

ヒトゲノム解析から発見した注目すべき8個の新規遺伝子の機能解析と医療への応用

●清水 信義¹⁾ ◆塩濱 愛子^{1,2)} ◆清水 厚志²⁾

1) 慶應義塾大学先端研 GSP センター 2) 慶應義塾大学医学部分子生物学教室

<研究の目的と進め方>

本研究では、永年のヒトゲノム解読から申請者らが発見した約300個の新規遺伝子の中から、予備的な性状解析に基づいて特に興味深い遺伝子8個を選抜し、それらの機能および発現調節の分子機構を解析して新たな知見を得ることを目的とする。

対象の8個の遺伝子とは、特徴的な組み合わせでドメイン・モチーフをもつ6個の遺伝子 DGCR8 (WW + DSRB)、SGSM1 (RUN + TBC)、ZNF295 (BTB/POZ + ZF + HFC-1 + HFC-2)、CSMD3 (CUB + SUSHI + TM)、FFER1L6 (C2)、PKHD1L1 (IPT + PbH1) およびドメイン・モチーフは全く不明だがユニークな発現を示す2個の遺伝子 DSCR4 (類人猿胎盤のみで発現) と VPS13B である。これらの遺伝子のうち DGCR8 は DiGeorge 症候群、CSMD3 は家族性てんかん、PKHD1L1 は多発性嚢胞腎、FFER1L6 は非症候群性難聴、DSCR4 はダウン症の関連遺伝子、VPS13B はコーエン症候群の原因遺伝子である。現在までに、いくつかの遺伝子についてはノーザンブロット法や *in situ* ハイブリダイゼーションなどによる発現解析も終了している。これらの基礎データを基に遺伝子ごとに最も有効である解析法をゲノム解析(遺伝子構造決定)、トランスクリプトーム解析(RT-PCR、ノーザンブロット解析)、プロテオーム解析(共免疫沈降法、マスペクトル解析)、フェノーム解析(ノックアウト、ノックダウン)などから選択し、それぞれ独自の解析を順次進めていく。

<2007年度の研究の当初計画>

各遺伝子についてヒトゲノム解読研究から発見した特に重要度の高い8個の新規遺伝子について以下のゲノム --> トランスクリプトーム --> プロテオーム --> フェノームにわたる統合的な解析を順次行う。

- 1) ゲノム解析：RT-PCRによるバリエーションを含む全長構造決定、比較ゲノム解析
- 2) トランスクリプトーム解析：MTCパネルや発生時期別cDNAパネルを用いたRT-PCRによる発現解析、ノーザンブロット解析、WISH法などによる発現動態の観察、レポーター遺伝子を用いたプロモーター解析
- 3) プロテオーム解析：ドメイン・モチーフ構造の解析、一本鎖抗体による細胞内局在解析、マスペクトル解析によるタンパク質修飾部位の同定、共沈殿物のマスペクトル解析による相互作用タンパクの同定
- 4) フェノーム解析：マウスを用いたノックアウト解析、メダカを用いたノックダウン解析
- 5) 上記解析結果は当研究室で独自に開発した遺伝子データベースに順次入力し、さらなる研究の促進に供する。また、各段階で得られた解析結果と疾患との関連については適宜検討する。

<2007年度の成果>

DGCR8 は1つの WW ドメインと2つの DSRB ドメインを併せ持ったヒトゲノム中に1つしかない極めて稀な構造を持つタン

パク質を産生する。マウス胎児を用いた発現解析で、DiGeorge/CAFS 症候群の発症組織(特に胸腺)に一致するパターンを認めため、我々は DiGeorge 症候群の原因遺伝子の一つとして提唱してきた。その後、pri-miRNA のプロセッシングに必須な RNase 活性をもつ Drosha との相互作用が報告され、RNA 干渉における役割がにわかに関与を浴びている。Drosha、Nucleolin とともに主に核小体近傍に局在した。Nucleolin、ILF3、XPO5 など11種類のタンパク質とヘテロ2量体を形成することを見出した。さらに、これらタンパクの強制発現と免疫染色の結果から、DGCR8 は RNA 干渉において miRNA 形成過程で重要な役割を果たすというモデルを提唱した (Shiohama, A. *et. al. Exp Cell Res.* (2007))。

SGSM1 は RUN と TBC モチーフを持つ SGSM 遺伝子ファミリーに属し、特に脳で強い発現を示しており機能が注目されている。そこでこれら SGSM ファミリーの全長構造決定、さらに全長クローニングを行い発現系の構築を行った。SGSM ファミリーはいずれも脳で強い発現を示し、特に SGSM1 は海馬のアンモン角や歯状回、小脳の Prukunje 細胞層や顆粒層、大脳皮質の神経細胞などで発現を確認できた。SGSM ファミリータンパク質の RUN モチーフを含む領域は RAP ファミリー蛋白質と相互作用を示すことが報告されている。そこで、SGSM1/2/3 と RAP ファミリーとの蛋白質相互作用解析を免疫沈降法により行った。その結果、新規 RAPID (RAP Interacting Domain) モチーフを発見し、RAP サブファミリー(4個)の全てと RAPID モチーフを介して、RAB サブファミリー(70個)の幾つかと TBC モチーフを介して結合することが明らかになった。さらに、内在性 SGSM1 タンパクがトランスゴルジネットワークに局在することが観察された。SGSM1 で代表される新規タンパクファミリーは小型 G タンパクの作用をさらに調節するモジュレーターと考えられ、新たに SGSM (Small G protein Signaling Modulator) 1/2/3 と命名した (Yang, H., *et. al. Genomics.* 90: 249-260 (2007))。

ZNF295 は転写抑制因子であるが、その転写抑制活性は BTB/POZ ドメイン以外に新規ドメイン (HFC-1 と命名) にも存在することが判った。またドーパミントランスポーター遺伝子の転写活性因子 ZFP161 と結合してその転写を抑制することを見出した (BBRC. 327:615-627, 2005)。さらに、メダカのオルソログを同定して、ノックダウン解析を行なったところ、メダカ脳全体の縮小と脳室の拡大が観察された。

FER1L6 は非症候群性難聴 DFNB9 の原因遺伝子 OTOFERLIN にもっとも類似していた。FER1L ファミリーは6メンバーから成り、総て2000aa程の巨大タンパク質であった。メダカをモデル生物として採用し、ノックダウン解析を行なったが、明瞭な表現型は得られなかった。

CSMD3 は類似の3遺伝子で CSMD ファミリーを構成している。連鎖解析により CSMD3 が良性的家族性てんかん (BAFME1/FAME) の原因候補遺伝子と考えている。一方、CSMD3 遺伝子

の配列は進化的に著しい変動を示したため脳の構築や機能の多様化に関与すると考えられた (Shimizu, A., et.al. *Biochem. Biophys. Res. Commun.*, 309:143-154 (2003))。メダカをモデル生物として採用し、ノックダウン解析を行なったが、明瞭な表現型は得られなかった。

PKHD1L1 は多発性嚢胞腎 (ARPKD) の原因遺伝子である PKHD1 のファミリー遺伝子である。モデル生物としてメダカを選択し、RT-PCR による全長構造決定を試み、75 エキソンの配列決定を終了した。モルフォリノアンチセンスオリゴによるノックダウン解析を行なったが、明瞭な表現型は得られなかった。

DSCR4 はヒト胎盤にのみ検出され、プロモーター領域の欠失変異体を作製しプロモーター活性を測定したところ、nt-800 付近で著しい活性の上昇と抑制を示す塩基配列が見出された。これら配列から推定された cis- エレメントから OLF-1, E47 様因子によって発現制御を受ける可能性が示唆された (Asai, S., et. al. *Biochim Biophys Acta.* (2008))。

VPS13B は本研究進行中に Cohen 症候群の原因遺伝子であることが報告されたが、我々も日本人患者で新規変異との相関を確認できた (*Clin. Genet.*, 67:270-272 (2005))。COH1 は VPS13 遺伝子ファミリーに属する VPS13B である。メダカをモデル生物として採用し、ノックダウン解析を行なったが、明瞭な表現型は得られなかった。

<国内外での成果の位置づけ>

米国の ENCODE プロジェクトでは特定の領域を対象に詳細なゲノム機能の解明を進めているが、これは申請者が進めてきたゲノム --> トランスクリプトーム --> プロテオーム --> フェノームにわたる統合的な解析が世界のスタンダードとして認知されていることを示している。

特に近年、表現型解析および新たな疾患モデルとしてゼブラフィッシュあるいはメダカが着目されているが、我々は研究当初からゲノムの基盤整備も含めて解析対象としてメダカを活用しており、我々の解析結果は国内外で高い評価を得ている。

<達成できなかったこと、予想外の困難、その理由>

本研究で対象としている遺伝子全てのメダカオルソログを同定し、順次モルフォリノアンチセンスオリゴを用いてノックダウン解析を進めたが、いくつかの遺伝子については明瞭な表現型が得られていない。そこでポジティブコントロールとして複数の遺伝子を選択し、ノックダウンを行った結果、配列によっては全く表現型がでないもの、また単独のモルフォリノでは表現型が確認できないものがあった。今後、遺伝子内の複数の箇所に対するアンチセンスオリゴを用いることで改善が見られると考えている。

<今後の課題>

引き続き 8 個の新規遺伝子について以下のゲノム --> トランスクリプトーム --> プロテオーム --> フェノームにわたる統合的な解析を順次行う。

1) ゲノム解析：マウス、メダカ組織由来 mRNA および臓器別 cDNA を鋳型にして RT-PCR を行いオルソログ cDNA を単離する。シーケンスを決定し、ゲノムシーケンスに照合することによって各遺伝子のモデル生物におけるゲノム構造を決定する。分子系統樹を作成して構造と機能の多様性を進化的に検証する

2) トランスクリプトーム解析：初期発生段階の解析にメダカ発生時期別 cDNA を鋳型に用いて発現解析を行う。1 の情報を特に様々なスプライシング、複数の転写開始点や終止点、その結果

生ずる転写産物の多様性に留意して、トランスクリプトバリエーションの全体像を解明する。ヒト組織、マウス成体及び胎仔、メダカ初期胚の組織に対してノーザンブロット法、*in situ* ハイブリダイゼーション法を用いて、各遺伝子の細胞・組織・発生レベルでの空間的・時間的発現パターンを徹底的に解析する。

3) プロテオーム解析：ドメイン・モチーフ構造の解析、抗体染色による細胞内局在解析、マスマスペクトル解析によるタンパク質修飾部位の同定、共沈殿物のマスマスペクトル解析による相互作用タンパクの同定を行う。

4) フェノーム解析：マウスを用いたノックアウト解析、メダカを用いたノックダウン解析を行う。

特に進展のあった DGCR8 と SGSM1 に関しては以下の解析を現在計画している。

DGCR8

現在作製中のノックアウトマウスを用いた表現型解析、及び幹細胞を用いた miRNA の発現解析など、DGCR8 と miRNA の発生・分化段階におよぼす影響に注目した解析を目指す。

SGSM1

SGSM ファミリーの機能解析に関しては、SGSM タンパクと特異的に反応する RAB ファミリーメンバーがまだ同定されていないため、mammalian 2-hybrid system を用いて、60 以上の RAB メンバーの中から SGSMS1/2/3 に特異的に相互作用する RAB タンパクを同定する予定である。また、SGSM タンパクが RAB を介して、細胞内小胞輸送に関与すると予測されるので、SGSM タンパクの強制発現およびノックダウン後の細胞内小胞輸送の変化を観察する予定である。

<成果公表リスト>

0710301210

Asai, S., Yamaki, A., Kudoh, J., Shimizu, N. and Shimizu, Y.; Analysis of the promoter region of human placenta-specific DSCR4 gene. *Biochim. Biophys Acta.* 1779: 40-50 (2008).

0710301208

Yang, H., Sasaki, T., Minoshima, S., and Shimizu, N.; Identification of three novel proteins (SGSM1, 2, 3) which modulate small G protein (RAP and RAB)-mediated signaling pathway. *Genomics.* 90: 249-260 (2007)

0710301206

Sasaki, T. and Shimizu, N.; Evolutionary conservation of a unique amino acid sequence in human DICER protein essential for binding to Argonaute family proteins. *Gene.* 396: 312-320 (2007).

0702141159

Shimizu, N., Ohtsubo, M. and Minoshima, S.; *MutationView/KMccancerDB*: a database for cancer gene mutations. *Cancer Sci.* 98: 259-267 (2007)

0710301454

Shiohama, A., Sasaki, T., Noda, S., Minoshima, S. and Shimizu, N.; Nucleolar localization of DGCR8 and identification of eleven DGCR8-associated proteins., *Exp Cell Res.* 313: 4196-4207 (2007)