

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 14 日現在

機関番号：32665

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25460703

研究課題名(和文)ボトムアッププロテオミクスを応用した多因子遺伝性疾患感受性遺伝子検索

研究課題名(英文)The research for the susceptibility genes of multifactorial genetic disease applied bottom-up proteomics

研究代表者

中山 智祥 (NAKAYAMA, Tomohiro)

日本大学・医学部・教授

研究者番号：00339334

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：主に本態性高血圧症、脳梗塞、心筋梗塞、糖尿病の78症例を収集した。同意書取得後に採血しゲノムDNAを抽出した。家族歴を正確に聴取することで、遺伝性の関与を検証した。臨床データは、性別、年齢、当該疾患家族歴、身長、体重、Body mass index、総コレステロール、中性脂肪、HDLコレステロール、尿酸、血糖、尿データ、心電図、喫煙歴、飲酒歴をインプットした。血漿はディープフリーザーに保存しMALDI-TOF MSによるプロテオミクス解析に使用した。ゲノミクス解析についてはエクソーム解析およびターゲットシーケンスに次世代シーケンサーを使用、個々の遺伝子の解析にはサンガー法を用いた。

研究成果の概要(英文)：We have collected the seventy-eight samples with essential hypertension, cerebral infarction, myocardial infarction, diabetes. After getting a written consent, we extracted genomic DNA from white blood cells. We inspected participants whether they have family history of the disorders by hearing exact information about family careers. The clinical data such as sex, age, family history, height, weight, body mass index, total cholesterol, triglyceride, HDL cholesterol, uric acid, blood sugar, urine acids, electrocardiogram, a smoking history, drinking career. The DNA samples were stocked in deep freezers until analysis, and the isolated plasma were analyzed for proteomics analysis by the MALDI-TOF MS.

A next-generation sequencer was used for the genomics analysis especially for exome analysis. In addition, Sanger method was applied for the individual genetic analysis. SNPalyze software was used for haplotype-based case-control study.

研究分野：臨床遺伝学

キーワード：遺伝性疾患 プロテオミクス 感受性遺伝子 疾患マーカー

1. 研究開始当初の背景

- (1) ヒトゲノム解析計画が終了したことによって、臨床遺伝学の研究は単なる塩基配列の解読から遺伝性疾患の原因・病態解明へと移行した。すなわち遺伝性疾患の原因遺伝子がヒトゲノム上に存在する約 22,000 個の遺伝子のどれかという事のみならず、原因変異の同定や変異と疾患の重症度(表現型)などとの関係が究明課題として残された。
- (2) 単一遺伝子疾患の原因変異の多くが家系を用いた連鎖解析による遺伝子クローニングによって同定された。一方多因子遺伝性疾患の感受性遺伝子同定法では今まで様々な方法論が採られてきており、候補遺伝子内・近傍の遺伝子マーカーを用いた関連解析、一塩基多型 (single nucleotide polymorphism: SNP) を用いて網羅的検索を行う genome-wide association study: GWAS と広がってきている。そして次世代シーケンサーと呼ばれる大きなスペックを持つ解析機器が導入されるようになった。
- (3) GWAS では関連があるとのデータが出た SNP を中心に周辺の他の SNP との連鎖不平衡解析を通して各 SNP が同一ハプロタイプに存在しているかを確認した上で塩基配列決定をしてゆく方法が主となっている。また網羅的に全ゲノムシーケンシングでは家系内、あるいは同じ疾患を有する群で全ゲノムあるいは exon 領域のみを解析するエクソーム解析にて塩基配列を決定し、コントロールと比較して感受性遺伝子に到達する。しかし、未だに多因子遺伝性疾患の疾患感受性遺伝子の同定には至らないことが多い。
- (4) 本(成果)報告書の中で用いているトップダウンとボトムアップという単語はもともと情報や知識の順序付けなどの設計戦略に使用される用語である。トップダウン設計は、段階的に詳細にしていく設計技法で、最初にシステム全体を定式化し個々の詳細には立ち入らないが、その後システムの個々の部分の設計を詳細化していく。一方、ボトムアップ設計では、最初にシステムを構成する個々の部品を細部まで設計し部品群を組み合わせより大きな部分を作っていく、最終的にシステム全体が構成される。ソフトウェア工学など様々な分野で使われる用語である。

プロテオミクス研究ではタンパク質を酵素消化し小さなペプチドにした後に質量分析を行いタンパク質同定するボトムアッププロテオミクスと、酵素消化なしで質量分析しタンパク質同定するというトップダウンプロテオミクスが行われる。後者はサンプル調整に工夫

が必要で技術的に困難な面が大きいいため、前者が広く行われる。ゲノム研究であるゲノミクスとプロテオミクス研究を結びつける研究は多くの研究室で発想されることであるが、反面実際に遂行される例はほとんど無いに等しい。これは研究者間の連携の問題や研究方向性の問題などがあると考えられる。

- (5) 実験手技としてのボトムアッププロテオミクスという意義も包含し、プロテオミクス、ゲノミクス、臨床データベース解析が個々に解析され、最終的に全体のデータベースとして疾患感受性同定に至るという研究戦略そのものを指すボトムアップ戦略は今まで例がない。これは研究者間の連携の問題や研究方向性の問題などがあると考えられる。

2. 研究の目的

多因子遺伝性疾患の感受性遺伝子検索に関して今まで様々な方法論が採られてきており、候補遺伝子内・近傍の遺伝子マーカーを用いた関連解析、一塩基多型 (single nucleotide polymorphism: SNP) を用いて網羅的検索を行う genome-wide association study: GWAS そして次世代シーケンサーを用いて網羅的に全ゲノムシーケンシングが行われるようになった。しかし多因子遺伝性疾患の疾患感受性遺伝子の同定には至らないことが多い。やみくもに塩基配列決定することはコストや労力の浪費につながるため、効率的な方法が望まれる。本申請者のこれまでの様々な疾患感受性遺伝子同定法の経験を生かして、新たな方法を試みたい。本研究の目的は独自のボトムアッププロテオミクス手法(ボトムアップ戦略)で大まかにターゲットを絞り、exon シーケンス(エクソーム)解析を行い、効果的に感受性遺伝子を同定する方法を確立することである。

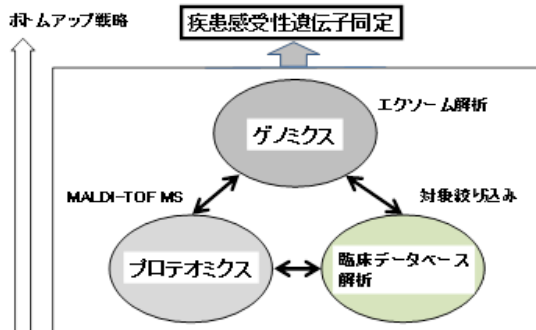
3. 研究の方法

(1) 概要

本研究では多因子遺伝性疾患として申請者が今まで解析してきたものの中で本態性高血圧症、脳梗塞、心筋梗塞、糖尿病について、やみくもに解析するのではなく、臨床データを活用しターゲットを絞り込む。その後に対象者の血清を採取しボトムアッププロテオミクスを実施する。特に重要なデータとして家族歴、薬剤反応性(治療前後でのデータ変動やアレルギー)に着目したい。絞られた蛋白質情報を基に遺伝子に着目し、家系内サンプル DNA を基にエクソーム解析をし、疾患感受性遺伝子および感受性多型・変異を同定する。今までの方法論である case-control study (関連解析) を行い本方法論の正当性を検証する。

本研究のボトムアップ戦略はボトムアッププロテオミクスを包含し、プロテ

オミクス、ゲノミクス、臨床データベース解析を統合しその情報をもとに疾患感受性遺伝子同定に至る新しい研究戦略を意味する。



<図の説明>

プロテオミクス、ゲノミクス、臨床データベース解析の3つの方法から出た結果を統合し、最終的に疾患感受性遺伝子の同定に至るというボトムアップ戦略の流れを示す。

(2) 臨床データベース解析

これまで記録してきた臨床データで多因子遺伝性疾患感受性遺伝子検索に必要なものである性別、年齢、当該疾患家族歴、身長、体重、Body mass index、総コレステロール、中性脂肪、HDL コレステロール、尿酸、血糖、尿データ、心電図所見、家族歴、喫煙歴、飲酒歴聴取はそのまま継続する。特に家族歴と薬剤反応性の違いについて詳細に検討し、同一家系で疾患の有無、薬剤応答性の相違で分類してボトムアッププロテオミクス用の対象を絞り込む。

(3) サンプル収集

対象として今までに本申請者の所属施設で収集した本態性高血圧症、脳梗塞、心筋梗塞、糖尿病の血液検体で関連解析に使用可能なものは合計6,000例に及ぶ。これらのDNAは使用可能な状態で保存されている。上記1で絞り込まれた対象について新たに血清を採取する。血液はEDTA塩入り試験管とプレーン試験管とにとりわけ前者はDNA抽出と血漿用、後者は血清用とする。収集と同時にDNAを抽出する。血漿および血清は蛋白質解析用に凍結保存しておく。

(4) ボトムアッププロテオミクス

上記で収集したサンプルについてMALDI-TOF MSでボトムアッププロテオミクスを実施する。蛋白質データベースなど必要機器、ソフトウェアはすでに揃っている。

(5) 遺伝子エクソーム解析

本解析は絞られた対象についてターゲット遺伝子について次世代シーケンサーによって遂行する。次世代シーケンサーはすでに研究室に搬入された(株)イルミナ社のMiSeqを使用する。

(6) 遺伝子型の決定

同定した原因遺伝子内の多型を用い、初年度から本年度までに収集したサンプルを関連解析用に遺伝子型決定する。本態性高血圧症を対象とした際には血管収縮・拡張、増殖に関与する遺伝子が同定される可能性が期待できる。

(7) ハプロタイプを用いた関連解析

遺伝子型決定されたデータをコンピューターにインプットし、遺伝子型のデータをもとにハプロタイプを用いた関連解析を行う。ハプロタイプ構築についてはSoftwareとして申請者がすでに所有しているSNPAlyze software (Dynacom社 横浜)が利用する。

ハプロタイプを用いた関連解析の利点

コントロール群			疾患群		
C/T	A/T	G/A	C/T	A/T	G/A
-C	-A	-G	-C	×	-T
-T	-T	-G	-T	-T	-G
-C	-A	-G	-C	-T	-A
-C	-T	-A	-C	×	-T
-T	-T	-G	-T	-A	-A
-C	-T	-A	-C	-A	-G

<図の説明>

C/T, A/T, G/Aの3個のSNPを用いてcase-control studyを行なった場合C、G、T、いずれのアレル頻度も66.7%で患者群とコントロール群とで差がない。このようにいずれのSNPも一般集団において、ありふれた型である場合、個々のSNPに関して関連解析を行っても統計学的な有意差は得られないことが多い。そこでハプロタイプを用いて解析した結果、疾患の感受性変異・多型(×で示す)と連鎖不平衡状態にあるC-T-Gハプロタイプは疾患群にしか認められないことがわかる。

(8) 感受性変異・多型とハプロタイプとの連鎖の確認

ハプロタイプを用いた関連解析でcase群にのみ認められたハプロタイプに連鎖している変異・多型が存在すると考えられる。原因変異は常染色体優性遺伝形式の場合基本的には罹患者に1個、常染色体劣性遺伝形式の場合、ホモ接合体の場合1個、複合ヘテロ接合体の場合2個である。その観点から原因変異に連鎖するハプロタイプを判別する。

(9) 疾患感受性遺伝子同定

プロテオミクス、ゲノミクス、臨床データベース解析による疾患感受性遺伝子検索を実施し、疾患感受性遺伝子を同定する。これら3つのデータから出されたものから疾患感受性遺伝子、またその多型・変異を同定していく手法であり、全体としてプロテオミクスに限らずポ

トムアップ戦略となる。こうした手法は斬新であると考え。

#### 4. 研究成果

本研究でのターゲットになっている多因子遺伝性疾患のうち本態性高血圧症、脳梗塞、心筋梗塞、糖尿病について、末梢血サンプル収集が進んでいる。研究者によって研究説明し、同意書にサインをいただいた後に採血、パフイーコートの白血球からゲノムDNAを抽出した。一部のDNAは濃度調整し、今後すぐに使用可能な状態にある。これらのサンプルについては特に家族歴を正確に聴取することで、遺伝性が関与しているか否かを検証した。臨床データベースのインプットについてはこれまで記録してきた臨床データで多因子遺伝性疾患感受性遺伝子検索に必要なものである性別、年齢、当該疾患家族歴、身長、体重、Body mass index、総コレステロール、中性脂肪、HDLコレステロール、尿酸、血糖、尿データ、心電図所見、喫煙歴、飲酒歴聴取はそのまま継続している。

プロテオミクス解析実施に用いるMALDI-TOF MSはすでに運用されており、蛋白質データベースなど必要機器、ソフトウェアを用いている。分離された血漿は分析までは-80のディープフリーザの中に保存し、必要時にプロテオミクス解析に使用した。

ゲノミクス解析については次世代シーケンサーが初年度の途中に配備され、運用が開始となった。特にエクソーム解析実施および特定領域における多検体解析に力を発揮できる。また、個々の遺伝子解析についてはキャピラリーシーケンサーを用いた。

今後もサンプル収集を継続していくとともに過去に収集したサンプルについてデータのインプット作業をこなし、更に解析可能なサンプル数を増やしていく予定である。ハプロタイプ関連解析にはSNPAlyze softwareを用いた。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 1 件)

- (1) 中山智祥、Jiang Jie、相馬正義、青井則子：SMTN 遺伝子と本態性高血圧症とのハプロタイプを用いた関連解析。第 36 回 日本分子生物学会年会、神戸国際会議場、兵庫県神戸市、2013.12.3-12.6 (Poster 12.5)

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年月日：  
国内外の別：

取得状況(計 0 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年月日：  
国内外の別：

〔その他〕  
ホームページ等

#### 6. 研究組織

##### (1) 研究代表者

中山 智祥 (NAKAYAMA, Tomohiro)  
日本大学・医学部・教授  
研究者番号：00339334

##### (2) 研究分担者

久保田 健夫 (KUBOTA, Takeo)  
山梨大学・総合研究部医学・教授  
研究者番号：70293511

野村 文夫 (NOMURA, Fumio)  
千葉大学・医学(系)研究科(研究院)・教授  
研究者番号：80164739

##### (3) 連携研究者

なし