

S+Arrayanalyzer による生物統計解析

石井一夫^{1,4}、青木幹雄²、木村徹²、柏崎安男²、松本美佐子^{3,4}、瀬谷司^{3,4}

1 東京理科大学 DDS 研究センター、2 大日本住友製薬株式会社、

3 北海道大学大学院医学研究科、4CREST

1. はじめに

ヒトゲノム解析終了後、DNA マイクロアレイ、SNPs 多型解析、MALDI-TOF-MASS など生体成分をゲノムレベルで網羅的に解析する技術が開発され、臨床での予後診断や治療方針の決定に応用されつつある。このような生体成分の解析は、大量のデータを用いた統計的解析が必須であり、各種の商用解析ソフトが利用されている。しかし、これらは、解析アルゴリズムが公開されていない場合も少なくなく、決められたフォーマット、出力形式を強いられることもあり、臨床での柔軟な解析に不向きな場合も少なくない。S+Arrayanalyzer は、オープンソースのゲノム解析ソフト Bioconductor の商用版であり、このような柔軟な解析に対応することを可能にする。

大阪府立成人病センター研究所では、肺癌患者の癌組織摘出手術後の再発防止、延命効果を目的として、癌患者に BCG-CWS (ウシ結核菌弱毒化ワクチン株細胞骨格成分) の単独投与による免疫アジュバント療法をおこなってきた。この効果は個人差があり、その判定として血液中の IFN γ の濃度を指標にしているが、十分な判定法ではない。そこで、今回、大日本住友製薬株式会社の協力のもと、BCG-CWS の標的である樹状細胞での、DNA マイクロアレイ (Gene Chip) による遺伝子発現の網羅的な解析を行うことにより、より有効な診断法を構築することを試みた。

まず、1) 健常者より血液を採取し、これより BCG-CWS の標的である樹状細胞を調製し、BCG-CWS に応答性のある遺伝子を解析した。さらに、2) 癌患者による BCG の応答性の違いを調べるため、癌患者より、末梢血単核球(PBMC)を採取し、RNA を精製した後、遺伝子発現解析をおこなった。今回、これらのうち、1) について S+Arrayanalyzer を用いてそのデータ解析をおこなったので報告する。

2. 方法

2.1 データの収集

BCG-CWS に応答性のある遺伝子を解析するため、健常者より血液を採取し、これより BCG-CWS の標的である樹状細胞を調製した。これに BCG - CWS、LPS、polyIC の各種アジュバントを投与し、一定時間培養後、その RNA を精製した。その後 Gene Chip™ (Human Genome U133A) による遺伝子発現解析をおこなった。

2.2 データ解析

データの解析は S-PLUS6.2 および S+Arravanalyzer2.1 (または、R2.3.0 および Bioconductor1.8) を用いておこなった。データの解析は、データの標準化 (スケール化)、フィルタリング (クオリティ評価、有意差検定)、変動解析、注釈づけ (アノテーション)、各種データ解析 (決定木、クラスター解析とヒートマップによる可視化) の順に行った。

1) データのフィルタリング

解析データは、Affymetrix™ MAS 5.0 (Affymetrix) により、標準化 (スケール化) をおこなった後、各解析のプロブセットの統計データの結果に基づき、すべての解析データが Absent Call と判定されたものを除いた (13758 プロブセットとなった。)。その後、各解析群について、コントロール

群に対して Wilcoxon の順位和検定を行い $p < 0.05$ のプローブセットを選択した (10772 プローブセットとなった。)。

2) データの変動解析

各データの変動は、各解析群の median 値とコントロール群の median 値の比で表現した。

3) 各種データ解析

各解析群に対して、決定木によるグループ分け、クラスター解析とヒートマップによる可視化、主成分分析による傾向の解析などを行い診断に有用な遺伝子の評価をおこなった。

4) 各プローブの注釈づけ (アノテーション)

解析した各プローブセットの注釈づけ (アノテーション) は、Affymetrix 社のウェブサイト (<http://www.affymetrix.com/index.affx>) から情報を入手した。

3. 結果

1) 遺伝子の発現変動解析

今回、健常者における BCG-CWS に応答性のある遺伝子の解析の結果を以下に紹介する。

表 1 には、各薬剤 (BCG - CWS, LPS (細菌の細胞壁成分), polyIC (RNA ウィルスに類似の化合物)) を処理した各時間での発現が変動した遺伝子のプローブセットで、その変動比が 2 倍以上、5 倍以上、2 分の 1 以下、5 分の 1 以下であったものの数を示した。このうち、いずれかの時間で 5 倍以上の変動が見られた 550 の遺伝子のプローブセットを以後の解析に用いた。

表1 各アジュバントの各時間における変動遺伝子プローブ数

2倍以上上昇したプローブセット					1/2倍以下減少したプローブセット					
Time	3h	8h	24h	Total	Time	3h	8h	24h	Total	
BCG	432	1441	781	1948	BCG	336	863	865	1450	
LPS	732		1299	1586	LPS	873		1615	2174	
polyIC	550		1318	1494	polyIC	426		1634	1898	
				Total					Total	3036
5倍以上上昇したプローブセット					1/5倍以下減少したプローブセット					
Time	3h	8h	24h	Total	Time	3h	8h	24h	Total	
BCG	85	143	118	238	BCG	25	112	140	230	
LPS	215		291	391	LPS	86		356	422	
polyIC	147		287	340	polyIC	22		315	332	
				Total					Total	604

2) 各種データ解析

今回、選択した 550 の遺伝子プローブセットについて、段階的クラスター解析おこないヒートマップで表示した結果を図 1 に示した。この図は横軸に各薬剤とその反応時間を、縦に 550 の遺伝子プローブセットを配置している。この図のうち暖色であるほど遺伝子の発現量が多いことを示す。このように、LPS と polyIC は、似たようなパターンを示すことがわかる。

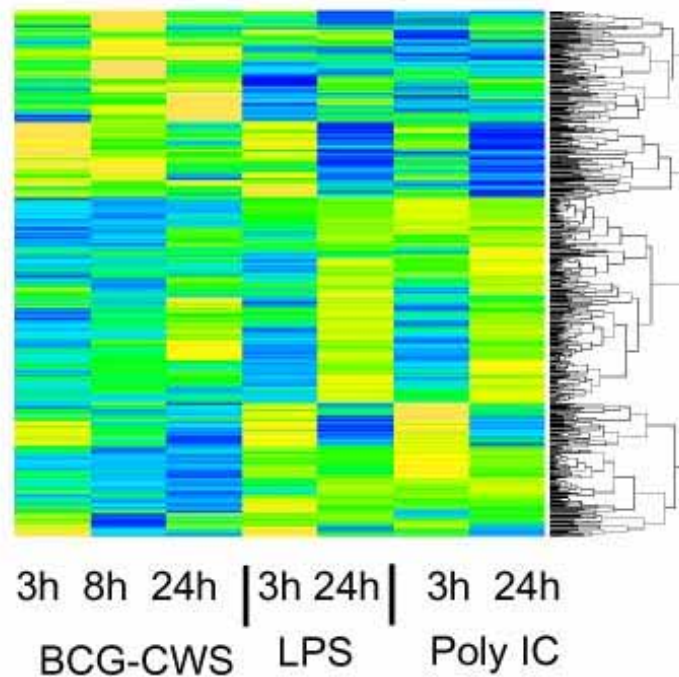


図1 各種の試薬で各時間のいずれかで5倍以上発現変動のみられた遺伝子のプローブセットの階層的クラスタ解析とそのヒートマップ。LPSとpolyICは似たようなパターンを示すことがわかる。

そのほか、今回データは示さないが、決定木によるグループ分け、主成分分析の解析もおこなった。これらのLPSとpolyICとの発現変動との比較による解析の集計により、BCG-CWSの樹状細胞での遺伝子発現は図2に示すようなシグナル伝達経路と発現パターンを示すことが示された。

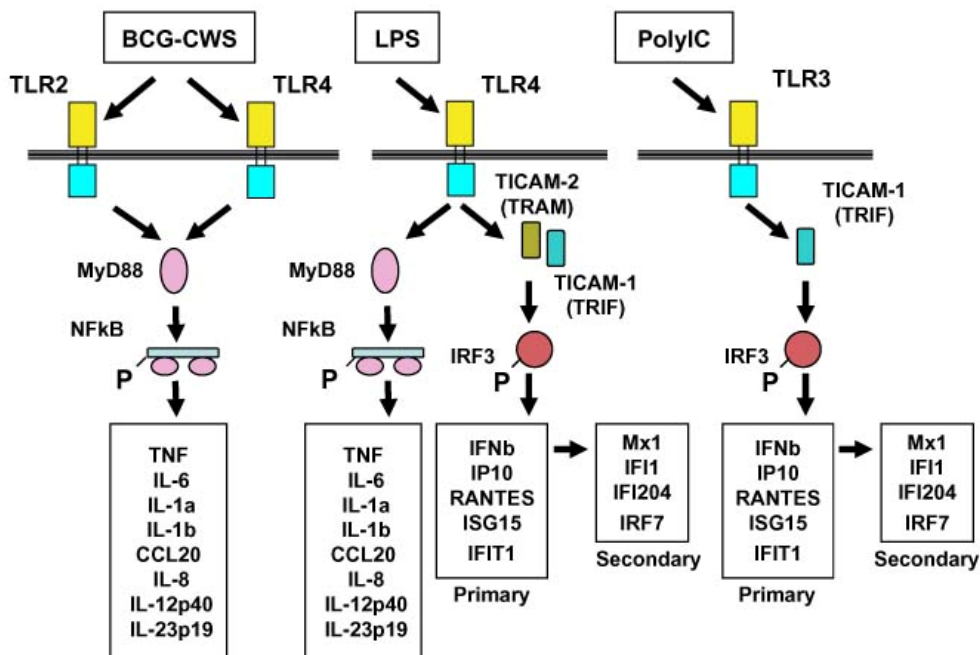


図2 各薬剤の細胞内でのシグナル経路と発現遺伝子

すなわち、BCG-CWS は、Toll 様受容体^{注)} (TLR)2 または TLR 4 による刺激を介して NF- κ B の活性化により、TNF、IL6、IL1 などの免疫活性化遺伝子が強く発現が誘導されたが、IFN- γ 、IP10、IFIT1 など (IFN 誘導遺伝子と呼ばれる) LPS や polyIC で発現が活性化された遺伝子の発現は強くなかった。これらの発現パターンは、BCG-CWS が LPS や polyIC にくらべてアナフィラキシーショックなどの副作用が弱く、より抗癌免疫剤として適していることを示している。

今回、Affymetrix 社の Gene ChipTM によるマイクロアレイ解析データを、S-PLUS と、S+Arravanalyzer で解析を行い、その解析に有用であることが示された。我々は、さらに、この解析システムを用い、癌患者より採取した末梢血単核球(PBMC)での遺伝子発現解析をおこない、癌患者に対する BCG-CWS の抗癌効果判定のより有効な診断法マーカーの解析も進めており、近日中に発表する予定である。

4. 謝辞

本研究は、独立行政法人科学技術振興機構 CREST「免疫難病・感染症等の先進医療技術」の事業としておこなわれた。

参考文献

1. Begum N.A., Ishii K., Kurita-Taniguchi M., Tanabe M., Kobayashi M., Moriwaki Y., Matsumoto M., Fukumori Y., Azuma I., Toyoshima K., Seya T., Mycobacterium bovis BCG cell wall-specific differentially expressed genes identified by differential display and cDNA subtraction in human macrophages. *Infect Immun.* 2004, **72**, 937-948.
2. Ishii K., Kurita-Taniguchi M., Aoki M., Kimura T., Kashiwazaki Y., Matsumoto M., Seya T., Gene-inducing program of human dendritic cells in response to BCG cell-wall skeleton (CWS), which reflects adjuvancy required for tumor immunotherapy., *Immunol Lett.*, 2005, **98**, 280-290.
3. Ishii, K., Kurita-Taniguchi, M., Aoki, M., Kimura, T., Kashiwazaki Y., Kato, A., Homma T., Saito H., Matsumoto, K., Matsumoto, M. and Seya T., The Global Scale Scenario of Dendritic Cell-Maturation Signaling and Toll-like Receptor-ligand Specific Innate Immunity Depicted by Microarray Analysis., submitted.
4. Gentleman, R., Carey, V., Huber, W., Irizarry, R., Dudoit, S. editors *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, 2005.
5. Gentleman R.C., Carey V.J., Bates D.M., Bolstad B., Dettling M., Dudoit S., Ellis B., Gautier L., Ge Y., Gentry J., Hornik K., Hothorn T., Huber W., Iacus S., Irizarry R., Leisch F., Li C., Maechler M., Rossini A.J., Sawitzki G., Smith C., Smyth G., Tierney L., Yang J.Y., Zhang J. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004, **5**, R80.

^{注)}用語説明 ; Toll 様受容体 (TLR) 免疫をつかさどる細胞の表面に存在し、病原菌やウイルスと反応して免疫反応を引き起こすたんぱく質。