

長いエクソン群が支える液-液相分離転写ネットワーク

増田 章男, 河地 利彦

1. はじめに

細胞内には、さまざまな機能性画分が存在する。たとえば、核、ミトコンドリア、リソソーム、小胞体といった膜で囲まれた細胞小器官が代表的であるが、これらに加え、核内スペckルやストレス顆粒といった膜を持たない機能性画分もあり、それぞれ重要な細胞機能を担っている。膜に依存しない画分の形成には、近年、水と油に代表される液-液相分離 (liquid-liquid phase separation: LLPS) が中心的な役割を果たすことが明らかとなり、注目を集めている¹⁾。

液-液相分離は多様な分子間相互作用で形成されうるが、中でも、タンパク質内にある天然変性領域 (intrinsically disordered region: IDR) がその形成に重要な役割を持っている。天然変性領域は、数種程度の限られたアミノ酸を中心に構成され、整った立体構造をとれず、ひも様に揺らぐタンパク質領域を指す。いわゆるドメインがさまざまなアミノ酸から構成され、複雑で固定した立体構造をとるのと対照的である。以前から、天然変性領域には強い転写活性や分子間相互作用があることが知られていたが、近年、その本態に液-液相分離による液滴形成があることが見いだされ、盛んに研究が行われている。

天然変性領域は、特に高等生物に顕著で、そのプロテオームの30~50%を占めるほどである²⁾。単純で偏った配列特徴を持つ天然変性領域は、ドメインとは相当に異なる機構で維持・機能発揮されている、と考えられるが、その詳細はまだ研究途上である。本稿では、最近のトランスクリプトームワイド解析の知見を交えて紹介していきたい。

2. 天然変性領域のアミノ酸配列

天然変性領域の一例として、転写調節因子BRD4の天然変性領域を図1に示す。一見してわかるように、プロリン(P)が頻出し、グルタミン酸(Q)やセリン(S)がそれに続く。天然変性領域には、P, E, S, K, Q, H, Dといったアミノ酸が頻出することが知られており³⁾、数種のアミノ酸が延々と並び、ところどころ他のアミノ酸が入っている、というのが典型的である。アミノ酸組成の偏りは、物性 (電荷、疎水性など) の偏りにつながり、揺らぐ構造と相まって液-液相分離の誘引となっている。異なる性質の天然変性領域は、水と油のようにはじき合い、親和性のある領域どうしなら混和して、液滴を形成する。こうして、液-液相分離を介して必要な分子群が局所に集まることで、機能性画分が生まれる、というのが液-液相分離モデルである。

天然変性領域のアミノ酸残基は、リン酸化やメチル化などの翻訳後修飾をよく受けている²⁾。これらの修飾は、液-液相分離の性質を変化させ、天然変性領域を持ったタンパク質の機能性画分への参画を動的に制御することに役立っている。たとえば、高い液-液相分離能を持つことで有名なRNA結合タンパク質FUSでは、天然変性領域内のセリンやトレオニンのリン酸化⁴⁾や、チロシンとカチオン- π 相互作用するアルギニンのメチル化⁵⁾が、その相分離能を制御している。

3. 天然変性領域をコードする塩基配列の偏り

天然変性領域の限られたアミノ酸構成は、そのmRNA塩基配列の偏りにもつながっている (図1B)。RNA代謝の観点からみると、この塩基配列の偏りは由々しきことで、スプライシングをはじめとするmRNA代謝に必要な制御配列 (たとえば、スプライスサイト配列) の確保が困難になっている⁶⁾。実際、天然変性領域をコードするエクソンは、エクソンスキッピングとなってしまう機会が多く⁷⁾、こうしたスプライス制御配列の不十分さがその一因と考えられる。

スプライス制御配列の不十分さを補うため、天然変性領域をコードするエクソンにはスプライシング促進配列 (exon splicing enhancer: ESE) が頻繁に埋め込まれている。偏った天然変性領域のアミノ酸コドン、当然、これら

名古屋大学大学院医学系研究科先端応用医学部門神経遺伝情報分野 (〒466-8550 愛知県名古屋市昭和区鶴舞町65)

Liquid-liquid phase separation of transcription factors are sustained by large exons in vertebrates

Akio Masuda and Toshihiko Kawachi (Division of Neurogenetics, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, 65 Tsurumai, Showa-ku, Nagoya, Aichi 466-8550, Japan)

本論文の図版はモノクロ (冊子版) およびカラー (電子版) で掲載。

DOI: 10.14952/SEIKAGAKU.2022.940706

© 2022 公益社団法人日本生化学会

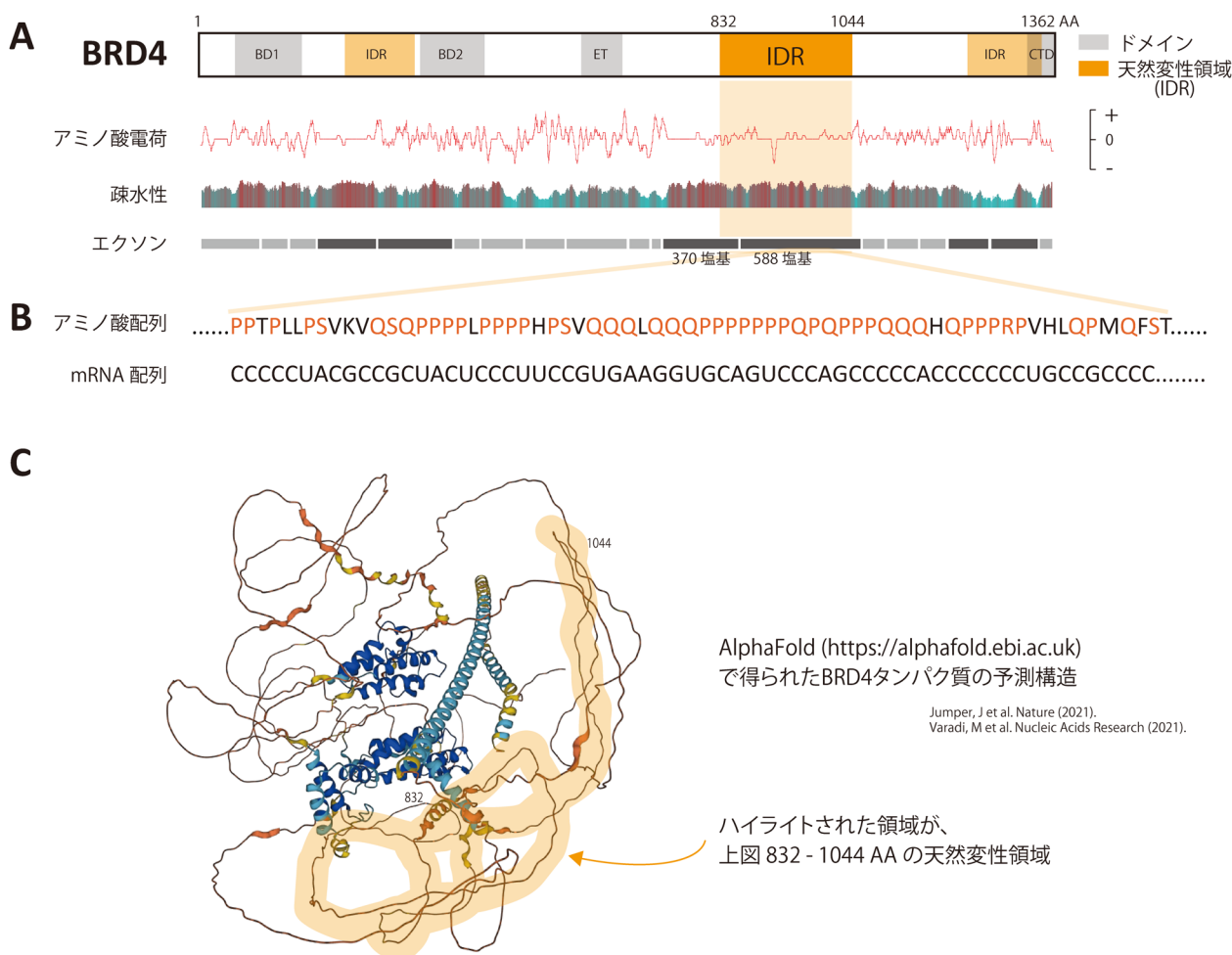


図1 転写調節因子BRD4のタンパク質構造

(A) ドメイン／天然変性領域 (IDR), アミノ酸電荷, 疎水性, エクソン構造. IDR (832-1044 AA) では, 電荷が0付近で, 疎水性に傾いている. また, 天然変性領域をコードするエクソン (濃灰色) は, ドメインをコードするエクソン (薄灰色) に比べ長い. BD: bromodomain, ET: extra-terminal domain, CTD: C-terminal domain. 天然変性領域は, PONDR (VL-XT, <http://www.pondr.com>)¹³⁾ を使って予測した. 連続100アミノ酸以上の天然変性領域を示す. (B) 天然変性領域のアミノ酸配列, mRNA 配列. 図示した領域の配列を示す. P, Q, S アミノ酸とC塩基に著しく偏っている. (C) BRD4タンパク質の予測構造. IDR (832-1044 AA) は, ほとんど構造をとらず, ひも状になっている.

のESE配列の確保をも制約するが, たとえば, グリシンが豊富な天然変性領域では, そのコドンGGNの集積がESEとなり, GGG配列に好んで結合するRNA結合タンパク質SRSF1の作用によりスプライシングが促進されている. これが, プロリンならコドンはCCNでCCC配列を好むSRSF3 (図2B参照), リシンならコドンはAARでTRA2といった具合で, それぞれの偏ったコドンに対応したRNA結合タンパク質の活用により, 天然変性領域エクソンのスプライシングが確保されている^{8,9)}.

4. エクソン長と天然変性領域

天然変性領域をコードするエクソンは, 塩基配列に加え

その長さにも特徴がある. 図2Aは, ヒトのすべてのエクソンを用いて, エクソン長別に天然変性領域をコードする割合を示したものである. 天然変性領域が, とりわけ長いエクソンに集積していることがわかる. ヒトにおけるエクソン長は, 中央値で約120塩基, 95%以上のエクソンが300塩基以下であることを考えると, 天然変性領域をコードするエクソンの特殊さが想像できると思う.

これら長いエクソンには, プロリンやセリンからなる天然変性領域が集積している. また, 特にプロリンのコドンによって, 塩基レベルならC塩基に大きく偏っている (図1). 一般的に, 長いエクソン長はスプライシングの障壁となるが, 上述のように, 集積したC塩基により活性化されるSRSF3がそのスプライシングを保証している (図

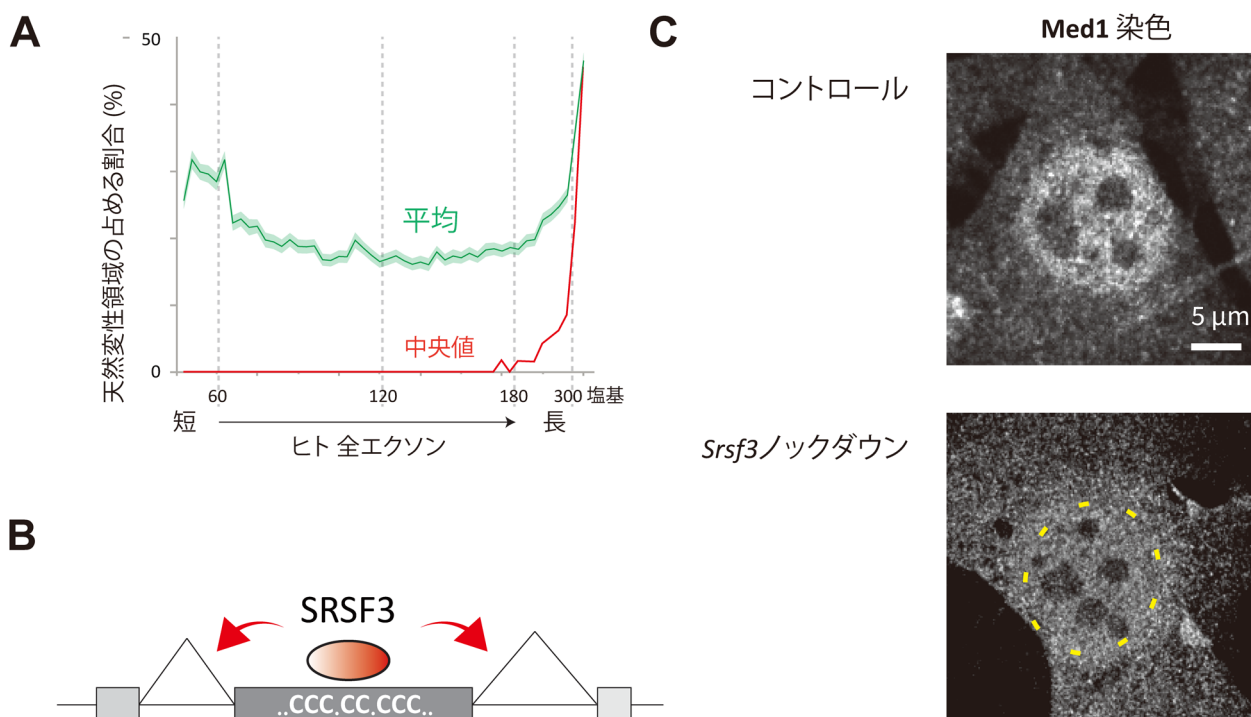


図2 SRSF3が天然変性領域をコードするエクソンのスプライシングを保証する

(A)長さ別にヒトの全エクソンがコードする天然変性領域の割合。mRNAの最初と最後のエクソンを除いた、ヒトの全エクソンを長さ順に並べた後、50分割し、それぞれの画分における天然変性領域の平均含有率を示す。文献9より改変。(B) SRSF3による長いエクソンのスプライシング模式図。SRSF3がC塩基配列を認識してmRNAに結合することで、長くC塩基に偏ったエクソンのスプライシングが促進される。(C) *Srsf3* ノックダウンC2C12細胞の抗Med1抗体による免疫染色写真。コントロール細胞では、Med1は核内に集積し、顆粒状にメディエーター複合体を形成する。*Srsf3* ノックダウン細胞では、これらが消失する(破線で囲んだ部分が核領域)。

2B)。SRSF3のスプライシング促進作用は非常に強力で、CCC配列を好む他のRNA結合タンパク質(hnRNP Kなど)のスプライシング抑制作用を完全に打ち消してしまうことが実験的に確認されている⁹⁾。遺伝子オントロジー(GO)解析では、これら大きなエクソンを持つ遺伝子は、転写関連因子に集中しており、転写との深い関係が示唆される。

逆に、短いエクソンでは、やや天然変性領域をコードする傾向がみられるものの(図2A)、特に3~15塩基のきわめて短いエクソンの場合、むしろドメインをコードする領域に存在することが知られている¹⁰⁾。これらは、神経特異的因子のタンパク質-タンパク質結合部位に相当する位置を占めることが多く、神経発達への関与が報告されている。

5. SRSF3による転写制御機構の維持

天然変性領域をコードする長いエクソンは、転写因子(CREBBP/EP300, MEF, NFAT, SP, SMAD)、基本転写因子(TAF1, 4, 6, 12, SUPT2L)、メディエーター複合体(MED1, 12~16, 25, CDK8)、プロモドメイン複合体(BRD1~4, 8)、BAF複合体(ARID1A, 1B, SMARCC1, C2)など、クロマチ

ン制御の中核をなす転写関連因子群に幅広く認められる。実験的にSRSF3をノックダウンさせた細胞では、これら長いエクソンのスプライシングができなくなり、大量の遺伝子(2000個以上)のmRNAで著しいエクソンスキッピングが生じる⁹⁾。結果、多くの転写関連タンパク質で天然変性領域が消失し、転写複合体の崩壊へと進み、激しい転写障害から急速な細胞死が生じる(図2C)。天然変性領域の利用により転写機構は高度な発達を遂げているが、SRSF3というたった一つのRNA結合タンパク質が、そのmRNAレベルを支えていることが示唆される。

ちなみにSRSF3はさまざまながん細胞でコピー数が増加しており、がん遺伝子としても知られている。SRSF3の発現増加は、がん細胞で活発化しているmRNA産生を支えているのであろう。しかし、SRSF3に変異が入ったがん細胞はまれであり、こうした異常細胞にとってもSRSF3の機能異常は受け入れがたいことであることを示している。

6. 天然変性領域の進化上の獲得

上述のように、天然変性領域の塩基配列やエクソン長

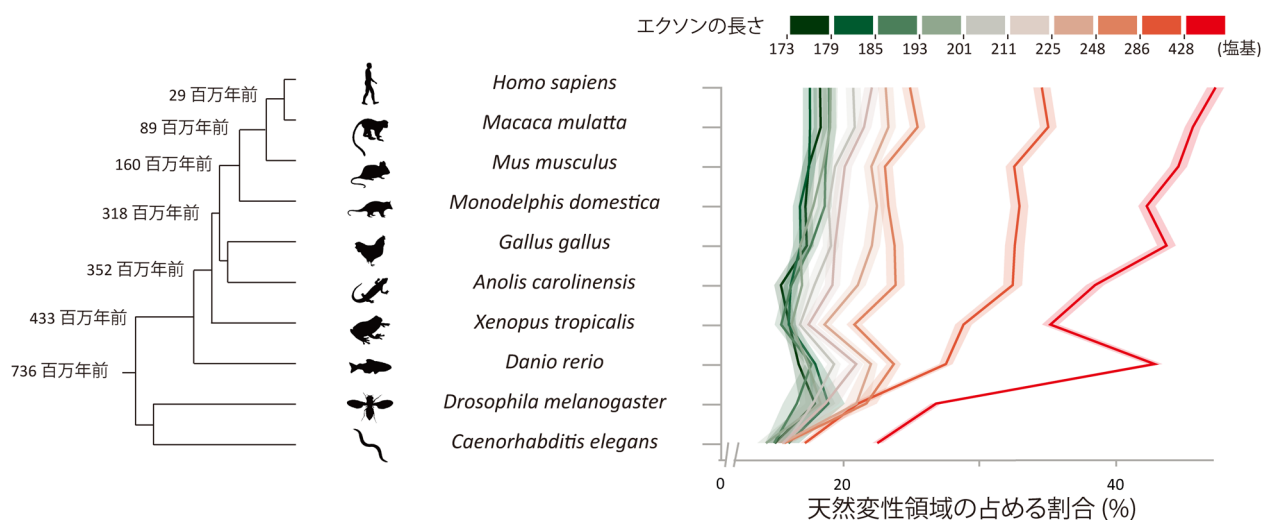


図3 長いエクソン (mRNAの最初と最後のエクソンを除く) がコードする天然変性領域の割合種ごと (縦軸) にエクソン長別 (上横軸) に示す。進化した種ほど、天然変性領域のコードが長いエクソンに集中する。文献9より改変。

は、通常と大きく異なった特徴を示すが、これらは進化上、どのように獲得されてきたのだろうか。天然変性領域は、原核生物のタンパク質領域の数%を占めるにすぎないが、真核生物では30%以上を占めるに至っている¹¹⁾。さらに、真核生物でも比較的遅くに出現した遺伝子により多く認められるなど、真核生物で迅速な進化を遂げてきたことが明らかとなっている¹²⁾。

長いエクソン領域に目を移すと、天然変性領域の長いエクソンへの集積は、線虫やショウジョウバエでは目立たず、魚類で急に出現し、その後、霊長類に至るまで緩徐に進行していく⁹⁾ (図3)。この進化は、長いエクソンのC塩基やプロリン・セリンのコードの増加、転写関連因子遺伝子への偏在を伴って進行し、また、SRSF3遺伝子の出現経過とよく一致している。C塩基を介したSRSF3のスプライシング保証作用を獲得したおかげでプロリン・セリンに富んだ天然変性領域の拡張が可能となり、脊椎動物の転写制御機構が進化したのではないかと考えている。

7. おわりに

我々ヒトを含む脊椎動物は、大きなゲノム変化の結果として、それまでの生物とは一線を画す進化を果たしている。その変化の一つに、転写因子の多様なネットワーク活動の発展があげられ、その源にはエクソンレベルでの天然変性領域の獲得があったことが、今回の研究から示唆される。天然変性領域には、プロリン・セリン以外にも、さまざまなアミノ酸に偏った種類が存在している。それらの領域は、どのようなエクソンから構成され、どのような細胞機能を担っているのだろうか？ 疑問は尽きない。今後と

も、さらなる研究を進めていきたい。

謝辞

本研究は科研費 (18K06058, 21H02476), AMED (21ek0109497h0001) の助成を受けたものである。

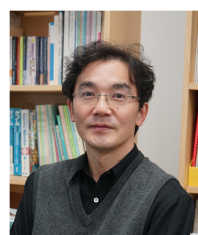
文 献

- 1) Misteli, T. (2020) The self-organizing genome: Principles of genome architecture and function. *Cell*, **183**, 28–45.
- 2) Oldfield, C.J. & Dunker, A.K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.*, **83**, 553–584.
- 3) Campen, A., Williams, R.M., Brown, C.J., Meng, J., Uversky, V.N., & Dunker, A.K. (2008) TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept. Lett.*, **15**, 956–963.
- 4) Monahan, Z., Ryan, V.H., Janke, A.M., Burke, K.A., Rhoads, S.N., Zerze, G.H., O'Meally, R., Dignon, G.L., Conicella, A.E., Zheng, W., et al. (2017) Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.*, **36**, 2951–2967.
- 5) Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Strohl, F., et al. (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation- π Interactions. *Cell*, **173**, 720–734.e715.
- 6) Smithers, B., Oates, M.E., & Gough, J. (2015) Splice junctions are constrained by protein disorder. *Nucleic Acids Res.*, **43**, 4814–4822.
- 7) Ellis, J.D., Barrios-Rodiles, M., Çolak, R., Irimia, M., Kim, T., Calarco, J.A., Wang, X., Pan, Q., O'Hanlon, D., Kim, P.M., et al. (2012) Tissue-specific alternative splicing remodels protein-protein interaction networks. *Mol. Cell*, **46**, 884–892.
- 8) Fontrodona, N., Aube, F., Claude, J.B., Polveche, H., Lemaire, S.,

- Tranchevent, L.C., Modolo, L., Mortreux, F., Bourgeois, C.F., & Auboeuf, D. (2019) Interplay between coding and exonic splicing regulatory sequences. *Genome Res.*, **29**, 711–722.
- 9) Kawachi, T., Masuda, A., Yamashita, Y., Takeda, J.I., Ohkawara, B., Ito, M., & Ohno, K. (2021). *EMBO J.*, **21**, e107485.
- 10) Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallieres, M., Tapial, J., Raj, B., O'Hanlon, D., et al. (2014) A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell*, **159**, 1511–1523.
- 11) Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., & Jones, D.T. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.*, **337**, 635–645.
- 12) Afanasyeva, A., Bockwoldt, M., Cooney, C.R., Heiland, I., & Gossmann, T.I. (2018) Human long intrinsically disordered protein regions are frequent targets of positive selection. *Genome Res.*, **28**, 975–982.
- 13) Liu, J., Perumal, N.B., Oldfield, C.J., Su, E.W., Uversky, V.N., & Dunker, A.K. (2006) Intrinsic Disorder in Transcription Factors. *Biochemistry*, **45**, 6873–6888.

著者寸描

●増田 章男 (ますだ あきお)



名古屋大学大学院医学系研究科先端応用医学部門神経遺伝情報分野准教授。医学博士。

■略歴 1994年名古屋大学医学部卒業。同年一宮市立市民病院内科医師。98年名古屋大学医学部大学院。2001年名古屋大学医学部付属病態制御研究施設生体防御分野助手。03年名古屋大学大学院医学系研究科先端応用医学部門神経遺伝情報分

野助教。12年より現職。

■研究テーマと抱負 幅広い生命現象制御の分子基盤であるRNA代謝について、生化学や細胞生物学の手法にバイオインフォマティクスを融合させて、明らかにしていきたいと考えています。

■ウェブサイト https://www.med.nagoya-u.ac.jp/medical_J/laboratory/neurological-diseases-cancer/dept-neuroscience/neurogenetics/

■趣味 登山。