

## 科学研究費助成事業 研究成果報告書

平成 26 年 5 月 27 日現在

機関番号：32653

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23592108

研究課題名(和文)脳動脈瘤家系のパーソナルゲノムの全塩基配列決定：特にゲノム構造多型の解析

研究課題名(英文)Next generation sequencing of a large pedigree with intracranial aneurysms

研究代表者

糟谷 英俊 (KASUYA, HIDETOSHI)

東京女子医科大学・医学部・教授

研究者番号：50169455

交付決定額(研究期間全体)：(直接経費) 4,000,000円、(間接経費) 1,200,000円

研究成果の概要(和文)：我々は85の脳動脈瘤家系を用いた罹患同胞対ゲノムワイド連鎖解析により、染色体7q11領域に疾患との連鎖を報告している。これら罹患同胞対家系の長期フォローにより、子の世代に新たに罹患者を生じた家系が認められた。本研究ではこの家系を対象に次世代シーケンサーによる網羅的遺伝子配列解析を行い、感受性遺伝子の特定を目指した。大量の配列データから目的の変異を絞り込むため、本家系の連鎖解析も加え効率よく候補の絞り込みを行った。その結果、染色体7q11近傍およびその他の新規連鎖領域から、有望な候補遺伝子が特定された。これらは脳動脈瘤の成因としてかねてから予想されていた分子生物学的機序に関連する遺伝子であった。

研究成果の概要(英文)：Recent advances in sequencing technologies, in particular the next generation sequencing (NGS), have led us to focus on rare variant mapping together with a renewed interest in a familial linkage analysis. In the present study, we successfully identified susceptibility genes for intracranial aneurysms (IA) using NGS in combination with a linkage analysis of a large IA pedigree. These susceptibility genes were associated with molecular causes of atherosclerosis or polycystic kidney disease, both known to be initiating IA formation.

研究分野：医歯薬学

科研費の分科・細目：外科系臨床医学・脳神経外科

キーワード：脳動脈瘤 感受性遺伝子 次世代シーケンサー

## 1. 研究開始当初の背景

脳動脈瘤は頭蓋内の主幹動脈の分岐部に生じる囊状の血管の病的な膨らみで、その破裂によりクモ膜下出血を生じる。クモ膜下出血を来した患者の転機は、その半数以上が死亡または重大な後遺症を残す重篤なもので、診断法・治療法が進歩した現在においても、医療上重要な疾患のひとつと考えられている。脳動脈瘤の成因としては、高血圧症、喫煙、血行力学的要因などの環境要因と、家族集積性・脳動脈瘤を合併しやすい遺伝性疾患の存在から遺伝要因が想定され、これらの複合的作用により発生すると考えられている。このような遺伝的背景の存在から、成因の解明を目的に遺伝要因の解析が盛んに行われてきた。その解析手法として、かつては罹患家系を用いた連鎖解析が行われ、我々も当時は脳動脈瘤では前例がなかったゲノムワイド連鎖解析を行い、第 5、7、14 番染色体に連鎖領域を特定した(Onda et al. *Am J Hum Genet.* 69:804-819,2002)。これを皮切りに国内外で多数の連鎖解析や関連解析が相次いだ。サンプル数が少なく結果の信頼性が低いものも多かった。現在では高密度マイクロアレイを用いた大規模コホートでのゲノムワイド関連解析 (genome-wide association study: GWAS) が主流となり、脳動脈瘤においても我々とエール大学、フィンランド、オランダの他施設共同研究により再現性・信頼性の高い感受性領域がいくつか特定されている (Bilguvar et al. *Nature Genet.* 40:1472-7, 2008、Yasuno et al. *Nature Genet.* 42:420-5 2010)。いずれも機能的にも有望な遺伝子座ではあるが、脳動脈瘤発生のメカニズムとの分子生物学的関わりについてはほとんど検証されていないのが現状である。効果サイズが小さく、リスクアレルによりもたらされる生物学的効果が軽微であるのも一因であろう。脳動脈瘤のような多因子疾患においては、GWAS のアプローチでは頻度が低

く効果サイズが高いバリエーション (いわゆる rare variant) や、構造多型 (structural variant: SV) 検出し得ない問題も指摘されている。近年、次世代型シーケンサーの登場により、個人の全ゲノム配列までも決定することが可能となった。これを利用して次々に単一遺伝子疾患の原因遺伝子が特定されており、多因子疾患においても GWAS では捉えることのできない遺伝要因の同定にその応用が期待されている。

## 2. 研究の目的

本課題では、非常に多数のサンプル数を解析する GWAS のような大規模関連解析から、今一度、遺伝的背景が濃厚な家系サンプルに立ち返った。我々は先に行った罹患同胞対連鎖解析において、ゲノム全域を対象としたマイクロサテライトマーカーの遺伝子型タイピングを完了している家族性脳動脈瘤家系を多数保有している。この中から、特に家系内に多数の罹患者が認められ、かつ先の罹患同胞対連鎖解析で特定された領域と連鎖している家系に注目した。このような遺伝的背景が濃厚で、既に疾患候補遺伝子領域がある程度しぼり込まれているサンプルを対象にして、次世代シーケンサーを用いた網羅的遺伝子配列解析を行う計画を立案した。これにより、病態解明や治療に直結するような新規感受性遺伝子の特定を目指した。

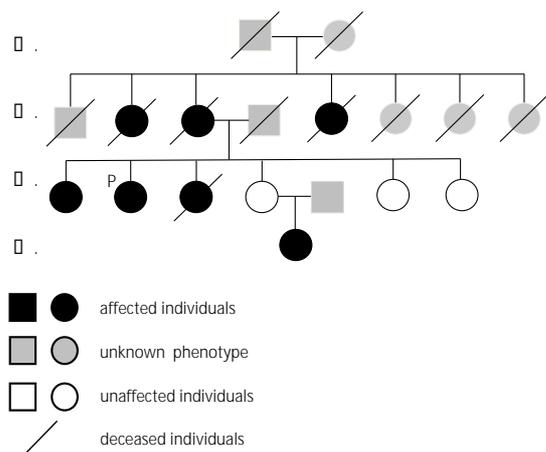
## 3. 研究の方法

### 解析対象の選定

我々は先のゲノムワイド連鎖解析でも使用した家族性脳動脈瘤 90 家系のサンプルを有している。このうち、非罹患同胞も含めてサンプリングできている家系が 18 家系あった。脳動脈瘤は中年期以降に発症してくる場合も多いことから、注意深く家系追跡を行ったところ、子の世代に新たに脳動脈瘤患者を生じた家系が観察された。我々の罹患同胞対

連鎖解析で最も強い連鎖を認めた染色体 7q11 とも連鎖している家系であり、渉猟しうる限り家系内で 7 名の罹患者を有していた。本研究課題に最も適した家系と考えられたので、子の世代の新たな罹患者からもサンプリングを追加し、解析に供した。家系図を下に示す(図 1)。

図 1. 本課題で使用した脳動脈瘤家系



#### 次世代シーケンシング

世代 0 の 6 姉妹と、その子の世代 -1 からサンプリングが出来た。これら 7 検体について本学の SOLiDs システム(Lifetechnologies 社)を用いた次世代シーケンス解析を行った。発端者 -1 については全ゲノムシーケンスと全エクソームシーケンスを、残りの 6 検体については全エクソームシーケンスを行った。全ゲノムシーケンスでは、フラグメントライブラリに加え、Lifetechnologies 社が提唱しているメイトペアライブラリの作成も行い、比較的大きな欠失や挿入等 SV の検出も行えるようにした。具体的な作成方法としては、物理的に断片化しサイズ選択したサンプル DNA 端に内部アダプターを結合する。両端の内部アダプターを結合することにより環状化し、次いで酵素反応により内部アダプター両端にサンプル DNA 端が均等に含まれるよう切り出しを行う。このように切り出された DNA 断片の両端に P1、P2 アダプターをさらに結合させ、エマルジョン PCR により増幅してメイトペアライブラリとする。全エクソ

ームについては、アジレント・テクノロジー社の SureSelect ターゲットエンリッチメントシステムを用いて全エクソン領域をキャプチャーした。

さらに、本課題では次世代シーケンスデータのクオリティコントロールと家系での連鎖解析を目的に、次世代シーケンスに供した同一のサンプルで Omni1-quad SNP array (イルミナ社)による全ゲノム領域およそ 100 万塩基多型 (single nucleotide polymorphism: SNP) のジェノタイプングも行った。

以上により、本家系においては全ゲノムにわたって、一塩基置換 (single nucleotide variant: SNV)、小さな欠失・挿入 (small indel) さらに SV も含めて網羅的に解析が行える体制を整えた。

#### 4. 研究成果

##### 家系情報より想定される遺伝様式

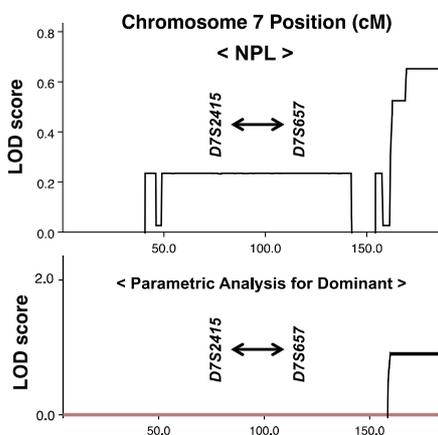
本家系は渉猟しうる限り血族婚はなく、優性遺伝形式が想定される。脳動脈瘤はいわゆる common disease であり、不完全浸透や表現模写も当然考慮に入れなくてはならない。様々なパターンの組み合わせが考えられるが、本課題では特に、最も考えうる次のパターンをまず想定して解析を進めた。1) -4 は保因者で罹患者全員が同じ疾患リスクアレルを共有している、2) -4 は疾患リスクアレルのない非罹患者で -1 の遺伝的リスクは父方由来。

##### 第 7 染色体連鎖領域の解析

先の罹患同胞対連鎖解析で最も高い連鎖を認めたのは、染色体 7q11 を中心とした領域であり、1 lod drop interval はマーカー D7S2415 (7q11.22) から D7S657 (7q21.3) までのおよそ 21Mb に相当する。本家系世代 0 の 3 姉妹は、この領域でハプロタイプを共有していた。まず、新たに罹患者 -1 が加わったので、高密度 SNP データを用いて連鎖解析を行

った。連鎖不平衡にあるマーカーの pruning を行い、第 7 染色体全体を 2199 個の SNP でカバーした。浸透率 2/3 (第 世代の非罹患者のうち 2 名での不完全浸透を許容) の優性遺伝形式でパラメトリック連鎖解析と、遺伝形式を規定しないノンパラメトリック連鎖解析を行い複雑な遺伝形式を検討できるようにした。計算には MERLIN (Abecasis et al. Nat Genet. 30:97-101, 2002) を用いた。単一家系で、サンプルは近親者に限られているため、高いロッド値は期待できない。そこで、次世代シーケンスデータと組み合わせるときにはロッド値が正の領域を採用することが一般的であるが、さらに精度を高めるため、IBD (identity by descent) セグメント推定も加えて行っている。本家系 7 例の SNP アレイデータを国際 HapMap プロジェクト (<http://hapmap.ncbi.nlm.nih.gov/>) で公開されている日本人 89 人分の全ゲノム SNP 遺伝子型データ (Phase III) と結合して相決定を行うことにより、家系内でどのメンバーが領域のハプロタイプを共有しているかがわかる。計算には BEAGLE version 3.3.2 (<http://faculty.washington.edu/browning/beagle/beagle.html>) を用いた。ゲノム全域にわたって検討してみると、パラメトリック連鎖解析で正の連鎖を示した領域は、すべて罹患患者全員で IBD セグメントが共有されている領域であった。第 7 染色体における連鎖解析の結果を次に示す (図 2)。

図 2. 第 7 染色体連鎖解析



ノンパラメトリック解析では、先の罹患同胞対連鎖解析での 1 lod drop interval に一致して正の連鎖がみられた。しかし、パラメトリック解析の方では 1 lod drop interval に連鎖のシグナルはない。このことから、本家系は第 7 染色体において、世代 の罹患同胞はアレルを共有しているが、子の世代の罹患患者には伝達されていないことが明らかになった。

そこで、この連鎖領域において世代 姉妹罹患患者で共有される病的変異を次世代シーケンスデータにより検索した。その結果、明らかに pathogenic な遺伝子変異 (フレームシフト変異) が、マイクロサテライトマーカーによるハプロタイプ共有パターンと完全に一致して観察された。これはサンガー法によるバリデーションでも正しく確認され、その他の家系からも同一の変異が検出された。この変異は db SNP ID が付与されていない、1000 人ゲノムプロジェクト等でも検出されていない稀な変異であった。論文未発表であるため、ここで遺伝子名は述べられないが、この変異遺伝子は血管内非細胞にも発現が認められており、動脈硬化の形成において重要な役割を果たすことが明らかにされている。近年、肥満や冠動脈疾患との関連も報告されており、非常に有望な候補遺伝子であると考えられた。

先行連鎖解析の枠をこえた網羅的検討

第 7 染色体の検索と並行し、先行研究の結果に限定せず、その他の領域にも可能性を広げて検索を行った。その際、よりインパクトの高い成果が期待される「 -4 は保因者で罹患患者全員が同じ疾患リスクアレルを共有している」というパターンを想定して検索を進めた。膨大な領域とデータの検討になるが、前出の IBD セグメント解析とパラメトリック連鎖解析の組み合わせが有用であった。先にも述べたように、全ゲノム領域のパラメトリック連鎖解析で LOD 値が正の連鎖領域は、IBD

セグメント解析により罹患者および保因者全員が同一のハプロタイプを共有していた。次世代シーケンスデータで非罹患2名も含めて参照し、IBD 共有パターンと矛盾なく観察される変異を抽出することで、高精度なフィルタリングを行うことが出来た。いくつかの有望なミスセンス変異がサンガー法による検証を経て検出された。SV についても検索を行ったが、精度の高いものは同定されなかった。これらの有望な候補ミスセンス変異のうち、多発嚢胞腎症関連分子の結合パートナーとして同定された遺伝子が存在した。さらにこの遺伝子は、他施設からの脳動脈瘤連鎖解析により特定された連鎖領域に存在していた。多発嚢胞腎症は高率に脳動脈瘤を合併する疾患として知られ、かねてより脳動脈瘤発症機序解明の糸口になりうると考えられてきた。この変異は 1000 人ゲノムデータベースではアジア人特有にごく低頻度観察される変異であるが、我々の非脳動脈瘤コントロール 244 例では観察されず、種々のバイオインフォマティクスツールにより高い機能障害性が予測されている。さらに他の脳動脈瘤家系からも同一の変異や別の変異も検出されている。こちらの遺伝子についてもあわせて論文未発表のため詳細は述べることはできないが、病態解明のための非常に有望な候補遺伝子と考えられた。

#### 総括

GWAS が盛んに行われる以前は家系による連鎖解析が盛んに行われ、有望な遺伝子座が複数報告されていたが、検出される連鎖領域は数十メガ塩基におよぶこともあり、その後の遺伝子やその変異の特定は困難を極めた。しかし、近年の次世代シーケンサー登場により病因変異そのものを直接検出することが可能となり、その膨大なデータを効率よくフィルタリングしていくために連鎖解析の手法が再び注目されている。興味深いことに、過去の連鎖解析と GWAS 得られる座位はほと

んど一致していない。GWAS では捉えることの出来ない遺伝要因を捉えるためには、やはり小さなサンプルサイズであっても遺伝的背景が濃厚な家系解析が有効であると考えられる。本課題でも、このような家系に注目した解析により有望な候補遺伝子を特定するに至った。かねてより、脳動脈瘤発生の初期段階に深く関わりとされていた動脈硬化、さらに分子生物学的発症機序の重要な一端を担うと考えられていた多発嚢胞腎関連分子、これらに関連する遺伝子であったことは意義深いといえる。

#### 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1件)

Nakaoka F, Tajima A, Yoneyama T, Kasuya H, Mizutani T, Inoue I: Gene expression profiling reveals distinct molecular signatures associated with the rupture of intracranial aneurysm. Gene expression profiles of intracranial aneurysms. Stroke (in press)

#### 6 . 研究組織

##### (1)研究代表者

糟谷英俊 (KASUYA HIDETOSI)

東京女子医科大学・医学部・教授

研究者番号 : 50169455

##### (2)研究分担者

恩田英明 (ONDA HIDEAKI)

東京女子医科大学・医学部・非常勤講師

研究者番号 : 60185692

米山琢 (YONEYAMA TAKU)

東京女子医科大学・医学部・助教

研究者番号 : 90318105

赤川浩之 (AKAGAWA HIROYUKI)  
東京女子医科大学・医学部・准教授  
研究者番号: 60398807

(3)連携研究者  
なし