

2021年10月6日

報道関係者各位

 慶應義塾大学  
 大阪大学大学院医学系研究科

## より複雑で長い RNA の二次構造を高速に予測可能に

慶應義塾大学理工学部生命情報学科の佐藤健吾専任講師と大阪大学大学院医学系研究科の加藤有己助教からなる研究グループは、シュードノットと呼ばれる複雑な RNA 部分構造を予測可能な RNA 二次構造予測法 (IPknot++) を開発することに成功しました。従来、数千塩基を超える RNA 配列に対するシュードノットを含めた RNA 二次構造予測は計算量の観点から極めて困難でしたが、本手法はこれを克服して高速かつ高精度の予測を実現しました。メッセンジャー RNA (mRNA) やウイルス RNA などの長鎖 RNA 配列においてシュードノットが関与する遺伝子発現制御機構の解明などへの応用が期待されます。

本研究成果は、2021年10月2日にイギリスの科学雑誌 *Briefings in Bioinformatics* のオンライン版に掲載されました。

### 1. 本研究のポイント

- 必要な計算量を配列長の3乗から線形(1乗)に削減し、mRNA やウイルス RNA などの長い RNA 配列に対するシュードノットを考慮した RNA 二次構造予測を可能にした。
- あらかじめ予測精度を見積もる指標を導入することで適切なパラメータが選択可能となり、二次構造予測の精度と汎用性が向上した。

### 2. 研究背景

RNA (リボ核酸) は生物の細胞内に存在し、DNA のゲノム (遺伝情報) からタンパク質の情報をコピーしている鎖状の高分子です。近年、COVID-19 に対するワクチンで利用されるなど、その医療応用に向けた有用性が世界的に認知されています。この RNA の未知の機能を解明するには、強い相関関係がある構造の解析が不可欠で、コンピューターにより RNA の塩基配列データから、その折り畳み構造 (RNA 二次構造<sup>\*)</sup>) を予測するアプローチは、時間とコストがかかる構造解析実験技術を補完する有力なツールとされています。

シュードノットは、RNA 二次構造において重要なトポロジーの一つです (図 1)。任意のシュードノット構造を含む RNA 二次構造予測の厳密解法は、その計算量 (計算の複雑さ) が極めて複雑なクラスであることが証明されており、その実装は現実的ではないとされています。そのため、これまでの多くの RNA 二次構造予測法はシュードノット構造を無視した予測を行います。しかしながら、シュードノット構造は翻訳やスプライシングの制御、リボソームのフレームシフトなどに関与することが知られており、シュードノット構造を考慮した RNA 二次構造予測手法が求められています。

そこで、本研究グループは近似解法を用いて計算を高速化する IPknot を 2011 年に開発しました [Sato et al. 2011]。しかし、配列長に対して3乗に比例する計算時間を必要とする上に、配列長

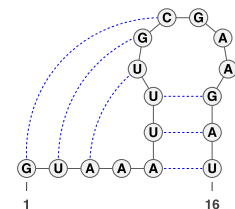


図 1 シュードノット

が 500 塩基を超えたあたりから予測精度が著しく低下するため、mRNA やウイルス RNA のような長い配列に適用することは困難でした。

### 3. 研究内容・成果

本研究では、IPknot を改良することによって、mRNA やウイルス RNA のような長い配列に対しても高速かつ高精度でシュードノット構造を含む RNA 二次構造を予測できる手法の開発を目指しました。高速化のために、ベースとなる計算モデルに解析手法高速化の点で良い近似を実現する LinearPartition モデル[Zhang et al. 2020]を採用しました。LinearPartition モデル自体はシュードノットを考慮した計算は行いませんが、これに IPknot による近似解法を組み合わせることによって、配列長に対して線形（1 次関数）の計算量でシュードノットを考慮した RNA 二次構造予測を実現しました。さらに、あらかじめ予測精度を見積もる指標 pseudo-expected accuracy をシュードノット構造に適用し、これに基づいて配列ごとに最適なパラメータを自動的に選択する方法を開発し、これによって高精度化を実現しました。

本研究で開発した IPknot++ は、網羅的なベンチマークにおいて幅広い条件でシュードノットを含まない RNA 二次構造予測と同等の計算速度でありながら（図 2）、シュードノットを含む配列に対しても良好な予測精度であることを示しました（表 1）。

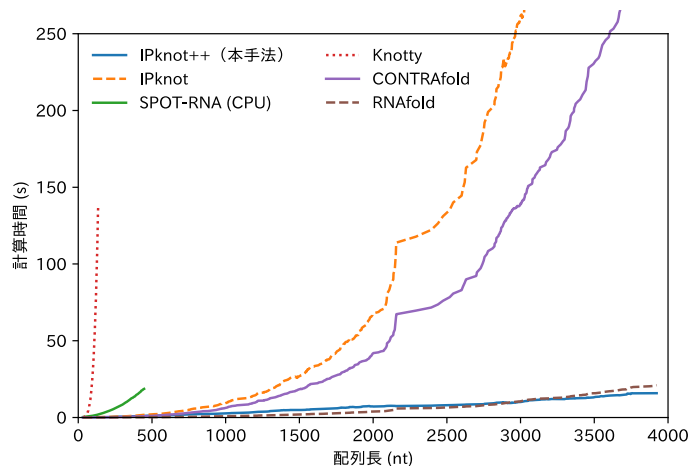


図 2 他手法との計算時間の比較

表 1 配列長ごとの他手法との F 値の比較

配列長	12-150 塩基		151-500 塩基		501-4381 塩基	
	PK なし	PK あり	PK なし	PK あり	PK なし	PK あり
IPknot++ (本手法)	0.681	0.552	0.492	<b>0.482</b>	<b>0.433</b>	<b>0.428</b>
IPknot	0.669	0.500	0.480	0.461	0.212	0.317
Knotty	0.641	0.550	-	-	-	-
SPOT-RNA	0.658	<b>0.621</b>	0.462	0.479	-	-
CONTRAfold	<b>0.682</b>	0.519	<b>0.500</b>	0.479	0.425	0.415
RNAfold	0.668	0.472	0.474	0.442	0.361	0.347

F 値：予測した塩基対が正解構造に含まれる割合（正解率; PPV）と正解構造に含まれる塩基対を予測できた割合（網羅率; SEN）の調和平均

PK なし：シュードノットを含まない配列の F 値

PK あり：シュードノットを含む配列の F 値

本手法 IPknot++はウェブサイトよりダウンロード可能なほか、ウェブサーバとして簡単に利用することができます。

ダウンロード用ウェブサイト <https://github.com/satoken/ipknot/>

ウェブサーバ <http://rtips.dna.bio.keio.ac.jp/ipknot++/>

#### 4. 今後の展開

RNA 二次構造予測は計算機による RNA 配列解析において最も基盤となる技術です。本研究で開発した IPknot++は、長い配列に含まれるシュードノット構造を予測できることを示しました。シュードノット構造は翻訳やスプライシングの制御などに関与することが知られています。特に近年、SARS-CoV-2 でリボソームのフレームシフトを引き起こす要因となっていることがわかり、創薬ターゲットとしても注目されています。このような RNA の機能を推定するために、本手法により実現した長い配列に対するシュードノットを考慮した RNA 二次構造予測が有用であると期待されます。

#### <参考文献>

[Sato et al. 2011] Sato, K., Kato, Y., Hamada, M., Akutsu, T. & Asai, K. “IPknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming.” *Bioinformatics* 27, i85-i93 (2011)

[Zhang et al. 2020] Zhang, H., Zhang, L., Mathews, D. H. & Huang, L. “LinearPartition: linear-time approximation of RNA folding partition function and base-pairing probabilities.” *Bioinformatics* 36, i258-i267 (2020)

#### <原論文情報>

Sato, K., Kato, Y. “Prediction of RNA secondary structure including pseudoknots for long sequences.” *Briefings in Bioinformatics*.  
doi: 10.1093/bib/bbab395

#### <用語説明>

##### ※1 RNA 二次構造

立体構造の骨格を成す塩基対の集合で、A-U、G-C、G-U の規則で結合する。

※ご取材の際には、事前に下記までご一報くださいますようお願い申し上げます。

※本リリースは文部科学記者会、科学記者会、大阪科学・大学記者クラブ、各社科学部等に送信させていただきます。

---

#### ・研究内容についてのお問い合わせ先

慶應義塾大学 理工学部 生命情報学科 専任講師 佐藤 健吾 (さとう けんご)

TEL : 045-566-1511 FAX : 045-566-1789 E-mail : [satoken@bio.keio.ac.jp](mailto:satoken@bio.keio.ac.jp)

#### ・本リリースの配信元

慶應義塾広報室 (澤野)

TEL : 03-5427-1541 FAX : 03-5441-7640

Email : [m-pr@adst.keio.ac.jp](mailto:m-pr@adst.keio.ac.jp) <https://www.keio.ac.jp/>

大阪大学大学院医学系研究科 広報室

TEL: 06-6879-3388 FAX: 06-6879-3399

Email: medpr@office.med.osaka-u.ac.jp