

「テーラーメイド医療を目指したゲノム情報活用基盤技術」
平成15年度採択研究代表者

井ノ上 逸朗

(東京大学 医科学研究所 客員助教授)

「sub-common disease の感受性遺伝子同定と個人型易罹患性診断への応用」

1. 研究実施の概要

SNPデータベースは急速に充実し、SNP解析による個人差医療のインフラは整備されている。疾患感受性、薬剤感受性に関与するSNPは個人差医療の重要な鍵を握る。しかしながら、300-1000万といわれる全ゲノムに存在するSNPすべてのタイピングは時間、労力、費用の面から現実的でない。また遺伝要因の弱い疾患では感受性遺伝子を遺伝統計学的に同定できても、その生物学的機能関与が捉えにくい。また個々人の将来予測に応用することは困難である。本研究では、比較的遺伝背景の強い、脳動脈瘤、後縦靭帯骨化症、そして家族性甲状腺癌といったsub-common diseaseを対象とし、ゲノム全域連鎖解析でマップされた領域から効率の良いポジショナル（候補）クローニングをおこない、疾患感受性遺伝子を同定する。そして遺伝要因、生活習慣、環境要因の関与を統合した量的形質モデルを構築し、集団から個人への関与を明らかにする。そして個人型易罹患性診断への応用を目指し、テーラーメイド医療の実現を目指す。

比較的遺伝背景の強いsub-common diseaseではあるが、遺伝様式、浸透度などパラメーターの事前決定はできないので、ノンパラメトリック連鎖解析をおこなう。脳動脈瘤において114罹患同胞対でのゲノム全域スクリーニングで5、7、14番染色体に連鎖を認めた。後縦靭帯骨化症においては141対のスクリーニングで21番染色体に連鎖を確認できている。また連鎖領域からの感受性遺伝子同定作業も進行中である。家族性甲状腺癌は310家系が登録されており、現時点で75対、135検体の収集ができ、予備的ゲノム全域スクリーニングを終了している。ノンパラメトリック連鎖解析ではおよそ10cMにおよぶ連鎖を認め、領域には100個以上の遺伝子が存在するので連鎖不平衡解析によるピンポイントが必要となる、そのためには効率的なSNPタイピング確立、統計解析が求められる。SNPタイピングはハイスループットであるTaqMan assay-on-demandを用いる。連鎖不平衡の検討、Permutation法を加えたハプロタイプ解析により疾患感受性遺伝子を同定する。すでに必要なプログラム (SNPalyze) はダイナコムと共同で開発し、膨大なSNP情報を迅速に処理できる。

これまでのSNPでの遺伝疫学研究は集団での易罹患性（リスク）を示しており、個人がどの程度病気に罹りやすいかにはまったくアプローチされていなかった。SNPデータ、生

活習慣、臨床検査データを組み合わせ、ベイズ理論を展開することにより個人の疾患予測をおこなう未来型診断システムを提案し、実用プログラム化を目指す。特に個人の易罹患性は連続的な変数を示すという量的形質とそれがある閾値を越えたとき発症するという閾値モデルの考えに基づき、variance component modelを展開することで遺伝情報、生活習慣、臨床データなどを組み込んだ易罹患性モデルを構築し、将来予測へ繋げる。

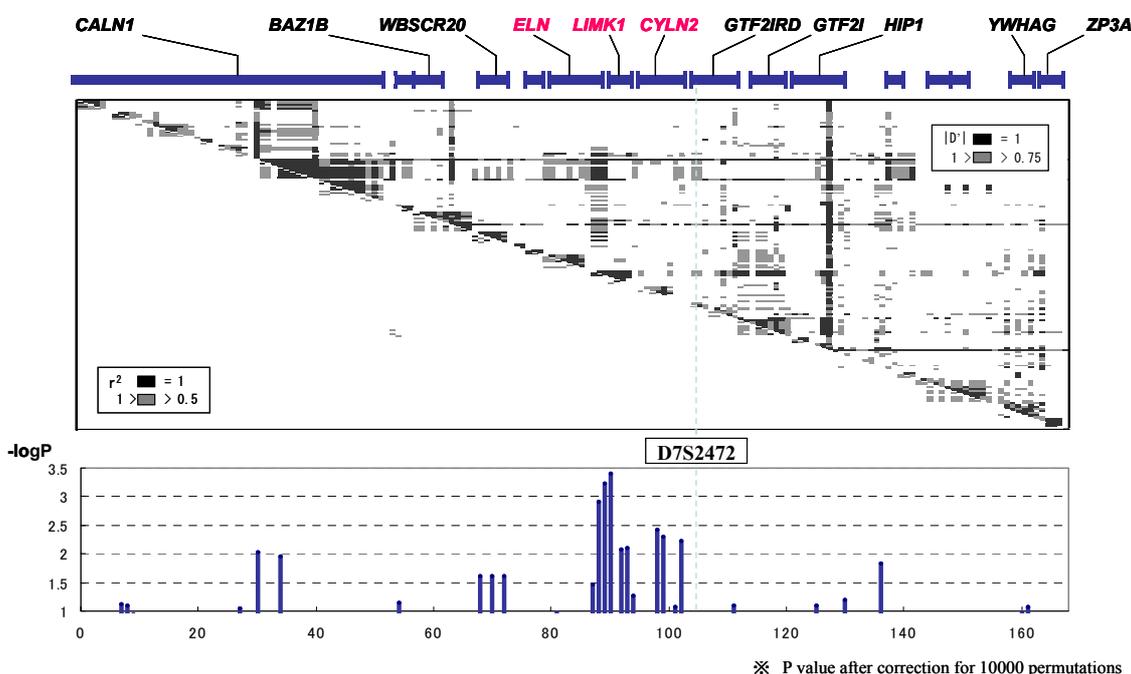
2. 研究実施内容

脳動脈瘤の責任遺伝子同定

脳動脈瘤の破裂によりくも膜下出血をきたし、生命の危険を伴う重篤な状態となる。脳動脈瘤そのものは無症状ながら破裂により重篤な状態となるため、破裂前診断が予防のために重要となる。近年、脳動脈瘤発生には比較的強い遺伝背景が存在することが知られてきた。疾患責任遺伝子の解明がその発生機序の解明、および疾患予測に重要である。すでにわれわれは110対の罹患同胞対連鎖解析によりゲノム全域スクリーニングをおこない、5、7、14番染色体に連鎖を認めている。最も強い連鎖、最大LOD=3.22,を得た染色体7番の連鎖ピークマーカーD7S2472を中心とし1 LOD dropに相当する約4.7 Mの領域を候補とし、連鎖不平衡解析によるポジショナル・クローニングを試みた。領域には30個の遺伝子が存在しており、168箇所のgene-based SNP (single nucleotide polymorphism: 1塩基多型)を用いた遺伝子スクリーニングを行った。SNPタイピングは直接シーケンス法もしくはTaqMan法を用いた。さらに、ゲノム領域の連鎖不平衡を詳細に検討し連鎖不平衡(LD)ブロック構造を明らかにし、かつ新たに開発したPermutation法でのアソシエーション・スタディにより、多検定補正をおこなった。単独SNPでの解析と3個の隣り合うSNPでハプロタイプを組みゲノム領域をスキャンするwindow-scan法で検討した。図1に示すよう、単独SNPでもハプロタイプ解析でも共通して有意差を示したのは、ピークマーカーから300 k塩基離れた領域であった。

図1. D7S2472近傍の連鎖不平衡解析

4.7 Mb, 168 SNPs



特定された領域について、さらに詳細なSNP解析をおこなったところ、もっとも強い有意差を認めたSNPsはひとつの連鎖不平衡ブロック内、エラスチン3'UTRとLIMK1 (LIM kinase 1) プロモーターにまたがる領域、に存在していた。すなわち、脳動脈瘤発症の原因となりうるSNPはこの領域に存在すると考えられ、疾患遺伝子ピンポイントに成功したといえる。今後、さらに詳細なSNPスクリーニングを進め原因となっているSNPを同定する。5番、14番染色体の候補領域についても同様の手法でスクリーニングを開始している。

閾値モデルを用いた個人型易罹患性診断用ツールの開発

ヒト疾患には、単一の遺伝子における変異がその発症の有無を決定しているもの他、生活習慣病等のように、複数の遺伝子および環境要因によって支配されている多因子疾患も存在する。Wright (1934) は、量的遺伝学の理論に基づいた、閾値モデルの理論を提唱し、発症/非発症といった、表面上不連続な形質の発現に、連続型変数（通常は正規分布にしたがうとする）であるliability（易罹病性）が関与すると想定した解析法を提唱している。liabilityは複数の遺伝要因および環境要因から成り立ち

$$(liability_i = \mu + \sum_{j=1}^N \alpha_j x_{ij} + e)$$

、ある閾値を超えるか否かで、表現型が決定されるとする。

閾値モデルによる新たな個人型易罹患性診断システムの開発に向けての第一歩として、基本的な計算プログラムを作成し、シミュレーションによる検討を行った。さらに、実際のデータへの適用例として、前立腺がんの臨床記録を用いた計算を試みた。

・閾値モデル

図中の式における x_{ij} は、疾患要因である項目 j における被験者 i の標準化された測定値を、 α_j は、liabilityの項目 j に対する偏回帰係数を表す。 α_j の値が大きいほど、liabilityに対する項目 j の影響が大きいことを示す。 α_j は、liabilityの項目 j に対する偏回帰係数を表す。そして、 i が発症に至る ($w_i = 1$) か至らない ($w_i = 0$) かは、 i のliabilityが閾値 t を超えるか否かにより決定される。

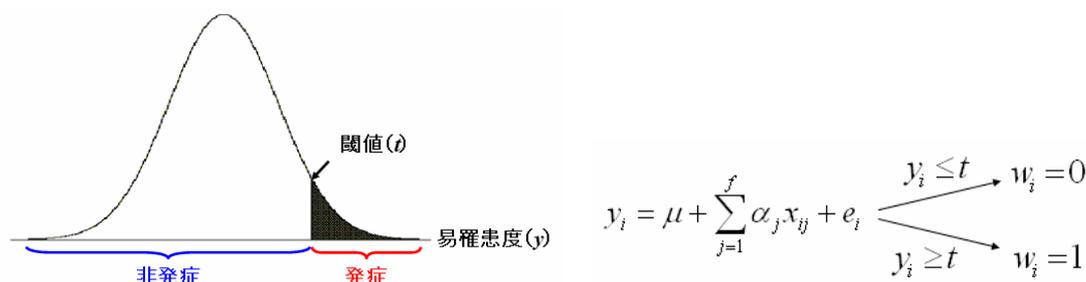


図1. 閾値モデルの概略図

・シミュレーション実験

以下に示すように、2種類のケースを想定し、閾値モデルにしたがって5つの独立変数と1つの二値データを各々600件発生させ、今回開発したツールにより、モデルに考慮された変数、および α_j を推定した。

表1. 各ケースにおけるliabilityの各独立変数に対する偏回帰係数の真値

| | α_1 | α_2 | α_3 | α_4 | α_5 |
|------|------------------|---------------|---------------|----------------|---------------|
| ケース1 | 1.00 (30.7) * | 0.50 (7.7) | 0.05 (0.1) | 1.00 (30.7) | 0.05 (0.1) |
| ケース2 | 1.00 (43.1) | 0.25 (2.7) | 0.05 (0.1) | 0.50 (10.8) | 0.05 (0.1) |

*カッコ内の値は、その変数の分散が易罹患度全体のそれに占める割合を表す。

なお計算には、MCMC (Markov chain Monte Carlo) アルゴリズムによるベイズ推定法を用いている。

・データ解析

総合病院（千葉県旭市）でのデータ（年齢（初診時）、PSA（初診時）、F-PSA（初診時）、F/T（%、初診時）、echo重量、PSAD、DRE、硬結、触診によるがんの有無の計9項目）794件のうち、欠測値のあるデータは削除し、最終的に653件を解析に用いた。なお、各項目の基本統計量を以下に示す。

表2. 各項目の基本統計量

| | 平均 | 標準偏差 | 最小値 | 最大値 |
|------------|-------|-------------------|-------|---------|
| 年齢（初診時） | 70.34 | 8.26 | 48.00 | 91.00 |
| PSA（初診時） | 79.94 | 531.18 | 0.17 | 7875.70 |
| F-PSA（初診時） | 15.68 | 129.36 | 0.06 | 2481.70 |
| F/T（%、初診時） | 21.75 | 13.03 | 0.31 | 100.00 |
| echo重量 | 38.87 | 24.39 | 4.00 | 280.00 |
| PSAD | 2.07 | 12.79 | 0.01 | 210.57 |
| | - | | + | |
| がんの有無 | 460 | | 193 | |
| DRE | 465 | | 188 | |
| 硬結 | 465 | Lt-73、Rt-60、Bl-55 | | |

・ 結果

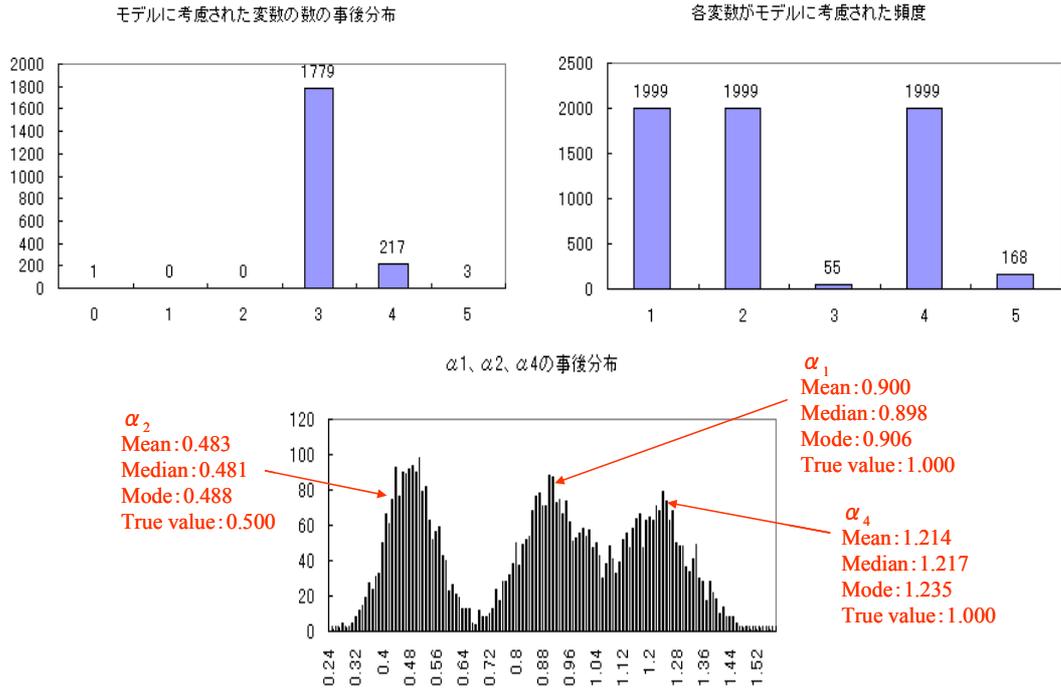


図2. シミュレーション集団 (ケース1) における結果

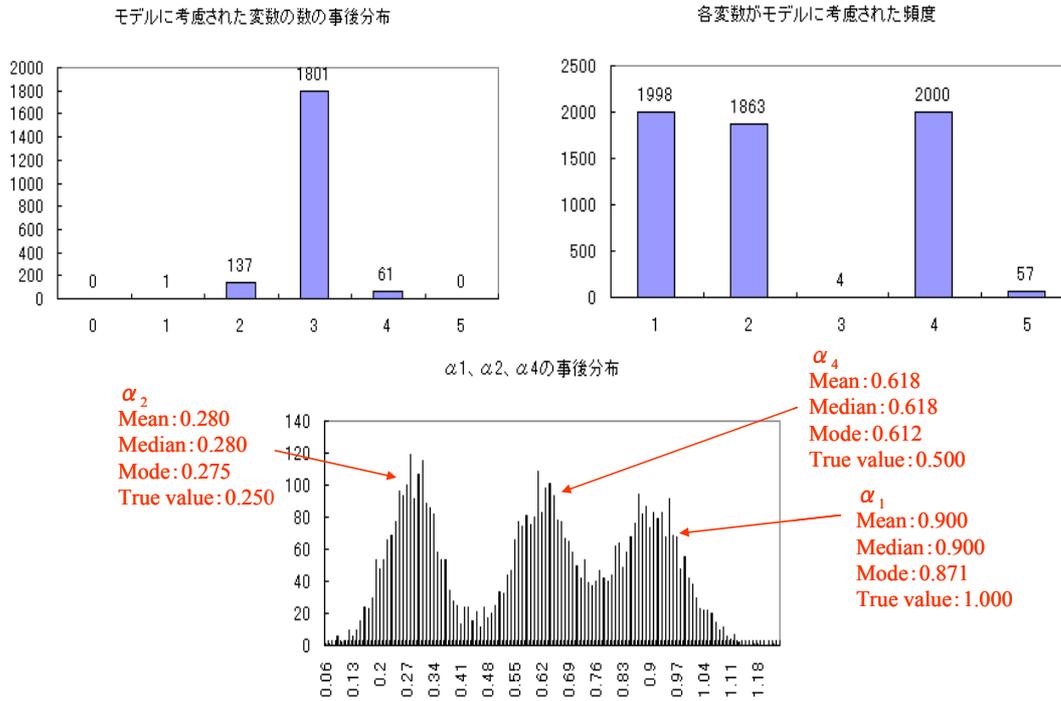


図3. シミュレーション集団 (ケース2) における結果

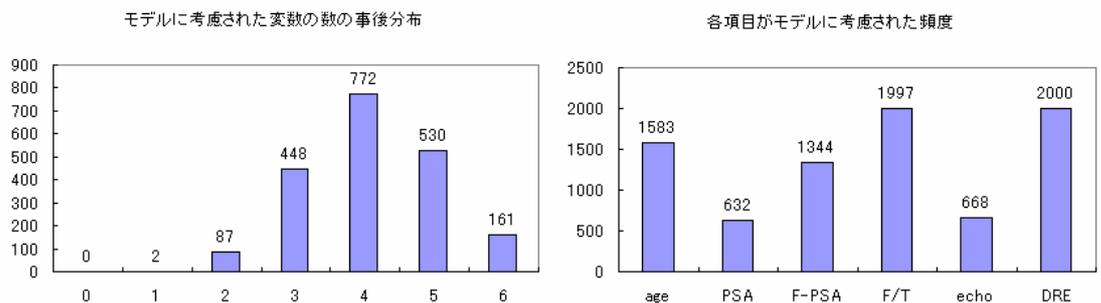


図4. 前立腺がんデータにおける結果（モデルに考慮された変数の数の事後分布、および各項目がモデルに考慮された頻度のヒストグラム）

図2、3に示す通り、いずれのケースにおいても、強く関与するものと設定した変数1、2、4のみがモデル内に取り込まれた。また易罹患度のこれらの変数に対する偏回帰係数も、ほぼ真値と一致した結果が得られており、適切な計算がなされていることが分かる。これを踏まえ、実際に前立腺がんのデータを対象として行った解析では、上位3モデルから、主に年齢、F/T、DREの3要素が強く関与しているとする結果が得られた（図4および表3参照）。

以上の結果から、今回提唱した閾値モデルの理論に基づいた手法が、様々な疾患、特に生活習慣病等の多因子疾患の発症要因の解明、そして将来における発症の可能性を、個人レベルで予測するためのシステムの開発に大きく貢献し得る可能性が示唆された。

表3. 前立腺がんデータにおける結果（尤度の高い上位3モデルにおける項目リスト、偏回帰係数、および各モデルのIntensity）

| モデルに考慮された項目 | 各項目における偏回帰係数 | | | | | | Intensity |
|------------------------------|----------------|----------------|------------------|----------------|-----------------|----------------|-----------|
| | α_{age} | α_{PSA} | α_{F-PSA} | $\alpha_{F/T}$ | α_{echo} | α_{DRE} | |
| (age, F-PSA, F/T, DRE) | 1.648* | — | 12.161 | -7.539 | — | 0.255 | 0.2270 |
| | 1.646 | — | 12.194 | -7.517 | — | 0.254 | |
| | 1.564 | — | 12.494 | -7.212 | — | 0.255 | |
| (age, F-PSA, F/T, echo, DRE) | 1.735 | — | 13.169 | -7.168 | -3.460 | 0.251 | 0.1340 |
| | 1.721 | — | 12.832 | -7.264 | -3.452 | 0.250 | |
| | 1.715 | — | 11.847 | -7.342 | -3.149 | 0.248 | |
| (age, F/T, DRE) | 1.648 | — | — | -7.352 | — | 0.258 | 0.1200 |
| | 1.631 | — | — | -7.334 | — | 0.258 | |
| | 1.810 | — | — | -7.261 | — | 0.266 | |

*上から平均値、中間値、最頻値を表す。

3. 研究実施体制

疾患遺伝子研究グループ

- ① 研究分担グループ長：井ノ上 逸朗（東京大学 医科学研究所、客員助教授）
- ② 研究項目：
 - 1) 脳動脈瘤、後縦靭帯骨化症の遺伝素因の解明
 - 2) variance component モデルの構築

家族性甲状腺癌解析グループ

- ① 研究分担グループ長：小島 俊男（理化学研究所、チームリーダー）
- ② 研究項目：家族性甲状腺癌の遺伝子同定

易罹患性予測プログラム開発グループ

- ③ 研究分担グループ長：羽田 明（千葉大学 医学部、教授）
- ④ 研究項目：個人型易罹患性予測プログラムの開発

4. 主な研究成果の発表（論文発表および特許出願）

(1) 論文発表

- Ikeda R, Yoshida K, Tsukahara S, Sakamoto Y, Tanaka H, Furukawa-K, Inoue I. PLZF promotes osteoblastic differentiation of human mesenchymal stem cells as an upstream regulator of CBFA1. *J Biol Chem*, 280, 8523-8530, 2005.
- Yamaguchi T, Park S-B, Maki K, Inoue I. Genome-wide linkage analysis of mandibular prognathism in Korean and Japanese patients. *J Dent Res*, 84, 255-259, 2005.
- Jinnai N, Sakagami T, Sekigawa T, Kakihara M, Nakajima T, Yoshida K, Goto S, Hasegawa T, Koshino T, Hasegawa Y, Inoue H, Suzuki N, Sano Y, Inoue I. Polymorphisms in the prostaglandin E2 receptor subtype 2 gene confer susceptibility to aspirin-intolerance asthma: a candidate gene approach. *Hum Mol Genet*, 13, 3203-3217, 2004.
- Akagawa H, Kasuya H, Onda H, Yoneyama T, Choi K-Y, Lee W-J, Hong S-W, Yang T-K, Kim C-J, Hori T, Inoue I. Endothelial nitric oxidase synthase T-786C single nucleotide polymorphism is not a factor influencing aneurysm size. *J Neurosurgery*, 95, 622-625, 2005.
- Yoneyama T, Kasuya H, Onda H, Akagawa H, Hori T, Inoue I. Choi K-Y, Lee W-J, Hong S-W, Yang T-K, Kim C-J. Absence of alpha 1-antitrypsin deficiency alleles (S and Z) in Japanese and Korean patients with intracranial aneurysm. *Stroke*, 35, e376-378, 2004.
- Ishikawa C, Ozaki H, Nakajima T, Ishii T, Kanai S, Anjo S, Shirai K, Inoue I. A frameshift variant of CYP2C8 was identified in a patient who suffered from rhabdomyolysis after administration of cerivastatin. *J Hum Genet*, 49,

582-585, 2004.

- Imaizumi T, Yagihashi N, Hatakeyama M, Yamashita K, Ishikawa A, Taima K, Yoshida H, Inoue I, Fujita T, Yagihashi S, Satoh K. Expression of retinoic acid-inducible gene-I in vascular smooth muscle cells stimulated with interferon-gamma. *Life Sci*, 75, 1171-1180, 2004.
- Yoshida K, Inoue I. Expression of MCM10 and TopBP1 is regulated by cell proliferation and UV irradiation via the E2F transcription factor. *Oncogene*, 23, 6250-6260, 2004.
- Sakagami T, Witherspoon DJ, Nakajima T, Jinnai N, Wooding S, Jorde LB, Hasegawa T, Suzuki E, Gejyo F, Inoue I. Local adaptation and population differentiation at the interleukin 13 and interleukin 4 loci. *Genes Immun*, 5, 389-397, 2004.
- Sasaki T, Kasuya H, Onda H, Sasahara A, Hori T, Inoue I. Role of p38 mitogen-activated protein kinase on cerebral vasospasm following subarachnoid hemorrhage. *Stroke*, 35, 1466-1470, 2004.
- Yoshida K, Inoue I. Peptide binding to Geminin and inhibitory for DNA replication. *Biochem Biophys Res Commun* 317, 218-222, 2004.
- Nakajima T, Wooding S, Sakagami T, Emi M, Tokunaga K, Tamiya G, Ishigami T, Umemura S, Munkhbat B, Jin F, Guan-jun J, Hayasaka I, Ishida T, Saito N, Pavelka K, Lalouel J-M, Jorde LB, Inoue I. Natural selection and population history in the human angiotensinogen gene (AGT): 736 AGT sequencing in worldwide chromosomes. *Am J Hum Genet*, 74, 898-916, 2004.
- Yoshida K, Inoue I. Regulation of Geminin and Cdt1 expression by E2F transcription factors. *Oncogene*, 23, 3802-3812, 2004.
- Ohmori H, Ando Y, Makita Y, Onouchi Y, Nakajima T, Saraiva MJM, Terazaki H, Sulu O, Sobue G, Nakamura M, Yamaizumi M, Munar-Ques M, Inoue I, Uchino M, Hata A. Common origin of the Val30Met mutation responsible for the amyloidogenic transthyretin type of familial amyloidotic polyneuropathy. *J Med Genet*, 41, e51, 2004.
- Yoneyama T, Kasuya H, Onda H, Akagawa H, Hashiguchi K, Nakajima T, Hori T, Inoue I. Collagen type I α 2 (*COL1A2*) is the susceptible gene for intracranial aneurysms. *Stroke* 35, 443-448, 2004.
- Matsui H, Suzuki K, Ohtake N, Nakata S, Takeuchi T, Yamanaka H, Inoue I. Genomewide linkage analysis of familial prostate cancer in Japanese population. *J Hum Genet* 49, 9-15, 2004.
- Yoshida K, Oyaizu N, Dutta A, Inoue I. The destruction box of human Geminin is critical for proliferation and tumor growth in human colon

cancer cells. *Oncogene* 23, 58-70, 2004.

- Kobashi G, Hata A, Ohta K, Yamada H, Hirayama E, Minakami H, Fujimoto S, Kondo K. A1166C variant of angiotensinogen II type 1 receptor gene is associated with severe hypertension in pregnancy independently of T235 variant of angiotensinogen gene. *J Hum Genet*, 49, 182-186, 2004.
- Onouchi Y, Onoue S, Tamari M, Wakui K, Fukushima Y, Yashiro M, Nakamura Y, Yanagawa H, Kishi F, Ouchi K, Terai M, Hamamoto K, Kudo F, Aotsuka H, Sato Y, Nariai A, Kaburagi Y, Miura M, Saji T, Kawasaki T, Nakamura Y, Hata A. CD40 ligand gene and Kawasaki disease. *Eur J Hum Genet*, 12, 1062-1068, 2004.
- Yamazaki K, Takazoe M, Tanaka T, Ichimori T, Saito S, Iida A, Onouchi Y, Hata A, Nakamura Y. Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. *J Hum Genet*, 49, 664-668, 2004.
- Abe K, Noguchi H, Tagawa K, Yuzuriha M, Toyoda A, Kojima T, Ezawa K, Saitou N, Hattori M, Sakaki Y, Moriwaki K, Shiroishi T. Contribution of Asian mouse subspecies *Mus musculus molossinus* to genomic constitution of strain C57BL/6J, as defined by BAC end sequence- SNP analysis. *Genome Res.*14, 2439-2447, 2004.
- Kojima T, Kamei H, Aizu T, Arai Y, Takayama M, Nakazawa S, Ebihara Y, Inagaki H, Masui Y, Gondo Y, Sakaki Y, Hirose N. Association analysis between longevity in the Japanese population and polymorphic variants of genes involved in insulin and insulin-like growth factor 1 signaling pathways. *Exp Geronto.* 39, 1595-1598, 2004.
- Murakami K, Kojima T, Sakaki Y. Assessment of clusters of transcription factor binding sites in relationship to human promoter, CpG islands and gene expression. *BMC Genomics* 5, 16, 2004.

(2) 特許出願

平成16年度出願件数：1件（CREST研究期間累積件数：1件）