

大規模多型・発現プロファイルからの多要因データマイニングによる疾患機序経路の解明

●角田 達彦¹⁾ ◆古山 洋一²⁾

1) 理化学研究所遺伝子多型研究センター 2) 東京大学医科学研究所ヒトゲノム解析センター

研究の目的と進め方

疾患の多要因・大規模かつ不完全な多型・発現プロファイル、および細胞への薬剤投与・放射線照射・遺伝子導入による遺伝子発現変化のデータから柔軟に推論することのできるデータマイニング手法を確立し、それを実際に収集した疾患・発現変化データに適用して疾患機序経路および個人の薬剤応答性を解明するのが本研究の目的である。従来の相関解析を進め、遺伝的要因に環境要因を加えた多要因の現象を一度に扱うデータマイニングアルゴリズムと仮説推論の理論的枠組を構築する。次に現象の中から最も有意な部分空間から順に切り出し仮説を組み立てるアルゴリズムを構築する。以上を実際の疾患の遺伝子多型・発現パターンに適用し疾患原因遺伝子探索・疾患原因ネットワークを解明していく。

2001年度の研究の当初計画

- ・多値あるいは連続値、欠損値を扱える遺伝子ネットワークの枠組に必要な要件を検討し構築する。
- ・推定の枠組みを疾患の個性や薬理作用の差を観測する系に適用し、それらのマーカーや作用点となる遺伝子を抽出する。
- ・転写制御の知識ベースを整備する。
- ・多要因相関解析の枠組の構築と、多型と発現を連動させるためのデータ整備を行なう。

2001年度の成果

- ・マイクロアレイからの発現プロファイルの解析に必要な遺伝子ネットワーク推定の枠組の検討。
- ・癌の薬剤感受性の個人差を見分けるマーカー遺伝子群の抽出と、薬剤感受性の診断のためのスコアシステムの提案・構築。
- ・ヌードマウス上のヒト癌細胞の培養実験からの発現プロファイルと薬剤感受性の表現型との相関の解析によるマーカー遺伝子の抽出と、スコアシステムの提案・構築。
- ・Permutation testによる遺伝子スクリーニングの統計アルゴリズムの構築と、大腸癌発現プロファイルの解析および大腸癌肝転移症例マーカー遺伝子抽出への適用。
- ・AXINやPTEN遺伝子導入による発現プロファイルの変化による下流遺伝子候補の抽出と相関解析。
- ・転写因子結合予測アルゴリズムの整備とヒトプロモータ、植物プロモータへの適用。
- ・遺伝子多型データベース構築および相関解析自動化のためのハイスループットシステムの構築。
- ・日本人集団のSNPタイピングデータによるゲノムワイドな連鎖不平衡地図およびハプロタイプブロックの構築。

国内外での成果の位置づけ

いずれも世界で初めてかつ独自の試みである。発現プロファイルの解析は当初より完全に自前で考案したアルゴリズムを使用しており、追従を許さない。遺伝子多型解析は、これだけのデータが出つつあるのは日本だけであるし、世界のゲノムデータベースを吸収しつつ多面的な解析を行っている点からも世界で無比のものである。

今後の課題

- ・多要因の関係を推定するアルゴリズムの考案と実際の系への適用。
- ・統計的・生物学的な意味の有無によるデータの取舍選択を行うアルゴリズムの考案。
- ・神経培養細胞へ薬剤を投与した系での発現プロファイルの解析による薬剤作用点・パスウェイの推定と解析。

成果公表リスト

1. T. Takei, A. Iida, K. Nitta, T. Tanaka, Y. Ohnishi, R. Yamada, S. Maeda, T. Tsunoda, S. Takeoka, K. Ito, K. Honda, K. Uchida, K. Tsuchiya, Y. Suzuki, T. Fujioka, T. Ujiie, Y. Nagane, S. Miyano, I. Narita, F. Gejyo, H. Nihei, and Y. Nakamura. Association between single-nucleotide polymorphisms in selectin genes and IgA nephropathy. *American Journal of Human Genetics*, 2002, Vol. 70, pp. 781-786.
2. H. Zembutsu, Y. Ohnishi, T. Tsunoda, Y. Furukawa, T. Katagiri, Y. Ueyama, N. Tamaoki, T. Nomura, O. Kitahara, R. Yanagawa, K. Hirata, and Y. Nakamura. Genome-wide cDNA Microarray Screening to Correlate Gene Expression Profiles with Sensitivity of 85 Human Cancer Xenografts to Anticancer Drugs. *Cancer Research*, 2002, Vol. 62, pp. 518-527.
3. M. Nakamura, T. Tsunoda, J. Obokata. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds light through an initiator. *The Plant Journal*, 2002, Vol. 29, pp. 1-10.
4. M. Fujita, Y. Furukawa, T. Tsunoda, T. Tanaka, M. Ogawa, and Y. Nakamura. Up-Regulation of the Ectodermal-Neural Cortex (ENCL) Gene, a Downstream Target of the β -Catenin/T-Cell Factor Complex, in Colorectal Carcinomas. *Cancer Research*, 2001, Vol. 61, pp. 7722-7726.
5. C. Kihara, T. Tsunoda, T. Tanaka, H. Yamana, Y. Furukawa, K. Ono, O. Kitahara, H. Zembutsu, R. Yanagawa, K. Hirata, T. Takagi, and Y. Nakamura. Prediction of Sensitivity of Esophageal Tumors to Adjuvant Chemotherapy by cDNA Microarray Analysis of Gene-Expression Profiles. *Cancer Research*, 2001, Vol. 61, pp. 6474-6479. (equally contributed).
6. R. Yanagawa, Y. Furukawa, T. Tsunoda, O. Kitahara, K. Murata, O. Ishikawa, and Y. Nakamura. Genome-wide Screening of Genes Showing Altered Expression in Liver Metastasis of Human Colorectal Cancers by cDNA Microarray. *Neoplasia*, 2001, Vol. 3, pp. 395-401.
7. H. Ishiguro, Y. Furukawa, T. Tsunoda, T. Tanaka, Y. Fujii, and Y. Nakamura. Identification of AXUD1, a novel human gene induced by AXIN1 and involved in human carcinogenesis of the lung, liver, colon and kidney. *Oncogene*, Vol. 20, 2001, pp. 5062-5066.
8. Y. M. Lin, K. Ono, S. Satoh, H. Ishiguro, M. Fujita, N. Miwa, T. Tanaka, T. Tsunoda, K. C. Yang, Y. Nakamura, and Y. Furukawa. Identification of AF17 as a Downstream Gene of the β -catenin/Tcf Pathway, and its Involvement in Colorectal Carcinogenesis. *Cancer Research*, 2001, Vol. 61, pp. 6345-6349.
9. O. Kitahara, Y. Furukawa, T. Tanaka, C. Kihara, K. Ono, R. Yanagawa, M. E. Nita, T. Takagi, Y. Nakamura, and T. Tsunoda. Alterations of Gene Expression During Colorectal Carcinogenesis Revealed by cDNA Microarrays after Laser-Capture Microdissection of Tumor Tissues and Normal Epithelia. *Cancer Research*, Vol. 61, 2001, pp.3544-3549.
10. M. Matsushima-Nishiu, M. Unoki, K. Ono, T. Tsunoda, T. Minaguchi, H. Kuramoto, M. Nishida, T. Satoh, T. Tanaka, and Y. Nakamura. Microarray Analysis of Gene-Expression Profiles in Endometrial-Cancer Cells That Was Induced Exogenous PTEN. *Cancer Research*, Vol. 61, 2001, pp. 3741-3749.
11. Y. Suzuki, T. Tsunoda, J. Sese, H. Taira, J. Mizushima-Sugano, H. Hata, T. Ota, T. Isogai, T. Tanaka, Y. Nakamura, A. Suyama, Y. Sakaki, S. Morishita, K. Okubo, and S. Sugano. Identification and Characterization of the Potential Promoter Regions of 1031 Kinds of Human Genes. *Genome Research*, Vol. 11, 2001, pp. 677-684.
12. Y. Suzuki, H. Taira, T. Tsunoda, J. Mizushima-Sugano, J. Sese, H. Hata, T. Ota, T. Isogai, T. Tanaka, S. Morishita, K. Okubo, Y. Sakaki, Y. Nakamura, A. Suyama, and S. Sugano. Diverse Transcriptional Initiation Revealed by Fine, Large-Scale Mapping of the mRNA Start Sites. *EMBO reports*, Vol. 2, 2001, pp. 388-393.
13. T. Kato, S. Satoh, H. Okabe, O. Kitahara, K. Ono, C. Kihara, T. Tanaka, T. Tsunoda, Y. Yamaoka, Y. Nakamura, and Y. Furukawa. Isolation of a novel human gene, MARKL1, homologous to MARK3 and its involvement in hepatocellular carcinogenesis. *Neoplasia*, Vol. 3, 2001, pp.4-9.
14. H. Okabe, S. Satoh, T. Kato, O. Kitahara, R. Yanagawa, Y. Yamaoka, T. Tsunoda, Y. Furukawa, and Y. Nakamura. Genome-Wide Analysis of Gene Expression in Human Hepatocellular Carcinomas Using cDNA Microarray: Identification of Genes Involved in Viral Carcinogenesis and Tumor Progression. *Cancer Research*, Vol. 61, 2001, pp.2129-2137.
15. R. Yamada, T. Tanaka, M. Unoki, T. Nagai, T. Sawada, Y. Ohnishi, T. Tsunoda, M. Yukioka, A. Maeda, K. Suzuki, H. Tateishi, T. Ochi, Y. Nakamura, K. Yamamoto. Association between a Single-Nucleotide Polymorphism in the Promoter of the Human Interleukin-3 Gene and Rheumatoid Arthritis in Japanese Patients, and Maximum-likelihood Estimation of Combinatorial Effect That Two Genetic Loci Have on Susceptibility to the Disease. *American Journal of Human Genetics*, Vol.68, 2001, pp.674-685.
16. T. Tsunoda, R. Yamada, T. Tanaka, Y. Ohnishi, and N. Kamatani. Environmental Factor Dependent Maximum Likelihood Method for Association Study Targeted to Personalized Medicine. *Genome Informatics*, 2000, pp.96-105.
17. M. Kato, T. Tsunoda, T. Takagi. Inferring Genetic Networks from DNA Microarray Data by Multiple Regression Analysis. *Genome Informatics*, 2000, pp.118-128.
18. K. Ono, T. Tanaka, T. Tsunoda, O. Kitahara, C. Kihara, A. Okamoto, K. Ochiai, T. Takagi, and Y. Nakamura. Identification by cDNA Microarray of Genes Involved in Ovarian Carcinogenesis. *Cancer Research*, Vol. 60, 2000, pp.5007-5011.