

臨床データとゲノム情報の統合を基礎とした疾患のシステムの理解と医療への応用

●永井 良三 ◆原 一雄 ◆森田 啓行
 東京大学医学部

<研究の目的と進め方>

日本人の主要な疾患である心血管疾患と糖尿病を含む生活習慣病は、生活習慣などの環境因子と遺伝素因の各々が相互作用し、その結果システムとしての生命の営みに破綻を来している状態と捉えることができる。本研究計画では「生命システムの破綻」の表現型としての臨床情報を日常臨床の中から網羅的・体系的に収集することと、ゲノム研究・モデル動物などを用いた疾患の *in vivo* 研究の知見を統合することによって、生命・疾患のシステムの理解を基盤とする個人への最適で、安全かつ高効率な医療の提供を目指す。本研究計画では、1) 臨床用途に耐えるゲノム-臨床情報統合データベースシステムの構築と臨床での運用を推進する。2) データベースとゲノム情報の統合により生命システム破綻を引き起こす要因を抽出・同定し、さらに発症メカニズム解明のために *in vivo* の遺伝子機能解析と遺伝子ネットワーク研究を中心とした基礎的研究を展開する。3) 臨床情報データベースの導入と運用による循環器・糖尿病領域におけるテララーメード医療の実用化を推進することを目的とする。

<2008年度の研究の当初計画>

(1) 臨床用途に耐えるゲノム-臨床情報統合データベースシステムの構築と臨床での運用

循環器疾患データベースシステムについては、2007年度までに病名や薬剤マスター、診断クライテリアの確立、ICD-10やMEDISコードへのマッピング、病院診療システムからのHL7プロトコルを介した情報移行、外来情報へ対応できるデータベース構造の組み直しを終了し、2000名分のデータクレンジングを完了した。また、院内ホームページMULINSを介して、簡便なデータ解析システムを公開した。

今後の展開には、複数の医療機関における情報収集が必要である。多忙な臨床現場を考慮したとき、循環器の部門システムとして必須の心臓カテーテル検査の画像表示・情報入力・閲覧・レポート作成・データ管理を一括して行えるシステムを開発し、このシステムを活用してもらうことでデータ収集を行えるようにすることが現実的である。今までに開発を進めてきた、目下当院で運用中の心臓カテーテル検査データベースシステムに関し、情報の収集・解析機構の組み込みが可能かを検討する。現在、日本心血管インターベンション学会で臨床情報のレジストリシステムが検討されていることから、そこで必要とされる情報を基礎としてデータの選別・解析法の検討を行う。同時に、病院間の情報共有化の方法について検討する。

糖尿病のデータベースシステムについても、複数病院での使用を前提として、現有システムの活用について検討する。

(2) データベースとゲノム情報の統合による、生命システム破綻を引き起こす要因の抽出・同定と遺伝子ネットワークの解析

2007年度までの研究により、心血管系モデリングの鍵分子であるKLF5のプロモーター領域SNPが冠動脈硬化重症度に関連することを示した。2008年度はこのSNPがMEF2の結合配列となっていることに着目し、KLF5への情報伝達のネットワークを実験的に明らかにする。またKLF5と相互作用する転写因子およびKLF5と同様に心血管系・代謝系のストレス応答において重要な遺伝子について、転写因子を中心としたSNP解析を行う。既に、ARID5B、MEF2などの解析を開始している。また、分子

間ネットワークに基づいて解析候補遺伝子を拡大していくためには効率的な手法の確立が必要と考える。テキストデータベースや、two-hybrid解析データベースの応用を検討する。

さらに、上記解析で得られた遺伝子ネットワークについては、個体レベルでの解析を行う。KLF5については、平滑筋、心筋、骨格筋などで特異的な過剰発現・ノックアウトマウスモデルを作成して検討を行う。

(3) 臨床情報データベースの日常臨床への導入

既に、簡便な解析が行えるシステムを院内ネットワークで公開しており、患者教育に有効な情報を提供できるようにしているが、多忙な外来診療中に操作する時間がないのが現状である。システムの改善普及を図る。一方、心臓カテーテル検査データベースシステムに関しては、部門システムとして現に運用中であることから、情報提供の目的を明確化し、さらに臨床現場で使いやすいシステムにするため改善を試みる。

<2008年度の成果>

臨床情報統合データベースシステムを日々更新改良する一方、データマネージャーを養成し、入力データの正確性確保につとめた。既に、3,500症例以上のデータの蓄積を行った。2008年度は、心臓カテーテル検査データベースシステムと連動させ、データ入力の迅速化と正確化を図った。以上のように、院内の情報を収集・解析するシステムとしては当初の目的を達成した。

構築した臨床情報データベース及び心臓カテーテル検査データベースを用いてSNP解析を行った。我々は転写因子KLF5を心血管系の外的ストレスに応じた組織モデリングを制御する鍵分子として報告してきた。本研究では、まずKLF5転写ネットワークに関わる遺伝子群と、KLF5と同様に組織モデリングに関わる遺伝子群の解析を行った。その結果、KLF5の5'隣接領域に冠動脈疾患発症と関連するSNPを見いだした。2種のgenotypesでKLF5プロモーター活性基礎値に差は見られなかったが、冠動脈疾患発症率の低いgenotypeでは、アンジオテンシンII刺激によるKLF5プロモーター活性化が著明に減弱しており、ストレスに応じたKLF5発現制御にとってこのSNPを含む領域が重要であると考えられた。このSNPはMEF2結合配列中に存在しており、冠動脈疾患発症が少ないgenotypeでは、MEF2結合が見られなくなる。また、アンジオテンシンIIからMEF2、KLF5へと進む情報伝達経路に活性酸素のシグナルが重要であることも明らかにした。

ARID5Bは血管平滑筋の分化・脱分化、さらには脂肪細胞分化にもかかわる転写因子である。本遺伝子の4つのSNPsがハプロタイプを構成し冠動脈疾患発症と有意に相関するというデータを得た。さらに2型糖尿病とハプロタイプとの有意な相関、血中アディポネクチン濃度とSNPとの有意な相関を確認できた。

MnSOD(Manganese superoxide dismutase) 遺伝子の非同義置換Ala16Valが冠動脈疾患発症と有意に相関することを新たに見出した(Val/Val vs non Val/Valでの補正後オッズ比は1.49)。Val/Val型ではMnSOD酵素活性が低下することを反映して、Val/Val型症例マクロファージでは酸化LDL処理後にアポトーシスを起こす比率が高いことを確認した。

一方、2型糖尿病に関連する遺伝子多型としてAMPK α サブユニット遺伝子、HNF-4A 遺伝子、TCF7L2 遺伝子、HHEX 遺

伝子の SNP 解析をおこない成果を公開している。

<国内外での成果の位置づけ>

国内でも類をみないゲノム研究対応型臨床情報統合データベースシステムを確立した。またゲノム解析結果をふまえて個体内でのメカニズムを探るといった進め方はユニークであり、いくつかのシーズを得始めている。

<達成できなかったこと、予想外の困難、その理由>

心筋梗塞発症と関連する SNP に関しては約 200 例での予備的検討では有意相関が疑われるものが相当数存在するが、そのほとんどは 500 例までスケールアップして検証すると再現性が得られなかった。当初予想していたよりも心筋梗塞のマーカーとなる遺伝子多型は少ないことがわかった。遺伝子変異解析の方向性を見直す必要がある。それをうけて今後の遺伝子解析方法に変更を加える。

また本研究はあくまで東大病院内のゲノム・臨床情報統合データベース構築を目的とし、事実それを実行しているが、将来的には病院ごとのネットワークを病院間でつないでデータを共有するという構想を描いている。そのためにはデータの匿名化が大前提であるが、技術的・制度的なハードルも多い。個人情報の取り扱い、倫理的問題に関する議論も必要である。また疾患の定義にしてもその微妙なズレは共有したデータの価値を損なうことから、ICD-10 に準拠した疾患のコード分類を見直し試験運用している。このように共有データへのバイアスを最小限にとどめるためのデータ「標準化」作業をさらに進める必要がある。

院内の診療端末より情報の閲覧・解析が可能なシステムを提供しているが、その情報がどの程度日常臨床に反映されているか検証する必要がある

<今後の課題>

入力症例数を増やす一方、詳細な時系列データの解析を行えるようにデータベースの改良を逐次行う。心臓カテーテル検査データベースシステムとの連動に合わせて、解析システムに今後どのような機能を実装するか検討を行う。

また、複数医療機関での臨床情報共有化のためのデータ取得法、データセンターの設計などを進める。ゲノム解析に関しては、従来おこなってきた SNP 解析およびその機能解析と並行して Rare variant 検索も開始する。従来、「単一遺伝子病は Rare variant、多因子疾患は Common variant による」というパラダイムの下、解析が進められてきたが、2008 年度までのわれわれの解析でも心血管病と「真に」相関する遺伝子多型は少ないこと、相関する場合でもその相対危険度は 1.2-1.5 と低いことがわっている。最近、多因子疾患においても、より稀な遺伝子変異 (Rarer variant) が疾患発症進展に関与している可能性が議論されている。SNP と比べて Rare variant の病態とのリンクはより強く、優秀な疾患マーカーとして期待される。頻度は遺伝子多型に比べて稀であるが、生物学的には強いインパクトをもたらす病態とのリンクも強い。個の医療に大きく貢献することが期待される。必ずしも Rare variant に限定するのではなくその遺伝子内の SNP も含めて網羅的に解析する。相当数の症例を対象に、候補遺伝子を絞って、頻度は低いインパクトの強い Rare variant とインパクトはより小さいが頻度が高い SNP を同時に解析することで従来のパラダイムを検証する。したがって解析方法としてはダイレクトシークエンス法を併用し網羅的解析をおこなう

一方、従来おこなってきた遺伝子多型解析についても継続予定である。転写因子を各種生理活性物質の上流に位置する鍵分子と捉えたと転写因子遺伝子は有望な候補として期待できる。現在転写因子 ARID5 と冠動脈疾患、糖尿病との相関を見出し、検討を進めている。この ARID5 は血管平滑筋の分化・脱分化、脂肪細胞分化に関与する転写因子であることが細胞レベルの実験から明らかであり、今後の成果が期待される。上記遺伝子解析と併行して細胞レベル、実験動物レベルでのメカニズム解析をおこなう。

<成果公表リスト>

0901071444 Kohro T, Hayashi D, Okada Y, Yamazaki T, Nagai R;

JCAD Investigators. Demographics and changes in medical/interventional treatment of coronary artery disease patients over a 3.5-year period in Japan: the Japanese Coronary Artery Disease Study: trend examination. *Circ J.* 72(9):1397-1402 (2008)

0901071443 Nagai R, Izumi T, Kurabayashi M, Daida H, Tojo T, Hasegawa A, Miyauchi K, Hayashi D, Kohro T, Okada Y, Yamazaki T; JCADII Investigators. Rationale and design of a study to examine lower targets for low-density lipoprotein-cholesterol and blood pressure in coronary artery disease patients. *Circ J.* 72(4):515-520 (2008)

0901071441 Monzen K, Ito Y, Naito AT, Kasai H, Hiroi Y, Hayashi D, Shiojima I, Yamazaki T, Miyazono K, Asashima M, Nagai R, Komuro I. A crucial role of a high mobility group protein HMGA2 in cardiogenesis. *Nat Cell Biol.* 10(5):567-574 (2008)

0901071438 Kubota N, Kubota T, Itoh S, Kumagai H, Kozono H, Takamoto I, Mineyama T, Ogata H, Tokuyama K, Ohsugi M, Sasako T, Moroi M, Sugi K, Kakuta S, Iwakura Y, Noda T, Ohnishi S, Nagai R, Tobe K, Terauchi Y, Ueki K, Kadowaki T. Dynamic functional relay between insulin receptor substrate 1 and 2 in hepatic insulin signaling during fasting and feeding. *Cell Metab.* 8(1):49-64 (2008)

0901071432 Wang G, Watanabe M, Imai Y, Hara K, Manabe I, Maemura K, Horikoshi M, Kohro T, Amiya E, Sugiyama T, Fujita T, Kadowaki T, Yamazaki T, Nagai R. Genetic variations of Mrf-2/ARID5B confer risk of coronary atherosclerosis in the Japanese population. *Int Heart J.* 49(3):313-327 (2008)

0806251314 Oishi Y, Manabe I, Tobe K, Ohsugi M, Kubota T, Fujiu K, Maemura K, Kubota N, Kadowaki T, Nagai R. SUMOylation of Krüppel-like transcription factor 5 acts as a molecular switch in transcriptional programs of lipid metabolism involving PPAR-delta. *Nat Med.* 14(6): 656-666 (2008)

0806251309 Kada N, Suzuki T, Aizawa K, Munemasa Y, Matsumura T, Sawaki D, Nagai R. Acyclic retinoid inhibits functional interaction of transcription factors Krüppel-like factor 5 and retinoic acid receptor-alpha. *FEBS Lett.* 582(12):1755-1760 (2008)

0806251307 Okamoto M, Ohara-Imaizumi M, Kubota N, Hashimoto S, Eto K, Kanno T, Kubota T, Wakui M, Nagai R, Noda M, Nagamatsu S, Kadowaki T. Adiponectin induces insulin secretion in vitro and in vivo at a low glucose concentration. *Diabetologia* 51(5):827-835 (2008)

0806251304 Nishimura S, Manabe I, Nagasaki M, Seo K, Yamashita H, Hosoya Y, Ohsugi M, Tobe K, Kadowaki T, Nagai R, Sugiura S. In vivo imaging in mice reveals local cell dynamics and inflammation in obese adipose tissue. *J Clin Invest.* 118(2):710-721 (2008)

0806251302 Munemasa Y, Suzuki T, Aizawa K, Miyamoto S, Imai Y, Matsumura T, Horikoshi M, Nagai R. Promoter region-specific histone incorporation by the novel histone chaperone ANP32B and DNA-binding factor KLF5. *Mol Cell Biol.* 28(3):1171-1181 (2008)

0806251258 Ishizaka N, Ishizaka Y, Toda E, Koike K, Yamakado M, Nagai R. Are serum carcinoembryonic antigen levels associated with carotid atherosclerosis in Japanese men? *Arterioscler Thromb Vasc Biol.* 28(1):160-165 (2008)

0806251255 Horikoshi M, Hara K, Ito C, Shojima N, Nagai R, Ueki K, Froguel P, Kadowaki T. Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia* 50(12):2461-2466 (2007)