

審査の結果の要旨

氏名 村上大勇

ディファレンシャル・ディスプレイ (以下 DD) は、mRNA より作成した cDNA (群) に、任意のプライマーと、3'-末端の Poly-A 配列に対応するプライマーとで PCR を行う実験法であり、遺伝子の機能解析と新規遺伝子を取得するのに、非常に有用である。また、近年、ゲノムプロジェクトの進展により、日々大量の遺伝子塩基配列データが蓄積され、巨大なコンピューター・データベースが作成されてきている。DD での泳動パターンを予測しておき、実際の DD の結果と比較できれば、実験の効率を飛躍的に向上させられる。そこで既知遺伝子についての DD パターンを遺伝子データベースより予測するシステムを考案し、“*in silico* Differential Display (*isDD*)” と名付けた。実際の DD パターンと、*isDD* パターンの比較により、特定バンドが既知のいずれの遺伝子に対応するか、また新規遺伝子であるか予測可能とすることを目的とする。本論文について以下の結果を得ている。

1. ヒト、マウス、ラットでは UniGene のそれぞれの DNA 塩基配列データベースファイルから、DD の実際の実験で検出出来る cDNA 3'-端 1000 塩基または 1500 塩基以内のコンプライト (全長) CDS および 3'-末端を含む cDNA エントリーを抜き出して FASTA 形式で保存し、*isDD* データベースとした。ヒト 8043 (平均長 934.3 nt)、マウス 5000 (平均長 946.2 nt)、ラット 3099 (平均長 938.2 nt) のエントリーの塩基配列が使用可能となった。

isDD の実証実験用に、パン酵母で推定 ORF 配列と EST 配列の結合を行い、*isDD* データベースを作成した。酵母推定 ORF 配列全体の 37.48%、2220 個の塩基配列を使って *isDD* を実行することが可能となった。

2. 通常 DD 実験ではランダムなプライマー配列が使われるが、DD 実験での電気泳動バンド数を多くし新規遺伝子の取得効率を最大限に上げるため、*isDD* データベース中でより多く現れるプライマー塩基配列をコンピュータープログラムにより総当りで調べ列挙した。その際、パンドローム禁止、3つ以上続いた T 配列を含まない、より多い T 配列を含んだものは除外、という条件付けをした。名古屋市立大学医学部医動物研究室の鈴木高史助手に、単一神経細胞から DD 実験を行っていただいたところ、市販キット (Stratagene) に添付されているプライマーよりも電気泳動バンド数が多く得ることができると示された。

3. より多くの cDNA の発現を確かめられるプライマーの組み合わせを、遺伝的アルゴリズム (GA) によって選択した。GA による選択では 99.70% 分の *isDD* データベース中のエントリーをカバーできたのに対し、上位 20 のプライマーをランダムに選択する場合 98.0% のカバー率に留まることが示された。

4. 選択されたバンドがどのような電気泳動パターンを示すかを予測するプログラムを作成した。*isDD* バンドビューは Windows ベースのプログラムで、フルスクリーンの DirectDraw 方式を使い

高速描画を行う。プログラムは実際の電気泳動ゲルの画像を白黒 256 階調の BMP (ウィンドウズ・ビットマップ)、フルカラー (24bit) のビットマップに変換したものを読み込み、シミュレーションのゲルイメージと重ね合わせられる。

5. WWW ベースのバンドビューワーシステムも用意した。isDD コモン・ゲートウェイ・インターフェース (CGI) は WWW サーバー上で実行される。クライアントの WWW ブラウザーでの表示は、ダイナミック HTML (DHTML) を使用したインタラクティブなアプローチを取っており、ウィンドウズ版と同様に isDD エントリー名をマウスカーソルが横切ると、対応するバンド位置が赤色で示される。エントリーをクリックするとそのエントリーの塩基配列が別ウィンドウで表示され、カット & ペーストすることで次のプログラムまたは Web サイトを使用して解析可能となる。
6. 市販酵母 RNA (Stratagene) を使用して実際の DD 実験を行った。合計 20 個のバンドをゲルより取得し、TA ベクターにサブクローニング後、それぞれのバンドについて 20 個の cDNA クローンを蛍光シーケンシング法で塩基配列を読んだところ、isDD データベースに含まれる ORF と EST のエントリー上で 5 塩基以上の精度で実際に存在したものが相当するクローン中の 4 個あり、さらに 2 クローンは EST のエントリー上でのみ一致した。それ以外の isDD の結果に出てこない 12 のクローンのうち少なくとも 6 個は Poly-A 配列を持った新規遺伝子で、酵母染色体上の推定 ORF 配列と推定 ORF 配列の 1000 塩基以上離れた場所に存在した。その 12 クローンのうち 9 はオープン・リーディング・フレームを欠いており、リボソーマル RNA のような翻訳されないで使用される RNA であると考えられる。以上の結果から isDD でのシミュレーションの効率は isDD で予測できるものにおいては 5 割、ORF 全体では約 20% 程度であるが、isDD に使用可能な 3'-端を含む遺伝子データベースが充実するか、推定 ORF の精度が上がる、EST の数が増える等の手法により改善可能であると考えられる。

以上、本論文は DD 実験における効率を高め、isDD でその遺伝子の名称を絞り込んで予測することができる。DD のゲルイメージで変化が観察されたバンドは取得して解析する価値があり、isDD で見られないバンドは新規遺伝子である可能性が高い。多数の遺伝子の発現の有無を、統一的に確認することが出来る。数本のプライマーで例えば酵母の場合、6 塩基配列のプライマー 8 種類の組み合わせで、全 Genbank と EST の isDD エントリー中の 99% 以上の発現を調べられると考えられる。これは、従来の莫大な量 (数千) のプライマーを用いて PCR を行い、DNA アレイ等を使用して調べる方法よりも、非常に安価で簡便に、さまざまな条件下での遺伝子発現についての知見が得られる。将来的には、シグナル伝達系、代謝系等のカテゴリー別のエントリー群が、どのようなバンドパターンを描くか予測して示すことが考えられる。また、オリゴ DNA の選択は DNA アレイに応用も可能である。このようにさまざまな応用が期待され、精度で優位性がある DD で単一神経細胞等での遺伝子発現に関する知見が得られると期待され、遺伝子発現に関する研究に重要な貢献をなすものと考えられ、学位の授与に値すると思われる。