

## 論文の内容の要旨

論文題目 遺伝子発現プロファイルを用いた  
腫瘍の分類に関する研究

指導教官 五十嵐 隆

東京大学大学院医学系研究科

平成 10 年 4 月入学

医学博士課程

生殖・発達・加齢医学専攻

氏名 堤 修一

### <背景>

ヒトゲノムの配列解読の進展や cDNA の解読に伴い、ヒトの遺伝子数はスプライシングを考慮しなければ、3 万から 4 万と推定されるに至った。また、ヒトの組織の遺伝子発現情報を網羅的に取得する手法として、マイクロアレイ技術は広い分野で使用されるようになってきた。特に高密度オリゴヌクレオチドアレイは対照検体を必要とせず、鋭敏な測定ができる反面、データの信頼性の解析が必要となる。しかし、このような基礎的研究は多くはない。腫瘍組織の遺伝子発現プロファイルは多ければ多いほど、情報として有用となるはずであるが、十分な種類のヒトの組織や腫瘍の遺伝子発現プロファイルデータがあるとは言えない状況である。また多量のデータはその量が増大するにしたがって、取り扱いも苦慮することになる。今研究は高密度オリゴヌクレオチドアレイを使用して、ヒト組織の個々の遺伝子発現プロファイルと混合された RNA から得られた遺伝子発現プロファイルを比較した。その結果、遺伝子発現の測定スコアと遺伝子量

について、ある程度の線形性が認められることを示した。また、遺伝子発現を示すスコアの信頼できる区間を示した。多量のプロファイルデータから注目すべき遺伝子群抽出のための一つのアルゴリズムを示した。その上で、白血病細胞株や固形腫瘍など腫瘍組織の発現プロファイルを得て、プロファイル情報のみからの腫瘍組織の判定のために遺伝子抽出を行った。

### <材料と方法>

1.ヒトの正常肝・肝硬変・肝細胞癌各 8 例のトータルRNAから Affymetrix の GeneChip 法を用いて約 12600 の遺伝子発現プロファイルを得た。また正常肝・肝硬変・肝細胞癌・C 型肝炎ウイルス(HCV)感染肝硬変・HCV感染肝細胞癌の各グループについてトータルRNAを等量ずつ混合させたプールサンプルを作製し、プールサンプルからも遺伝子発現プロファイルを得た。

2.加えて、肝芽腫 8 例、正常肝 2 例、慢性肝炎 6 例、肝硬変 2 例、肝細胞癌17例を得て同様に GeneChip 法により遺伝子発現プロファイルを得た。

3.クラスター解析とは、類似したパターンを隣接するように分配する方法である。類似度に応じて階層または樹形図を形成する階層型クラスター解析を 1.および 2.の個々のプロファイル全てを対象に施行した。

4.肝硬変・肝炎の個々の遺伝子発現プロファイルを HCV 感染群、HBV 感染群の 2 群に分け、各群に特異的に発現している遺伝子を抽出し、その統計的有意性について検証した。

5. t(4;11)を有する急性リンパ球性白血病(ALL)細胞株 3 種、t(11;19)を有する ALL 細胞株 3 種から GeneChip 法を用いて約 6000 の遺伝子発現プロファイルを得た。また、同型の microarray を使用している Whitehead の公開データから得られた ALL27 例・急性骨髄性白血病(AML) 12 例の遺伝子発現プロファイルを元にして、ALL 細胞株を未知のサンプルとして遺伝子発現からの判定を行った。

## <結果>

### 1. 混合サンプルにおける発現プロファイルの解析

GeneChip 法では遺伝子発現の値として、単一のスコア (Avg Diff)

が得られる。個々の発現プロファイルは同じ組織型の中でもある程度のばらつきを生じていた。個々のプロファイルからグループ毎に線形平均をとり、平均プロファイルを作成し、プールサンプルのプロファイルと比較した。平均プロファイルとプールサンプルのプロファイルは非常に良い相関を示した。閾値以下として切り捨てる Avg Diff 値が示され、加えて Avg Diff 値の精度を上げるための新たな Avg Diff 値の算出法を示した。

### 2. 多種類サンプルの発現プロファイル解析

遺伝子発現のパターンを元にサンプルの類似性から分類すると、腫瘍と非腫瘍で最も大きく分かれた。また、同じ組織型の腫瘍は隣接して分類される傾向を認めた。慢性肝炎・肝硬変群のみでサンプルをクラスター解析すると、HCV 感染と HBV 感染で主にクラスターが分かれ、統計的にも有意であった。しかし、肝細胞癌の遺伝子発現プロファイルは HCV 感染群と HBV 感染群では、非腫瘍部ほど有意な遺伝子発現の差は認められなかった。

### 3. 遺伝子発現パターンによる未知のサンプルの判定

Whitehead より得られた 39 例の遺伝子発現プロファイルと判定のための遺伝子リストを用いた ALL 細胞株の判定は、1 例が判定不能、5 例が ALL という結果であった。遺伝子発現情報だけから組織を判別する可能性が示唆された。また、対照サンプルを必要としない、GeneChip 法ではデータの互換性があることも示唆された。

## <考察>

相補的な DNA や RNA が互いに結合する原理を利用したマイクロアレイ法は、多数の遺伝子の発現を同時に測定する方法として開発された。測定されたデータも他方法にて検証され、遺伝子発現量への相関性が認められるに至っている。GeneChip 法に代

表される高密度オリゴヌクレオチド法はマイクロアレイの一種であるが、6～16対の 25mer プローブの組と独自の計算方法により、実験に対照サンプルを必要とせず単一な数値が得られる。データの信頼性としてそれぞれの遺伝子について、発現の差が 2 倍以上を有意な差として解析されることが一般的であった。またクラスター解析など、データの類似性を元にして腫瘍を分類する手法は、広く行われるようになってきた。

### 1. アレイデータの信頼性

GeneChip 法は、Avg Diff 値は、遺伝子の発現量に応じて相関する値となることが一部のプローブで報告されている。様々な値をとるプローブの Avg Diff 値の線形平均と、平均量存在していると考えられる遺伝子の Avg Diff 値が非常に近い値をとることは、プローブが対応する遺伝子の発現量に線形性をもって Avg Diff 値を示すことを示唆している。また、精度を上げるためにプローブの組を構成している 25mer プローブ対への新たな処理方法を示した。

### 2. 多種類サンプルの遺伝子発現情報解析

クラスター解析により、概ね組織型にしたがってサンプルが分類された。慢性肝炎・肝硬変群内においては、HCV 感染と HBV 感染で遺伝子発現のパターンが統計的に異なっていた。また HCV で高く発現している遺伝子はインターフェロン関連遺伝子群で、過去の報告と矛盾しない結果となった。

### 3. 遺伝子発現プロファイルによる未知のサンプルの判定

細胞株であり、かつ MLL 再構成のある今回のサンプルは、判定にはやや不適當であったが、6例中 5 例が ALL と判定された。また通常のマイクロアレイ実験では難しいとされるデータの互換性も示された。

マイクロアレイ法より得られる遺伝子発現プロファイルには、精度や解析方法で多くの改良が今もなされているが、本研究により一つのデータ取り扱い方法が示されたと考えられる。