

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 16 日現在

機関番号：12601

研究種目：基盤研究(A) (一般)

研究期間：2013～2015

課題番号：25240044

研究課題名(和文) RNA・タンパク質相互作用の網羅的予測と検証

研究課題名(英文) Comprehensive prediction of RNA-protein interactinos

研究代表者

浅井 潔 (Asai, Kiyoshi)

東京大学・大学院新領域創成科学研究科・教授

研究者番号：30356357

交付決定額(研究期間全体)：(直接経費) 36,700,000円

研究成果の概要(和文)：RNAとタンパク質の相互作用の有無及び複合体構造予測を行うことを目指した。PDBのRNA-タンパク質複合体の情報解析から、2次構造で塩基対を形成せずアミノ酸と水素結合を形成する塩基は、2次構造での塩基対形成もアミノ酸との水素結合もない塩基に比べて、低い塩基対確率を持つことがわかった。RNA 2次構造を多面的に解析するため、参照構造からの2次構造のハミング距離毎の構造全ての確率の合計の分布を高速に計算する方法を開発した。RNAとタンパク質の物理化学的性質を考慮した力場を導入したRNA-タンパク質複合体のドッキング手法を開発し、既存の方法よりも大幅に性能が改善することを明らかとなり、論文発表した。

研究成果の概要(英文)：The aim of the research was to predict the RNA-protein interactions for non-coding RNAs and function-known proteins. Our analysis of RNA-protein complex in PDB showed that the nucleotides that do not form base-pairs in RNA 2D structures but form hydrogen bond with amino acids have lower base-pairing probabilities than the nucleotides that form neither base-pairs or hydrogen bonds with amino acids. We developed a new method to understand the landscape of the distribution of RNA 2D structures, by efficiently calculating the probabilities of all the structures with specific Hamming distances from the canonical structures. In order to predict the joint structure of RNA-protein complex, we performed rigid body docking simulations. After revising the force field for RNAs, our docking simulations showed better accuracy than previous methods, and we reported that in a peer reviewed journal.

研究分野：バイオインフォマティクス

キーワード：RNA 2次構造 RNA-タンパク質相互作用 ドッキング計算

1. 研究開始当初の背景

タンパク質をコードしない非コード RNA (non-coding RNA) が多数見つっているが、その大部分の機能は未知である。マイクロ RNA などの機能が既知の非コード RNA ファミリーでは、RNA 間相互作用を予測することが標的分子の同定に必要であるが、機能未知の非コード RNA に関しては、機能的知見が蓄積しているタンパク質との相互作用を同定することが重要である。

RNA とタンパク質の相互作用を情報処理によって予測するためには、(1)タンパク質の立体構造、(2)RNA の立体構造、(3)それら 2 つの相互作用と(4)複合体の構造、の全てについて考慮する必要がある。しかし、(1)、(2)を予測することは、それ自体が挑戦的な課題となっている。

RNA 構造予測に関しては、立体構造予測は未だに挑戦的な課題であるものの、2 次構造の予測は近年大きく進展しており、我々も世界最高レベルの 2 次構造予測ソフトウェア、CentroidFold を開発して高い評価を受けている。

2. 研究の目的

機能未知の非コード RNA の生体内での役割の解明のため、非コード RNA、及び構造既知タンパク質に関して、情報科学・物理学・実験科学の手法を融合し、RNA とタンパク質の相互作用の有無及びその複合体構造の予測を、高速・高精度に行う。

背景で述べたように、タンパク質の立体構造、RNA の立体構造を予測すること自体が、別の挑戦的な課題である。一方、機能と立体構造が明らかとなったタンパク質の数は増加しており、RNA の立体構造はほとんど明らかになっていない。そこで、立体構造が既知のタンパク質と、構造は未知で配列のみが与えられた RNA という組み合わせについて、相互作用と複合体構造を求めることを目的とする。

3. 研究の方法

以下のように、必要な理論とアルゴリズムの研究を行いながら、RNA-タンパク質相互作用の予測と、複合体の立体構造予測を行う。

(1) 相互作用界面に関する解析

RNA 配列情報とタンパク質立体構造情報から相互作用の有無を判別する手法を開発するため、RNA・タンパク質の相互作用界面(相互作用する残基・塩基)に関する 2 次構造、配列の統計的特徴について解析を行う。

(2) RNA 2 次構造分布の数理的解析

RNA が取りうる 2 次構造は膨大な数に上り、個々の 2 次構造の自由エネルギーのボルツマン分布から計算される存在確率は、最小自由エネルギー (MFE) 構造であっても極めて小さい。このような 2 次構造の揺らぎの分布

を捉えることは、RNA のタンパク質の相互作用を考える上でも重要である。そのため、様々な角度から、RNA 2 次構造の分布を解析・可視化する手法を開発する。

(3) RNA-タンパク質複合体の構造予測

粗視化モデルで表現した RNA・タンパクによる分子動力学計算を行い、高速に複合体構造を予測する。さらに、得られた複合体構造を基に全原子モデル MD 計算を行うことにより、相互作用と複合体構造の精密な評価を行う。

4. 研究成果

研究の方法で示した項目について、以下のような研究成果が得られた。

(1) 相互作用界面に関する解析

RNA 2 次構造の確率分布の周辺分布の一つである塩基対確率(任意の 2 つの塩基ペアに注目した時、2 次構造中でその塩基ペアが塩基対を形成する確率)に注目し、RNA-タンパク質相互作用との関係を、PDB (Protein Data Bank) に登録されている RNA-タンパク質複合体の立体構造を解析した。

その結果、RNA-タンパク質複合体の RNA 側で 2 次構造としての塩基対を形成せず、タンパク質のアミノ酸と水素結合を形成する塩基は、2 次構造としての塩基対の形成もアミノ酸との水素結合もない塩基に比べて、明らかに低い塩基対確率を持つことが明らかとなった。一方、立体構造中で 2 次構造としての塩基対を形成している塩基に関しては、アミノ酸との水素結合の有無で、塩基対確率に違いは見られなかった。この結果は、RNA 2 次構造中で塩基対を形成している塩基がタンパク質との相互作用に関与するかどうかを判別する指標の一つとして、塩基対確率用いることができることを示しており、学会及び査読付論文誌に発表した(学会発表 11、雑誌論文 9)。

(2) RNA 2 次構造分布の数理的解析

RNA 2 次構造の周辺分布は、塩基対確率以外にも、2 次構造のすべての要素(ステム、バルジ、ヘアピンループなど)についても考えることができるが(雑誌論文 8)、RNA 2 次構造の確率分布を、理解しやすい形で可視化することは難しい。MFE 構造や セントロイド構造(CentroidFold が予測する構造)を参照構造として、参照構造からの 2 次構造のハミング距離(塩基対を作る組み合わせの違いの数)毎の構造全ての確率の合計の分布を高速に計算する方法を開発した。さらに、2 つの参照構造それぞれからのハミング距離毎の確率の分布の合計(2 次元)を高速に計算する方法も開発し、学会と査読付き論文誌に発表した(学会発表 6,8,12、雑誌論文 7)。また、この成果も含め、RNA 2 次構造解析ツール群をまとめて利用できる Web サーバを構築

し、学会と査読付き論文誌に発表した(学会発表7, 雑誌論文2)。

また、近年ヒトなどで数多く発見されている機能未知の長鎖非コード RNA (long non-coding RNA) について、2次構造や分子間相互作用の情報学的な解析に必要なツールについての総説を執筆した(雑誌論文4)。

(3) RNA-タンパク質複合体の構造予測

粗視化モデルを用いた分子動力学計算(MD)による RNA 単体の立体構造予測, 及び RNA-タンパク質複合体の構造予測に関して, いくつかの複合体については予測を実施したが, 網羅的に予測を行う段階に至らなかった(学会発表3,10)。そこで, Protein Data Bank(PDB)に登録されている RNA とタンパク質の単体の立体構造を用いて, RNA-タンパク質複合体の立体構造を剛体ドッキングで予測することを試みた。タンパク質間相互作用におけるドッキングで実績のある ZDOCK に, MD で用いられる RNA の力場(AMBER 及び CHARMM)を導入することにより, RNA とタンパク質の物理化学的性質を考慮した RNA - タンパク質複合体のドッキング手法を提案した。この手法を様々な RNA-タンパク質複合体構造から構成されるベンチマークデータへ適用し, 予測精度の検証を行った結果, 従来手法より高い精度で予測できることが明らかになり, 査読付き論文誌に発表した(学会発表2, 雑誌論文1)。

5. 主な発表論文等

(研究代表者, 研究分担者及び連携研究者には下線)

[雑誌論文](計10件)

1. Junichi Iwakiri, Michiaki Hamada, Kiyoshi Asai, Tomoshi Kameda (2016) Improved Accuracy in RNA-Protein Rigid Body Docking by Incorporating Force Field for Molecular Dynamics Simulation into the Scoring Function, *J. Chem. Theory Comput.*, 12 (9), pp. 4688-4697. (査読有)
DOI: 10.1021/acs.jctc.6b00254

2. Michiaki Hamada, Yukiteru Ono, Hisanori Kiryu, Kengo Sato, Yuki Kato, Tsukasa Fukunaga, Ryota Mori, Kiyoshi Asai (2016) Rtools: a web server for various secondary structural analyses on single RNA sequences, *Nucleic Acids Res.* 44 (W1): pp. W302-W307. (査読有)
DOI: 10.1093/nar/gkw337

3. Goro Terai, Junichi Iwakiri, Tomoshi Kameda, Michiaki Hamada, Kiyoshi Asai (2016) Comprehensive prediction of lncRNA-RNA interactions in human transcriptome, *BMC Genomics* 17(Suppl 1):12. (査読有)

DOI: 10.1186/s12864-015-2307-5

4. Junichi Iwakiri, Michiaki Hamada, Kiyoshi Asai (2015) Bioinformatics tools for lncRNA research, *Biochimica et biophysica acta* 1859(1), pp. 23-30. (査読有)

DOI: 10.1016/j.bbagr.2015.07.014

5. Shun Sakuraba, Kiyoshi Asai, Tomoshi Kameda (2015) Predicting RNA Duplex Dimerization Free-Energy Changes upon Mutations Using Molecular Dynamics Simulations, *J. Phys. Chem. Lett.*, 6 (21), pp. 4348-4351. (査読有)

DOI: 10.1021/acs.jpcllett.5b01984

6. Haruka Yonemoto, Kiyoshi Asai, Michiaki Hamada (2015) A semi-supervised learning approach for RNA secondary structure prediction, *Computational Biology and Chemistry* 57, pp. 72-79. (査読有)

DOI: 10.1016/j.compbiolchem.2015.02.002

7. Ryota Mori, Michiaki Hamada, Kiyoshi Asai (2014) Efficient calculation of exact probability distributions of integer features on RNA secondary structures, *BMC Genomics* 15(Suppl 10):S6. (査読有)

DOI: 10.1186/1471-2164-15-S10-S6

8. Tsukasa Fukunaga, Haruka Ozaki, Goro Terai, Kiyoshi Asai, Wataru Iwasaki, Hisanori Kiryu (2014) CapR: revealing structural specificities of RNA-binding protein target recognition using CLIP-seq data, *Genome Biology* 15:R16. (査読有)

DOI: 10.1186/gb-2014-15-1-r16

9. Junichi Iwakiri, Tomoshi Kameda, Kiyoshi Asai, Michiaki Hamada (2013) Analysis of base-pairing probabilities of RNA molecules involved in protein - RNA interactions, *Bioinformatics* 29 (20): pp. 2524-2528. (査読有)

DOI: 10.1093/bioinformatics/btt453

[学会発表](計15件)

1. Michiaki Hamada, Yukiteru Ono, Hisanori Kiryu, Kengo Sato, Yuki Kato, Tsukasa Fukunaga, Ryota Mori, Kiyoshi Asai, Rtools: a web server for various secondary structural analyses on single RNA sequences, RNA2016, 2016年6月29,30日, 京都国際会議場。

2. Junichi Iwakiri, Michiaki Hamada, Kiyoshi Asai, Tomoshi Kameda, Improved accuracy in RNA-protein rigid body docking by incorporating force field for molecular

dynamics simulation into the scoring function, RNA2016, 2016年6月29,30日, 京都国際会議場。

3.由良敬, 岩切淳一, 浜田道昭, 浅井潔, 亀田倫史, 分子動力学計算を用いた蛋白質・RNA複合体立体構造予測, 第17回日本RNA学会年会; 2015年07月15日~17日, 札幌パークホテル

4.岩切淳一, 寺井悟朗, 亀田倫史, 浅井潔, 浜田道昭; RNA発現量を用いた組織特異的lncRNA-mRNA相互作用予測; 第17回日本RNA学会年会; 2015年07月15日~17日, 札幌パークホテル

5.寺井悟朗, 岩切淳一, 亀田倫史, 浜田道昭, 浅井潔; ヒトトランスクリプトームにおける網羅的lncRNA-RNA相互作用予測; 第17回日本RNA学会年会; 2015年07月15日~17日, 札幌パークホテル

6.Ryota Mori, Michiaki Hamada, Kiyoshi Asai, Efficient calculation of exact probability distributions of integer features on RNA secondary structures, GIW/ISCB-Asia2014. 日本科学未来館(東京), 2014年12月15~17日

7.浅井潔, 浜田道昭, 小野幸輝, RNA2次構造情報解析のための統合ウェブ, 第16回日本RNA学会年会, 2014年7月23~25日, ウィンク愛知(名古屋)

8.森遼太, 浜田道昭, 浅井潔, Credibility Limitは推定二次構造の定量的な信頼度を示す, 第16回日本RNA学会年会, 2014年7月23~25日, ウィンク愛知(名古屋)

9.岩切淳一, 亀田倫史, 浅井潔, 浜田道昭; RNAタンパク質相互作用予測法の開発, 第16回日本RNA学会年会, 2014年7月23~25日, ウィンク愛知(名古屋)

10.由良敬, 岩切淳一, 浜田道昭, 浅井潔, 亀田倫史; 分子動力学計算を用いた蛋白質・RNA複合体立体構造予測, 第16回日本RNA学会年会, 2014年7月23~25日, ウィンク愛知(名古屋)

11.Ryota Mori, Michiaki Hamada, Kiyoshi Asai, A method for calculating stability of RNA secondary structure, JSBi2013, タワーホール船堀(東京), 2013年10月29~31日

12.岩切淳一, 亀田倫史, 浅井潔, 浜田道昭, タンパク質-RNA相互作用におけるRNA2次構造認識機構:塩基対確率に基づく解析, 第15回日本RNA学会年会, 2013年7月24~

26日, ひめぎんホール(松山)

〔図書〕(計1件)

1. Kiyoshi Asai, Michiaki Hamada (2014) RNA structural alignments, part II: non-Sankoff approaches for structural alignments, Clifton, RNA Sequence, Structure, and Function: Computational and Bioinformatic Methods, Methods in molecular biology, 1097, pp. 291-301

〔その他〕

ホームページ等 <http://rtools.cbrc.jp>

6. 研究組織

(1) 研究代表者

浅井 潔 (ASAI, Kiyoshi)
東京大学・新領域創成科学研究科・教授
研究者番号: 30356357

(2) 研究分担者

浜田 道昭 (HAMADA, Michiaki)
早稲田大学・理工学術院・准教授
研究者番号: 00596538

亀田 倫史 (KAMEDA, Tomoshi)
(国研)産業技術総合研究所・人工知能研究センター・主任研究員
研究者番号: 40415774

石黒 亮 (Akira Ishiguro)
東京大学・医科学研究所・特任准教授
研究者番号: 70373264
(平成25年度のみ研究分担者)

(3) 連携研究者

由良 敬 (YURA, Kei)
お茶の水女子大学・人間文化創成科学研究科・教授
研究者番号: 50252226