

研究テーマ:マイクロアレイによるクラスタリングと逆相関解析

081 番 PHAN THI THU HANG

1 研究目的

今回の研究には、マイクロアレイデータでクラスタ解析や逆相関解析を通してバイオインフォマティクスの背景とデータ解析方法の理解を目的とする。さらに、自分で課題を見つけ、文献を探索・勉強し、実践的に実験を行うことにより研究という概念を正確に理解した上、研究方法を身につけるのも一つの目的である。

2 研究内容

本研究では、6000 個のマクロアレイデータで階層型クラスタ解析するプログラムを作成し、実験する。このプログラムに、クラスタ間距離やデータ間距離が自由に選べるように設計し、実験結果を表示ソフト TreeView3.0 で見れるように工夫する。さらに、マイクロアレイデータで特定の遺伝子と逆相関の発現パターンを示す遺伝子群を検出し、発現パターンが逆相関である遺伝子群とその遺伝子との抑制性などのような機能について考察する。

2.1 マイクロアレイの階層型クラスタ解析

マイクロアレイデータ解析においてクラスタ解析は有効的なツールの一つである。本研究で用いるクラスタ解析方法は集積的なアプローチの階層型クラスタ解析である。このクラスタ解析法の手順は次のように実現される。

1. 遺伝子間の距離行列を計算
2. 最も似た遺伝子/クラスタ A、B を探し
3. A と B を新クラスタ C にまとめ
4. C と他の全てのクラスタ間との距離を計算
5. ステップ 2 から 4 が全てのオブジェクトが 1 クラスタになるまで繰り返し

2.2 マクロアレイデータの逆相関分析

作成したプログラムの遺伝子間距離の結果を用いて、各々 RNA 安定性を制御する機能を持っている遺伝子とそれとある閾値以上に逆相関である遺伝子群を検出する。その後、検出した遺伝子群について GO Term Finder で共通機能を探し出して、見つけた機能とものと遺伝子の関連付けを考察する。

3 実験結果と考察

3.1 マイクロアレイの階層型クラスタ解析

6178 個の遺伝子発現パターンの階層型クラスタ解析した結果を TreeView3.0 で表示できた。Euclidean 距離での結果には、似た発現パターンがクラスタに収まらずに、ばらついた画像を表示したが、コサインと Pearson 相関で求めた結果は有意なクラスタを見つかることができた。

製作したプログラムと Cluster3.0 のクラスタリング結果を比較する。同じ条件、データ間の距離を Pearson 距離、クラスター間の距離を最近距離法の下でクラスタ解析を行った。その結果から、一つのクラスタを取り上げて

みる。図 1 と図 2 を見ると、クラスタに含まれる遺伝子が完全に一致したと分かった。製作したプログラムのクラスタ解析が正確できたと考えられる。

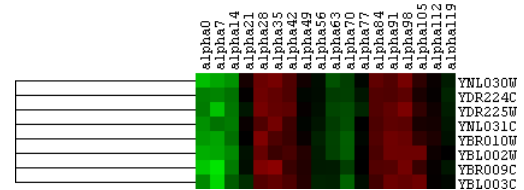


図 1: 作成したプログラムの結果 (Pearson 距離)

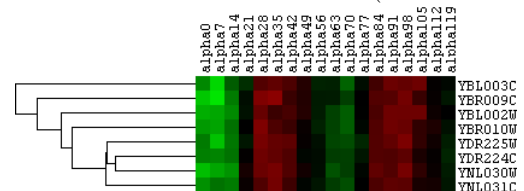


図 2: Cluster3.0 の結果 (Pearson 距離)

3.2 逆相関解析

逆相関解析の結果の一つを取り上げる。遺伝子 *YHR086W* から抽出された逆相関遺伝子群 *YEL032W*, *YER020W*, *YLR274W* を発見した。遺伝子 *MCM3*, *GPA2*, *CDC46* について GO Term Finder で見つけた有意な GO Term の意味から、*MCM3* と *CDC46* が細胞周期の *S* 期、DNA が複製された時に発現が強く、DNA 複製開始を促進させる働きを持つことが分かる。

それに対して RNA 結合機能を持つ遺伝子 *YHR086W* は *MER2* pre-mRNA のスプライシングの間に形成される DSB の形成に関わり、細胞の減数分裂の始めの段階に関わりがある。*YHR086W* の発現抑制ピークは *G1* 期であるので、*YHR086W* が DNA 複製開始に関わる遺伝子のスプライシングを介して DNA 複製開始を調節している可能性がある示唆された。

4 結論 - 感想

本研究では、マイクロアレイデータを階層型クラスタ解析するプログラムを製作し、6000 個のデータに対してクラスタ解析を行った。製作したプログラムは Cluster3.0 に比べてほぼ同じ結果出したが、計算時間の問題がある。逆相関解析には、逆相関である遺伝子群の有意性のある共通 GO Term から、mRNA のスプライシングを制御する遺伝子 *MRE2* は DNA 複製開始を調節している可能性があると予測した。

今回の自主課題研究を通して、研究をするには、課題の探りから課題の背景を理解した上、見つけた課題に対して何をすべき、どの方向を進むかなどを少し分かった。そして、データマイニング方法のクラスタリング解析、遺伝子情報科学の授業で多少紹介されたマイクロアレイについて理解を深めることができた。