

博士論文

GC-rich 配列を介した

UPF1 依存的 mRNA 分解メカニズムの解明

今町 直登

## 目次

第 I 部 序論 .....	2
第 II 部 本論 .....	11
第 III 部 総括と展望 .....	50
第 IV 部 材料と方法 .....	58

参考文献

謝辞

# 序論

生命の遺伝情報は、DNA から RNA に転写され、タンパク質に翻訳される。この一連の遺伝情報の伝達をセントラルドグマと呼んでいる。しかし、遺伝子発現の制御という観点からすると、セントラルドグマの概念は物質（RNA やタンパク質）の生成段階にしか着目していない。そのため、物質の分解段階、つまりは RNA やタンパク質が分解される過程については、いわばセントラルドグマから外れた影の存在であり、セントラルドグマを提案したクリックらの時代では重要視されていなかったメカニズムである。しかし、遺伝子発現の量的制御を考えたときに、例えば、RNA の発現量は RNA の転写レベルと分解レベルのバランスによって制御されていることから、RNA やタンパク質の生成段階だけでなく、分解段階に着目することも生理機能の制御、ひいては生命を理解する上で重要な要素であるといえる。実際に、選択的に特定の RNA 集団を認識して分解するために、RNA ヘリカーゼや RNA 結合タンパク質、microRNA が、それらの標的 mRNA の特徴的な構造や配列を認識して、RNA 分解を誘導していることが解明されてきた（Wu and Brewer 2012; Balagopal et al. 2012）。

さらに、RNA 分解は、異常な mRNA を排除する「RNA 品質管理機構」とも密接に関係している。RNA の品質管理機構として最もよく知られているのが、ナンセンス変異依存 mRNA 分解（NMD）である。この分解経路によりナンセンス変異（PTC; premature termination codon）を持つ異常な mRNA を生体内で分解している（Kurosaki and Maquat 2016; Lykke-Andersen and Jensen 2015; Schweingruber et al. 2013）（図 0-A）。ヒト細胞では、スプラ

イシングを受けた mRNA では、exon-exon boundary の上流 20-24nt の地点に、EIF4A3、Y14 (RBM8A)、MAGOH、CASC3 からなる exon junction complexes (EJCs)が結合している。また、ほとんどのヒトの遺伝子では、最終エキソン上に終止コドンが存在する。そのため、通常は終止コドンよりも上流に EJC が位置している。EJC が終止コドンの >50-55 nt 下流に位置していた場合には、そのような終止コドンは PTC として認識され、異常な mRNA として NMD 経路によって分解される。現在のモデルでは、UPF1 と SMG1 の複合体が翻訳終結因子である eRF1 (ETF1)と eRF3 (GSPT1)を介して PTC を認識しているとされている (Kashima et al. 2006; Schoenberg and Maquat 2012)。また、UPF2 が EJC にリクルートされ、EJC と UPF1 に結合する。結果として、EJC-UPF2/3-UPF1-eRF1/3 の複合体が形成されることで、SMG1 により UPF1 がリン酸化される (Lykke-Andersen and Jensen 2015; Isken et al. 2008) (図 0-B)。UPF1 のリン酸化が引き金となり、RNA 分解因子が RNA 上にリクルートされることで RNA 分解を誘導する。このように、UPF1 は NMD 経路において中心的な役割を果たしている。

ここで、STAU1 という RNA 結合タンパク質が一群の mRNA を分解することで発現量を制御していることが知られている (STAU1 依存 mRNA 分解 (SMD 経路) と呼ばれる)。近年の研究から、SMD 経路をはじめとする正常な mRNA の分解にも UPF1 が関与していることが明らかとなっている。例えば、SMD 経路では、UPF1 は直接 STAU1 と結合し、正常な mRNA の分解を誘導している (Kim et al. 2005)。このとき、NMD 経路に関わる UPF2 は SMD 経路には関与していない。

UPF1 は酵母からヒトに至るまでさまざまな真核生物で高度に保存されたタンパク質である (Culbertson and Leeds 2003; Imamachi et al. 2012)。UPF1 のノックアウトマウスは胎生致死となり、その原因は UPF1 を介した RNA 分解が破綻することにより、mRNA の分解が正常に行われなくなることで個体が死に至ると考えられている (Medghalchi et al. 2001)。UPF1 を介した RNA 分解により、1-10%の正常な mRNA の遺伝子発現制御が行われていることが知られている (Tani et al. 2012)。UPF1 は 3'UTR に結合する傾向があり (Hurt et al. 2013; Zünd et al. 2013; Gregersen et al. 2014; Kurosaki et al. 2014; Lee et al. 2015)、3'UTR への UPF1 の結合が RNA 分解を誘導していると考えられている (Hurt et al. 2013)。このことから、3'UTR 中に存在する特徴配列が UPF1 を介した RNA 分解に重要な役割を果たしていると考えられる (図 0-A)。しかしながら、どのような特徴配列が UPF1 を介した RNA 分解を誘導しているのかはわかっていなかった。そこで、本研究では、RNA の分解速度と RNA と UPF1 タンパク質の相互作用をそれぞれ測定することにより UPF1 の標的 mRNA を同定し、どのような配列特徴が UPF1 を介した正常な mRNA の分解経路における基質認識に関わっているか明らかにすることを目的とした (図 0-C)。

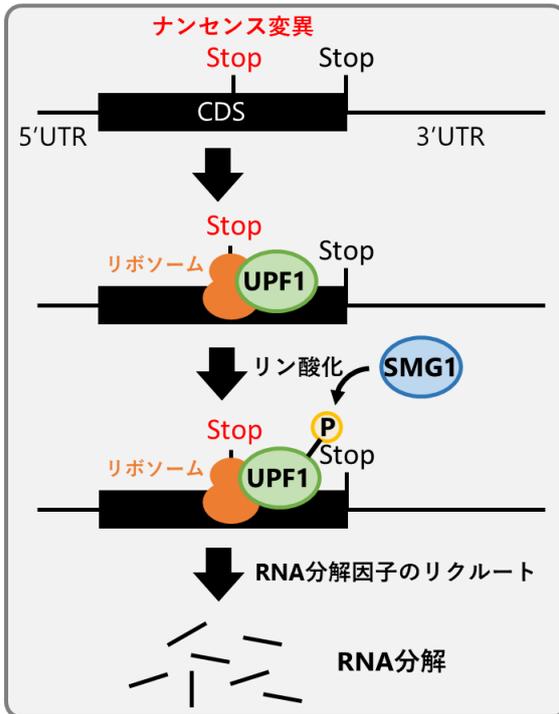
先行研究では、UPF1 をノックダウンしたときに発現量が上昇する mRNA を標的 mRNA として着目する方法が取られてきた (Mendell et al. 2004; Viegas et al. 2007)。ただ、RNA の発現量は RNA の転写と分解のバランスによって決められていることから、この方法では RNA の発現量の上昇が RNA 分解の抑制による影響であるかどうか判断することは難しい。例えば、特定の転写因

子は RNA 分解による遺伝子発現の制御を受けており (Maekawa et al. 2015)、UPF1 のノックダウンによる影響は転写の亢進による RNA の発現量の上昇に起因している可能性がある。以上のことから、先行研究の方法では False-positive にあたる遺伝子が多く取れてきてしまうという問題があった。そこで、UPF1 をノックダウンしたときの RNA の発現量ではなく、RNA の分解速度・半減期を直接測定した (Tani et al. 2012)。また、UPF1 タンパク質は、RNA と直接結合するタンパク質であることから、RNA 免疫沈降法を利用することで、RNA と UPF1 タンパク質の相互作用の有無を評価した。RNA 分解速度を測定するために、5' -bromouridine (BrU) immunoprecipitation chase-deep sequencing analysis (BRIC-seq)と呼ばれる次世代シーケンサーを用いた網羅的な測定方法を利用した (Tani et al. 2012; Imamachi et al. 2014)。BRIC-seq 法を概説すると、まず、新規に転写される mRNA を BrU によりパルスラベル標識する。次に、BrU でラベルされた RNA を経時的にサンプリングし、ラベルされた RNA の減少量を次世代シーケンサーにより定量することで、RNA の分解速度を測定する (図 0-D)。RNA と UPF1 タンパク質の相互作用を測定するために、RIP-seq と呼ばれる網羅的な測定方法を利用した (図 0-E)。

さらに、リン酸化 UPF1 タンパク質に対する RNA フットプリントデータ (Kurosaki et al. 2014)

(図 0-F) のデータを用いて、UPF1 標的 mRNA 中に含まれる特徴配列を特定した。

### 異常なmRNAの分解



### 正常なmRNAの分解

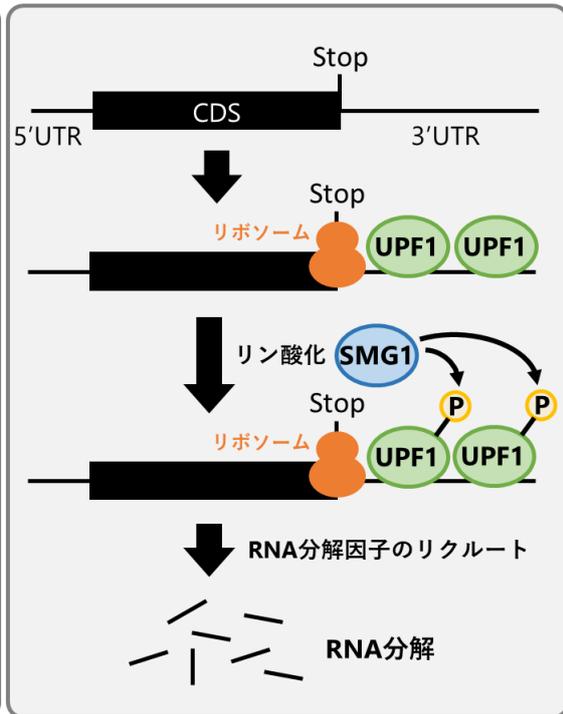


図 0-A. UPF1 を介した RNA 分解経路の概略

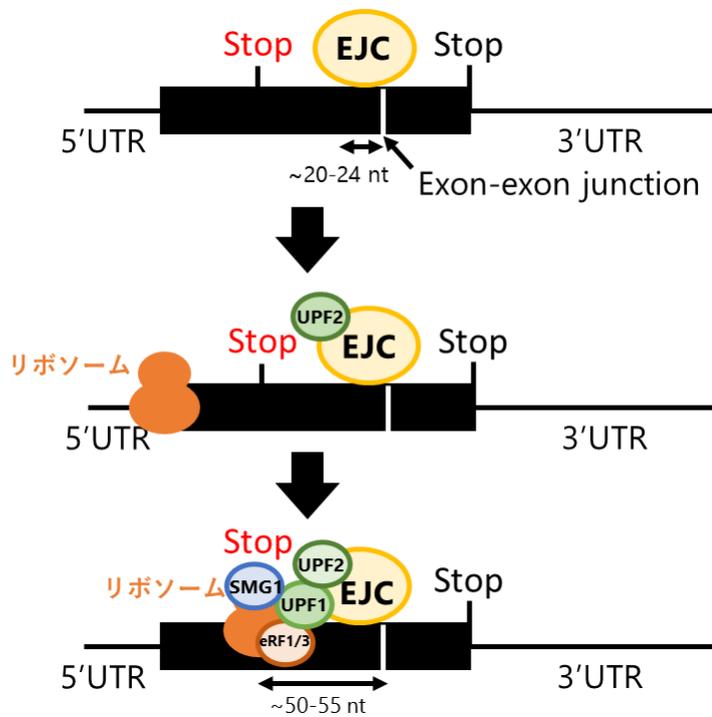
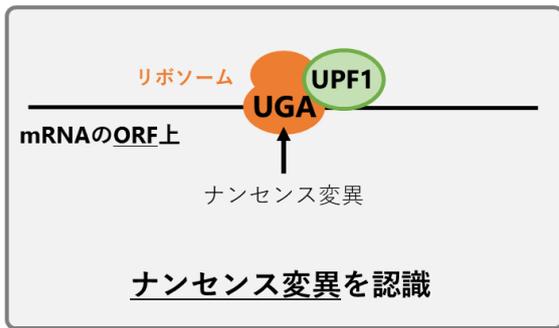


図 0-B. NMD 経路におけるナンセンス変異 (PTC) の認識

### 異常なmRNAの分解



### 正常なmRNAの分解

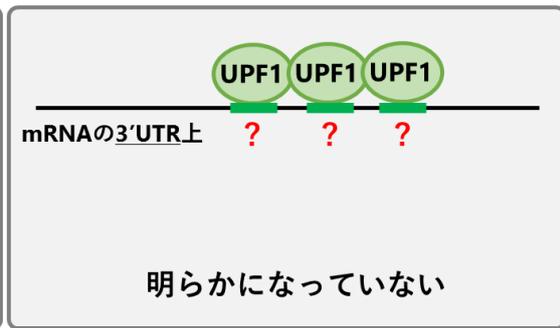


図 0-C. UPF1 を介した RNA 分解経路における基質認識

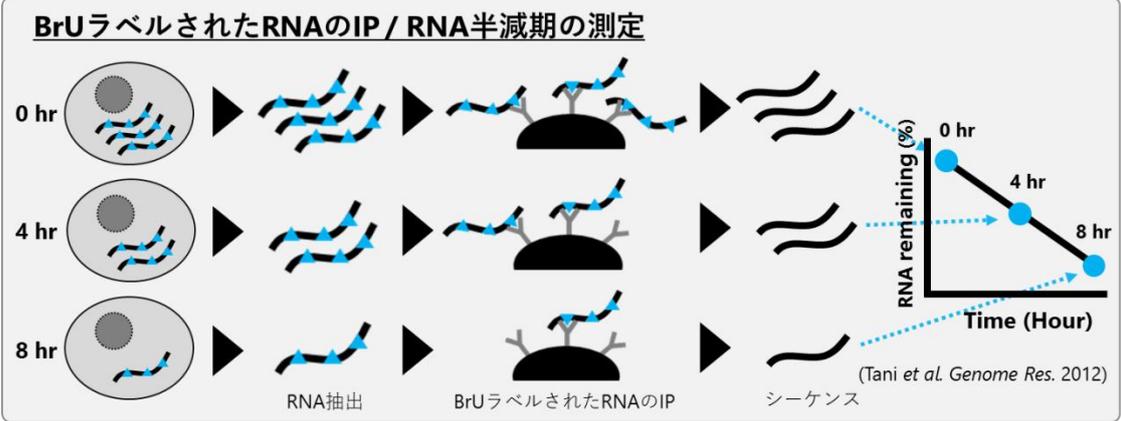
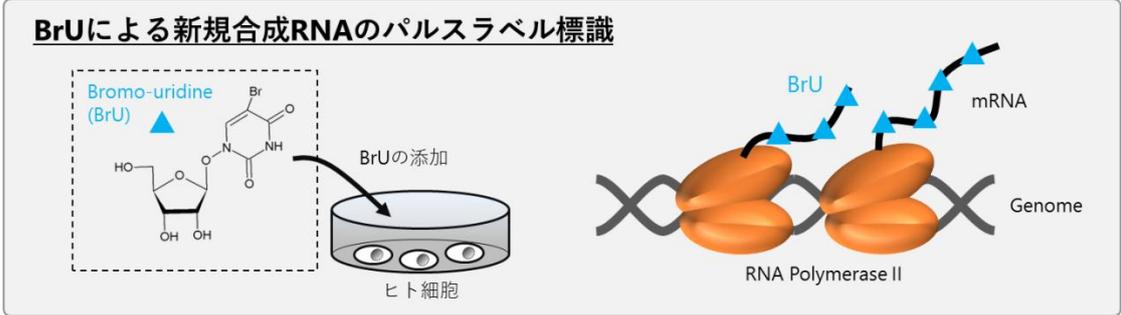


図 0-D. BRIC-seq 法

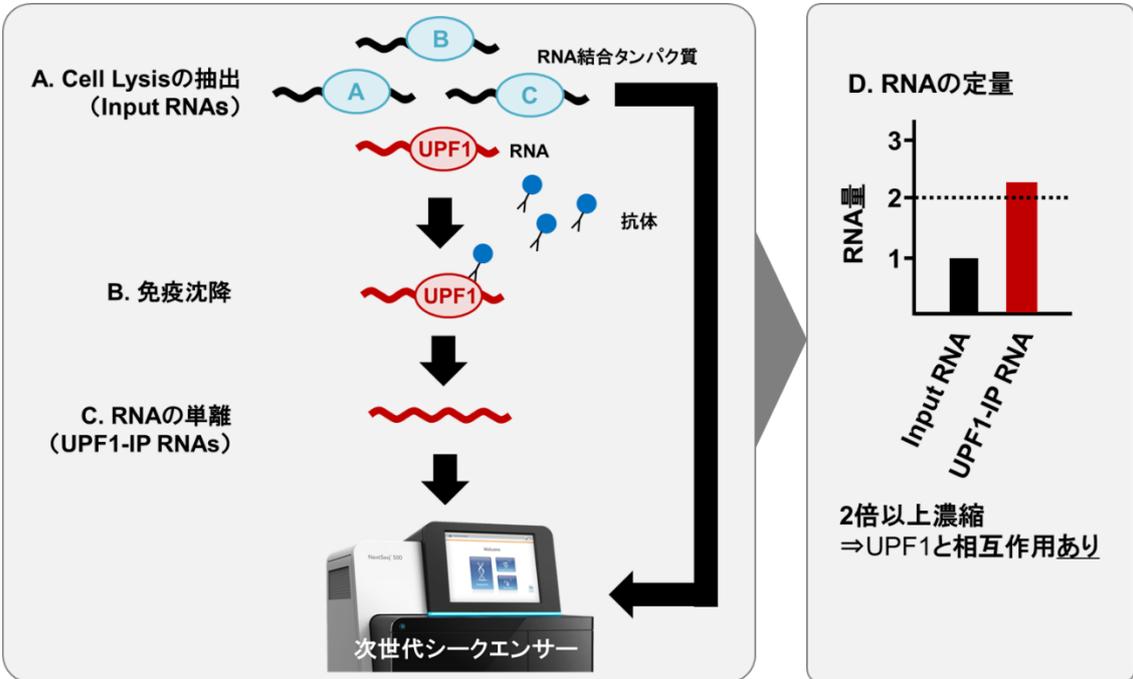


図 0-E. RIP-seq 法

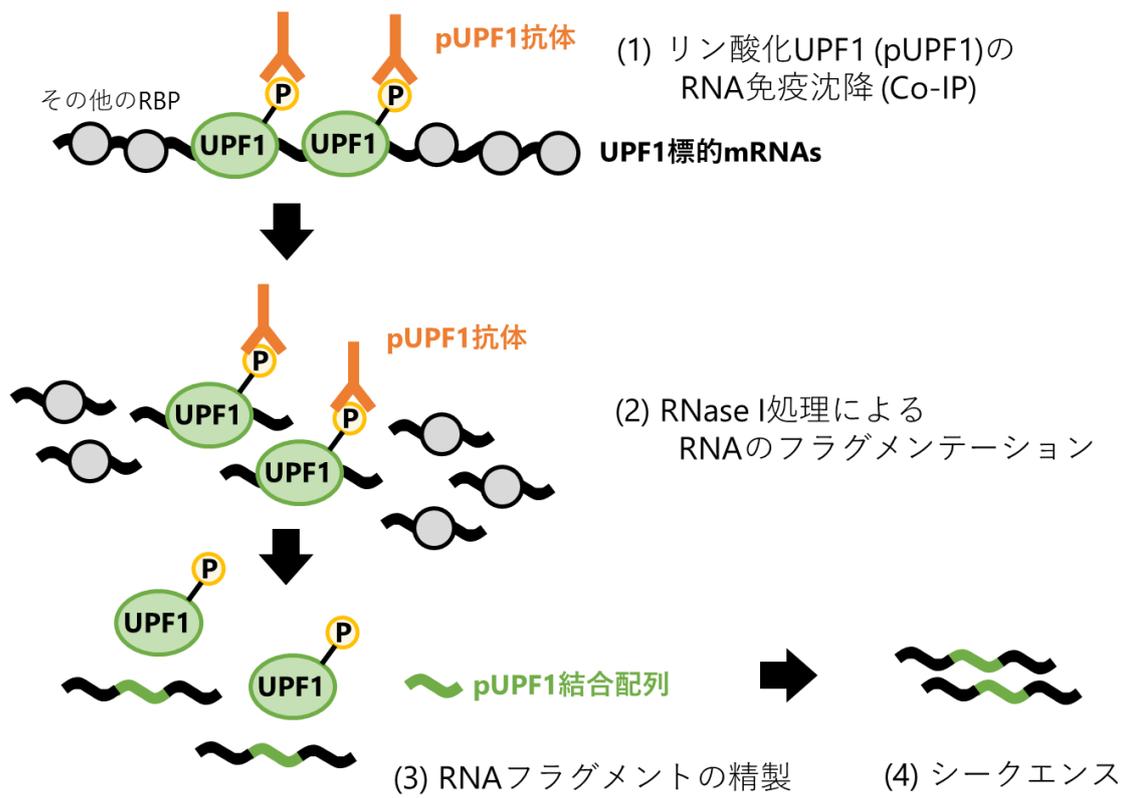


図 0-F. 次世代シーケンサーを用いた網羅的な RNA フットプリント法

## BRIC-seq による、PTC を含む異常な mRNA の分解促進の検出

選択的スプライシングなどの転写後制御を介して、多くの遺伝子では複数の isoform がコードされる。Isoform の中には、PTCが生じるために NMD 経路によって速やかに分解されるものが存在する。そのため、PTC を持つ異常な isoform を分解するためには、当然 UPF1 が必須となる。そこでまず、BRIC-seq 法で UPF1 の分解標的 RNA を検出できることを確認するため、PTC を含む異常な isoform mRNA の分解促進を検出できるか否かを検証した。方法としては、コントロールと UPF1 ノックダウン条件下における、PTC を含む異常な mRNA (isoform) の RNA 半減期をそれぞれ実測した。その結果、TUBA1B、PRMT1、RPS9 遺伝子について、UPF1 ノックダウンにより PTC を含む異常な isoform の RNA 半減期が延長することがわかった (図 1A)。

近年、GENCODE プロジェクトにより次世代シーケンサーから得られたデータを用いて PTC を含む異常な isoform (mRNA) の同定が行われてきた。そこで、GENCODE プロジェクトから得られた PTC を含む異常な mRNA のアノテーション情報をもとに、BRIC-seq のデータから PTC を含む異常な mRNA、つまりは UPF1 の標的 RNA を網羅的に抽出することができるかどうか評価を行った。定常状態時では、PTC を含む isoform の RNA 半減期 ( $T_{1/2} \sim 3.9\text{h}$ ) は PTC を含まない isoform の RNA 半減期 ( $T_{1/2} \sim 7.2\text{h}$ ) よりも短い傾向にあることがわかった。しかしながら、UPF1 をノックダウンしたときでは、PTC を含む異常な isoform の RNA 半減期の分布がより長い半減期へとシフトすることがわかった (図 1B)。次に、NMD 経路の重要な因子である UPF1、UPF2、SMG1 をそれぞれノックダウンした条件下において、PTC を含む異常な isoform と PTC を含まない正常な isoform の

発現量の変化を測定した。その結果、PTCを含む異常な isoform ではノックダウンにより RNA の発現量が上昇したのに対し、PTC を含まない正常な isoform では RNA の発現量に変化はなかった（[図 2A-D](#)）。以上の結果から、PTC を含まない正常な mRNA に比べて、PTC を含む異常な mRNA の分解が速いことを BRIC-seq で検出できると判断した。すなわち、UPF1 の標的となる異常 mRNA（PTC を有する mRNA）を BRIC 法で検出することが可能であると判断した。

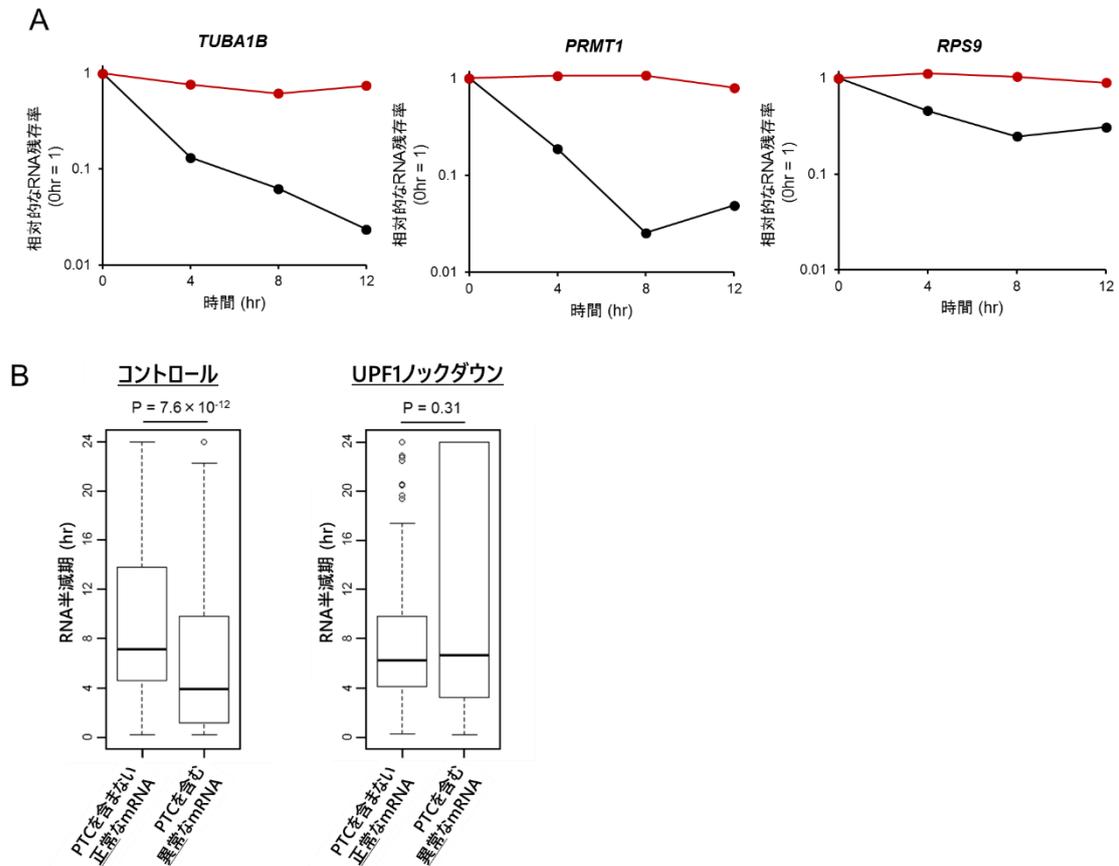


図 1. PTC を含む異常な mRNA と PTC を含まない正常な mRNA の分解速度の比較

(A) コントロール (黒線) と UPF1 ノックダウン (赤線) 条件下における、PTC を含む異常な isoform の分解速度の比較。RNA 分解速度については、BRIC-seq を用いて実測した。

(B) コントロール (黒線) と UPF1 ノックダウン (赤線) 条件下における、PTC を含む異常な isoforms (N = 292) と PTC を含まない正常な isoforms (N = 292) に対する RNA 半減期の分布。P-value は Wilcoxon 順位和検定により計算した。コントロール条件下における、PTC を含む異常な isoforms と PTC を含まない正常な isoforms の RNA 半減期の平均はそれぞれ 7.2h、3.9h だった。UPF1 ノックダウン条件下では、それぞれ 6.2h、6.8h だった。

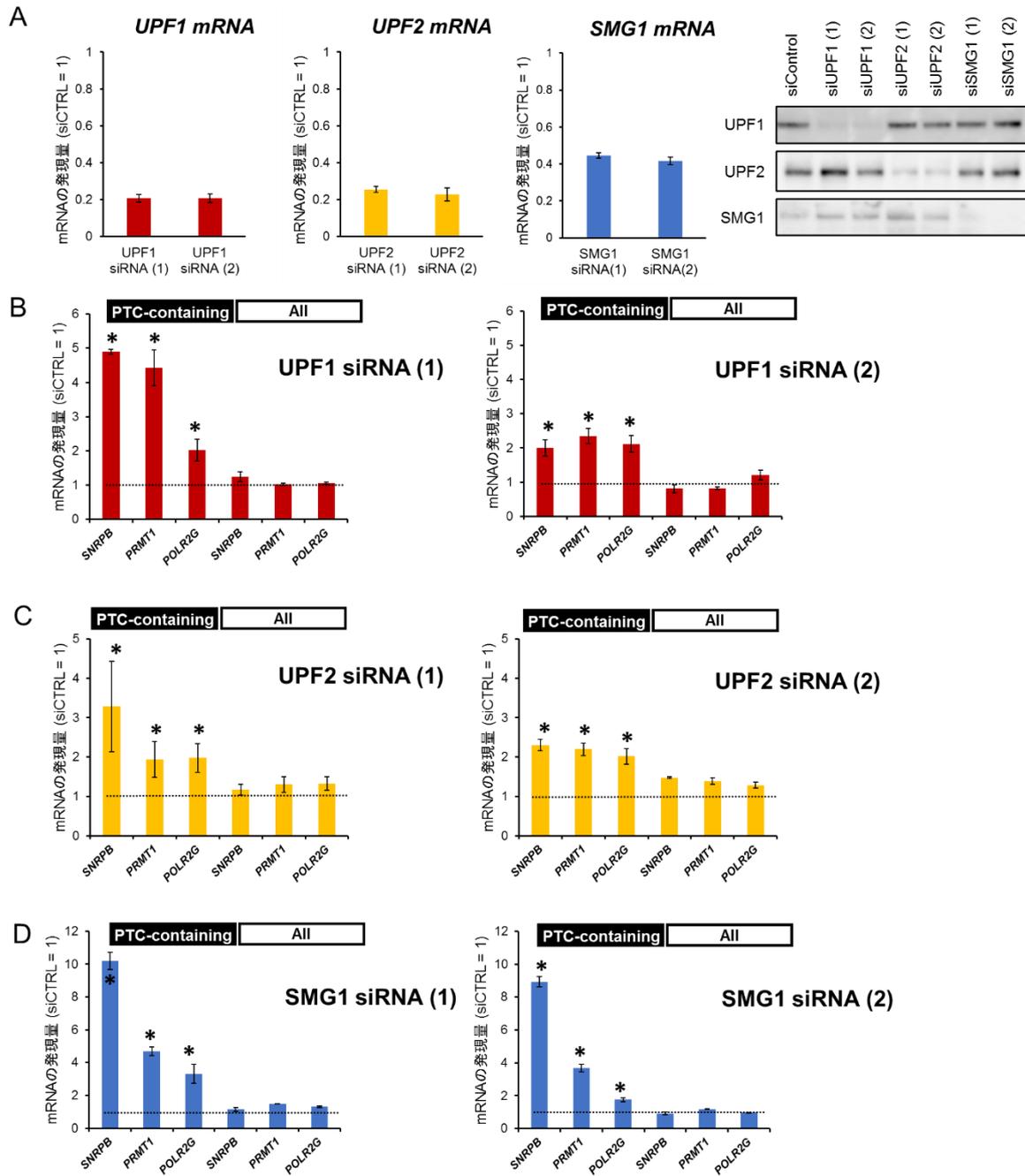


図 2. UPF1、UPF2、SMG1 ノックダウン時における PTC を含む異常な mRNA と PTC を含まない

正常な mRNA の発現量の変化

(A) UPF1、UPF2、SMG1 のノックダウン効率。RNA の発現量の変化は qRT-PCR により定量し

確認した。発現量のノーマライズには GAPDH mRNA を用いた。エラーバーは 3 回の実験間での誤差

を示している。タンパク質の発現量の変化はウエスタンブロットにより確認した。UPF1、UPF2、SMG1 はそれぞれ 2 種類の異なる siRNA を用いてノックダウンされた。

(B-D) UPF1、UPF2、SMG1 のノックダウン時における、PTC を含む異常な mRNA と PTC を含まない正常な mRNA の発現量の変化。RNA の発現量の変化は qRT-PCR により定量し確認した。発現量のノーマライズには GAPDH mRNA を用いた。エラーバーは 3 回の実験間での誤差を示している。P-value は Student's t-test 検定により計算した。

## UPF1 標的 mRNA の同定

先の研究によって、UPF1 の標的となる異常 mRNA を BRIC 法で検出できることが分かったので、次に、BRIC 法を用いて、UPF1 によって分解される正常 mRNA（この場合、PTC を含まない正常 mRNA）の同定を行った。すなわち、BRIC-seq を用いて HeLa 細胞で UPF1 ノックダウンしたときの RNA 半減期の変化を測定し、UPF1 ノックダウンによって安定化する mRNA の同定を行った。BRIC-seq について概説すると、BrU でパルスラベルした RNA の減少量を、次世代シーケンサーにより経時的（0h、4h、8h、12h）に定量することで RNA の分解速度・半減期を計算した。結果として、8,426 遺伝子の RNA 半減期を実測することができた。コントロールと比較して、UPF1 ノックダウン時に 619 遺伝子の mRNA について、RNA 半減期が 2 倍以上延長していることが確認された（[図 3A, B](#)）。

UPF1 標的 mRNA ならば、UPF1 と相互作用することが期待される。そこで次に、HeLa 細胞に HA タグを有する UPF1 を発現させ、HA に対する抗体で免疫沈降を行い UPF1 と共沈降する RNA を次世代シーケンサーを利用して網羅的に同定した。結果として、インプットと比較して、共沈降したサンプルで 2 倍以上エンリッチした mRNA を 2,426 遺伝子の mRNA を UPF1 と相互作用のある mRNA として同定した。

RIP-seq 法は UPF1 と間接的に結合する RNA も同定される。そこで、RIP-seq により得られた UPF1 と相互作用のある mRNA が UPF1 と直接結合しているかどうか CLIP 法を用いて確認した（[図 3A, B](#)）。CLIP 法とは、UV クロスリンクにより UPF1 と直接結合する RNA と UPF1 の間に架

橋反応を起こした上で、UPF1 を免疫沈降で回収するときに同時回収される RNA を同定することで、UPF1 と RNA との直接結合を評価する手法である (図 4)。結果として、UPF1 標的 mRNA として同定された GADD45B、PEA15、ABHD4、DNAJB2、DDIT3、SMG5 mRNA は、インプットと比較して共沈降したサンプルで 2 倍以上エンリッチしていることから、これらの mRNA は UPF1 と直接結合していると判断した (図 4)。

最後に、BRIC-seq と RIP-seq の結果を比較し、UPF1 ノックダウンにより RNA 半減期が 2 倍以上延長し、UPF1 と相互作用のある mRNA を 264 遺伝子同定した (図 3A, B)。これらの mRNA を UPF1 の標的 mRNA として以下の解析に用いた。

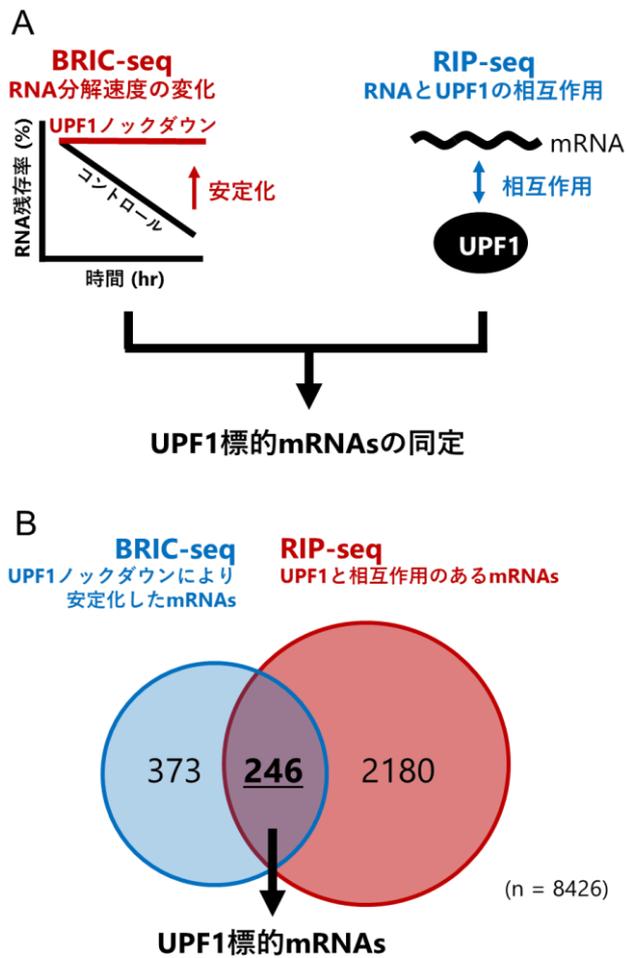


図3. UPF1 標的 mRNA の同定

(A) RNA 分解速度の変化 (BRIC-seq) と RNA とタンパク質の相互作用 (RIP-seq) を測定する複合解析。

(B) BRIC-seq と RIP-seq の結果を比較したベン図。UPF1 ノックダウンにより RNA 半減期が2倍以上延長した mRNA 群 (BRIC-seq, n = 619)。HA-UPF1 の共沈降により2倍以上エンリッチした mRNA 群 (RIP-seq, n = 2,444)。ベン図の共通項を UPF1 の標的 mRNA として同定した (n = 246)。

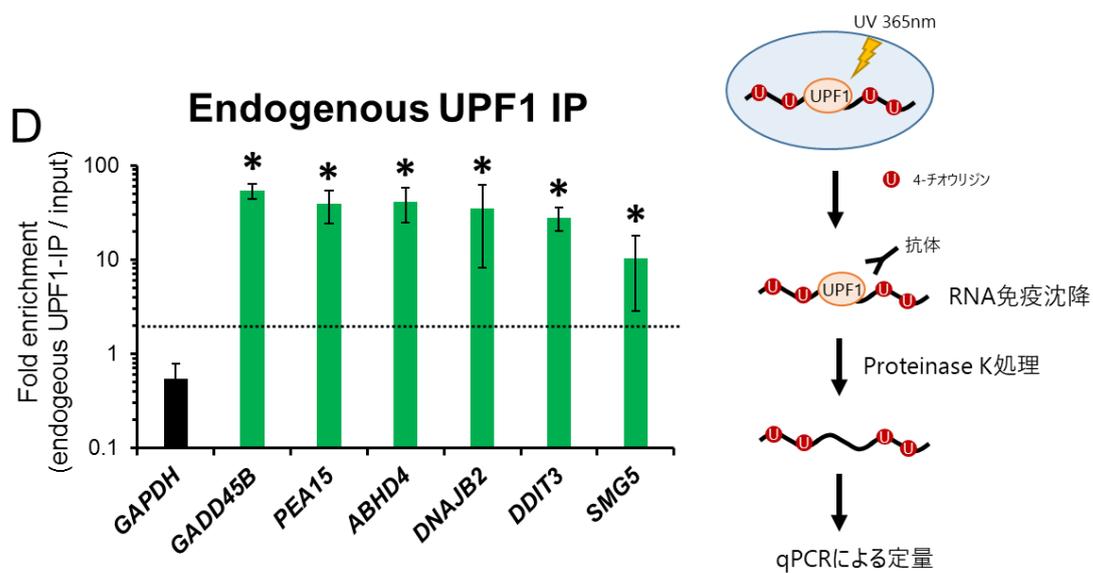


図4. UPF1 タンパク質と UPF1 標的 mRNA との直接的な結合

In vivo cross-linking (CLIP) 法により、UPF1 タンパク質と UPF1 標的 mRNA との直接的な結合を確認した。インプット RNA 量 (Total RNA 中に含まれる特定の遺伝子の mRNA の量) と比較したときの、内在性の UPF1 と共沈降したサンプル中に含まれる RNA 量の Fold enrichment を qRT-PCR により測定した。発現量のノーマライズには PGK1 mRNA を用いた。エラーバーは 3 回の実験間での誤差を示している。P-value は Student's t-test 検定により計算した。

## UPF1 標的 mRNA の中に転写因子や細胞周期の制御に関わる遺伝子がエンリッチされていた

UPF1 標的 mRNA の生理的な機能について評価するために、遺伝子オントロジー解析を行った。その結果、UPF1 標的 mRNA の中に転写因子 (GO: 0006350, transcription P =  $6.7 \times 10^{-6}$ , Benjamini-adjusted P =  $8.5 \times 10^{-3}$ ) や細胞周期の制御 (GO: 00511726, regulation of cell cycle P =  $3.4 \times 10^{-5}$ , Benjamini-adjusted P =  $2.1 \times 10^{-2}$ ) に関わる遺伝子がエンリッチされていることがわかった。先行研究において、UPF1 をノックダウンした細胞において、ATR 依存的な DNA 障害への応答を介して、S 期のアレストが誘導されることが報告されている (Azzalin and Lingner 2006)。今回の私の研究結果から、UPF1 ノックダウンによる S 期のアレストが、細胞周期の制御に関わる UPF1 の標的遺伝子の発現上昇によって引き起こされた結果であると予測される。

## STAU1 タンパク質によって一部の UPF1 標的 mRNA は制御されている

先行研究において、STAU1 が UPF1 を介した正常な mRNA の分解の基質認識に関与しているという報告がある (Kim et al. 2005)。しかしながら、UPF1 の制御下にある標的 mRNA のうち、STAU1 によって制御されている mRNA がどの程度含まれているかはわかっていなかった。そこで、STAU1 の標的となっている UPF1 標的 mRNA を同定するために、BRIC-seq を用いて HeLa 細胞で STAU1 ノックダウンしたときの RNA 半減期の変化を測定した。その結果、UPF1 標的 mRNA の 246 遺伝子のうち、STAU1 ノックダウンによって RNA の半減期が 2 倍以上延長する mRNA として 15 遺伝子 (全体の 6%) が同定された (図 5)。以上の結果から、UPF1 標的 mRNA のうち、STAU1 の標的 mRNA である遺伝子はごく少数であることがわかった。これはつまり、UPF1 を介した正常な mRNA の分解における基質認識に関して、STAU1 タンパク質の寄与は大きくないことを示唆している。すなわち、今回同定した UPF1 標的 RNA の分解は既知の SMD とは異なる経路で分解されていると考えた。

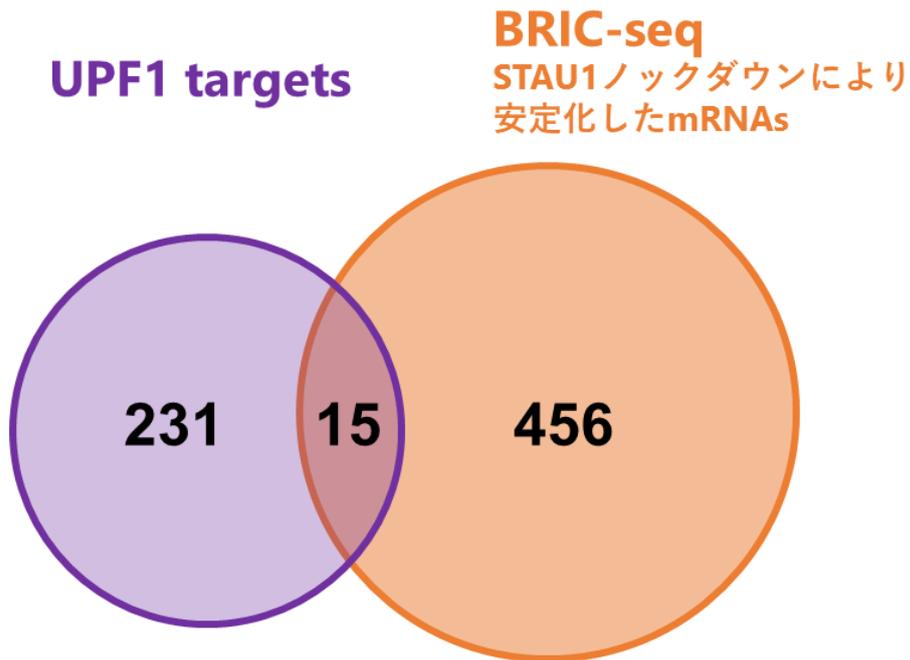


図5. UPF1を介した正常な mRNA 分解に対する STAU1 タンパク質の寄与

UPF1 標的 mRNA と STAU1 ノックダウンでの BRIC-seq データの比較。UPF1 標的 mRNA (n = 246)。STAU1 ノックダウンにより RNA 半減期が 2 倍以上延長した mRNA 群 (BRIC-seq, n = 471)。ベン図の共通項は、UPF1 標的 mRNA のうち、STAU1 の標的 mRNA である遺伝子を指している (n = 15, 全体の 6%)。

### UPF1、UPF2、SMG1 ノックダウンにより、UPF1 標的 mRNA の発現量は上昇する

NMD 経路において重要な遺伝子が、UPF1 を介した正常な mRNA の分解にも必要な因子であるかどうか確認するために、UPF1、UPF2、SMG1 をそれぞれノックダウンしたときに、今回同定した UPF1 標的 mRNA の発現量が変化するかどうか測定した。その結果、これらの因子をノックダウンした条件下において、確認した 6 種類すべての UPF1 標的 mRNA で発現量の増加が確認された（[図 6](#)）。この結果は、NMD 経路に重要な因子である UPF2 や SMG1 が正常な mRNA 分解においても関与していることを示唆している。

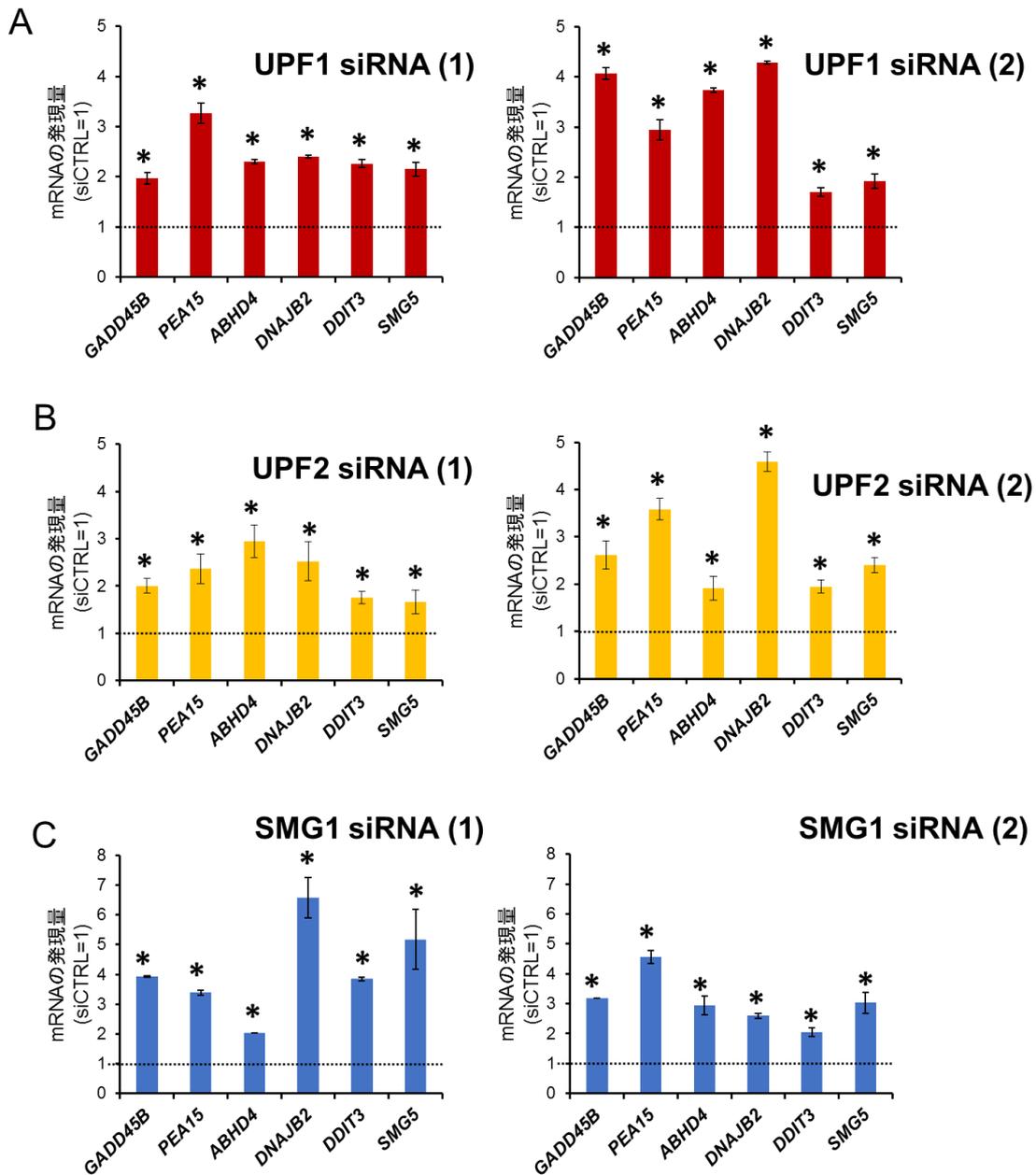


図 6. UPF1、UPF2、SMG1 ノックダウン時における UPF1 標的 mRNA の発現量の変化

(A-C) UPF1、UPF2、SMG1 のノックダウン時における、UPF1 標的 mRNA の発現量の変化。

RNA の発現量の変化は qRT-PCR により定量し確認した。発現量のノーマライズには GAPDH mRNA を用いた。エラーバーは 3 回の実験間での誤差を示している。P-value は Student's t-test 検定により計算した。

## UPF1 標的 mRNA の 3'UTR 中に含まれるリン酸化 UPF1 の結合配列の予測

UPF1 のリン酸化は PTC を含む mRNA を排除する NMD 経路において、RNA 分解を誘導するための重要なステップである (Schweingruber et al. 2013) (図 7)。リン酸化 UPF1 は mRNA の 3'UTR 上に結合する傾向があり (Kurosaki et al. 2014)、SMG6 や SMG5-SMG7 のリクルートを介して RNA 分解が誘導されることが知られている。そのため、mRNA 上に結合するリン酸化 UPF1 は、UPF1 によって制御されている mRNA を同定するためのマーカーとして使えると考えられる。リン酸化 UPF1 が、先に同定した UPF1 標的 mRNA 上に結合しているかどうか確認するために、過去の研究で得られたリン酸化 UPF1 に対する RNA フットプリントデータ (Kurosaki et al. 2014) を NCBI GEO から取得し解析を行った。この解析には、UPF1 の標的 mRNA を同定する際に解析対象とした 8,426 遺伝子のうち、RNA フットプリントのデータから定量できる遺伝子に絞って解析を行った。結果として、5,409 遺伝子を解析対象とする全遺伝子とし、そのうち今回の解析で同定した UPF1 標的遺伝子は 160 遺伝子含まれている。まず、UPF1 標的 mRNA が他の mRNA と比較してリン酸化 UPF1 と結合している割合が高いことがわかった (図 8A)。加えて、UPF1 標的 mRNA の 3'UTR 上にリン酸化 UPF1 が結合している傾向にあることがわかった (図 8B、C)。また、UPF1 標的 mRNA の 3'UTR 上でのリン酸化 UPF1 の結合領域では、グアニン (G) とシトシン (C) の含有率が高いことがわかった (図 8D)。さらに、MEME (Bailey et al. 2006) を用いて UPF1 標的 mRNA 配列中に存在するリン酸化 UPF1 の結合サイトを予測した。その結果、3'UTR 中に CUG を中心とした GC-rich なモチーフ配列が予測された (図 8E)。以上の結果から、3'UTR 中に存在する GC-

rich なモチーフ配列が、UPF1 の基質認識に関わる特徴配列であると考えた。

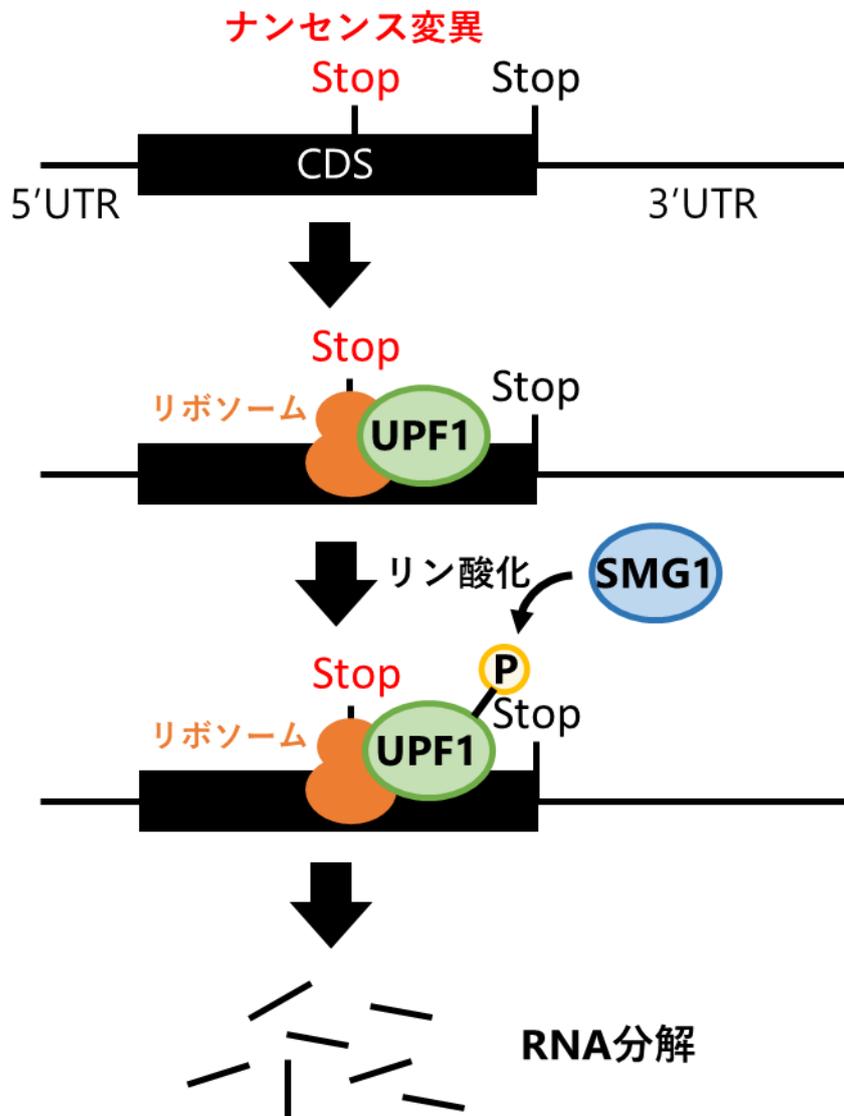


図 7. UPF1 のリン酸化による PTC を含む異常な mRNA の分解 (NMD 経路)

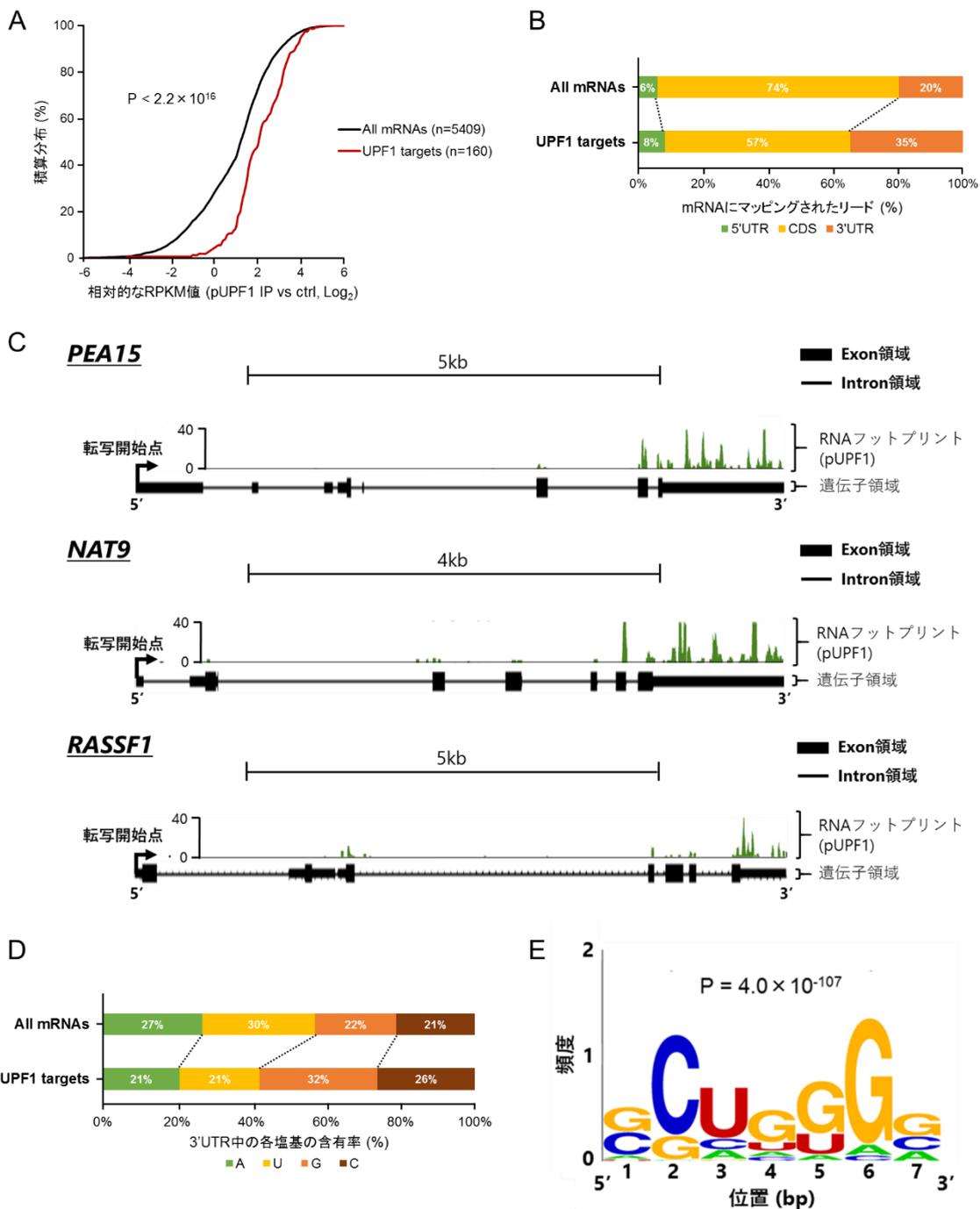


図 8. UPF1 標的 mRNA の 3'UTR 中に存在するリン酸化 UPF1 の結合配列の予測

(A) リン酸化 UPF1 と mRNA との相互作用の程度 (インプットと比較して、リン酸化 UPF1 で共沈降した RNA サンプル中での mRNA の濃縮の程度) に関する積算分布。黒線が全 mRNA、

赤線が UPF1 標的 mRNA をそれぞれ指している。P-value は Wilcoxon の順位和検定により計算した。

(B) 全 mRNA と UPF1 標的 mRNA の 5'UTR、ORF、3'UTR 上にそれぞれマッピングされたリード数の割合。

(C) UPF1 標的 mRNA に対する RNA フットプリントデータの例。縦軸はリード数を示している。

(D) mRNA の 3'UTR 上での、リン酸化 UPF1 の結合領域における各塩基の含有率。

(E) MEME を用いて予測されたリン酸化 UPF1 のモチーフ配列。モチーフの長さは 5-8bp に限定して計算した。最も低い p-value ( $P = 4.0 \times 10^{-107}$ ) を持つモチーフを示した。

## UPF1 標的 mRNA 中での配列的特徴

AU-rich エlement (ARE) は mRNA の 3'UTR 中に含まれる RNA 分解モチーフとして知られており、例えば、サイトカインやインターロイキンなどの遺伝子の mRNA の制御に関わっている。この ARE に特定の RNA 結合タンパク質が結合することで、RNA の安定性を制御している。また、ARE を介した RNA 分解経路において、ARE の周辺の RNA 配列の AU 含有量が高いことも基質認識に重要であることが明らかとなっている (Spasic et al. 2011)。このように、RNA 結合タンパク質の基質認識には、特定のモチーフ配列以外にも、周辺の RNA の配列的特徴が重要であることがわかっている。これらの知見をもとに、予測された GC-rich なモチーフ配列以外に UPF1 標的 mRNA の 3'UTR 中どのような配列的な特徴が含まれているか解析を行った。その結果、UPF1 標的 mRNA の 3'UTR の GC 含有率が、他の mRNA と比較して 8-12% 高いことがわかった (図 9A)。また、GC 含有率が高い 3'UTR を持つ mRNA 群は、他の mRNA と比較して、UPF1 ノックダウンによる RNA 半減期の延長の程度が大きいこと (図 10A)、UPF1 およびリン酸化 UPF1 との相互作用も強い傾向にあることがわかった (図 10B、C)。次に、UPF1 標的 mRNA の 3'UTR の GC 含有率がマウスでも高い傾向にあるかどうか、つまり、GC 含有率の高さが種間で保存されているかどうか確認した。マウスとヒト間での各 mRNA の GC 含有率を比較するために、RefSeq データベースに収載されているオーソログな遺伝子の mRNA のペアを用意した (Fong et al. 2013)。解析の結果、ヒトの UPF1 標的遺伝子に対する、マウスのオーソログな遺伝子の mRNA の 3'UTR 配列も、ヒトと同様に GC 含有率が高いことがわかった (図 9B)。以上の結果から、UPF1 標的 mRNA の 3'UTR 中の GC 含有率

が高いことも、UPF1 を介した RNA 分解の基質認識に関わる重要な特徴の 1 つであると考えられる。

一方で、3'UTR 中の GC 含有率が高いにも関わらず、UPF1 の標的とならない mRNA も一定数存在する。そのため、GC 含有率以外にも UPF1 標的 mRNA であるかどうかを決める要素があると考えた。そこで、mRNA の 3'UTR 中のモノヌクレオチドおよびジヌクレオチドの含有率に着目して解析を行った。まず、mRNA の 3'UTR の GC 含有率が 55% (UPF1 標的 mRNA の GC 含有率の中央値) 以上である mRNA 群を抽出し、UPF1 標的 mRNA 群とその他の mRNA 群に分類分けした。それぞれの RNA 群における 3'UTR 中でのモノヌクレオチドおよびジヌクレオチドの含有率を調べた結果、グアニン (G) の含有率と GA もしくは AG のジヌクレオチドの含有率が UPF1 標的 mRNA では高いことがわかった。一方で、シトシン (C) の含有率と CC のジヌクレオチドの含有率が UPF1 標的 mRNA では低いことがわかった (図 11A、B)。以上の結果から、3'UTR 中での GC 含有率の高さに加えて、グアニン (G) の含有率と GA・AG のジヌクレオチドの含有率が高いことが、UPF1 を介した RNA 分解の基質認識に関わる重要な特徴であると考えられる。

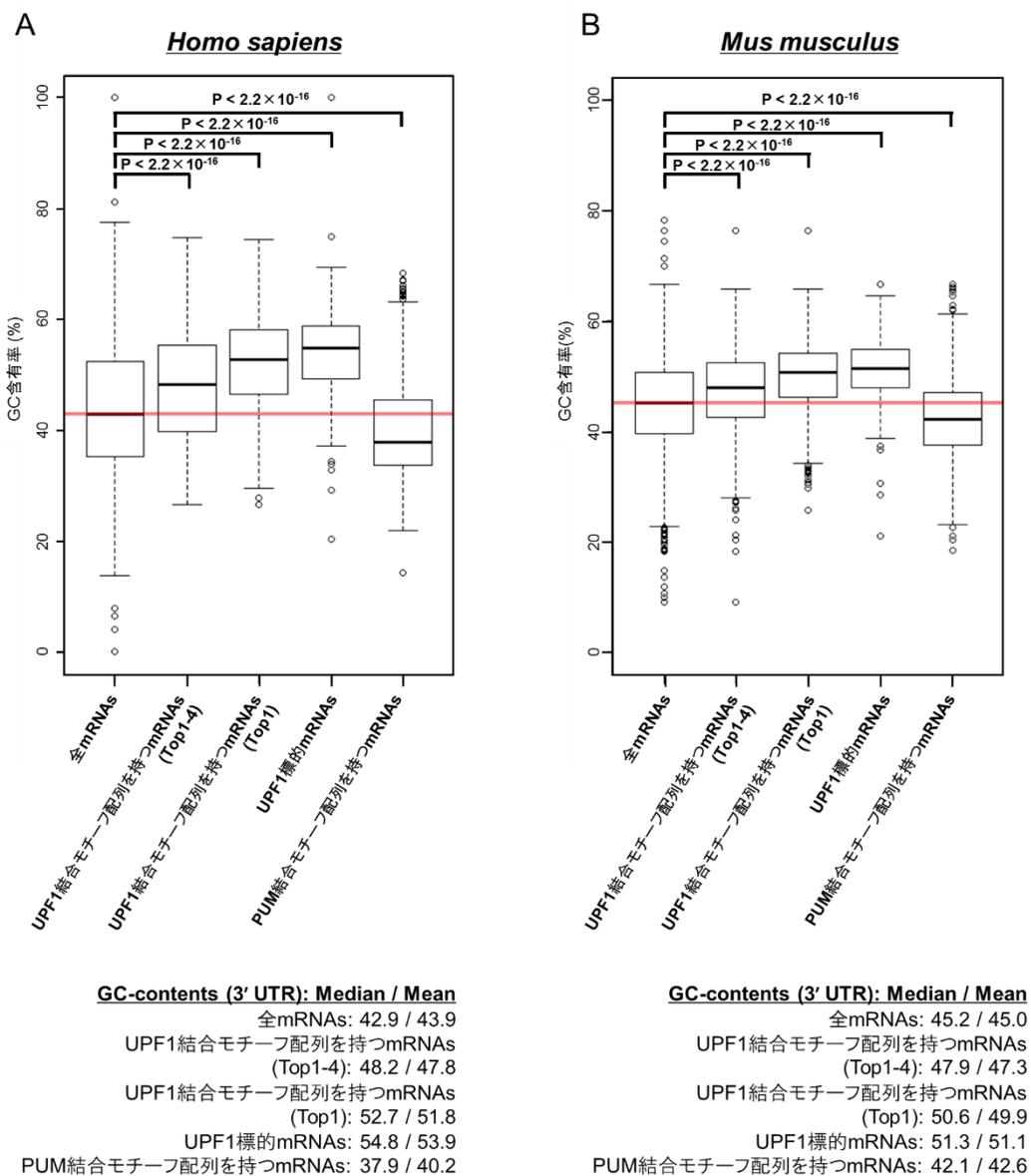


図9. UPF1 標的 mRNA の 3'UTR 中の GC 含有率

(A, B) Homo sapiens (A) と Mus musculus (B) について、全 mRNA、UPF1 結合モチーフ配列 (Top1-4: CCUGGGG, CCUGGGA, CCUGGAA, CCUGAGA) を持つ mRNAs、UPF1 結合モチーフ配列 (Top1: CCUGGGG)、UPF1 標的 mRNAs、PUM 結合モチーフ配列を持つ mRNAs の 3'UTR 中の GC 含有率。P-value は Wilcoxon の順位和検定により計算した。赤線は全 mRNAs の GC 含有率の中央値を示している。

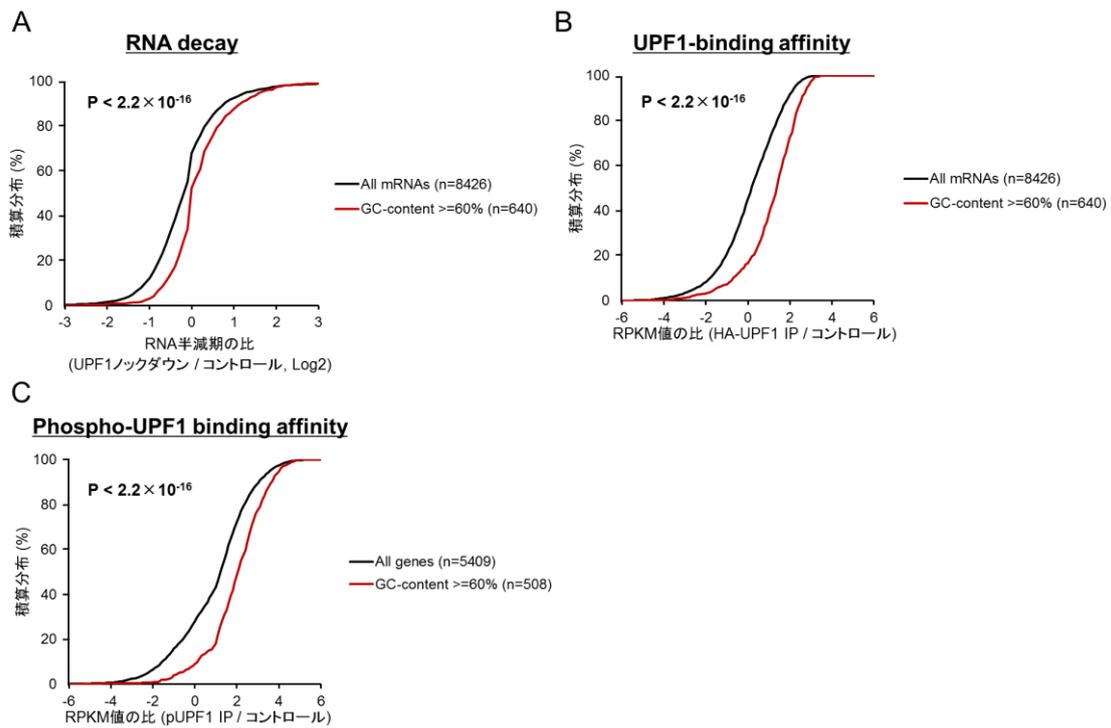


図 10. 3'UTR 中の GC 含有率と UPF1 を介した RNA 分解及び UPF1 との相互作用の関係

(A-C) UPF1 ノックダウンによる RNA 半減期の変化 (A)、UPF1 (B) 及びリン酸化 UPF1 (C) の RNA に対する共沈降の効率 (インプットと比較したときの、UPF1 もしくはリン酸化 UPF1 で IP した RNA サンプルでの各 mRNA の濃縮の程度) に関する積算分布。

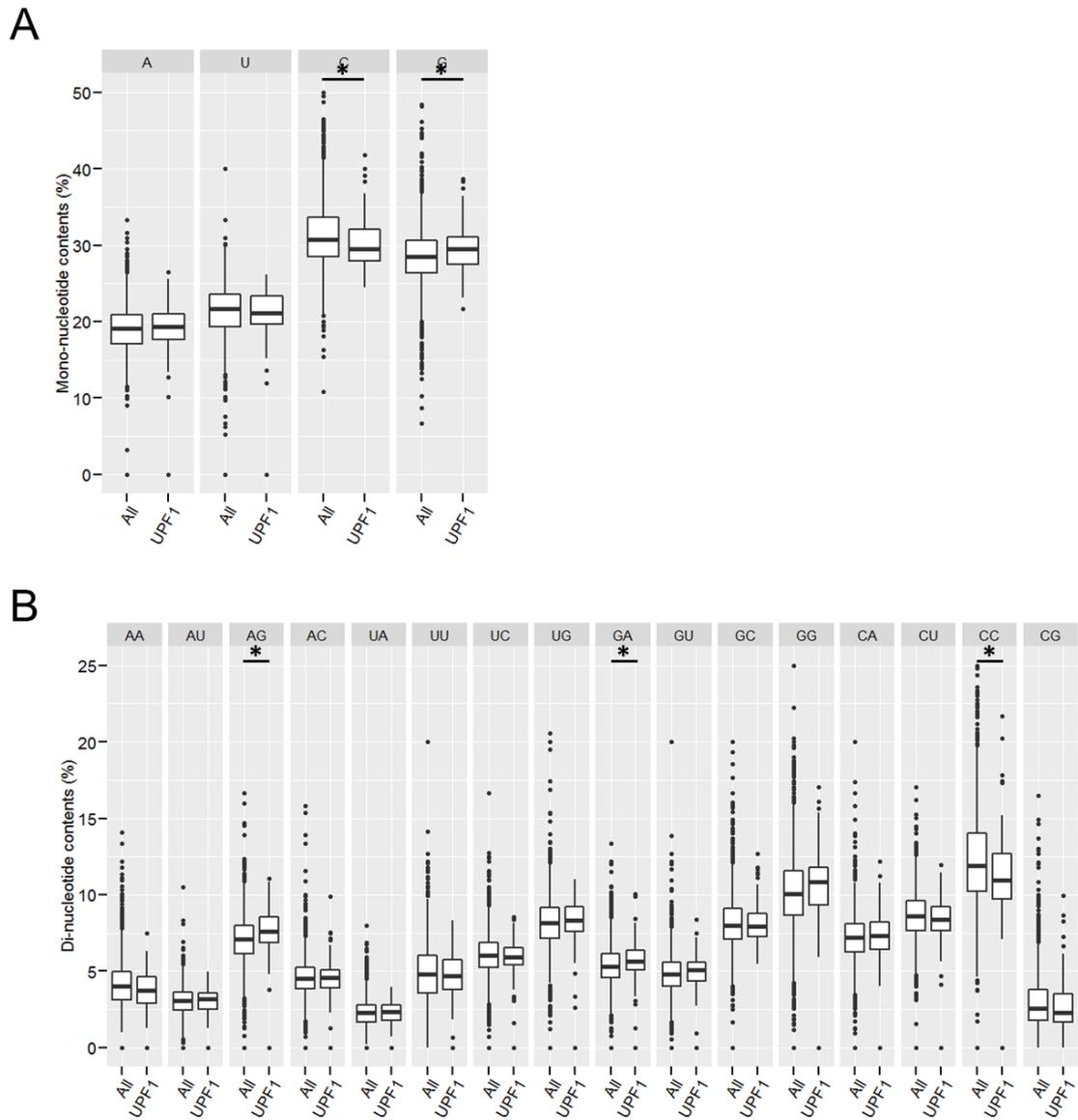


図 11. UPF1 標的 mRNA の 3'UTR 中のモノヌクレオチド及びジヌクレオチドの含有率

(A) UPF1 標的 mRNA とその他の mRNA の 3'UTR 中における、モノヌクレオチド (A, U, C, G) とジヌクレオチド (AA, AU, AG, AC, UA, UU, UC, UG, GA, GU, GC, GG, CA, CU, CC, CG) の含有率。P-value は Wilcoxon の順位和検定によって計算した。

## 転移因子 (SINE、LINE、LTR) の配列は UPF1 標的 mRNA の 3'UTR 中にエンリッチしていない

mRNA の 3'UTR 中に含まれる long terminal repeat (LTR) 、 long interspersed element (LINE) 、 Short Interspersed Element (SINE) のような転移因子の配列が RNA の発現制御に重要な役割を果たしていることが近年明らかとなっている (Feschotte 2008) 。SINE に分類される *Alu* エlementは、UPF1 を介した正常な RNA 分解の 1 つである STAU1 依存的な mRNA 分解 (SMD) の標的となる mRNA の 3'UTR 中にふくまれており、*Alu* エlementと相補鎖となる配列を持つ mRNA もしくは long non-coding RNA と 2 本鎖 RNA の構造 (*Alu* エlementを介した mRNA-mRNA もしくは mRNA-lncRNA の分子間での相補対) を形成する。*Alu* エlementを介した RNA の相補対 (RNA の 2 本鎖構造) は STAU1 タンパク質によって認識され、SMD 経路によって分解が誘導されることが知られている (Gong and Maquat 2011; Gong et al. 2013) 。加えて、UPF1 は LINE 配列に対して結合し、LINE 配列を持つ mRNA の発現量を低下させることが知られている (Taylor et al. 2013) 。そこで、先行研究の結果を踏まえて、SINE や LINE などの転移因子の配列を持つ mRNA が UPF1 の分解標的になっているかどうか確認した。その結果、LTR や SINE に分類される mammalian-wide interspersed repeat (MIR) エlementを持つ mRNA が UPF1 の標的 mRNA の中に 4-8%程度しか含まれていないことがわかった (図 12) 。以上の結果から、UPF1 を介した RNA 分解での基質認識に転移因子の配列は大きくは関与していないと判断した。

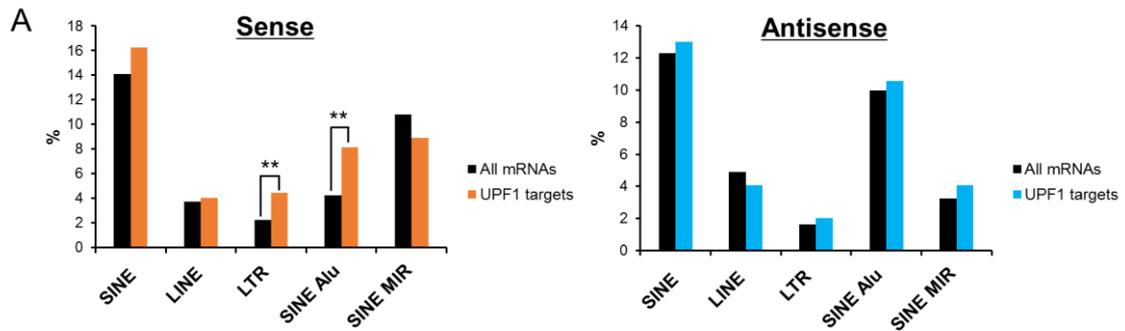


図 12. 3'UTR 中に転移因子の配列を含む UPF1 標的 mRNA の割合

(A) SINE (MIRとAlu)、LINE、LTR、SINE (MIRのみ)、SINE (Aluのみ) を含むUPF1標的 mRNA 群とその他の mRNA 群。左では、それらの転移因子のセンス配列を3'UTR中に持つ割合、右には、転移因子のアンチセンス配列を3'UTR中に持つ割合をそれぞれ示している。mRNAの3'UTR中に含まれる転移因子は RepeatMasker 4.0.3 を用いて同定した。P-value は Fisher's exact test により検定した (\*\*: p-value < 0.001)。

## リン酸化 UPF1 の結合配列を持つ mRNA は UPF1 依存的な RNA 分解に関与する

3'UTR 中に含まれる GC-rich なモチーフ配列が、UPF1 依存的な RNA 分解に関与しているかどうか評価を行った。GC-rich なモチーフ配列を持つ mRNA 群と持たない mRNA について、UPF1 ノックダウン時における RNA 半減期の変動 (BRIC-seq データ)、UPF1 との共沈降の効率 (RIP-seq データ)、リン酸化 UPF1 との共沈降の効率 (リン酸化 UPF1 の RNA フットプリントデータ) に違いが見られるかどうか解析を行った。解析には、リン酸化 UPF1 の結合配列として見つかった上位 4 種類の GC-rich なモチーフ配列を持つ mRNA 群と持たない mRNA 群を利用した。解析の結果、GC-rich なモチーフ配列を持つ mRNA 群では、持たない mRNA 群と比較して、UPF1 ノックダウン時における RNA の安定化の程度が大きいことがわかった (図 13A)。また、UPF1 及びリン酸化 UPF1 の共沈降の効率 (インプットに対する共沈降した RNA サンプルでの濃縮の程度) に関しても GC-rich なモチーフ配列を持つ mRNA のほうが高いことがわかった (図 13C、E)。さらに、GC-rich なモチーフ配列の数が多いほど、それらの影響 (RNA 分解、UPF1 との相互作用) が大きくなることがわかった。一方で、ネガティブコントロールとして、RNA 分解に関わる RNA 結合タンパク質として知られる PUM (Morris et al. 2008) の結合配列を持つ mRNA に対しても同様な解析を行った結果、そのような変化はみられなかった (図 13B、D、F)。

一般的に、RNA 分解に関わるモチーフ配列を持つ mRNA は RNA の半減期が短くなる傾向にある。そこで、リン酸化 UPF1 の結合配列 (GC-rich なモチーフ配列) や PUM の結合配列を持つ mRNA がその他の mRNA と比較して、RNA 半減期が短いかどうか確認した。その結果、リン酸化

UPF1 もしくは PUM の結合配列を持つ mRNA 群は、他の mRNA 群と比較して、RNA の半減期が短いことがわかった (図 14A)。GC-rich なモチーフ配列を持つ mRNA 群では、UPF1 ノックダウンにより、RNA 半減期の分布がコントロールと比較して長くなる傾向があった。一方で、PUM のモチーフ配列を持つ mRNA 群では、RNA 半減期の分布に変化は見られなかった (図 14B)。以上の結果から、予測された GC-rich なモチーフ配列が UPF1 を介した RNA 分解に関与していると考えられる。

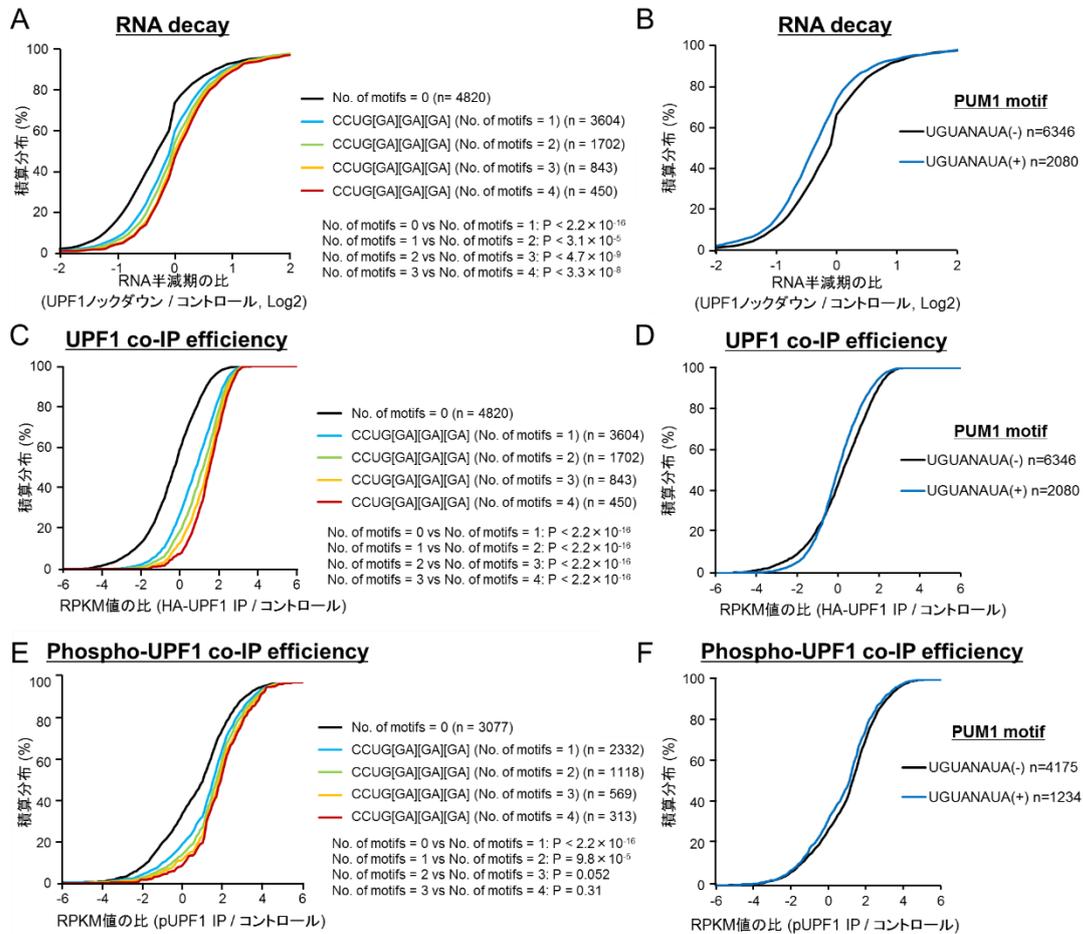


図 13. リン酸化 UPF1 の結合配列を持つ mRNA の UPF1 依存的な RNA 分解への寄与

(A-C) UPF1 ノックダウンによる RNA 半減期の変化 (A)、UPF1 (B) 及びリン酸化 UPF1 (C) の RNA に対する共沈降の効率 (インプットと比較したときの、UPF1 もしくはリン酸化 UPF1 で IP した RNA サンプルでの各 mRNA の濃縮の程度) に関する積算分布。リン酸化 UPF1 の結合配列 (CCUGGGG, CCUGGGA, CCUGGAA, CCUGAGA; 図中では CCUG[GA][GA][GA]と表記) と PUM 結合配列 (UGUANAUA) を持つ mRNA 群と持たない mRNA 群を比較した。P-value は Wilcoxon の順位和検定によって計算した。

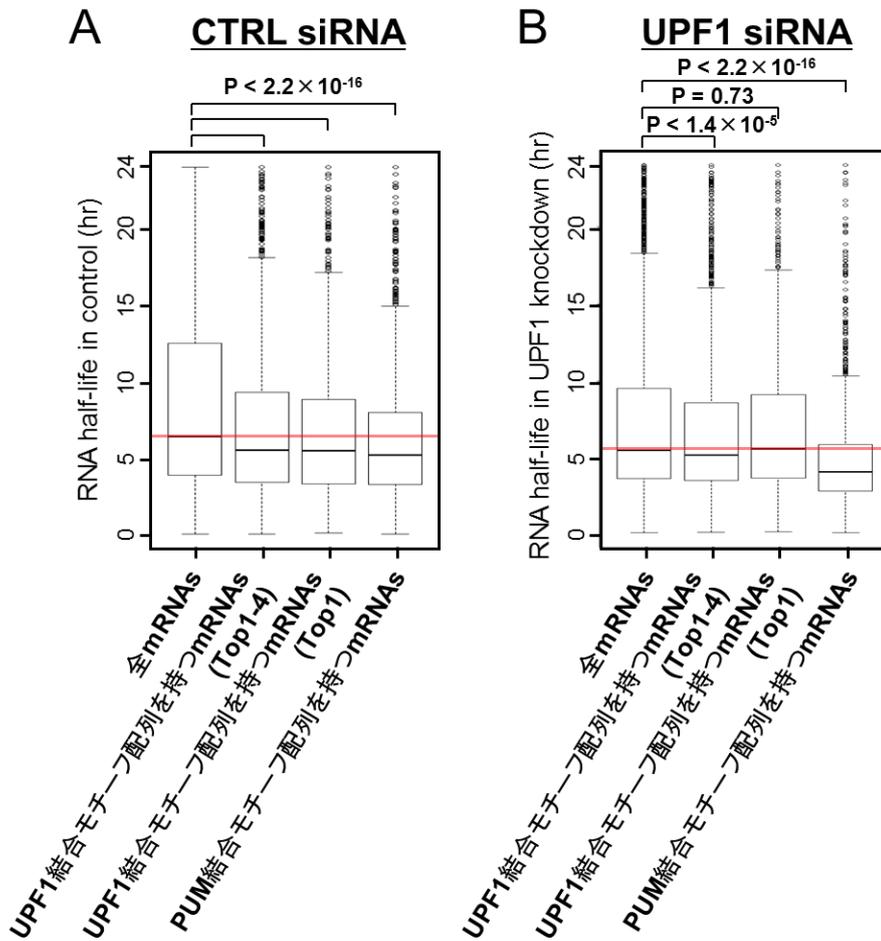


図 14. リン酸化 UPF1 の結合配列を持つ mRNA 群の RNA 半減期の分布

左から、(i) 全 mRNA 群、(ii) リン酸化 UPF1 のモチーフ配列（上位 4 位のモチーフ配列である CCUGGGG、CCUGGGA、CCUGGAA、CCUGAGA）を持つ mRNA 群、(iii) リン酸化 UPF1 のモチーフ配列（上位 1 位のモチーフ配列である CCUGGGG）を持つ mRNA 群、(iv) PUM のモチーフ配列（UGUANAUA）を持つ mRNA 群に対する RNA 半減期の分布。

(A) コントロール条件下における、各 mRNA 群の RNA 半減期の分布。

(B) UPF1 ノックダウン条件下における、各 mRNA 群の RNA 半減期の分布。

P-value は Wilcoxon の順位和検定によって計算した。

## GC-rich なモチーフ配列を介して UPF1 依存的な RNA 分解が誘導される

UPF1 標的 mRNA の 3'UTR 中に含まれている GC-rich なモチーフ配列が、UPF1 依存的な RNA 分解を誘導しているかどうか実験的な検証を行った。方法としては、Tet-off system を利用したレポーターアッセイにより、RNA の分解速度の測定を行った。まず、テトラサイクリンで誘導されるプロモーターを持ち、UPF1 の標的 mRNA である GADD45B の 3'UTR の配列を Rabbit の  $\beta$ -グロビンの ORF の下流に繋いだレポーターベクター (pTET-BBB+GADD45B 3'UTR) を作成した。上記のベクターを用いてレポーターアッセイを行った結果、HeLa tet-off 細胞において、GADD45B の 3'UTR の配列を組み込んだ  $\beta$ -グロビン mRNA は、もとの  $\beta$ -グロビン mRNA と比較して、RNA の分解速度が速いことがわかった (図 15A)。また、この GADD45B の 3'UTR を介した RNA の不安定化は、UPF1 ノックダウンによりキャンセルされることがわかった (図 15A)。さらに、NMD 経路において重要な因子である UPF2 タンパク質と、UPF1 のリン酸化に関わる SMG1 タンパク質をそれぞれノックダウンした条件下においても、GADD45B の 3'UTR を介した RNA の不安定化はキャンセルされた (図 15B、C)。次に、UPF1 を介した RNA 分解は翻訳依存的に起こることが知られているため、シクロヘキシミドによりタンパク質の翻訳を阻害したときに、RNA の分解速度に変化が見られるか確認した。その結果、翻訳阻害により GADD45B の 3'UTR を介した RNA の不安定化はキャンセルされた (図 15E)。一方で、UPF1 を介した正常な mRNA 分解経路の 1 つである SMD 経路に必要な STAU1 タンパク質をノックダウンした場合には、GADD45B の 3'UTR を介した RNA の不安定化はキャンセルされなかった (図 15D)。以上の結果から、UPF1 の標的 mRNA である GADD45B の 3'UTR 配列を介

して、UPF1 を介した RNA 分解が誘導されることが示された。また、SMD 経路ではない別のメカニズムによって RNA 分解が誘導されていることも示唆された。

上述のレポーターアッセイに用いたベクター（pTET-BBB+GADD45B 3'UTR）に使用した、GADD45B 3'UTR の配列中には GC-rich なモチーフ配列が 4 箇所存在することがわかった。そこで次に、それら GC-rich なモチーフ配列を欠いたときに RNA 分解が抑制されるかどうか確かめた（[図 16A](#)）。その結果、GC-rich なモチーフ配列を 1、2 箇所削ったレポーター-RNA では、RNA の分解速度にはほとんど変化は見られなかった。一方で、GC-rich なモチーフ配列を 3、4 箇所削ったレポーター-RNA では、もとのレポーター-RNA と比較して RNA の不安定化がキャンセルされた。また、モチーフ配列を削り 3'UTR が短くなったことによる RNA 分解への影響を考慮して、GC-rich なモチーフ配列をすべてアデニン（A）の配列に置換したレポーター-RNA も用意し、同様の実験を行った。その結果、GC-rich なモチーフ配列を置換したレポーター-RNA でも RNA の不安定がキャンセルされた（[図 16B](#)）。さらに、UPF1 をノックダウンした条件下において、GC-rich なモチーフ配列を置換したレポーター-RNA の分解速度は変化しないことがわかった（[図 16C](#)）。以上の結果から、3'UTR 中に含まれる GC-rich なモチーフ配列を介して、UPF1 依存的な RNA 分解が誘導されていると示唆された。

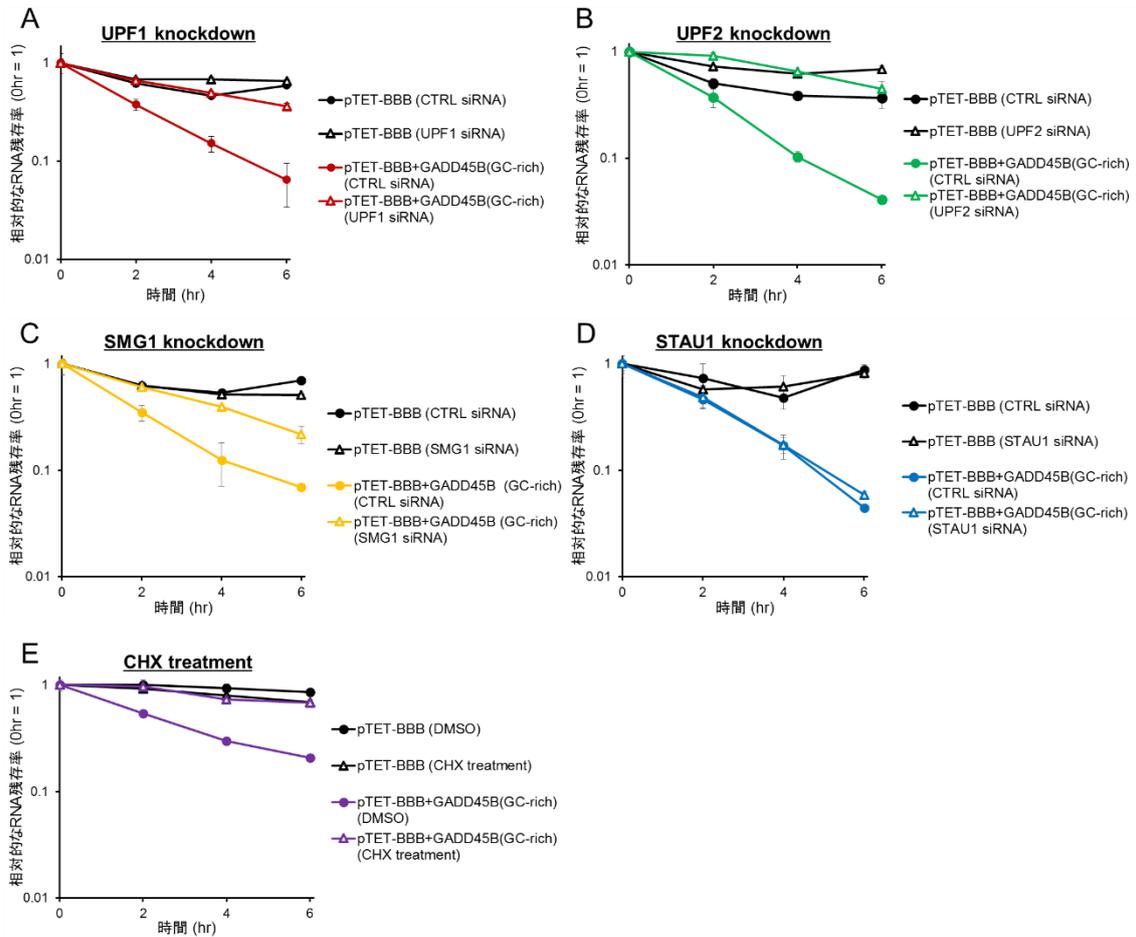


図 15. UPF1 標的 mRNA の 3'UTR 配列を介した UPF1 依存的な RNA 分解

(A-E) HeLa Tet-off 細胞に pTET-BBB もしくは pTET-BBB+GADD45B 3'UTR ベクターをトランスフェクションした。pGL4.13 luciferase 発現ベクターを内部コントロールとして合わせてトランスフェクションした。レポーター遺伝子の転写を止めるために、ドキシサイクリンを培地に添加した。ドキシサイクリン添加後、Total RNA を経時的 (0、2、4、6 時間) にサンプリングした。Luciferase と  $\beta$ -グロビンの mRNA の発現量はそれぞれ qRT-PCR により定量し、0 時間を 1 としたときの相対的な RNA の残存量 (RNA の分解速度) を求めた。各時間における  $\beta$ -グロビンの mRNA の残存量は、Luciferase の mRNA の発現量でノーマライズを行った。pTET-BBB もしくは pTET-BBB+GADD45B

3'UTR ベクターからコードされる  $\beta$ -グロビンの mRNA の分解速度は、UPF1 (A)、UPF2 (B)、SMG1 (C)、STAU1 (D) をそれぞれノックダウンした条件下、もしくはシクロヘキシミドを添加 (E) した条件下で、それぞれコントロールと比較して測定を行った。

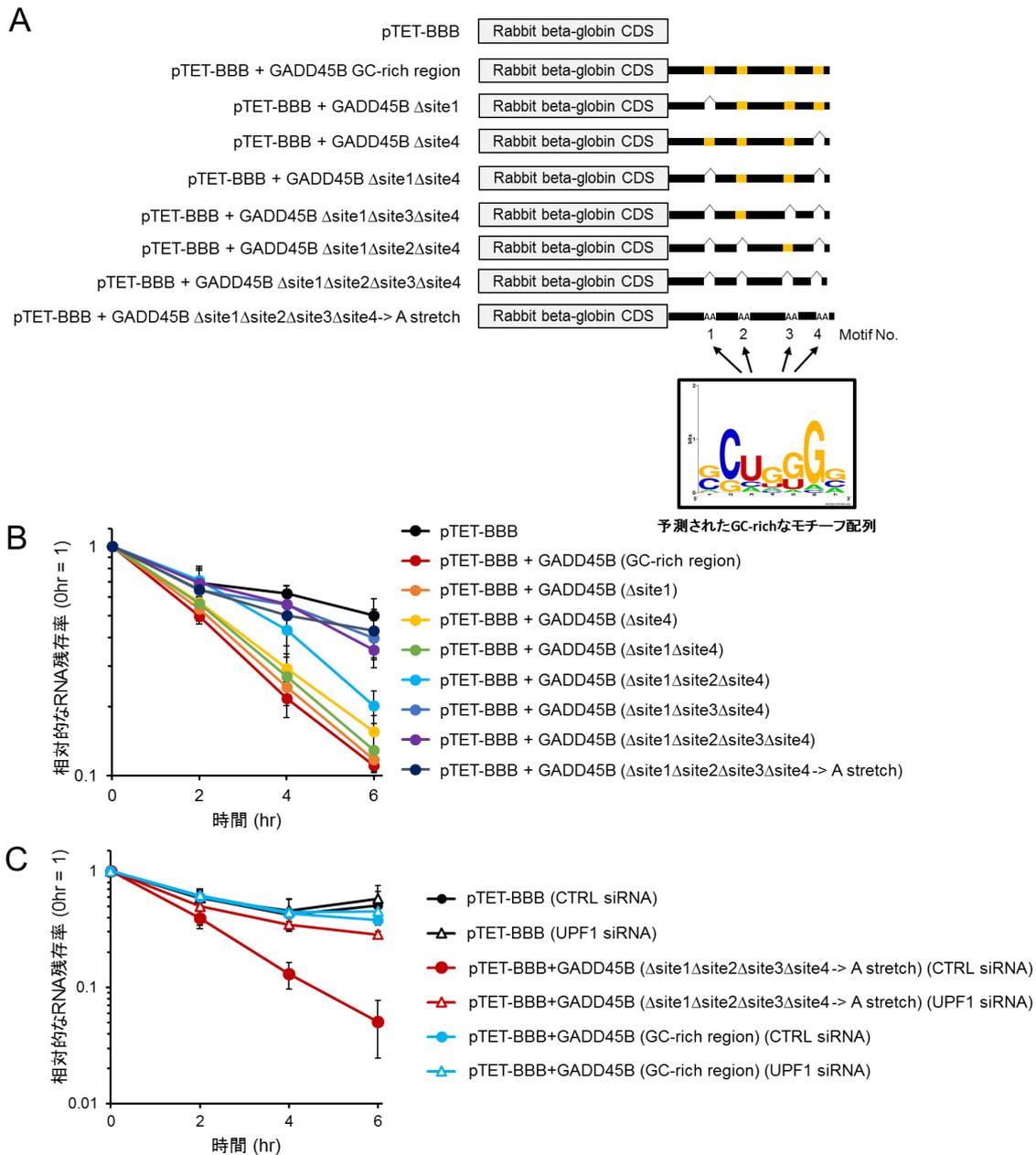


図 16. 3'UTR 中に含まれる GC-rich なモチーフ配列を介した UPF1 依存的な RNA 分解

HeLa Tet-off 細胞に pTET-BBB もしくは pTET-BBB+GADD45B 3'UTR ベクターをトランスフェクションした。pGL4.13 luciferase 発現ベクターを内部コントロールとして合わせてトランスフェクションした。レポーター遺伝子の転写を止めるために、ドキシサイクリンを培地に添加した。ドキシサイクリン添加後、Total RNA を経時的 (0、2、4、6 時間) にサンプリングした。Luciferase と β-グロビンの mRNA の

発現量はそれぞれ qRT-PCR により定量し、0 時間を 1 としたときの相対的な RNA の残存量（RNA の分解速度）を求めた。各時間における  $\beta$ -グロビンの mRNA の残存量は、Luciferase の mRNA の発現量でノーマライズを行った。エラーバーは 3 回の実験間での誤差を示している。

(A) GADD45B 3'UTR を組み込んだ  $\beta$ -グロビン mRNA の発現ベクター（pTET-BBB+GADD45B 3'UTR）。GADD45B 3'UTR 中には GC-rich なモチーフ配列が 4 箇所存在しており、それぞれのモチーフ配列を段階的に削った（ $\Delta$ site1、 $\Delta$ site4、 $\Delta$ site1 $\Delta$ site4、 $\Delta$ site1 $\Delta$ site2 $\Delta$ site4、 $\Delta$ site1 $\Delta$ site3 $\Delta$ site4、 $\Delta$ site1 $\Delta$ site2 $\Delta$ site3 $\Delta$ site4）、もしくはアデニンに置換した（ $\Delta$ site1 $\Delta$ site2 $\Delta$ site3 $\Delta$ site4  $\rightarrow$  A stretch）コンストラクトを作成した。

(B) GADD45B 3'UTR 中に含まれる GC-rich なモチーフ配列を削った、もしくはアデニンに置換したレポーター-RNA の分解速度の測定。

(C) UPF1 ノックダウン時における、GC-rich なモチーフ配列をアデニンに置換したレポーター-RNA の分解速度の変化。

## UPF1 は GC-rich なモチーフ配列から解離しにくい

近年の研究により、少なくとも NMD 経路においては、標的でない mRNA からの UPF1 タンパク質の解離が速いことが、UPF1 による基質認識に重要であると示唆されている (Lee et al. 2015)。

そこで私は、UPF1 標的 mRNA が選択的に分解されるメカニズムとして、標的 RNA からの UPF1 タンパク質の解離の違いに着目した。つまりは、標的 RNA 上で UPF1 の解離が遅くなることで、RNA 上に UPF1 が滞留しているのではないかと仮説を立てた。もしこの仮設が正しければ、GC-rich なモチーフ配列からの UPF1 タンパク質の解離が遅くなるはずだと考えた。RNA からのタンパク質の乖離は、RNA-タンパク質複合体を形成させた後に、一定時間経過した後の RNA-タンパク質複合体の残存量を調べることで評価できる (複合体の残存量が少ないほど RNA からタンパク質が乖離したと判断できる) (図 17)。

そこで、ゲルシフトアッセイ法を用いて、RNA-UPF1 複合体の崩壊 (すなわち、RNA からの UPF1 乖離) を評価することにした。まず、<sup>32</sup>P 標識した GADD45B の 3'UTR 配列について、元の GADD45B の RNA 配列 (WT) と GC-rich な配列をアデニンに置換した RNA 配列 (MUT) の 2 種類を用意した。それぞれの RNA を UPF1 タンパク質と In vitro で結合させ、その後、非標識 RNA と ATP を加えたときに、標識 RNA からの UPF1 の解離のされやすさを調べた。ゲルシフトアッセイの結果、それぞれの RNA (WT と MUT) はいずれも UPF1 タンパク質と結合していることを確認した (図 18、レーン 3 と 4)。

一方で、非標識 RNA を加えた後、MUT の RNA からの UPF1 タンパク質の解離は、WT の RNA と比較して速いことがわかった (図 18、レーン 5 と 6)。

また、ATP をさらに添加した条件下では、MUT の RNA からの UPF1 タンパク質の解離の程度が大きくなることがわ

かった（図 18、レーン 7 と 8）。RNA ヘリケースである UPF1 の RNA 上の運動には ATP が要求されるという考えに合致する結果と解釈した。以上の結果から、UPF1 は GC-rich な配列をアデニンに置換した RNA と比較して、GC-rich なモチーフ配列を含む RNA 上から解離しにくいことが示唆された。

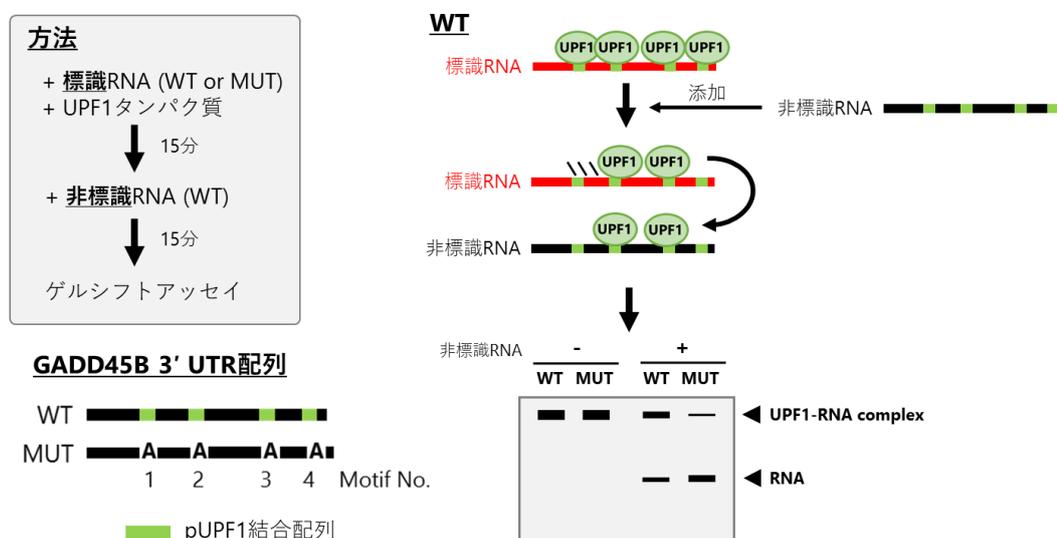


図 17. ゲルシフトアッセイによる RNA からの UPF1 乖離の評価

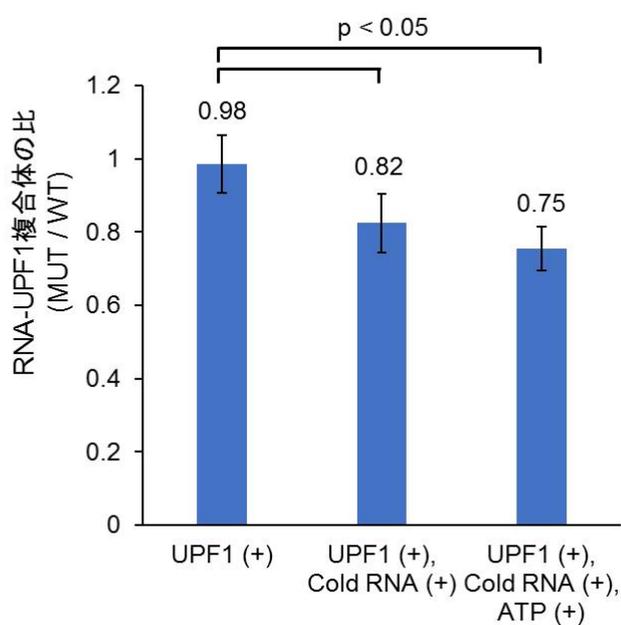
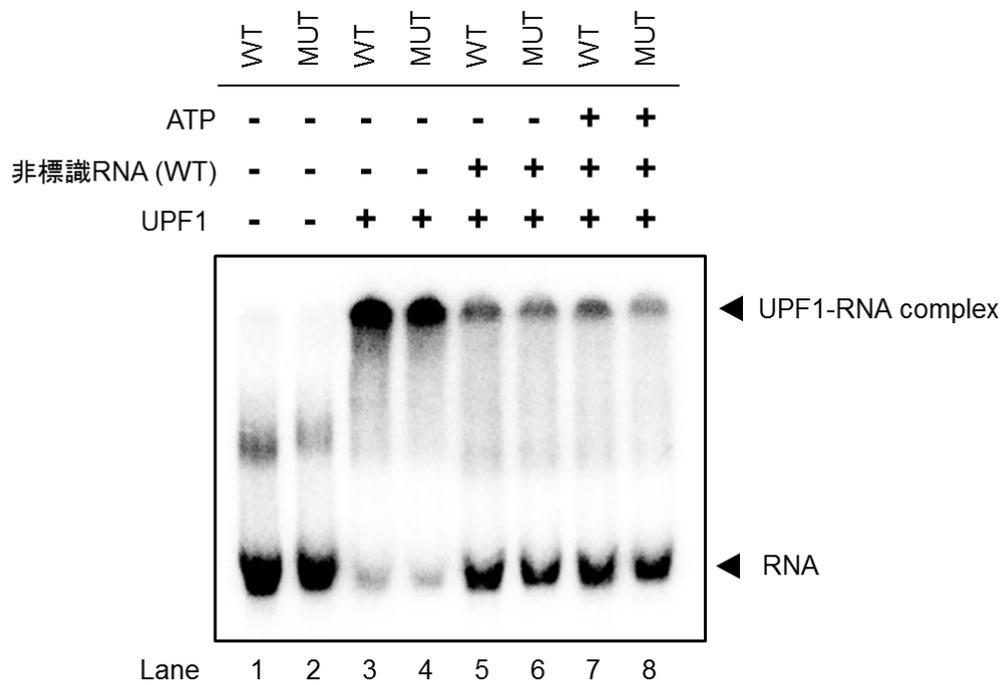


図 18. GC-rich なモチーフ配列を持つ GADD45B 3'UTR の RNA、もしくはモチーフ配列をアデニンに置換した RNA 上からの UPF1 タンパク質の解離

GADD45B 3'UTR の RNA 配列 (WT)、もしくはモチーフ配列をアデニンに置換した RNA 配列

(MUT) との UPF1 タンパク質の結合・解離をゲルシフトアッセイにより検出した。0.5 pmol  $^{32}\text{P}$  で標識された RNA のみ (レーン 1 と 2) と、標識した RNA と 500ng の UPF1 タンパク質とをインキュベートしたサンプル (レーン 3-8) 。レーン 5-8 に示したのは、1 pmol の非標識 RNA (WT) を上記の反応液に加え、30 min インキュベートしたサンプル (レーン 5 と 6) 、及び非標識 RNA とさらに ATP を添加してインキュベートしたサンプル (レーン 7 と 8) に該当する。標識 RNA と RNA と UPF1 の複合体のバンドの位置をそれぞれ矢印で示している。棒グラフは、RNA からの UPF1 タンパク質の相対的な解離の割合 (RNA-UPF1 複合体のバンド強度の比: MUT/WT) を示している。エラーバーは 3 回の実験間での誤差を示している。

## 総括と展望

## BRIC-seq と RIP-seq を用いた複合解析による UPF1 標的 mRNA の同定

UPF1 の標的 mRNA を同定する従来の方法では、UPF1 をノックダウンしたときに発現量が上昇する遺伝子群に着目していた。しかし、RNA の発現量は RNA の転写と分解のバランスによって決められていることから、mRNA の発現量の増加は、RNA 分解が抑制されたことに起因するとは限らない。例えば、UPF1 ノックダウンにより RNA の転写が亢進することで RNA の発現量が増加するケースも考えられる。実際に、転写因子をコードする mRNA は RNA の半減期が短い傾向にあり (Maekawa et al. 2015)、これは転写因子の遺伝子発現が、転写後調節 (特に RNA 分解) によって制御されていることを示唆している。また本研究においても、転写因子の遺伝子が UPF1 の標的 mRNA の中に多く含まれていることを確認している。これらの事実から、UPF1 ノックダウンにより転写因子の発現量が上昇し、結果として転写因子の標的となる RNA の発現量が副次的に上昇している可能性がある。また、先行研究において、Kurosaki ら (Kurosaki et al. 2014) はこのような発現量解析によって同定された UPF1 の標的 mRNA の候補 (Viegas et al. 2007) は、リン酸化された UPF1 と結合しないものも多いことを示している。以上のことから、RNA の発現変動量解析では、False-positive にあたる遺伝子群が多く含まれてしまうことが先行研究における問題点であると考えられる。一方で今回、BRIC-seq と RIP-seq を用いて、UPF1 ノックダウン時において RNA の半減期が延長し、かつ UPF1 と相互作用のある mRNA を同定して得られた UPF1 標的 mRNA は、他の mRNA と比較してリン酸化 UPF1 と結合している傾向にあることが示された (図 7A)。さらに、今回同定された UPF1 標的 mRNA には、細胞周期に関係する遺伝子も多く含まれており、この結

果は、UPF1 が細胞周期の制御に関わっているとする先行研究の結果（Azzalin and Lingner 2006）とも一致している。以上の結果から、BRIC-seq と RIP-seq を用いた複合解析が UPF1 の標的となる mRNA を正確に同定する上で有用な方法であることを示唆している。また、UPF1 タンパク質以外の RNA 結合タンパク質（RBP）の標的 mRNA を同定する上で、BRIC-seq と RIP-seq を用いた複合解析が、RBP の機能制御を標的遺伝子のデータセットから理解する上で有用な方法であると提案できる。

### UPF1 標的 mRNA の 3'UTR 中の配列的特徴

まず、リン酸化 UPF1 に対する網羅的な RNA フットプリントデータを用いて、UPF1 標的 mRNA の配列的特徴を解析した。その結果、リン酸化 UPF1 は標的 mRNA の 3'UTR 上に結合する傾向があることを明らかにし（図 7B、C）、UPF1 標的 mRNA の 3'UTR 中から CUG を中心とした GC-rich なモチーフ配列を予測した（図 7E）。さらに、この GC-rich なモチーフ配列を介して、UPF1 依存的な RNA 分解が誘導されることをレポーターアッセイにより実験的に検証した（図 15、16）。加えて、UPF1 の標的でない mRNA の 3'UTR 中にこの GC-rich なモチーフ配列を組み込んだとき、UPF1 依存的な RNA 分解が誘導されるかどうか確認を行った。その結果、GC-rich なモチーフ配列を 3'UTR 中に組み込むだけでは、RNA 分解は誘導されなかった（Data not shown）。これは RNA 分解を誘導する上で GC-rich なモチーフ配列の存在が必要な条件である一方、十分な条件でないことを意味している。UPF1 標的 mRNA の 3'UTR 配列をさらに解析すると、3'UTR 全体の GC 含有

量が他の mRNA と比較してとても高いことが示された。また、モノヌクレオチド・ジヌクレオチドのレベルでは、G、及び GA・AG の含有量が高いことが示された（図 11）。これらの結果は、UPF1 を介した RNA 分解において、GC-rich なモチーフ配列だけでなく、その周辺の配列的な特徴（G-rich、GA/AG-rich）が基質認識に重要であることを示唆している。

### 3'UTR を介した UPF1 依存的 RNA 分解経路のモデル

UPF1 は、一本鎖 RNA 上を数 kb にわたって移動するタンパク質である（Fiorini et al. 2015）。RNA 上を UPF1 が縦走している間に、UPF1 が持つ ATPase 活性やヘリカーゼ活性により、RNA から UPF1 が解離する（Franks et al. 2010）。In vitro において、UPF1 がポリ G もしくはポリ GC 配列と結合した場合、UPF1 の ATPase 活性は減弱することが知られている（Bhattacharya et al. 2000）。また、近年の研究から、UPF1 の持つ ATPase 活性の低下により、UPF1 が RNA 上に滞留することがわかっている（Lee et al. 2015）。これらの結果から、GC-rich な RNA 配列上では UPF1 の ATPase 活性が減弱することで、UPF1 タンパク質の RNA 上での移動や解離が抑制され、結果として UPF1 が RNA 上に滞留しているのではないかと考えられる。この考えに一致して、本研究において、GADD45B 3'UTR の RNA 配列からの UPF1 タンパク質の解離が、GC-rich なモチーフ配列をアデニンに置換した RNA 配列と比較して遅くなることを明らかにした（図 17）。加えて、UPF1 が RNA 上に滞留することで UPF1 がリン酸化され、結果として RNA 分解が誘導される（Kurosaki et al. 2014; Durand et al. 2016）ことから、GC-rich な配列上での UPF1 の滞留により、UPF1 標

的 mRNA が選択的に分解されていると考えられる。また、3'UTR 上での UPF1 の滞留は、翻訳依存的に起こっているとされている (Hurt et al. 2013; Zünd et al. 2013; Gregersen et al. 2014; Kurosaki et al. 2014)。これらの先行研究に結果と一致して、本研究においても UPF1 を介した RNA 分解が翻訳依存的に起こっていることを明らかにした (図 15D)。

先行研究の結果と本研究の結果を合わせて考えると、以下のようなモデルが考えられる (図 18)。まず、UPF1 タンパク質は 3'UTR 上を移動し、GC-rich な RNA 配列上で ATPase 活性が低下する。その結果、UPF1 標的 mRNA の 3'UTR 上に含まれる GC-rich な領域上に UPF1 が滞留する。mRNA 上で UPF1 が滞留することで、SMG1 により UPF1 がリン酸化される確率が増加し、結果として、SMG5/SMG7 もしくは SMG6 のリクルートを介して、RNA 分解が誘導される。一方で、GC-rich な配列を持たない mRNA (UPF1 の標的でない mRNA) 上では、UPF1 の解離が速い。そのため、GC-rich な配列を持たない mRNA は UPF1 依存的な RNA 分解を受けないと考えられる。

### RNA 結合タンパク質の基質認識と RNA 上の滞在時間

近年、ENCODE プロジェクトをはじめとする大規模なトランスクリプトーム解析が行われ、RNA の発現量だけでなく、特定の RNA 結合タンパク質とその標的となる mRNA 群を網羅的に同定することで、RNA 結合タンパク質の機能を理解しようという試みがなされている (Van Nostrand et al. 2016)。このように、今までの RNA 結合タンパク質の研究では、標的 RNA に RNA 結合タンパ

質が「基質である RNA に結合するかどうか」に着目した研究が多い。

一方で私は、本研究を通して、RNA 結合タンパク質による基質認識に関する新しいメカニズムを提案したい (図 19)。UPF1 タンパク質のように、RNA 上を移動するタイプの RNA 結合タンパク質では、RNA に単に結合するかどうかが重要ではなく、「RNA 上での滞留時間」が RNA に対する基質認識に重要である。すなわち、RNA 結合タンパクが RNA 上をゆっくりと運動することで、結果的に RNA 結合タンパク質が RNA 上に長く滞在することが RNA 結合タンパク質による標的 RNA の識別につながるという考えである。このように、RNA 結合タンパク質の RNA に対する結合速度だけでなく、解離速度を含めた機能解析が RNA 分解メカニズムを解くカギになると主張したい。

最後に、本研究が UPF1 のみならず、さまざまな RNA 結合タンパク質による RNA 分解メカニズムの解明の糸口になることを願いつつ、本稿を終えたい。

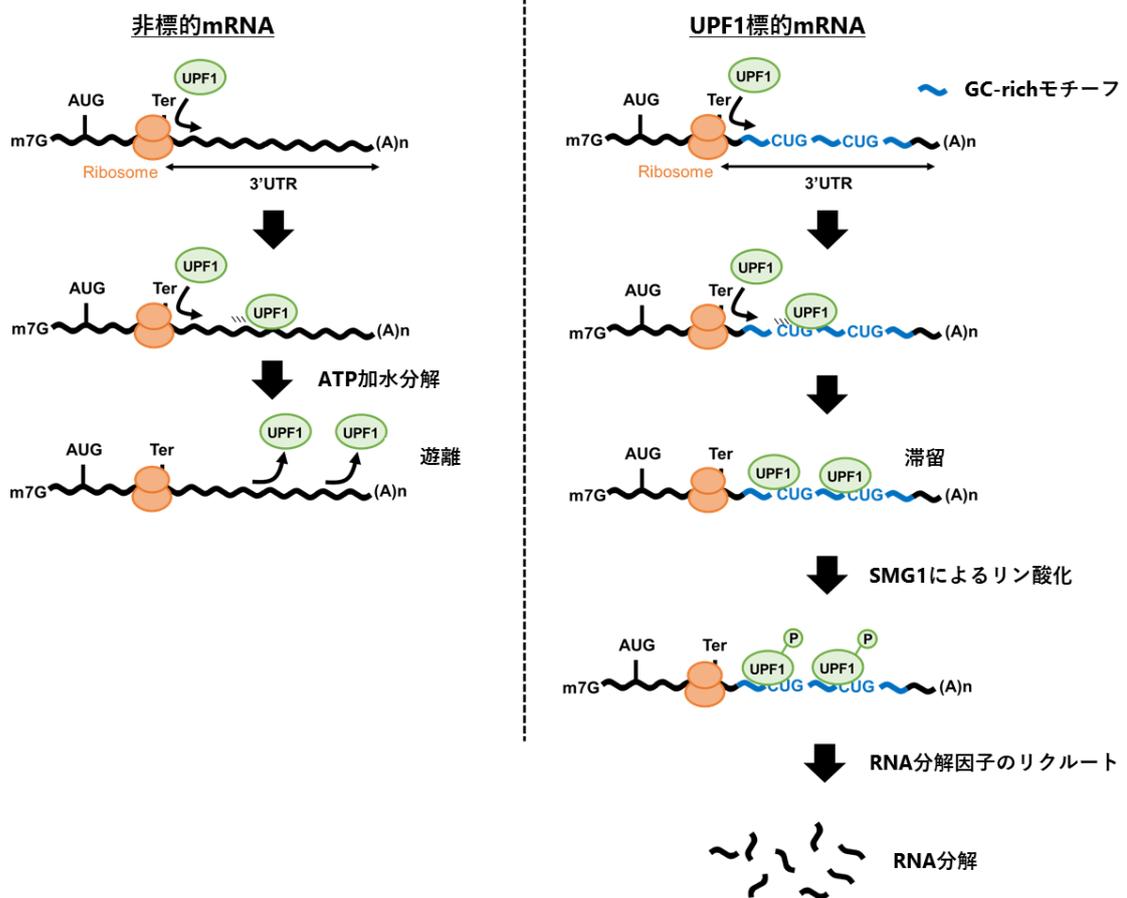


図 18. UPF1 を介した正常な mRNA 分解経路のモデル

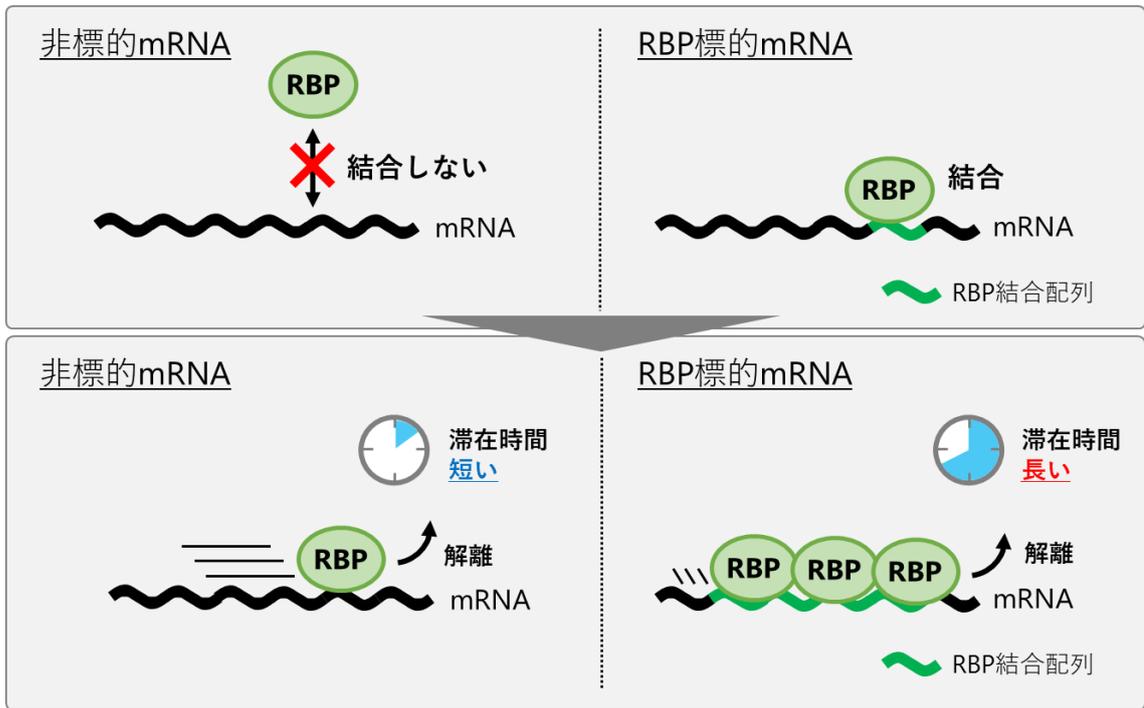


図 19. RNA 結合タンパク質の基質認識と RNA 上の滞在時間

## 材料と方法

## 1. 細胞培養

HeLa Tet-off 細胞 (Clontech, Mountain View, CA) を 37°C、5%の CO<sub>2</sub> インキュベーターで培養した。培地は、10% fetal bovine serum (Thermo Fisher Scientific, Waltham, MA) および penicillin/streptomycin を添加した Dulbecco's modified Eagle's medium (Wako, Osaka, Japan) を用いた。

## 2. siRNA トランスフェクション

siRNA は Lipofectamine RNAiMAX (Thermo Fisher Scientific) を用いて細胞にトランスフェクションした。終濃度 10 nM siRNA を細胞にトランスフェクションし、6 時間インキュベートさせた後、培地交換を行った。細胞は 48 時間培養した後、回収した。

siRNA の配列リストは以下のとおりである。

Name	Sence sequece (5'-3')	Antisense sequece (5'-3')
Control siRNA-1	GTACCTGACTAGTCGCAGAAG	TCTGCGACTAGTCAGGTACGG
Control siRNA-2	UUCUCCGAACGUGUCACGUTT	ACGUGACACGUUCGGAGAATT
UPF1 siRNA-1	GAUGCAGUUCGCUCCAUUdTdT	AAUGGAGCGGAACUGCAUCdTdT
UPF1 siRNA-2	GAGAAUCGCCUACUUCACUUU	AGUGAAGUAGGCGAUUCUCUU
siUPF2 siRNA-1	CUUUUGUCCCAGCCAUCUU	GAUGGCUGGGACAAAAGCC
siUPF2 siRNA-2	GACAUGUGGACGGUUUCUUUU	AAGAAACCGUCCACAUGUCUC
SMG1 siRNA-1	AAGUGUAUGUGCGCCAAAGUA	CUUUGGCGCACAUACACUUCA
SMG1 siRNA-2	GGCCUACUUAAGAAACGCAGC	UGCGUUUCUUAAGUAGGCCUG

## 3. プラスミドの構築

pTetBBB/GADD45B 3'UTR ベクターを構築するために、HeLa Tet-off 細胞から抽出した Total RNA と SuperScript II Reverse Transcriptase (Thermo Fisher Scientific) から一本鎖 cDNA を作成した。5'-GAAGATCTTGAGTTGCTGCCACAAACAA-3' (センス鎖) と 5'-GAAGATCTCCAGCTCCAGTCTCAAGGA-3' (アンチセンス鎖) のプライマーセット (下線は BgIII の制限酵素サイト) を用いて、PCR により目的の DNA 配列を増幅させた。PCR 産物を BgIII 酵素で処理し、pTetBBB ベクターに組み込んだ。

pTetBBB/GADD45B Δsite1 (858–862 bp) と Δsite4 (1113–1122 bp) と Δsite1Δsite4 (858–862 bp / 1113–1122 bp) をそれぞれ構築するために、両末端に BgIII の制限酵素サイトを持った GADD45B mRNA (NM\_015675.3, 742–1122 bp, Δsite1 と Δsite4 と Δsite1Δsite4) の合成 DNA を Eurofins Genomics から購入した。その DNA 配列を BgIII 酵素で処理し、pTetBBB ベクターに組み込んだ。

pTetBBB/GADD45B Δsite1Δsite2Δsite4 (858–862 bp / 921–927 bp / 1113–1122 bp) と Δsite1Δsite3Δsite4 (858–862 bp / 985–991 bp / 1113–1122 bp) をそれぞれ構築するために、pTetBBB/GADD45B Δsite1Δsite4 を PCR のテンプレートとして、site2 を削るために 5'-GCCGCCGATCAGATGGAGA-3' (センス鎖) と 5'-TCCTGGATCTCCGCCCTG-3' (アンチセンス鎖) のプライマーセットを、site3 を削るために 5'-CCTTGGGATGGAGCAGAAG-3' (センス鎖) と 5'-TTCGGGGTCTGTCTCC-3' (アンチセンス鎖) のプライマーセットをそれぞれ用いて、PCR で目的の配列を増幅させた。

TetBBB/GADD45B Δsite1Δsite2Δsite3Δsite4 (858–862 bp / 921–927 bp / 985–991 bp / 1113–1122 bp) を構築するために、pTetBBB/GADD45B Δsite1Δsite2Δsite4 を PCR のテンプレートとして、site3 を削るために 5'-CCTTGGGATGGAGCAGAAG-3' (センス鎖) と 5'-TTCGGGGGTCCTGTCTCC-3' (アンチセンス鎖) のプライマーセットを用いて、PCR で目的の配列を増幅させた。PCR 産物の 5'末端と 3'末端を T4 polynucleotide kinase (Toyobo, Osaka, Japan) を用いてリン酸化し、self-ligation させた。

pTetBBB/GADD45B の site1-4 をアデニンに置換したベクターを構築するために、5'末端と 3'末端に BglIII の制限酵素サイトを持った合成 DNA を Eurofins Genomics から購入した。その DNA 配列を BglIII 酵素で処理し、pTetBBB ベクターに組み込んだ。

pTET-BBB ベクターに AP2M1、EVL、HNRNPA2B1 の 3'UTR をそれぞれ組み込んだベクターと GC-rich なモチーフ配列をランダムに配置した 3'UTR の配列を組み込んだベクターをそれぞれ構築するために、合成 DNA を Eurofins Genomics から購入した。その DNA 配列を BglIII 酵素で処理し、pTetBBB ベクターに組み込んだ。

以下にクローニングで使用した合成 DNA の配列を示す。

#### HNRNPA2B1 3' UTR

AGATCTTAGGAGAGGATGAGAGCCCAGAGGTAACAGAACAGCTTCAGGTTATCGAAATAAC

AATGTTAAGGAAACTTATCTCAGTCATGCATAAATATGCAGTGATATGGCAGAAGACACC

AGAGCAGATGCAGAGAGCCATTTTGTGAATGGATTGGATTATTTAATAACATTACCTTACTGT

GGAGGAAGGATTGTAAAAAAAATGCCTTTGAGACAGTTTCTTAGCTTTTTAATTGTTGTTTC  
TTTCTAGTGGTCTTTGTAAGAGTGTAGAAGCATTCTTCTTTGATAATGTTAAATTTGTAAGTT  
TCAGGTGACATGTGAAACCTTTTTTAAGATTTTCTCAAAGTTTTGAAAAGCTATTAGCCAGG  
ATCATGGTGTAATAAGACATAACGTTTTTCCTTTAAAAAATTTAAGTGCGTGTGTAGAGTTA  
AGAAGCAGATCT

GC-rich なモチーフ配列を含む HNRNPA2B1 3' UTR (太文字は置換した配列)

AGATCTTAGGAGAGGATGAGAGCCCAGAGGTAACAGCCTGGGGTCAGGTTATCGAAATAAC  
AATGTCCTGGGGACTCTTATCTCAGTCATGCATAACCTGGGGTGATATGGCAGAAGACAC  
CAGACCCTGGGGCAGAGAGCCATTTTGTGAATGGACCCTGGGGATTTAATAACATTACCTTACT  
GTCCTGGGGGGGATTGTAAAAAAAATGCCTTTCCTGGGGTTTCTTAGCTTTTTAATTGTTGTCC  
TGGGGTAGTGGTCTTTGTAAGAGTGTAGCCTGGGGCCTTCTTTGATAATGTTAAATTTCCTGG  
GGTCAGGTGACATGTGAAACCTTTTCCTGGGGTTTTCTCAAAGTTTTGAAAAGCTCCTGGGG  
AGGATCATGGTGTAATAAGACATCCTGGGGTTCTTTAAAAAATTTAAGTGCCCCTGGGGGA  
GTTAAGAAGCAGATCT

AP2M1 3' UTR

AGATCTCTGCCACTAGGCAGCTAGCCCACCTCCCCAGCCACCCTCCTCCACAGGTCCAGGTG

CCGCTCCCTCCCCACCACACATCAGTGTCTCCTCCCTCCTGCTTTGCTGCCTTCCCTTTGCAC  
CAGCCCGAGTCTAGGTCTGGGCCAAGCACATTACAAGTGGGACCGGTGGAGCAGCCCCTGG  
GCTCCCTGGGCAGGGGAGTTCTGAGGCTCCTGCTCTCCCATCCACCTGTCTGTCCTGGCCTA  
ATGCCAGGCTCTGAGTTCTGTGACCAAAGCCAGGTGGGTTCCCTTTCCCTCCACCCCTGTG  
GCCACAGCTCTGGAGTGGGAGGGTTGGTTGCCCTCACCTCAGAGCTCCCCAAAGGCCAG  
TAATGGATCCCCGGCCTCAGTCCCTACTCTGCTTTGGGATAGTGTGAGCTTCATTTTGTACAC  
GTGTGACTTCGTCCAGTTACAAACCCAATAAACTCTGTAGAGTGGAGATCT

GC-rich なモチーフ配列を含む AP2M1 3' UTR (太文字は置換した配列)

AGATCTCTGCCACTAGGCAGCTAGCCCACCTCCCCACCTGGGGTCCTCCACAGGTCCAGGT  
GCCGCCCTGGGGCCCACCACACATCAGTGTCTCCTCCTGGGGGCTTTGCTGCCTTCCCTTTG  
CACCCTGGGGAGTCTAGGTCTGGGCCAAGCACACCCTGGGGTGGGACCGGTGGAGCAGCCC  
CTGCCTGGGGTGGGCAGGGGAGTTCTGAGGCTCCCTGGGGCCCATCCACCTGTCTGTCCTG  
GCCCCTGGGGCAGGCTCTGAGTTCTGTGACCAACCCTGGGGTGGGTTCCCTTTCCCTCCACCC  
CCTGGGGCACAGCTCTGGAGTGGGAGGGTTCCTGGGGCCTCACCTCAGAGCTCCCCAAAC  
CTGGGGAATGGATCCCCGGCCTCAGTCCCCCTGGGGGCTTTGGGATAGTGTGAGCTTCATCCT  
GGGGACGTGTGACTTCGTCCAGTTACACCCTGGGGTAAACTCTGTAGAGTGGAGATCT

EVL 3' UTR

AGATCTGGGGCCGGCCTCGCTGCGCTGATTCGTCGAGCCCATCCGGCGACAGAGGACAGCC  
AGAAGCCCAGCCAGCCCCAGACTCCAGTGCACCAGAGCACGCACAGGAGCCTGGGCGCGC  
TGCTGTGAAACGTCCTGACCTGTGATCACACATGACAGTGAGGAAACCAAGTGCAACTCCTG  
GGTTTTTTAGATTCTGCCTGACACGGAACACCAGGTCTGCTCGTCTTTTTTGTGTTTTATATT  
GCTTATTTAAGGTACATTTCTTTGGGTTTCTAGAGACGCCCTAAGTCACCTGCTTCATTAGAC  
GGTTCCAGGTTTTCTCCAGGTGACGCTGTTAGCGCCTCAGCTGGCGGTGACAGCCGGCCC  
AGCGTGGCGCCACCACACACCCGACAGAGCTGTCCAGGCACAGCTCCGTCCCCAGCGCTCATG  
GTGTTGAAACTGTCTGTCATGCACCACGGTGTCTGTGTCCACACAGTAATAAACGGTTTACTG  
TCCGCAGATCT

GC-rich なモチーフ配列を含む EVL 3' UTR (太文字は置換した配列)

AGATCTGGGGCCGGCCTCGCTGCGCTGATTCGTCGACCCTGGGGCGGCGACAGAGGACAGC  
CAGAAGCCTGGGGAGCCCCAGACTCCAGTGCACCAGCCTGGGGACAGGAGCCTGGGCGC  
GCTGCTGCCTGGGGTCTGACCTGTGATCACACATGACCCTGGGGGAAACCAAGTGCAACTC  
CTGGGTCCTGGGGGATTCTGCCTGACACGGAACACCCCCTGGGGGCTCGTCTTTTTTGTGTTTTA  
TATCCTGGGGTTTAAGGTACATTTCTTTGGGTTCCTGGGGACGCCCTAAGTCACCTGCTTCA  
CCTGGGGGTTTCCAGGTTTTCTCCAGGTGCCTGGGGTAGCGCCTCAGCTGGCGGTGACAC

CTGGGGCAGCGTGGCGCCACCACACACCGCCTGGGGGTCCAGGCACAGCTCCGTCCCCAC

CTGGGGTGGTGTGAACTGTCTGTCATGCCTGGGGGTGTCTGTGTCCACACAGTAATACCT

GGGGTACTGTCCGCAGATCT

#### 4. RNA 免疫沈降 (RIP-seq)

HeLa Tet-off 細胞に HA タグを付けた UPF1 を Lipofectamine 2000 (Thermo Fisher Scientific) を用いてトランスフェクションした。次の日、細胞を回収し、4°Cに冷やした PBS で洗う。ペレット化させた細胞を PIPA buffer[50 mM Tris-HCl (pH 8), 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100, 0.1% DOC, 0.5 U/μL RNasin plus RNase inhibitor (Promega, Madison, WI) and protease inhibitor cocktail (Sigma, St Louis, MO)]により再懸濁させ、Bioruptor UCD-250 (Cosmo Bio, Tokyo, Japan) を用いてソニケーションにかけた。RNA 免疫沈降に使うために、遠心上清と anti-HA High Affinity rat antibody (Roche, Mannheim, Germany, cat# 3F10) をそれぞれ用意した。このとき、抗体は Dynabeads protein G (Thermo Fisher Scientific) と1時間、4°Cの条件でインキュベートさせ、RIPA buffer で3回洗浄したものを使用した。遠心上清と抗体-ビーズの混合液を混ぜ合わせ、2時間、4°Cで転倒混和させた。抗体と RNA が結合したビーズを4°Cに冷やした RIPA buffer で3回洗浄した。RNA を回収するために、ビーズに ISOGEN LS (Nippon Gene, Tokyo, Japan) を加え、共沈降した RNA を抽出した。mRNA Seq Sample Preparation Kit (Illumina, San Diego, CA) と1μg の RNA を用いて、

RIP-seq 用の cDNA ライブラリを構築した。得られた cDNA ライブラリを次世代シーケンサーに供した。

Illumina HiSeq 2000 を用いて Single-end、36bp の RNA-seq のシーケンスリードを得た。

## 5. Quantitative real-time reverse transcription PCR (qRT-PCR)

抽出した Total RNA を PrimeScript RT Master Mix (Takara, Shiga, Japan) を用いて一本鎖 cDNA に逆転写した。特定の遺伝子に対するプライマーと SYBR Premix Ex Taq II (Takara) を用いて、cDNA を増幅させた。qRT-PCR の解析は、Thermal Cycler Dice Real Time System (Takara) を用いて行った。

使用した qRT-PCR のプライマーセットは以下のとおりである。

Name	Sence sequece (5'-3')	Antisense sequece (5'-3')
GAPDH	GCACCGTCAAGGCTGAGAAC	TGGTGAAGACGCCAGTGGA
UPF1	AGATCACGGCACAGCAGAT	TGGCAGAAGGGTTTTTCCTT
UPF2	GGATGTTGGGATCCACGTT	CGCCTCTGATTAATTTAGGTTG
SMG1	TGGTCATCTTGCACAACACC	CATGACTGCAGTAGATCTTGCAT
GADD45B	TTTGCAATATGACTTTGGAGGA	CATCCCCACCTTATCCAT
PEA15	ACTACTGGCAGTGCCTGGTT	GTTGTCTTTGTCCAGCTTGTG
ABHD4	GGCCGGATTCTATGTCC	CCTCCACCACAGCATTGAA
DNAJB2	CACGCAGAATCATGGAGAAC	GCCAGGTCATCTGGGACA
DDIT3	CAGATGAAAATGGGGG	TCAAGAGTGGTGAAGA
SMG5	TTGCGTCAATGGGGATGTA	GA CTCAGAGCCCTCCTCAGA
SNRPB	GAGGGGAGAATCTGGTCTCA	TGGA ACTCGAGCAATACCAGT
PRMT1	AAAGCAGTGAGAAGCCCAAC	TCGTCCTCAGCATCTCCTC
POLR2G	GGCCCATGTCTTGCTTCATC	TGGTGGGTTGGAGTTAGGAT
SNRPB (NMD target, ENST00000474384)	CTGTAGCCTAAAGGGCTCTCC	TCTGCTTGTTTGGAGTTCTTTG
PRMT1 (NMD target, NR_033397)	CTGTGGGCTGAGCTAGAGAC	GTTGGGCTTCTCACTGCTTT

POLR2G (NMD target, ENST00000525455)	GGACCTGCACAGGGAAAAT	CCTGAGAATCATCTTCTTAAAATGC
rabbit-globin	TTTGCTAAGCTGAGTGAAGTGC	TGCCAAAATGATGAGACAGC
Luciferase	TATGTGGCCAGCCAGGTTAC	GTCCACGAACACAACACCAC

## 6. In vivo cross-linking analysis (CLIP)

細胞は 150  $\mu$ m 4'-thio-uridine を添加した培地で 12 時間インキュベートした。次の日、細胞を回収し、4°Cに冷やした PBS で洗浄した。その後、氷上で 360 nm UV クロスリンクを行った。ペレット化させた細胞を 200  $\mu$ l lysis buffer [2% SDS, 50 mM Tris-HCl (pH 8), 1 mM EDTA, 1 mM DTT]に再懸濁させた後、5 分、95°Cでインキュベートさせた。4 倍量の Dilution buffer [1.25% NP-40, 0.625% DOC, 62.5 mM Tris-HCl (pH 8), 1.75 mM EDTA, 187.5 mM NaCl]を加え、3 回ソニケーションにかけ、4°C、14,000 $\times$ g で 90 分遠心する。遠心上清と anti-UPF1 goat antibody (Bethyl, Montgomery, CA, USA, A300-038A) をそれぞれ用意した。このとき、抗体は Dynabeads protein G (Thermo Fisher Scientific) と 3 時間、4 °Cの条件でインキュベートさせ、NET2 buffer [50mM Tris-HCl (pH 8), 100 mM NaCl, 0.05% NP40, 0.5 mM DTT]で 3 回洗浄したものを使用した。遠心上清と抗体-ビーズの混合液を混ぜ合わせ、2 時間、4°Cで転倒混和させた。抗体と RNA が結合したビーズを 4°Cに冷やした NET2 buffer で 3 回洗浄した。RNA を回収するために、ビーズに ISOGEN LS (Nippon Gene, Tokyo, Japan) を加え、共沈降した RNA を抽出した。RNA の定量には qRT-PCR を用いた。

## 7. mRNA 分解アッセイ

HeLa Tet-off 細胞に内在コントロールの RNA を発現するベクターとして 975ng pGL4.13 [luc2/SV40] (Promega) とレポーターRNA を発現するベクターとして 25 ng pTetBBB (Dr. Ann-Bin Shyu の好意によりいただいた) もしくは pTetBBB/GADD45B をそれぞれトランスフェクションした。pTetBBB と pTetBBB/GADD45B は、Rabbit の  $\beta$ -グロビン mRNA、および GADD45B 3'UTR 配列の付いた $\beta$ -グロビン mRNA をそれぞれコードするベクターである。プラスミドをトランスフェクションして 24 時間後に、1  $\mu$ g/ml ドキシサイクリンを培地に添加し、転写をストップさせる。ドキシサイクリン添加後を 0 時間として、0、2、4、6 時間の各時間に Total RNA を回収した。RNA 抽出は RNAiso Plus (Takara) を用いて行った。内部コントロールであるルシフェラーゼ mRNA の発現量と Rabbit の  $\beta$ -グロビン mRNA の経時的な減少量を qRT-PCR により定量した。各時間における Rabbit の $\beta$ -グロビン mRNA の発現量は、ルシフェラーゼ mRNA の発現量でノーマライズした。

## 8. GST-human UPF1 タンパク質の調整

1mM IPTG 存在下で 3 時間インキュベートし、Escherichia coli strain BL21-CodonPlus (DE3)-RIPL から GST-human UPF1 タンパク質を発現誘導させた。ペレット化させた大腸菌を、1mM PMSF, 2 mM DTT and 4 mg Lysozyme を加えた GST buffer [50mM Tris-HCl (pH = 8), 1M NaCl, 5% Glycerol] で再懸濁させ、ソニケーションした。得られた細胞溶解液の遠心上清に Glutathione Sepharose 4B (GE Healthcare, England, 17075601) を加え、1 時間 4°C でイン

キュベートさせる。ビーズは、1mM PMSF, 2 mM DTT を加えた GST buffer で 1 回、2 mM DTT を加えた GST buffer で 2 回洗浄した。GST-human UPF1 タンパク質は GST elution buffer [50 mM Tris-HCl (pH = 8), 150 mM NaCl, 5% Glycerol, 10mM reduced L-Glutathione (Sigma, St Louis, MO, G4251) and 1mM DTT]で溶出させた。溶出したリコンビナントタンパク質を dialysis buffer [50 mM Tris-HCl (pH = 8), 150 mM NaCl, 5% Glycerol and 1 mM DTT]で 2 時間透析させた。

## 9. RNA electrophoretic mobility shift アッセイ

GADD45B 3'UTR 配列、もしくは GC-rich なモチーフ配列をアデニンに置換した 3'UTR 配列をそれぞれ組み込んだ pGEM-T ベクターから、T7 RNA polymerase (Takara) と $[\alpha\text{-}^{32}\text{P}]$  UTP (800 Ci/mmol) (PerkinElmer, Waltham, MA, USA) を用いて  $^{32}\text{P}$  標識した RNA を in vitro 合成した。Binding buffer [20mM HEPES, 50mM KAc, 2.5mM Mg(Ac)<sub>2</sub>, 2mM DTT, 1.5% BSA, 10% Glycerol, pH7.0]と 0.5 pmol  $^{32}\text{P}$  標識した RNA、500ng UPF1 リコンビナントタンパク質の反応液を 10°C、15 分インキュベートさせた。その後、1 pmol 標識していない RNA を反応液に添加し、さらに 10°C、15 分インキュベートさせた。4.5% native PAGE ゲル (60:1 acrylamide:bisacrylamide in 0.5 × TBE containing 10% glycerol)を使った電気泳動により、RNA と UPF1 の複合体を分離させた。ゲルを乾燥させ、FLA9000 laser scanner (Fuji Film) を用いて標識された RNA を検出した。

## 10. 次世代シーケンサーのデータの前処理、マッピング、定量化

FASTX-toolkit を用いてクオリティの低いシークエンスリードと各塩基を除去し、10bp 未満のシークエンスリードを排除した。得られたリードをヒトゲノム (hg19) とトランスクリプトーム (RefSeq database、iGenomes (downloaded on 2nd July 2013)から取得) に Tophat (version 2.0.12) (Kim et al. 2013) を用いてマッピングした。マッピングされたリードは Cufflinks (version 2.1.1) (Trapnell et al. 2010) を用いて各遺伝子に対する RNA の発現量を定量化した。RNA の発現量は、各サンプル間の総リード数のバラつきと RNA 長の違いを考慮して、RPKM 値 (RNA1kb 長あたり、100 万リードあたりの相対的なリード数) として算出された。最終的に、各サンプルで RPKM が 1 以上である 8,426 遺伝子の mRNA を今回の解析の対象とした。

## 11. PTC を含む Isoform の解析

RefSeq と Gencode v19 のデータベースで "nonsense\_mediated\_decay" としてアノテーションされた、PTC を持つ mRNAs (isoforms) を抽出した。上記の候補の中から、さらに、Exon inclusion もしくは  $\geq 100$  bp 5'/3' Exon overhang に起因したフレームシフトによりナンセンス変異が ORF 上にできるタイプの isoforms を抽出した。mRNA のアノテーション情報のみを含む GTF ファイルに、上記で絞り込んだ PTC を含む isoforms の情報を加えた GTF ファイルを作成した。この GTF ファイルを用いて、Cufflinks により PTC を持つ isoform の発現量を定量化した。

## 12. GC 含有率の計算

RefSeq のデータ [<http://ftp.ncbi.nlm.nih.gov/refseq/supplemental/ProtCore/SM1.txt>]  
(Fong et al. 2013) からオルソログ遺伝子の情報を取得した。ヒトの遺伝子に対するマウスのオルソログ遺伝子を対応付けし、それぞれの GC 含有率を計算した。

## 13. 転移因子の配列解析

3'UTR に含まれる転移因子の配列を RepeatMasker 4.0.3 を用いて、RefSeq データベースから同定した。

## 14. GO term enrichment 解析

Functional Annotation Chart in DAVID Bioinformatics Resources 6.7 (<http://david.abcc.ncifcrf.gov/home.jsp>) (Huang et al. 2009) の Web サイトから GO term enrichment 解析を行った。

## 15. 5'-bromo-uridine immunoprecipitation chase-deep sequencing (BRIC-seq)の解析

UPF1 もしくは STAU1 をノックダウンしたときのそれぞれの BRIC-seq データは DDBJ Sequence Read Archive (DRA)の DRA000591 と DRA001215 からそれぞれダウンロードした。BRIC-seq データから RNA の半減期を計算した (Imamachi et al. 2014) 。

## 16. UPF1 の結合配列の予測

リン酸化 UPF1 に対する RNA フットプリントデータ、UPF1 に対する HITS-CLIP もしくは eCLIP データを Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) の GSE60045、GSE69586、GSE80039 からそれぞれダウンロードした。まず、アダプター配列を除き、クオリティの低いリードを除去した。Bowtie2 (version 2.2.4) (Langmead et al. 2012) を用いてヒトゲノム (hg19) にリードをマッピングした。リン酸化 UPF1 に対する RNA フットプリントデータに関しては、Bowtie2 の local mode (--local) を使ってマッピングした。各遺伝子の RNA の発現量は、Cufflinks を用いて定量した。Pyicoclip (version 2.0.6b) (Althammer et al. 2011) を用いて、各データから UPF1 の結合サイトに由来するピークを検出した。このとき、PCR duplicate に由来するブロック状のピーク (artifact) は除いた。UPF1 の結合モチーフを予測するために、UPF1 標的 mRNA の 3'UTR 中に検出された上記のピークの RNA 配列を使い、MEME によりモチーフ検索を行った (Bailey et al. 2006)。RPKM 値が 1 以上の 8,426 遺伝子の mRNA の 3'UTR 配列から、MEME の解析に用いる各塩基のバックグラウンドを取った。上述のデータ解析には、Perl、Python、R (R Core Team 2016) のカスタムスクリプトを用いた。

## 17. データアクセス

次世代シーケンサーを用いて得られたデータは、DDBJ Sequence Read Archive (DRA;

<http://trace.ddbj.nig.ac.jp/dra/>)上の DRA003675 からアクセスできる。データ解析に 用いたソースコードは、[https://github.com/Imamachi-n/UPF1\\_study\\_script\\_collection](https://github.com/Imamachi-n/UPF1_study_script_collection) からダウンロードできる。

Althammer S, González-Vallinas J, Ballaré C, Beato M, Eyra E. 2011. Pyicos: a versatile toolkit for the analysis of high-throughput sequencing data. *Bioinformatics* **27**: 3333-3340.

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166-169.

Azzalin CM, Lingner J. 2006. The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. *Curr Biol*. **16**: 433-9.

Bailey TL, Williams N, Misleh C, Li WW. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res*. **34**: W369-73.

Balagopal V, Fluch L, Nissan T. 2012. Ways and means of eukaryotic mRNA decay. *Biochim Biophys Acta*. **1819**: 593-603.

Bhattacharya A, Czaplinski K, Trifillis P, He F, Jacobson A, Peltz SW. 2000. Characterization of the biochemical properties of the human Upf1 gene product that is involved in

nonsense-mediated mRNA decay. *RNA* **6**: 1226-35.

Culbertson MR, Leeds PF. 2003. Looking at mRNA decay pathways through the window of molecular evolution. *Curr Opin Genet Dev.* **13**: 207-14.

Durand S, Franks TM, and Lykke-Andersen J. 2016. Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. *Nat Commun.* **7**:12434.

Eberle AB, Lykke-Andersen S, Mühlemann O, Jensen TH. 2009. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat Struct Mol Biol.* **16**: 49-55.

Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet.* **9**: 397-405.

Fiorini F, Bagchi D, Le Hir H, Croquette V. 2015. Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. *Nat Commun.* **6**: 7581.

Fong JH, Murphy TD, Pruitt KD. 2013. Comparison of RefSeq protein-coding regions in human and vertebrate genomes. *BMC Genomics*. **14**: 654.

Franks TM, Singh G, Lykke-Andersen J. 2010. Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. *Cell* **143**: 938-50.

Gong C, Maquat LE. 2011. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**: 284-8.

Gong C, Tang Y, Maquat LE. 2013. mRNA-mRNA duplexes that autoelicit Staufen1-mediated mRNA decay. *Nat Struct Mol Biol*. **20**: 1214-20.

Gregersen LH, Schueler M, Munschauer M, Mastrobuoni G, Chen W, Kempa S, Dieterich C, Landthaler M. 2014. MOV10 is a 5' to 3' RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3' UTRs. *Mol. cell* **54**: 573-585.

Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. 2012. GENCODE: the reference human genome annotation

for The ENCODE Project. *Genome Res.* **22**: 1760-74.

Huang W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocol.* **4**: 44-57.

Huntzinger E, Kashima I, Fauser M, Saulière J, Izaurralde E. 2008. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **14**: 2609-17.

Hurt JA, Robertson AD, Burge CB. 2013. Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Res.* **23**: 1636-50.

Imamachi N, Tani H, Akimitsu N. 2012. Up-frameshift protein 1 (UPF1): multitalented entertainer in RNA decay. *Drug Discov Ther* **6**: 55-61.

Imamachi N, Tani H, Mizutani R, Imamura K, Irie T, Suzuki Y, Akimitsu N. 2014. BRIC-seq: a genome-wide approach for determining RNA stability in mammalian cells. *Methods.* **67**: 55-63.

Isken O, Kim YK, Hosoda N, Mayeur GL, Hershey JW, Maquat LE. 2008. Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell* **133**: 314-327.

Kashima I, Yamashita A, Izumi N, Kataoka N, Morishita R, Hoshino S, Ohno M, Dreyfuss G, Ohno S. 2006. Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.* **20**: 355-367.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology.* **14**: R36.

Kim YK, Furic L, Desgroseillers L, Maquat LE. 2005. Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell* **120**: 195-208.

Kurosaki T, Li W, Hoque M, Popp MW, Ermolenko DN, Tian B, Maquat LE. 2014. A post-

translational regulatory switch on UPF1 controls targeted mRNA degradation. *Genes Dev.* **28**: 1900-16.

Kurosaki T, Maquat LE. 2016. Nonsense-mediated mRNA decay in humans at a glance. *J Cell Sci.* **129**: 461-7.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**: 357-359.

Lee SR, Pratt GA, Martinez FJ, Yeo GW, Lykke-Andersen J. 2015. Target discrimination in nonsense-mediated mRNA decay requires Upf1 ATPase activity. *Mol. Cell* **59**: 413-425.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat. Rev. Mol. Cell Biol.* **16**: 665-677.

Maekawa S, Imamachi N, Irie T, Tani H, Matsumoto K, Mizutani R, Imamura K, Kakeda M, Yada T, Sugano S, et al. 2015. Analysis of RNA decay factor mediated RNA stability contributions on RNA abundance. *BMC Genomics.* **16**: 154.

McIlwain DR, Pan Q, Reilly PT, Elia AJ, McCracken S, Wakeham AC, Itie-Youten A, Blencowe BJ, Mak TW. 2010. Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. *Proc Natl Acad Sci*. **107**: 12186-12191.

Medghalchi SM, Frischmeyer PA, Mendell JT, Kelly AG, Lawler AM, Dietz HC. 2001. Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum Mol Genet*. **10**: 99-105.

Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. 2004. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet*. **36**: 1073-1078.

Morris AR, Mukherjee N, Keene JD. 2008. Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets. *Mol Cell Biol*. **28**: 4093-4103.

Schoenberg DR, Maquat LE. 2012. Regulation of cytoplasmic mRNA decay. *Nat Rev Genet.* **13**: 246-59.

Schweingruber C, Rufener SC, Zund D, Yamashita A, Muhlemann O. 2013. Nonsense-mediated mRNA decay—mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim Biophys Acta.* **1829**: 612-623.

Spasic M, Friedel CC, Schott J, Kreth J, Leppek K, Hofmann S, Ozgur S, Stoecklin G. 2011. Genome-wide assessment of AU-rich elements by the AREScore algorithm. *PLoS Genet.* **8**: e1002433.

Sugimoto Y, Vigilante A, Darbo E, Zirra A, Militti C, D'Ambrogio A, Luscombe NM, Ule J. 2015. hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature.* **519**: 491-494.

Tani, H. Mizutani R, Salam KA, Tano K, Ijiri K, Wakamatsu A, Isogai T, Suzuki Y, Akimitsu N. 2012. Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome Res.* **22**: 947-56.

Tani H, Imamachi N, Salam KA, Mizutani R, Ijiri K, Irie T, Yada T, Suzuki Y, Akimitsu N. 2012. Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA Biol.* **9**: 1370-9.

Taylor MS, LaCava J, Mita P, Molloy KR, Huang CR, Li D, Adney EM, Jiang H, Burns KH, Chait BT, et al. 2013. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* **155**: 1034-48.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**: 511-515.

Trcek T, Sato H, Singer RH, Maquat, LE. 2013. Temporal and spatial characterization of nonsense-mediated mRNA decay. *Genes & development*, **27**: 541-551.

Unterholzner L, Izaurralde E. 2004. SMG7 acts as a molecular link between mRNA

surveillance and mRNA decay. *Mol Cell*. **16**: 587-96.

Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K, et al. 2016. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. methods*, **13**: 508-514.

Viegas MH, Gehring NH, Breit S, Hentze MW, Kulozik AE. 2007. The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the nonsense mediated decay pathway. *Nucleic Acids Res*. **35**: 4542-4551.

Wu X, Brewer G. 2012. The regulation of mRNA stability in mammalian cells: 2.0. *Gene* **500**: 10-21.

Zünd D, Gruber AR, Zavolan M, Mühlemann O. 2013. Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. *Nat Struct Mol Biol*. **20**: 936-43.

## 謝辞

本研究を遂行するにあたり、素晴らしい実験環境や解析環境を与えてくださり、貴重な時間を割いて研究について、懇切丁寧にご指導頂きました東京大学アイソトープ総合センター 教授 秋光信佳先生に心から感謝いたします。

東京大学大学院新領域創成科学研究科 教授 鈴木穰先生には次世代シーケンサーを用いた解析を行う上で、さまざまなご指導をいただきました。鈴木先生のご指導・ご協力のもと、次世代シーケンサーを用いたトランスクリプトーム解析を実現することができました。心から感謝いたします。

さらに、Dr. Kazi Abdus Salam には、RNA 免疫沈降法を教えていただき、次世代シーケンサーと組み合わせることにより、RNA とタンパク質の相互作用を網羅的に解析することができました。

また、私が所属する東京大学アイソトープ総合センター 秋光研究室の諸先輩方・後輩の皆さんには、日々の研究活動の中で、有意義なご意見、ご助言、ご批判をいただきました。深くお礼申し上げます。

最後に、私をここまで育ててくださり、精神面でも支えてくれた両親をはじめ、家族に心から感謝します。