNA」を初めて予測することに成功し、情報の逆転という、遺伝子発現制御の新たなルールの発見に寄与した。

SPLITS は既に、国内外の複数の研究プロジェクトにおいて採用して頂いた実績があり、実際に新規 tRNA 遺伝子のアノテーション等に貢献している [10-13]。今後もこれまで以上の勢いで多種生物ゲノムが解読されていくことが予想される中、SPLITS はイントロン介在型 tRNA、Split tRNA、及び Permuted tRNA 遺伝子を高精度に検出できるユニークなソフトウェアとして、ゲノム解析プロジェクトに広く貢献できるものと信じている。

生物の DNA 配列には、私たちが思いもつかないような機能や秘密がまだまだ沢山隠されているのだろう。それらは生命科学のみならず、医学や新素材開発等の幅広い分野へ応用されるものと考えられる。今回 tRNA をモデルに見出された遺伝情報制御の新たな知見が、ゲノムから生命情報を読み解こうと試みる学問分野において、斬新な視点を与えるものになると期待する。

5 謝 辞

本研究は慶應義塾大学の富田勝教授、金井昭夫教授のご指導のもと行われました。また、Permuted tRNA 遺伝子の発見は、立教大学の関根靖彦准教授、相馬亜希子博士らとの共同研究によるものです。本成果は両氏の許可のもと本論文に掲載させて頂いています。この場をお借りして上記の方々に感謝申し上げます。

6 引用文献

- [1] Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964.
- [2] Marck C, Grosjean H (2003) Identification of BHB splicing motifs in intron-containing tRNAs from 18 archaea: evolutionary implications. *RNA* 9:1516–1531.
- [3] Randau L, Münch R, Hohn MJ, Jahn D, Söll D (2005b) Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5′- and 3′-halves. *Nature* 433:537–541.
- [4] Sugahara J, Yachie N, Sekine Y, Soma A, Matsui M, Tomita M, Kanai A (2006) SPLITS: a new program for predicting split and intron-containing tRNA genes at the genome level. *In Silico Biol* 6:411–418.
- [5] Sugahara J, Kikuta K, Fujishima K, Yachie N, Tomita M, Kanai A (2008) Comprehensive analysis of archaeal tRNA genes reveals rapid increase of tRNA introns in the order thermoproteales. *Mol Biol Evol* 25:2709–2716.
- [6] Sugahara J, Yachie N, Arakawa K, Tomita M (2007) In silico screening of archaeal tRNA-encoding genes having multiple introns with bulge-helix-bulge splicing motifs. *RNA* 13:671–681.
- [7] Sugahara J, Fujishima K, Morita K, Tomita M, Kanai A (2009) Disrupted tRNA gene diversity and possible evolutionary scenarios. *J Mol Evol* 69:497–504.
- [8] Fujishima K, Sugahara J, Kikuta K, Hirano R, Sato A, Tomita M, Kanai A (2009) Tri-split tRNA is a transfer RNA made from 3 transcripts that provides insight into the evolution of fragmented tRNAs in archaea. *Proc Natl Acad Sci U S A* 106:2683–2687.
- [9] Soma A, Onodera A, Sugahara J, Kanai A, Yachie N, Tomita M, Kawamura F, Sekine Y (2007) Permuted tRNA genes expressed via a circular RNA intermediate in Cyanidioschyzon merolae. *Science* 318:450–453.
- [10] Hallam SJ, Konstantinidis KT, Putnam N, Schleper C, Watanabe Y, Sugahara J, Preston C, de la Torre J, Richardson PM, Delong EF (2006) Genomic analysis of the uncultivated marine crenarchaeote Cenarchaeum symbiosum. *Proc Natl Acad Sci USA* 103:18296–18301.
- [11] Maruyama S, Sugahara J, Kanai A, Nozaki H (2009) Permuted tRNA genes in the nuclear and nucleomorph genomes of photosynthetic eukaryotes. *Mol Biol Evol* 27:1070–1076.
- [12] Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, Chee GJ, Hattori M, Kanai A, Atomi H, Takai K, Takami H (2010) Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res* doi:10.1093/nar/gkq1228.
- [13] Abe T, Ikemura T, Sugahara J, Kanai A, Ohara Y, Uehara H, Kinouchi M, Kanaya S, Yamada Y, Muto A, Inokuchi H (2011) tRNADB-CE 2011: tRNA gene database curated manually by experts. *Nucleic Acids Res* 39:D210–213.