

論文の内容の要旨

論文題目 高出力次世代シーケンス技術を応用した小胞体ストレス応答性 mRNA の網羅的解析

氏 名 奥田 哲夫

真核細胞はグルコース飢餓などの栄養枯渇や低酸素状態、細胞内カルシウムの恒常性および酸化還元状態を乱すような環境ストレスへの曝露や遺伝子変異の存在により小胞体内に折りたたみ構造の異常なタンパク質が蓄積される。このようなストレスは小胞体ストレスと呼ばれ、近年、神経変性疾患、2型糖尿病、動脈硬化や癌といった疾病への関与が示されてきた。それ故、小胞体ストレス応答の分子メカニズムの解明が疾病の発症メカニズムの解明に繋がり、新たな治療法開発に貢献することが期待されている。

細胞内では小胞体ストレスに対する生体防御機構として2つの主要経路が機能する。第一に、小胞体内の異常タンパク質を正常化するために、小胞体シャペロンや酵素を転写レベルで誘導する IRE1 α -XBP1 経路が存在する。この経路はターゲットである小胞体シャペロン BiP の増加に伴い、負のフィードバック制御により不活性化される。第二に、全般的な翻訳抑制により小胞体への新生タンパク質のトランスロケーションを低下させる PERK-eIF2 α 経路が存在する。この PERK-eIF2 α 経路は翻訳開始複合体の形成阻害を通して翻訳の開始段階を抑制するが、5'UTR 内に uORF を有する mRNA に関しては逆に翻訳を促進することが知られ、その代表が ATF4 である。ATF4 はストレス後のリカバリーに関わる遺伝子群の転写を誘導するだけでなく、アポトーシス誘導に関与する転写調節因子 CHOP の発現も誘導する。

このように、小胞体ストレス応答は全般的な翻訳抑制状態において転写誘導を行う特異な経路である。最近、IRE1 α -XBP1 経路の不活性化後も PERK-eIF2 α 経路が恒常的に活性化していることが、小胞体ストレス下の細胞運命の決定と関係していることが報告された。そこで我々は、IRE1 α -XBP1 経路の活性状況下と不活性状況下それぞれにおいて、翻訳抑制の網を掻い潜って、タンパク質に翻訳されて機能していると考えられる、Polysome 形成 mRNA の網羅的発現解析を行うことで、細胞運命の決定に関与する遺伝子を同定する計画を立てた。

従来、ゲノムワイドな発現解析はマイクロアレイを用いて行われてきた。しかしながら、伝統的なハイブリダイゼーション技術に依存した解析手法ゆえに、事前に鋳型としての配列情報が必

要になるばかりでなく、非相同配列間におけるミスハイブリダイゼーションに由来するノイズによって高いバックグラウンドが生じてしまう。そのため、既知の配列情報においてのみ解析が可能であり、また、高コピー数の遺伝子の発現変化を検出することには適しているが、低コピー数の遺伝子の発現変化を検出することは困難であった。この問題を解決する方法として、近年、注目されているのが、全 RNA シークエンスと呼ばれる手法である。この解析手法を可能としたのは、従来のサンガー法に依存した電気泳動による配列決定とは全く原理を異にする、複数の次世代シークエンス技術である。その 1 つである Illumina 社製 1G Genome Analyzer (通称 Solexa) は 1 Run につき最大約 10 億塩基を読むことが可能な高出力を誇る。この Solexa の原理は、8 サンプル分のレーンを有するプレート上で、ブリッジ PCR により形成されたクラスターを可逆的ターミネーター法により 1 塩基ずつ発光させ、その画像を取り込む作業を繰り返し、約 36 塩基の配列情報を 1 つのタグとして決定することで、1 レーンにつき約 400~600 万タグを読む、というものである。決定された配列は Eland という software により、ヒトゲノム配列上にマップされることで、ゲノム上での位置が決定される。

Eland 解析の過程は、プローブ (読まれた配列) を鋳型 (ゲノム配列) に電子的にハイブリダイゼーションさせていることになるため、未知の転写産物 (non-coding RNA など) の同定が可能となるばかりではなく、読まれた配列がゲノム配列にマップされる頻度をカウントすることで、転写産物の発現量の変動を定量的に解析することが可能である。この定量法はマイクロアレイのようなバックグラウンドが生じないため、低発現量の転写産物の変化を検出することが可能となる。この原理に基づき、未知の転写産物を含めた、各転写産物の UTR、exon、intron、および splice junction などの詳細な解析を行う技術を全 RNA シークエンスと呼ぶ。これまで全 RNA シークエンスは細胞全体の RNA から分離された mRNA を原材料として実行されてきた。しかしながら、現時点において Polysome 形成 mRNA のランダムシークエンスは報告されていない。核内 mRNA を除外した細胞質の RNA や Polysome 形成 mRNA の発現解析を行うことはこれまでの報告との対比を行う上でも重要である。

ヒト大腸腺癌由来の細胞株 HT29 を、小胞体ストレス誘発剤として作用する N 結合型糖鎖形成阻害剤 Tunicamycin (Tm) で処理し、IRE1 α -XBP1 経路の活性化および不活性化の条件をそれぞれ 4 時間および 16 時間と決定した。各時点で HT29 を界面活性剤で可溶化後遠心により除核した。得られた細胞質成分および超遠心により分離した Polysome 画分より mRNA を抽出後、二本鎖 cDNA を合成し、圧力をかけて断片化した。各断片の末端を平滑化したのち、Solexa 用アダプターを両端に付加し、電気泳動により約 200bp の断片を分離した。この断片を PCR 増幅後濃度測定し、15pg (平均 mRNA 長を 2,000nt とした場合、低く見積もって約 1 細胞あたりの総 mRNA に相当する量) を Solexa 用プレート 1 レーンに対してアプライした。Solexa により 1 レーンあたり約 500 万タグが読まれ、Eland 解析により、ミスマッチなくマップされたタグの割合は、Refseq (タンパク質コード遺伝子の代表 mRNA/cDNA 配列) に対しては約 45%、ゲ

ノム配列に対しては約 50%であった。また、Refseq に対する Eland 解析で得た RefseqID (遺伝子の分類名の一つ) の数と比較し、ゲノム配列に対して Eland 解析を行った場合の Refseq ID の数は約 10%多かった。これは、細胞質由来の転写産物中にイントロンを含んだ転写産物もしくはイントロン由来の転写産物が含まれている可能性を示唆していた。この結果は、ストレスに依存せず、細胞質由来ばかりではなく Polysome 由来の cDNA においても得られた。従来、Splicing は核内で起こり、不完全な産物を核外に輸送させない mRNA 品質管理機構が存在するので、大部分の未成熟 mRNA は核内に存在するか分解され、細胞質中には成熟 mRNA が濃縮されていると考えられている。今回、タンパク質合成の場である Polysome にイントロンを含んだ転写産物が含まれる可能性があるという興味深い結果が得られた。

次に 18,001 種類の遺伝子に注目し、Refseq を用いた発現プロファイルを実行した。各遺伝子 (cDNA) の配列中にタグがマップされた場合に 1 カウントとして、遺伝子ごとに合計して頻度 (ppm で表示) を算出した。さらに各遺伝子の ppm を各遺伝子の cDNA 長を 200 で割った値で割ることによりカウントを mRNA についてノーマライズした。このとき、1 カウントは 1 細胞あたりに mRNA が 1 コピー存在することに相当する。この結果、1 細胞あたり 1 コピー以上 mRNA が存在する遺伝子は 7,394 種類、mRNA の総コピー数は約 10 万コピーであった。

ショットガンシーケンスの性格上、mRNA 長が長いほど、断片が増加することでカウントが増加し、その影響を受け、短い mRNA のカウントが減少してしまうことが懸念された。しかしながら、mRNA 長を 1,000bp ごとにクラス分けし、各クラスに存在する遺伝子数が全遺伝子数に占める割合と、各クラスの総コピー数が全体のコピー数に占める割合を比較した結果、短い mRNA 長のクラスのコピー数の減少は見られなかった。

また、Polysome は mRNA に複数のリボソームが結合して形成されるため、mRNA 長が長いほど結合するリボソーム数が増加すると考えられる。Polysome の分離は重さに依存して行うため、結合リボソーム数が多い長鎖 mRNA が Polysome に濃縮されることが予想された。このような実験の性質上起こりうるバイアスによって、本来の機能的な mRNA を同定することが難しくなれば、人工的な結果を招く恐れがある。しかしながら、細胞質において見られた各長さのクラスの発現量の分布と Polysome における各クラスの発現量の分布は類似していた。以前、mRNA 長が長いほど結合するリボソーム数は増加するが、その結合密度に注目すると、mRNA 長が短い程密度が高く、mRNA 長が長くなるに伴って密度が減少することが報告された。発現ライブラリーを作成するにあたって、各 Polysome 画分を 1 つに合わせて調整したため、上述の現象が反映し、細胞質で観察されたクラスの分布と類似の分布が観察されたと考えられた。

実際に小胞体ストレスによって 2 倍以上に発現量が増加した遺伝子に注目した場合、これらの遺伝子が所属するクラスは、細胞質においては短～中鎖クラスに属する遺伝子の増加が見られ、Polysome においては中～長鎖クラスに属する遺伝子の増加が見られた。細胞質と Polysome 両方で増加している遺伝子に注目すると、細胞質における分布が反映し、短～中鎖クラスに属する

遺伝子が抽出された。Polysome において見られた長鎖クラスの増加は、遺伝子機能の反映というよりも、翻訳の開始反応が阻害されていることで、短い転写産物がリボソームに取り込まれにくく、翻訳に時間のかかる長い転写産物に結合した状態のリボソームが乖離せずに残っていたことにより、見掛け上、増加しているように観測されと考えられた。

我々は Solexa 解析の出力が正当なものであるのかを評価するために発現の上昇した 93 遺伝子（既知の小胞体ストレス応答性遺伝子 BiP、ATF4 や CHOP を含む）について定量的リアルタイム PCR を実行した。解析対象を選出する基準として細胞質および Polysome 両方で 2 倍以上に発現量が増加していること、および Polysome においてのみ発現量が 2 倍以上に増加している遺伝子とした。これは、小胞体ストレスが転写誘導と翻訳抑制を誘発するからであり、また、uORF に代表される RNA の構造によって転写誘導に依存せずに翻訳が亢進される場合があるためである。

Solexa の結果とリアルタイム PCR の結果との相関は $R=0.7$ で、完ぺきな相関とは言えなかったが、将来的な解析対象を選ぶことは可能である。この検定で、全体としては約 3 割の不一致が生じたが、細胞質および Polysome 両方で発現増加が見られた遺伝子グループでは 2 割前後の不一致だった。このことから、細胞質および Polysome 両方の発現プロファイルを組み合わせることで、実際に機能的な遺伝子を同定することが容易になることが示唆された。この 93 遺伝子の中から、低コピー数の遺伝子も含め、新たな小胞体ストレス応答性遺伝子の候補が 22 遺伝子同定された。

Solexa 解析全体の結果から、検出された 7,323 遺伝子の内、小胞体ストレスによって細胞質において 2 倍以上に発現量が増加した遺伝子を抽出すると、刺激後 4 時間で約 2.0%、16 時間で約 4.5% の遺伝子が増加し、4 時間から 16 時間まで共通して増加した遺伝子は約 0.6% であった。逆に言うと、これらの遺伝子の内、約 80% が Polysome において一定量以上に保たれていた。この結果から、小胞体ストレス下に、細胞の生死の決定に関与する分子を機能させるには、その転写量を増加しさえすれば良いことが示唆された。この解析で ATF4 は Polysome においてのみ増加が見られたが、同様の挙動を示した遺伝子グループ内に ATF4 のように uORF を有した遺伝子は存在しなかった。このグループの検定結果が悪かったこととの関係は不明であるが、ATF4 のような制御機構はレアケースなのかもしれない。

今回、全 RNA シークエンスを応用することによって、マイクロアレイを用いた研究によって報告されてこなかった新規候補遺伝子の同定に成功した。これは、電子的なハイブリダイゼーションから獲得した高分解能により、低コピー数の遺伝子変化の検出が可能となったことが寄与している。IRE1 など低コピー数でありながら、重要な働きを担う分子が存在していることから、低コピー遺伝子の探索は重要な意義がある。また、未成熟 mRNA の翻訳やその調節機構および、RNA の新機能が存在する可能性も見出した、この高出力な解析能力は、RNA 研究の新展開の端緒を開く可能性を有しており、今後のさらなる発展が期待される。