

精巣胚細胞腫瘍に関する遺伝子発現データの統計的解析

年森彰子¹，堀本勝久²，清水邦夫³

¹ 慶應義塾大学大学院，² 産業技術総合研究所，³ 慶應義塾大学

1 概要

精巣胚細胞腫瘍は，組織学的にいくつかのタイプに分類されており，それらの腫瘍細胞について様々な研究がなされている．中でも胎児性癌は初期の胚発生と似たような分化をすることで注目されており，この胎児性癌がさらに他のタイプの腫瘍に分化するのではないかと推測されている．本研究では5つのタイプの精巣胚細胞腫瘍の細胞と前癌性の細胞，正常な精巣の組織の細胞の計7種類の細胞タイプの遺伝子発現データ間の正準相関係数，偏正準相関係数を計算し，各細胞タイプ間の関係性の強さを推定した．

2 精巣胚細胞腫瘍データの解析

2.1 データ

本研究で解析に用いたデータは，Skotheim，R.I. らによる論文 [4] で扱われた遺伝子発現データである．論文 [4] によると，Skotheim，R.I. らは Agilent Human 1A oligomicroarray を使い，ヒトの精巣の正常な組織 (N : Normal testicular parenchyma，3 個体)，前癌性で非侵襲性 (健康な組織を侵さない) の組織 (I : Intratubular germ cell neoplasia (IGCN)，3 個体)，そして5タイプの精巣胚細胞腫瘍の細胞内の遺伝子発現量を測定している．5つの腫瘍のタイプはセミノーマ (S : Seminoma，3 個体)，胎児性癌 (E : Embryonal carcinoma，5 個体)，絨毛癌 (C : Choriocarcinoma，1 個体)，奇形腫 (T : Teratoma，4 個体)，卵黄嚢癌 (Y : Yolk sac tumor，4 個体) である．この遺伝子発現データは Gene Expression Omnibus の accession no. GSE1818 で参照することができる (<http://www.ncbi.nlm.nih.gov/geo/>)．さらにこの遺伝子発現データから，Significance analysis of microarrays [5] によりこれら7タイプの細胞それぞれで特異的な遺伝子を計 678 個抽出している．本研究では，この 678 の遺伝子発現データを用いることにする．

2.2 ソフトウェア

データの解析においては統計解析ソフト R 2.7.1 (<http://www.r-project.org/index.html>) を用いて行った．また，複数の遺伝子の GO Term を一度に検索する際には GO Term Finder[1] を用いた．

2.3 精巣胚細胞腫瘍の形成

精巣胚細胞腫瘍は，まず IGCN が Embryonal carcinoma と Seminoma に分化し，さらに Embryonal carcinoma が Choriocarcinoma，Teratoma，Yolk sac tumor に分化すると推測されている．図にあらわすと図1のようになる [4]．

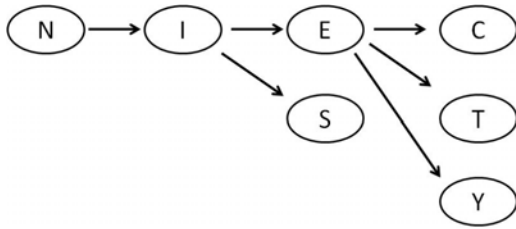


図 1: 精巣胚細胞腫瘍の形成

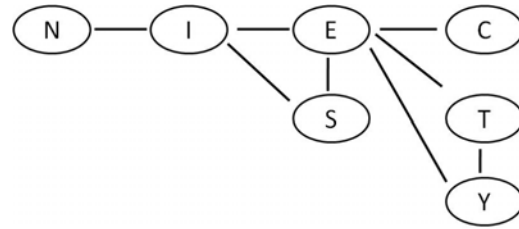


図 2: 直接的な関係性が強いことが示唆された細胞タイプを結んだグラフ

2.4 解析結果，考察

7 タイプの細胞の関係性を説明するために，それぞれのタイプに属する細胞を 1 つの群のデータと考え，各細胞タイプ間での発現データの正準相関係数，偏正準相関係数 [2, 3] を計算した．(偏) 正準相関係数の値が高く直接的な関係性が強いと示唆された細胞タイプ間を線で結んで描いたグラフは図 2 のようになった．図 1 と図 2 を見比べると，推測されている精巣胚細胞腫瘍の分化の過程でもつながりのある細胞タイプ間は，遺伝子発現データの (偏) 正準相関係数の値も高いことがわかった．また，推測されている分化の過程ではつながりがないが遺伝子発現データでは (偏) 正準相関係数の値の高かった細胞タイプ間もあった．それらの細胞で特異的に発現していた遺伝子の GO Term を調べると，GO Term のツリーの下位の方で近い位置にあるものがあり，分化の過程ではつながりがなくても細胞の機能が似ている可能性があることがわかった．様々な種類の細胞ごとの遺伝子発現データの正準相関係数，偏正準相関係数を計算することは，細胞タイプ間の機能的な関係性を測るひとつの方法として有用であると考えられる．

参考文献

- [1] Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, M. and Sherlock, G. (2004). GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes, *Bioinformatics*, **20**(18), 3710–3715.
- [2] Rao, B.R. (1969). Partial canonical correlations, *Trabajos de Estadística y de Investigación operativa*, **20**, 211–219.
- [3] 塩谷 實 (1990), 多変量解析概論, 朝倉書店.
- [4] Skotheim, R.I., Lind, G.E., Monni, O., Nesland, J.M., Abeler, V.M., Fossa, S.D., Duale, N., Brunborg, G., Kallioniemi, O., Andrews, P.W. and Lothe, R.A. (2005). Differentiation of human embryonal carcinomas in vitro and in vivo reveals expression profiles relevant to normal development, *Cancer Research*, **65** (13), 5588–5598.
- [5] Tusher, V.G., Tibshirani, R. and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response, *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5116–5121.