

「代謝調節機構に基づく細胞機能制御基盤技術」  
平成 18 年度採択研究代表者

磯辺 俊明

首都大学東京理工学研究科・教授

## RNA 代謝解析のための質量分析プラットフォームの開発

### 1. 研究実施の概要

本研究では、最新の質量分析法を基礎にした RNA 解析のプラットフォームを開発し、細胞の機能発現や制御に重要な低分子 RNA や RNA/蛋白質 (RNP) 複合体の実態を明らかにすることで、プロテオミクスと低分子 RNA 研究が融合した細胞機能ネットワーク解析の基盤作りを目指している。平成 20 年度の研究では、昨年度に引き続き「RNA のショットガン解析」のための LC-MS 法の基礎技術すなわち (1) RNase 消化したオリゴリボヌクレオチド断片の分離能の向上とエレクトロスプレーイオン化の効率化による分析の高感度化、(2) オリゴリボヌクレオチドを衝突解離 (CID) 法で解裂した際の断片化ルールについて検討するとともに、ここで得られたタンデム質量分析データからゲノム情報を検索して特定の RNA を同定するために開発したプロトタイプのソフトウェア (Ariadne と命名) の改良を行った。また、これらの方法を評価、検証するための RNA/蛋白質複合体の精製法や、RNA の分離・前処理法、RNase によるゲル内消化法などについて検討した。以上の研究によって、酵母やヒト培養細胞などから親和性タグなどを利用して分離精製したスプライセオソームやリボソーム先駆体などの RNP 複合体に含まれる U-snRNA や snoRNA などの主要な低分子 RNA を直接あるいは電気泳動後に LC-MS 解析して同定し、その塩基配列やメチル化などの転写後修飾を解析することができるようになった。特に、比較的小さなゲノムサイズをもつ酵母などの生物を対象とした解析では、RNP 複合体を構成する主要な RNA を直接染色体上にマップして同定することが可能になった。この方法の開発によって、原理的には、生体から分離した RNP 複合体を構成する RNA とタンパク質成分の種類や化学構造を、LC-MS 法を基盤とする共通のプラットフォームで解析することができるようになった (1)。今後はこの方法をより大規模な RNA 解析に適用できる汎用的なショットガン解析法に発展させるための技術開発を進めると同時に、国内外での共同研究を推進することで RNA/RNP 研究の進歩に貢献する成果をあげて行きたい。

### 2. 研究実施内容 (文中にある参照番号は 4.(1)に対応する)

## 1) 質量分析を基礎にした RNA 同定技術ならびに RNA 解析のための MS 技術の開発

本技術開発の目標は、生体試料に含まれる低分子 RNA とその代謝物を直接 LC-MS 法で分析することで、それぞれの RNA をゲノム情報に帰属し、転写後修飾も含めた化学構造を明らかにするための RNA 解析のためのプラットフォームを開発することである。この目標に向けた昨年度までの研究では、化学合成した siRNA や酵母 tRNA 標品、あるいは無細胞合成系で合成した mRNA など RNase T1 で消化して得られた混合物を (1) リン酸基に富む親水性化合物であるオリゴヌクレオチドに対して保持能力が高いシリカ系 C-30 逆相系充填剤を充填したエレクトロスプレー (ESI) カラムを利用したナノ LC で分離し、(2) ナノ LC から 50-100 nL/min の超微流速で送液されてくる微量の RNA 溶出液にオンラインで連続的にアセトニトリルを添加して負電荷をもつ RNA 成分を「ネガティブモード」で安定してイオン化できるハードウェアを設計・試作した。また、この基本的なナノ LC-MS 装置と理化学研究所で開発した RNA データベース検索エンジンを組み合わせたプロトタイプの RNA の LC-MS システムを利用することで、酵母細胞から調製した tRNA や、東京農工大学から供給されたヒト培養細胞由来のリボソーム先駆体などに含まれる rRNA や snoRNA などの主要な RNA 成分を RNase T1 で断片化し、それらの LC-MS/MS 解析で得られたマススペクトルを RNA データベースに対して検索することで、個々の RNA が同定できることを示した。

本年度の研究では、まずこのプロトタイプシステムの評価と実用化のための改良を行った。実際にこのプロトタイプシステムを酵母細胞やヒト培養細胞から分離した RNP 複合体の解析に適用してみると、複合体に含まれる RNA の分析感度はタンパク質の感度に比べて約 1/10 程度であることがわかった (論文 1)。そこで、このシステムの高感度化に向けた検討を行った。すなわち RNA 試料の RNase 消化で生じたオリゴヌクレオチド試料を ESI カラムに導入するために使用する「トラップカラム」の改良、オリゴヌクレオチドを効率よく分離してイオン化するためのナノ LC の分離溶媒の再検討、溶出したオリゴヌクレオチド断片を効率よくイオン化するための補助スプレー装置の改良を行った (補助スプレー装置については特許出願準備中)。また、培養細胞や RNP 複合体に含まれる RNA を電気泳動で分離後にゲル内で RNase 消化して LC-MS/MS システムに導入するための前処理法についても最適化を行った。これらの改良の結果、昨年度に比べて RNA の分析感度が約 10 倍向上し、培養細胞などから親和性タグを利用して分離した微量のスプライセオソームやリボソーム先駆体に含まれるフェムトモルレベルのタンパク質と RNA 成分をほぼ同程度の感度で同定し、その塩基配列や、メチル化などの修飾を含めた化学構造をタンパク質成分と同時に解析することができるようになった (論文 5)。

一方、このシステムでは、RNP 複合体に含まれる RNA 成分をあらかじめアクリルアミド電気泳動などで分離することが必要であった。そこで本年度の研究では、技術開発の次のステップとして、このシステムをプロテオミクスでは一般的な RNA のショットガン解析に適用するための検討を開始した。RNA の LC-MS 法の開発が困難である理由の 1 つは、RNA を構成するヌクレオチドが 4 種類のみで、その物理化学的性質も類似しているため RNase 消化で生じるオリゴヌクレオチド分子の化学的多様性が乏しく、タンパク質の消化で生じ

るペプチドに比べて複雑で精密な LC 分離が困難なことである。そこで今年度の研究では、試料中の RNA をまずそのままの状態でも分子全体として LC 分離し、一度分画したのちに RNase で断片化し、生成したオリゴヌクレチドをさらに LC で分離しながらオンラインで MS/MS 分析して同定する多次元 LC-MS/MS 法開発のための基礎検討を開始した。現在の段階では、1 段目の RNA 分離用 LC と 2 段目のオリゴヌクレチド分離用 LC のインターフェースとして簡易型の溶液操作ロボットを介在させることで RNA の分取と RNase 消化、2 段目 LC への試料の注入を行うためのシステム的设计を進めている。この技術開発は、RNA のショットガン解析のための一歩であるだけでなく、試料の分離から LC-MS による分析までの行程を自動化することでマニュアル操作による試料の損失や環境からの汚染を排除し、細胞などから調製した微量 RNA 成分をさらに高感度で効率よく分析するための汎用的なシステムの構築のためにも重要と考えている。

一方、従来の本研究では、試料 RNA を断片化する RNA 分解酵素として全てのグアニン (G) 塩基の 3' 末端側で切断する RNase T1 を採用した。この方法は、化学合成した純粋な RNA や、比較的単純な RNA 混合物の同定には効率的で有効な方法であったが、より複雑な混合物の直接的な解析を目指すショットガン解析では、多くの RNA に重複して存在する短い RNA 断片を生じる傾向があり、1 つ 1 つの RNA を一義的に特定できるユニークな配列をもつ断片が生じにくい問題点があった。そこで、本年度は、その問題を解決する手段として、農工大と共同で RNA をより限定的に切断できる RNA 分解酵素の探索を行い、まず GU の 2 塩基配列を認識して G の 3' 末端を切断する基質特異性をもつ大腸菌の ColicinE5 の C 末端ドメイン (ColicinE5 CTD) を分離精製し、その基質特異性について検証した。

## 2) RNA 質量分析データの処理と解析技術の開発

本研究では、質量分析法で得られるデータを基礎にして特定の RNA をゲノム上の特定の塩基配列へ帰属し、同時に転写後修飾基などを検出して同定するデータ解析法の開発を目指している。そのための重要な要素技術の 1 つは、プロテオミクス研究では既に一般的な RNA 解析のためのデータベース検索エンジンの開発である。本年度の研究では、(i) RNA の質量スペクトルデータから単同位体ピークおよびその価数を抽出するプログラムの開発、(ii) 昨年度の研究で開発したプロトタイプデータベース検索プログラムの発展・統合、(iii) 低分子 RNA のカタログ化とデータベースの構築の 3 項目を行った。

(i) RNA の質量スペクトルデータから配列から計算できる質量情報を持つ単一同位体ピークを抽出することはデータ解析の鍵となる過程である。本年度は Thermo-Fischer 社および Waters 社の装置から得られるデータから同一の分子に由来する同位体ピーク群を認識し単同位体ピークおよびその価数を抽出するプログラムを開発した。このプログラムを後述する検索エンジンと組み合わせて用いることで、同定感度、精度共に向上することを確認した。

(ii) 昨年度の研究で開発したプロトタイプデータベース (DB) 検索プログラムでは、通常のシーケンスイオンとそれ以外の内部イオンが同じ質量を与える場合には、誤同定が生じやすいというスコアリングの問題点が残った。本年度はこのスコアリングを見

直しシークエンスイオンに重みを付けることでより誤同定を生じにくくすることに成功した。同時に、プログラムの内部処理を見直し、効率化することにより酵母ゲノムに対して LC-MS/MS データを直接検索することが可能となった。また、理研に設置するサーバーコンピュータに対して首都大学から DB 検索可能なようにプロトタイププログラムを web アプリケーションとして統合した。これらの検索エンジンの改良の結果、首都大学で測定した未知試料 RNA のヌクレアーゼ消化物の LC-MS/MS データをインターネット経由で酵母 (*S. cerevisiae* および *S. pombe*) ゲノム DB に対して検索し、RNA を特異的に同定することが可能となった。この検索プログラム (Ariadne と命名) は、RNase 消化で生じたオリゴヌクレオチドのタンデム質量分析情報 (MS/MS 情報) からゲノム上で RNA を同定できる世界で初めてのゲノム検索エンジンとなった (論文 3)。今後は哺乳類ゲノムなどより大きな配列データベースに対して検索するため、より効率的な検索アルゴリズムを検討し、さらに混合物試料への適用を目指してスコアリングアルゴリズムを開発する予定である。

(iii) 昨年度整備した small RNA DB の拡張および維持を行った。公開されている低分子 RNA DB の調査、収集を継続して行い、収録データの重複を除去することで DB を整理した。これに一般的な公開 DB である NCBI nr などからキーワード抽出・配列相同性検索して得た低分子 RNA 配列を追加して独自の RNA カタログからなる DB を整備した。

### 3) RNA とプロテオームの機能的相関解析

本研究では、物理的相互作用解析に基づく低分子 RNA とタンパク質の対応づけを行うと同時に、1) の研究項目で実施される研究の過程で開発された技術を適宜細胞試料から得られる RNA や RNP 複合体の機能相関解析に適用し、その有効性を評価、検証することを目標としている。本年度は、低分子 RNA-タンパク質複合体の RNA 及びタンパク質成分の同定解析を実施し (論文 1)、同時に、1) の研究項目で必要とされる RNA 試料の安定的供給を行った。また、特にクロマチン結合 RNA 等通常の条件で抽出されない RNA の効率的回収法の開発を行った。RNA 分子の限定的断片化法の検討では、首都大グループと共同で二塩基認識断片化酵素 Colicin E5 を用いた断片化法を検討した。さらに、1) の研究項目で開発した手法を、リボヌクレアーゼ分解や転写後修飾を伴う細胞内での RNA 代謝経路の解析に適用し、1) で開発した手法の有効性の実証を試みた。研究の概要は以下の通りである。

#### 2-1) 低分子 RNA の調製法の開発

RNA の質量分析法の開発にあたっては、生体に存在する RNA を天然の化学構造を保ったまま効率良く分離して濃縮することが重要となる。今まで報告のある低分子 RNA の回収法や一般的に用いられている市販キットによる方法は、可溶化しやすい RNA に適用されるが、クロマチンや細胞内の不溶性画分に存在する RNA の回収にこれらの方法は適さないことが判明した。本年度は、従来法で抽出されていないと考えられる RNA の抽出法の検討を行った。その結果、高塩濃度での抽出に希釈操作を加えることで大量に混入してくる 28S や 18S rRNA などの高分子量 RNA を選択的に除去できること、DNA 分解酵素の共存化では断片化された DNA が低分子 RNA と混在してしまうが、低塩濃度のある条件では

断片化した DNA を混在させないで低分子 RNA を効率的に回収できることを見いだした。今後は、回収した RNA 種を同定することで、この RNA 回収法の有効性を評価する必要がある。標準 RNA の安定的供給に関しては、*in vitro* RNA 合成システムにより高純度かつ十分量の標準 RNA の合成が可能になり、1) の開発の標準となる RNA を安定的に供給することが可能になった。また、LC 法の各種溶離条件の検討によって、RNA の分離効率を向上させることができた。今後は、これを回収効率と純度を評価する技術として利用し、低分子 RNA をより簡便かつ効率的に回収する試料調製法として改良していく必要がある。

## 2-2) RNA 結合タンパク質を用いた RNA の回収と同定

本年度は、1) の研究項目で開発するシステムの高感度化に向けた性能評価体制を整えるための存在量が少なくかつより複雑な低分子 RNA 混合物の回収法と検出法の検討を行った。そのために、RNA polymerase I, II を含む転写関連タンパク質 (論文 2)、Drosha, Dicer, DGCR8, AGOs などマイクロ RNA の生成に関わるタンパク質、FBL, SMN1 など snRNA/snoRNA の生合成及び代謝に関わる様々なタンパク質を Bait として、RNA-タンパク質複合体の単離を行った。そして、これらのタンパク質成分の同定をプロテオミクスの手法で行うとともに (論文 4)、RNA 成分を 1) の研究項目で開発した質量分析-Ariadne データ検索の方法で解析した。その結果、まず、昨年度まで同定できなかった U7, U4 などに加え、非予見的に snRNA/snoRNA である Rn38 を同定した。Northern blot 法で検出された RNA の幾つかは LC-MS/MS-Ariadne データ検索システムで同定されなかったので、ゲル染色を指標とし、ゲル内 RNase 消化法と LC-MS/MS-Ariadne データ検索システムを組み合わせ、RNA の存在量と同定効率の関係について検討したところ、ゲル上で観察された主要染色バンドに含まれる RNA 分子種は LC-MS/MS-Ariadne データ検索システムでほぼ同定され、存在量と同定効率はほぼ相関していることが明らかになった。これらの結果は、本システムが予見的な知識を必要としない汎用的な RNA の同定法として有効であるが、存在量の少ない RNA の同定には本システムのよりいっそうの高感度化が必要であることを示している。一方で、SMN1 複合体に含まれる snRNA の解析評価において、本システムが転写後修飾を同定する手法として極めて有効であることが示された。この複合体には U1snRNA が含まれるが、従来の研究によって U1 snRNA の生合成過程では、5' 末端に存在する G が m<sup>7</sup>G と m<sub>3</sub>G の 2 種類のメチル化修飾を受けることが知られている。本研究で分離した SMN1 複合体には、5' 末端に m<sup>7</sup>G と m<sub>3</sub>G を持つ U1 が、1 対 20 の量比で存在することが明らかになった。この複合体には、既知の Gemins (3, 4, 5, 6, 8)、Sms (D1, D2, D3, F) や、その他の結合タンパク質 (unrip, fibrillarin, PRMT5, MEP50, pICln, U1 small nuclear ribonucleoprotein 70 kDa) が含まれていたことから、単離した SMN1 複合体は、細胞質内から核質および Cajal body で起こる U1 snRNP 複合体の生合成過程で形成される一連の前駆体を含んでいることが示された。また、この解析の過程で、SMN1-U1 snRNA と複合体を形成している新規の機能未知蛋白質を同定した。これらの結果は、本システムが、転写後修飾を含めた RNA の生合成経路や代謝経路の解析に有効であることを示している。m<sup>7</sup>G を持つ U1 snRNA は 3' 末端側で exonuclease による塩基配列の削除が行われ、m<sub>3</sub>G を持つ分子種へと変換されること

が知られているが、今後の研究では、本システムがこのような exonuclease による転写後修飾も含めた解析へ適用可能かどうかを検証していく必要があると考えられる。同時に、今回同定した新規の蛋白質が U1 snRNA あるいは関連する RNA の今までに知られていない代謝に関わっているかどうかを明らかにすることも、本研究で開発している方法論の汎用性を示すために有効と考えられる。本年度は、さらに、Drosha, Dicer, DGCR8, AGOs などを Bait として前駆体を含むマイクロ RNA の回収も行ったが、これらの同定解析には、RNA 試料の回収率の向上と同時に、本システムの分析感度のよりいっそうの向上が必要であることが判明した。

本年度の RNA とプロテオームの機能的相関解析を拡大してきたことに伴って浮上してきた重要な課題の一つは、ヒトなどのゲノムサイズが大きい種由来の細胞から回収した RNA について、質量分析でスペクトルが十分な強度で検出できているにも関わらず、Ariadne 検索で同定に至らない例が観察されるようになったという点である。Ariadne による検索性能は、ヒトゲノムのタンパク質を指定する全てのコード領域を対象とした場合には、現在プロテオミクスの分野で最も信頼性の高い Mascot の検索性能と比べても勝るとも劣らない。タンパク質をコードする領域はヒトのゲノムのわずか 1.5%程度であるが、非コード領域は遥かにその比率は大きく、しかも、転写される RNA の転写開始領域と終結領域に関する規則、転写される長さ等、不明な点が多い。これらの点の解明が進めば、より同定効率が向上されると考えられるが、当面は RNA データベースの充実を図っていくことが重要であると考えられる。1)で開発した質量分析-Ariade 検索システムは、解析対象が今まで構築した RNA データベースに存在する場合、RNA の同定と転写後修飾の解析には高い有効性を持つ。今後は、新規 RNA の同定とその機能解析を通して新たな知見を得ることで、本システムの真価を示すことが重要と考えられる。同時に、タンパク質と RNA との間の物理的相互作用が確認出来たものおよび既に報告がある相互作用について、Osprey などの相互作用解析ソフトを適用し、適宜、RNA-タンパク質相互作用ネットワークマップのデータとして蓄積し、その利用原理を構築していくことが必要と考えられる。

### 2-3) 質量分析のための RNA 断片化法の開発

RNase T1 は全ての G の 3' 末端側を切断する酵素であり、G 含量の多い RNA 領域は質量分析に適したサイズの RNA 断片が得られないこと、逆に G 含量の少ない領域は質量分析には大きすぎるサイズの RNA 断片を生じてしまうことから、本年度は RNase T1 とは異なる基質特異性を持つ RNase である colicin E の利用を検討した。colicin E は G/U の 2 塩基を認識するとの報告があったことから、この RNase の大腸菌による発現とその精製を行い、その基質特異性を確認し、RNase T1 に加え、質量分析の前処理法として、この colicin E を用いる方法を開発した。この酵素を用いた断片化によって、RNase T で回収できなかった RNA 断片を得ることができ、RNase T1 を相補する手法として利用できることが明らかになった。しかし、colicin E が切断できる塩基は、RNase T1 で細かく切断される領域でしかも GU 配列を持つ条件が必要であるという限定的なものであった。そこで、RNase T1 で回収できない RNA 断片を調製する別の断片化法として、RNase H を

用いる方法についても検討を開始した。この酵素は、DNA-RNA の最低 4 塩基対を認識し切断するが、転写後修飾が存在すると切断は起らなくなるという基質特異性を持つことから、転写後修飾の検出に用いられている酵素である。この酵素の DNA-RNA 塩基対を切断するという特性を利用し基礎的な切断の検討を行ったところ、モデル RNA の塩基配列に相補的な合成 DNA 断片を用いることで目的の箇所を切断できることが明らかになった。今後は、複数箇所での同時切断が可能かどうか、混合 RNA 試料に対しても有効かどうか、あるいは、重複の多い RNA 塩基配列を分解除去する手法として利用可能かどうか等々、LC-MS 分析の試料調製法として実用的かどうかを検討していく必要がある。

### 3. 研究実施体制

#### (1)「首都大学東京 磯辺」グループ

①研究分担グループ長： 磯辺 俊明(首都大学東京大学院、教授)

②研究項目

質量分析法を基礎とした RNA 同定技術の開発

#### (2)「東京農工大学 高橋」グループ

①研究分担グループ長： 高橋 信弘(東京農工大学大学院、教授)

②研究項目

RNA とプロテオームの機能的相関解析

#### (1)「理化学研究所 中山」グループ

①研究分担グループ長： 中山 洋((独)理化学研究所、専任研究員)

②研究項目

質量分析法を基礎とした RNA 同定技術の開発

### 4. 研究成果の発表等

#### (1) 論文発表 (原著論文)

1. Hayano, T., Yamauchi, Y., Asano, K., Tsujimura, T, Hashimoto, S, Isobe, T, and Takahashi, N. (2008) Automated SPR-LC-MS/MS System for Protein Interaction Analysis. *J. Proteome Res.* 7, 4183-4190 (2008).
2. Izumikawa, K., Yanagida, M., Hayano, T., Tachikawa, H., Komatsu, W., Shimamoto, A., Futami, K., Furuichi, Y., Shinkawa, T., Yamauchi, Y., Isobe, T., and Takahashi, N. (2008) Association of human DNA helicase RecQ5b with RNA polymerase II and its possible role in transcription. *Biochem. J.* 413:505-516 (2008).
3. Nakayama, H., Akiyama, M., Taoka, M., Yamauchi, Y., Nobe, Y., Ishikawa, H., Takahashi, N. and Isobe, T. (2009) Ariadne: a database search engine for

identification and chemical analysis of RNA using tandem mass spectrometry data. *Nucleic Acids Res.*; doi: 10.1093/nar/gkp099.

4. Matusnaga, K., Saitoh, T., Tabata, K., Omori, H., Satoh, T., Kurotori, N., Maejima, I., Shirahama-Noda, K., Ichimura, T., Isobe, T., Akira, S., Noda, T., Yoshimori, T. (2009) Two beclin 1-binding proteins, Atg14L and Rubicon, reciprocally regulate autophagy at different stages. *Nature Cell Biol.*, 11, 385-396. Epub 2009 Mar 8.
5. Taoka, M., Yamauchi, Y., Nobe, Y., Masaki, J., Nakayama, H., Ishikawa, H., Takahashi, N. and Isobe, T. "Analysis of oligoribonucleotides by direct nanoflow reversed phase liquid chromatography coupled with high-resolution LTQ-Orbitrap mass spectrometry". Submitted to *Journal of Proteome Research*.

(2) 特許出願

平成 20 年度 国内特許出願件数 : 1 件 (CREST 研究期間累積件数 : 1 件)