

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
東京大学医学部

＜研究の目的と進め方＞

本邦での主要疾患である心血管疾患および糖尿病などいわゆる生活習慣病は、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。本計画研究では「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、個人々への最適、安全かつ効率的な医療の提供を目指す。本計画研究では、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り循環器・糖尿病領域におけるテーラーメイド医療実用化を推進する。

＜2007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、2006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。2006年度は、同時に心臓カテーテルシステムを導入し、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。2007年度は、まず、未入力データの入力を行う。また、外来情報などへの対応のためデータベース構造の改良を随時行う。さらに心臓カテーテルシステムの展開に合わせて、今後の解析システムにどのような機能を実装するかを検討を行う。

今後の展開には、複数病院の情報収集が重要である。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・在庫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と考えられる。そこで、従来より開発を進めてきた心臓カテーテル検査システムについて、情報の収集・解析機構の組み込みが可能か検討する。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、現有システムの利用について検討する。

データベース情報の利用法として、ゲノムだけではなく、新規バイオマーカー(Ptx3等)の解析へ適用する。

(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の抽出・同定と遺伝子ネットワークの解析

2006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プロモーター領域のSNPが高血圧や冠動脈硬化重症度に関連することを示した。2007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。2007年度は、まず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で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をもとに解析候補遺伝子を拡大していくために、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。

また、上記のような解析で得られた遺伝子ネットワークについては、さらに、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が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。これに対して、心臓カテーテルシステムについては、部門システムとして確実に使われていることから、情報提供の目的を明確化して、さらに臨床現場で使われるシステムとするために改善を試みる。

＜2007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。

構築した臨床情報データベース及び心臓カテーテル情報データベースを用い、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リ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と同様に組織リ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'隣接領域に高血圧発症と関連する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、高血圧発症率の低い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、高血圧発症が少ない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を明らかにした。次に、KLF5に関連することが実験的に確認できた遺伝子であるMEF2、NF-ATについて、心血管代謝疾患との関連を解析した。同時に、血管の慢性炎症・組織リモデリングに重要なPTX3、MAPK9、ARID5B、Per1、Per3について解析を行った。

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4SNPがiPロタイプを構成し冠動脈疾患と有意に相関するというデータを得た。2型糖尿病との相関、血中アディポネクチン濃度との相関は目下解析中である。

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＞

国内外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う進め方はユニークであり、事実いくつかのシーズを得始めている。また生活習慣病という切り口から、分野横断的にシーズを拾い上げるという研究の姿勢は、病態の背景に存在する共通の分子メカニズムを解明する上で最適の方法である。今後ゲノム解析を進める上で、Rare variant 検索も併行するが、海外でも昨今 Re-sequencing がゲノム解析の一手法として精力的におこなわれており、本研究はその流れとも合致する。

<達成できなかったこと、予想外の困難、その理由>

500名の心筋梗塞患者および450名の対照群のゲノムDNAをもちいて約40種類の心血管作動物質遺伝子多型を候補として解析をおこなった。KLF5 遺伝子プロモーター領域の SNP およびアドレナリンβ2受容体遺伝子コーディング領域の SNP に関しては心筋梗塞発症との間に有意な相関を認めた。しかしながら他の SNP に関しては約200名の予備的検討により相関が疑われたものの、500名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。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。

<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、2007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と低いことがわかっていて、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。例えば、心肥大の遺伝要因を考えるときに、心肥大をきたす単一遺伝子病である肥大型心筋症をモデルとしてとらえ、その原因遺伝子(サルコメアタンパク遺伝子)を第一の解析候補にする。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ダイレクトシーケンシング法を併用し網羅的解析をおこなう。現在数種類の候補遺伝子の解析準備を完了した。今後は約300例のヒト心血管病症例を対象に、厳選した遺伝子群を対象に網羅的遺伝子解析をおこなう。現時点では細胞・臓器の老化に関与する遺伝子群を考えている。既に若干名の心筋梗塞患者ゲノムを用いて予備的解析をおこないプロモーター領域やエクソン領域に塩基配列の違いをみつけている。

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。ただし前述のように血管作動物質関連遺伝子を中心としたスクリーニングでは、有意なマーカーと断言できるものはアドレナリンβ2受容体遺伝子コーディング領域の SNP のみである。われわれには転写因子 KLF のプロモーター領域の SNP の有用性を検証し機序解明まで結び付けたという本研究での実績がある。

今後の遺伝子変異解析は転写因子遺伝子にもその範囲を広げる必要がある。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ると転写因子遺伝子は有望な候補として期待できる。実際、現在予備的検討で転写因子 Mrf-2/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 Mrf-2/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、実験動物レベルでメカニズム解析をおこなう。

<成果公表リスト>

<論文発表>

0802011009

Fujimoto H, Taguchi JI, Imai Y, Ayabe S, Hashimoto H, Kobayashi H, Ogasawara K, Aizawa T, Yamakado M, Nagai R, Ohno M. Manganese superoxide dismutase polymorphism affects the oxidized low-density lipoprotein-induced apoptosis of macrophages and coronary artery disease.

Eur Heart J 2007 Oct 29; [Epub ahead of print]

0710191037

Kubota N, Yano W, Kubota T, Yamauchi T, Itoh S, Kumagai H, Kozono H, Takamoto I, Okamoto S, Shiuchi T, Suzuki R, Satoh H, Tsuchida A, Moroi M, Sugi K, Noda T, Ebinuma H, Ueta Y, Kondo T, Araki E, Ezaki O, Nagai R, Tobe K, Terauchi Y, Ueki K, Minokoshi Y, Kadowaki T. Adiponectin stimulates AMP-activated protein kinase in the hypothalamus and increases food intake. *Cell Metab* 6(1):55-68 (2007).

0710191028

Hara K, Yamauchi T, Imai Y, Manabe I, Nagai R, Kadowaki T. Reduced adiponectin level is associated with severity of coronary artery disease. *Int Heart J* 48(2):149-53 (2007).

0705081217

Yamauchi T, Nio Y, Maki T, Kobayashi M, Takazawa T, Iwabu M, Okada-Iwabu M, Kawamoto S, Kubota N, Kubota T, Ito Y, Kamon J, Tsuchida A, Kumagai K, Kozono H, Hada Y, Ogata H, Tokuyama K, Tsunoda M, Ide T, Murakami K, Awazawa M, Takamoto I, Froguel P, Hara K, Tobe K, Nagai R, Ueki K, Kadowaki T. Targeted disruption of AdipoR1 and AdipoR2 causes abrogation of adiponectin binding and metabolic actions. *Nat Med* 13(3), 332-9 (2007).

0705081224

Horikoshi M, Hara K, Ito C, Nagai R, Froguel P, Kadowaki T. A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. *Diabetologia* 50(4), 747-51 (2007)

0608061258

Horikoshi, M., Hara, K., Ohashi, J., Miyake, K., Tokunaga, K., Ito, C., Kasuga, M., Nagai, R., Kadowaki, T. A polymorphism in the AMPKalpha2 subunit gene is associated with insulin resistance and type 2 diabetes in the Japanese population. *Diabetes* 55(4), 919-23 (2006).

0608061301

Hara, K., Horikoshi, M., Kitazato, H., Ito, C., Noda, M., Ohashi, J., Froguel, P., Tokunaga, K., Tobe, K., Nagai, R., Kadowaki, T. Hepatocyte nuclear factor-4alpha P2 promoter haplotypes are associated with type 2 diabetes in the Japanese population. *Diabetes* 55(5), 1260-4 (2006).

0608061304

Nishimura, G., Manabe, I., Tsushima, K., Fujiu, K., Oishi, Y., Imai, Y., Maemura, K., Miyagishi, M., Higashi, Y., Kondoh, H., Nagai, R. DeltaEF1 mediates TGF-beta signaling in vascular smooth muscle cell differentiation. *Dev Cell* 11(1), 93-104 (2006).

<データベース/ソフトウェア>

特になし