包括的な翻訳解析を可能にするリボソームプロファイリング

イルミナの大量並行シーケンスは、ゲノムワイドな*in vivo*タンパク質合成の解析を、個々の塩基レベルに推し進めます。

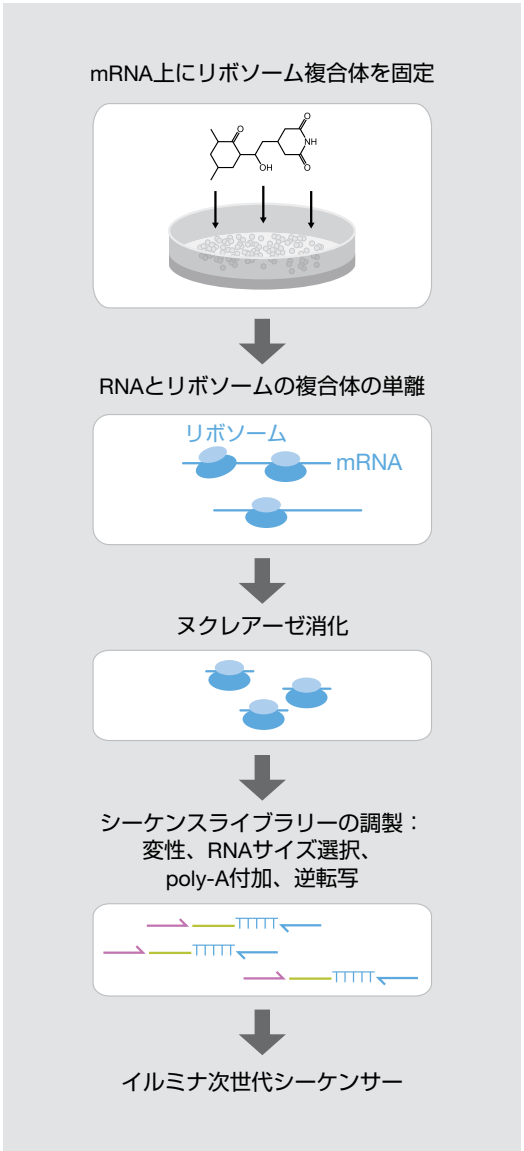
はじめに

マイクロアレイは、数千種類の遺伝子の発現レベルを同時に測定できます[1]。マイクロアレイはmRNA解析を変革したツールですが、これは細胞内で生じているタンパク質合成の実際の量を表すものではなく、その間接的な指標にすぎません。さまざまなレベルの転写調節および転写後調節が、プロテオーム解析を難しいものにしています。多様な調節イベントや翻訳効率に影響を及ぼす他の機構（mRNAスプライシング、リボソーム停止、非標準的な翻訳開始など）について検討するためには、遺伝子発現をタンパク質合成レベルで解析する直接的な手法が必要です。

リボソームフットプリンティングは、標的mRNAに結合した活性型リボソームの位置を同定するための確立された手法です。ヌクレアーゼ消化を用い、mRNAテンプレートの約30 ntの保護領域を解析することにより、リボソームの位置と翻訳されたmRNAを正確に決定できます[2]。mRNA断片をDNAに変換して大量並行シーケンスを行うと、活発に翻訳されているタンパク質を表す数百万の短いリードが得られます。図1に示すように、リボソームプロファイルのワークフローでは、まずシクロヘキシミド処理によりmRNAテンプレートの翻訳過程にあるリボソーム複合体を固定します。RNAとリボソームの複合体を回収してヌクレアーゼで消化し、余分なRNAを切り落とします。リボソームにより保護されたmRNA断片を抽出し、ゲル電気泳動でサイズ選択を行います。このようにして得られる5000万～1億の精製RNAプールは、翻訳過程にあるすべての細胞内のリボソームを表します。この28 bpの断片にpoly-Aを付加して逆転写を行い、イルミナ次世代シーケンサー用のcDNAライブラリーを調製します。シーケンスを決定することでリボソームにより読み取られているmRNA領域が正確に分かり、各RNAのリボソーム密度を調べることでタンパク質合成速度が測定されます。タンパク質合成速度は種々の条件下で測定することができ、タンパク質産生の様子を表す細胞のスナップショットが得られます。

最近Ingoliaらは、通常条件下とアミノ酸飢餓ストレス下で発育させた出芽酵母（*Saccharomyces cerevisiae*）におけるタンパク質合成を検討することにより、リボソームプロファイリングの原理実証を行いました[3]。この研究では、タンパク質の量に影響を及ぼす翻訳制御が広範になされていること、ペプチド伸長の前後期でリボソーム密度が異なること、非標準的な翻訳開始部位における翻訳開始が予想外に増加していることが観察されました。1回のランで数億ものシーケンスリードが得られるイルミナ次世代シーケンサーを用いれば、高スループットのシーケンスをトランスクリプトミクスやプロテオミクスに幅広く利用するという戦略的進歩が実現します。本稿に詳述されているように、リボソームプロファイリングは細胞全体のタンパク質産生のモニタリングに不可欠なツールとなり、生物学研究と臨床研究のほぼすべての領域に対して情報をもたらす可能性があります。

図1：リボソームプロファイリングのワークフロー



リボソームと結合したmRNAを細胞から精製し、サイズ選択、poly-A付加、逆転写を行って、イルミナシーケンスに使用できるライブラリーを調製します。

結果およびデータ解析

リーディングフレームの特異性の評価

リボソームプロファイリングアッセイと次世代シーケンサーを用いて出芽酵母から合計4200万リードが得られました。うち16%（約700万）が既知のCDS領域にアライメントされ、残りはリボソームRNA（rRNA）由来でした。著者らは次に、rRNAコンタミネーションの大半がサブトラクティブハイブリダイゼーションにより除去できることを示しました。どのmRNAについても、シーケンスで得られたリボソームの位置を調べてみると、転写産物のタンパク質コード領域のコドンから始まり、開始コドンから終止コドンにわたっていました。リボゾームが結合している位置は、各コドン内の塩基配列の位置に関して強い選択性を示し、リーディングフレームを統計学的に推測することが可能でした。

タンパク質合成速度の測定

リボソーム結合の強さ（密度）をmRNAの総量と比較してタンパク質の発現レベルを推定することにより、タンパク質合成速度を定量しました。種々の酵母遺伝子間の翻訳効率の差は約100倍の範囲であり、翻訳が行われていない遺伝子クラスも認められました。翻訳効率は、mRNA量の測定だけでは解明できない生物学的に情報量の多い現象である遺伝子発現のダイナミックレンジに大きく寄与していました。

タンパク質翻訳速度の計算値は、RNA量の測定値より質量分析によるタンパク質絶対量の測定値とより高い相関が得られます。したがって、全タンパク量の予測において、タンパク質合成速度はmRNAレベルの測定よりもよりよい指標となります（図2）。リボソームプロファイリングによるタンパク質合成速度の測定は、質量分析により測定される全タンパク量との関連で、制御されたタンパク質分解についての新たな情報をもたらすことができます[4]。

リボソーム密度の測定

これまでの報告で[5]、短いmRNA転写産物はリボソーム密度が高い（リボゾームの結合が強い）というエビデンスが得られていますが、何がこの傾向をもたらすのか、また実際の翻訳とどのように関連するのかは不明でした。リボソームプロファイリングにより、リボソームはすべての転写産物の5' 領域に選択的に局在し、短い遺伝子ではその全長に対する5' 領域の割合が高いことが示されました。特に、最初の30～40コドンにおけるリボソーム密度は3倍であり、その後は低下して転写産物の全長にわたり終止まで均一な密度を示しました。これは翻訳に一般的な特徴であり、コーディング配列の長さや翻訳レベルには依存しないと考えられます。この影響を考慮してタンパク質合成速度の推定値を補正すると、リボソーム密度とタンパク量の相関がかなり改善されます。

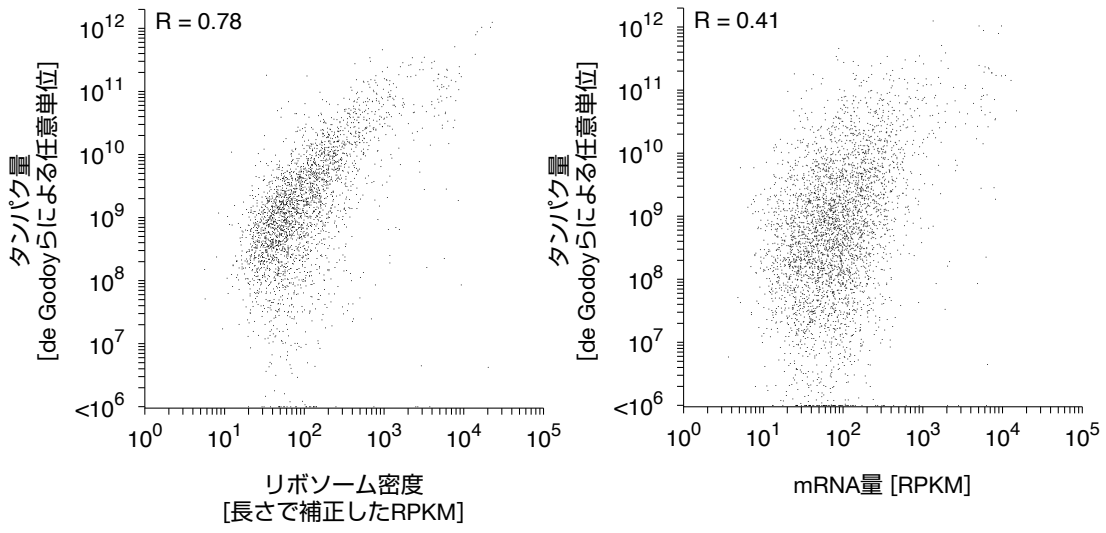
新規uORFの発見

酵母の実験では、全体の2%未満を占める説明のつかないリボソーム結合位置の大半が5' UTRに存在していました。5' UTRで生じている翻訳は、上流オープンリーディングフレーム（uORF）によるものと考えられました。翻訳調節に関与しうるCDSの上流の短い翻訳配列です。著者らは翻訳されているuORFを153個発見しましたが、先行研究などで知見があるものはそのうち30個のみでした。

非標準的な翻訳開始部位の解析

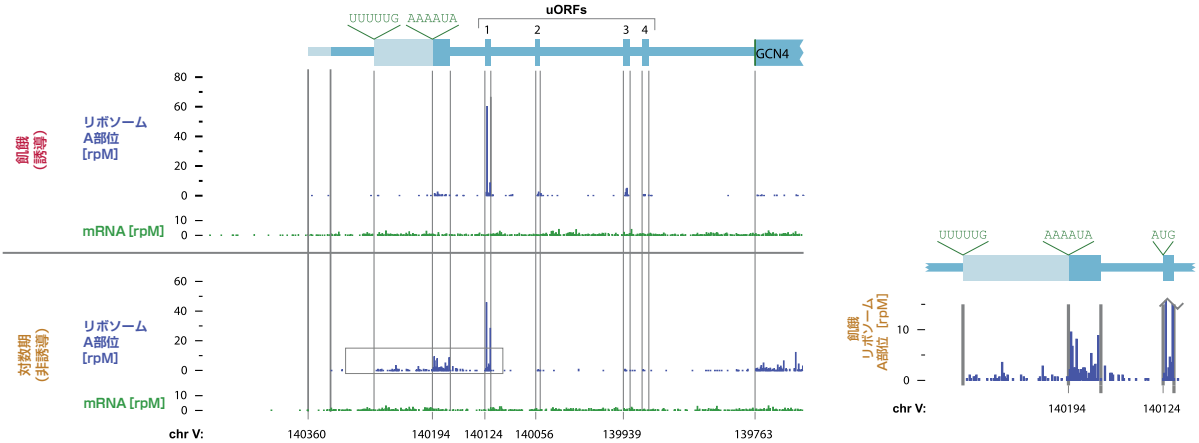
5' UTRにおける翻訳の別の重要な原因として、強力な翻訳開始配列に隣接する非標準的な部位における翻訳開始があります[6]。酵母ゲノムにおける既知の例から得られる配列情報を基に、他の非AUG翻訳開始部位の候補を同定できました。合計143個の非AUG ORFが5' UTR翻訳イベントに関与することが分かり、ある種の好都合な非AUG部位における翻訳開始が規則的に生じていることが示唆されました。

図2：タンパク量はmRNAよりもリボソーム密度とよく相関する



質量分析により測定したタンパク量（ハプロイドサンプルにおけるイオン強度の合計[4]）とリボソーム密度（左パネル、R=0.78）およびmRNA量測定値（右パネル、R=0.41）の相関の比較。

図3：出芽酵母 *S. cerevisiae* における飢餓条件中の5' UTR翻訳の変化



左パネル：抑制（対数期、上）および誘導（飢餓、下）条件下でのGCN45' UTRのリボソーム（青）およびmRNA（緑）の密度。4個の既知のuORFを示してあります。
右パネル：左パネルの枠囲み部分の拡大。上流翻訳開始部位と考えられる配列（UUUUUGおよびAAAAUA）におけるリボソーム密度を示します。
出典：Ingolia *et al.* 2009 Science 324: 218–223. AAASの許可を得て転載。

対数期および飢餓条件下での翻訳の解析

アミノ酸飢餓条件下の酵母では、全体的なタンパク質合成が大幅に低下して特定の標的遺伝子の発現が上昇するなど、転写および翻訳が大きく変化します[7,8]。リボソームプロファイリングにより、翻訳調節を受ける遺伝子としてよく研究されているGCN4で生じる7倍の翻訳誘導について詳細に解析することができました。GCN4の翻訳調節は4個の5' UTR uORFにより生じます。図3の左パネルに示すように、GCN4のuORF 1の翻訳が生じ、uORF2～4でも若干の翻訳が見られます。しかし、対数増殖期にはGCN4の主要なCDSの翻訳はほとんど認められませんでした。このリボソーム占有のパターンは、uORF 1が構成的に翻訳されるが下流での再開も許容されるという、GCN4 5' UTR機能の標準的なモデルと一致します。対数増殖期には、GCN4自体よりもuORF 2～4で再開が生じます。飢餓条件では、抑制的なuORFのリボソーム占有率の低下が観察され、CDSでの再開が増加し、翻訳抑制が解除されてタンパク質コード領域の翻訳が増加します（図3、右パネル）。対数増殖期には、解析したuORFの5' UTR上流でさらに翻訳が観察され、この翻訳は飢餓条件下で大幅に増強されました（図3）。翻訳は非標準的なAAAAUA部位で開始し、上流in-frameの非標準的なUUUUUG部位でも開始していました。この領域にまたがる配列はuORF 1の制御に重要であり[9]、新規の非AUG uORFの機能的役割を示唆しています。

結論と展望

アミノ酸飢餓条件下の酵母では、特に5' UTRにおいて、リボソーム結合領域の大幅な増加が認められました。この条件下では非飢餓条件下と比較して、非AUG翻訳開始部位を有するuORFでリボソーム密度が極めて大きく上昇しました。この現象は、幅広い翻訳制御因子の作用を介して、非標準的な翻訳開始部位の厳密性に全般的な変化が生じることが原因と考えられます。

リボソームプロファイリングは新しい手法であり、細胞の翻訳過程の系統的なモニタリングとタンパク量の予測を可能にします。極めて正確な位置情報に基づき、uORF、非標準的な翻訳開始部位における翻訳開始、タンパク質合成におけるその他のイベント（プログラムフレームシフト、終止コドンの読み過ごし、制御されたタンパク質分解、早期翻訳終了、スプライスバリアントの発現）などの翻訳制御要素を発見できます。高解像度の遺伝子特異的リボソーム密度プロファイルにより、翻訳速度、リボソーム停止、タンパク質合成、折り畳みの解析が推進されます。

イルミナの大量並行シーケンスを活用したリボソームプロファイリングにより、タンパク質産生の詳細かつ正確な*in vivo*解析が実現します。リボソームプロファイリングの手法は多くの生物種に適用でき、組織特異的な翻訳プロファイリングを含めるように改変することも可能です[10]。リボソームプロファイリングの将来的なアプリケーションには、遺伝子発現の翻訳制御に関する幅広い研究が含まれます。特に、mRNAのどの領域が翻訳されているのかを決定することは、複雑な生物種のプロテオームを定義したり、正常状態と疾患状態の分子生物学的解析を進めたりするのに役立つと考えられます[11]。

参考文献

1. Brown PO, Botstein D (1999) Exploring the new world of the genome with DNA microarrays. Nat Genet 21: 33–37.
2. Steitz JA (1969) Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. Nature 224: 957–964.
3. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman J (2009) Genome wide analysis *in vivo* of translation with nucleotide resolution using ribosomal profiling. Science 324: 218–223.

4. de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, et al. (2008) Comprehensive mass spectrometry-based proteome quantification of haploid versus diploid yeast. Nature 455: 1251–1254.
5. Arava Y, Wang Y, Storey JD, Brown PO, Herschlag D (2003) Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 100: 3889–3894.
6. Meijer HA, Thomas AA (2002) Control of eukaryotic protein synthesis by upstream open reading frames in the 5' untranslated region of an mRNA. Biochem J 367: 1–11.
7. Hinnebusch AG (1988) Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. Microbiol Rev 52: 248–273.
8. Hinnebusch AG (2005) Translational regulation of GCN4 and the general amino acid control of yeast. Annu Rev Microbiol 59: 407–450.
9. Grant CM, Miller PF, Hinnebusch AG (1995) Sequences 5' of the first upstream open reading frame in GCN4 mRNA are required for efficient translational reinitiation. Nucleic Acids Res 23: 3980–3988.
10. Heiman M, Schaefer A, Gong A, Peterson JD, Day M, et al. (2008) A translational profiling approach for the molecular characterization of CNS cell types.Cell 135: 738–748.
11. Holcik M, Sonenberg N (2005) Translational control in stress and apoptosis. Nat Rev Mol Cell Biol 6: 318–327.

イルミナ株式会社

〒108-0014
東京都港区芝5-36-7　三田ベルジュビル22階
Tel (03)4578-2800　Fax (03)4578-2810
www.illuminakk.co.jp

代理店

本製品の使用目的は研究に限定されます。

© 2013 Illumina, Inc. All rights reserved.

Illumina, illumina*Dx*, BaseSpace, BeadArray, BeadXpress, cBot, CSPro, DASL, DesignStudio, Eco, GAIIx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, NuPCR, SeqMonitor, Solexa, TruSeq, TruSight, VeraCode, the pumpkin orange color, the Genetic Energy streaming bases design は Illumina, Inc の商標または登録商標です。
その他の会社名や商品名は、各社の商標または登録商標です。予告なしに仕様を変更する場合があります。

Pub. No. 470-2009-J001 20JUN11

illumina®