

## 脳動脈瘤関連遺伝子の探索

●井ノ上 逸朗

東海大学医学部基礎医学系分子生命科学

### <研究の目的と進め方>

本研究では脳動脈瘤（くも膜下出血）の発症に関連する遺伝子同定を目指す。くも膜下出血はそのほとんどが脳動脈瘤の破裂を原因としており、致死率が高く、社会的損失の大きい疾患である。脳動脈瘤は高頻度に存在しており（50歳以上の3-6%に存在）、無症状であるものの、破裂に伴って50%程度が死に至る重篤な疾患であるため、易罹患性遺伝子診断の対象として最適な疾患と言えるだろう。疾患関連遺伝子が同定できれば、脳動脈瘤発生メカニズムの全容解明につながり、予防法や治療法の開発へ向けた大きな前進が期待できる。糖尿病や高血圧などと同様、脳動脈瘤もまた多因子性疾患の一つであり、複数の遺伝因子や環境因子が複雑な形でその発症リスクに関与していると考えられる。したがって、脳動脈瘤発生に関するメカニズムの全容解明のためには、従来の手法に加えて、ゲノム全域からの感受性遺伝子を検出する新たなアプローチを採用し、戦略を立てることが重要である。

ゲノム全域を対象とした罹患同胞対連鎖研究、および集団ベースのサンプルを用いた症例対照研究において遺伝子間相互作用を含む体系的な関連解析を行い、単一遺伝子のみでなく遺伝子間相互作用まで明らかにし病因解明を目指す。

### <2007年度の研究の当初計画>

#### ゲノム全域アソシエーション・スタディ (GWAS)

Affymetrix GeneChip Human Mapping 500K Array SetおよびIllumina Sentrix HumanHap 300 Genotyping BeadChipを用い、ゲノム網羅率を向上させた症例対照関連研究を試みる。前者には家族歴を有する症例を中心に脳動脈瘤患者200例、後者にはそれらを含む計300例と対照200例を1次スクリーニングに用いる。2次スクリーニングには、Illumina GoldenGate Genotyping Assayを用い、脳動脈瘤患者と対照それぞれ450例の遺伝子型を2,304個のSNPにわたって決定する。ジェノタイプングは基盤ゲノムの九州大学山本研にお願いする。

### <2007年度の成果>

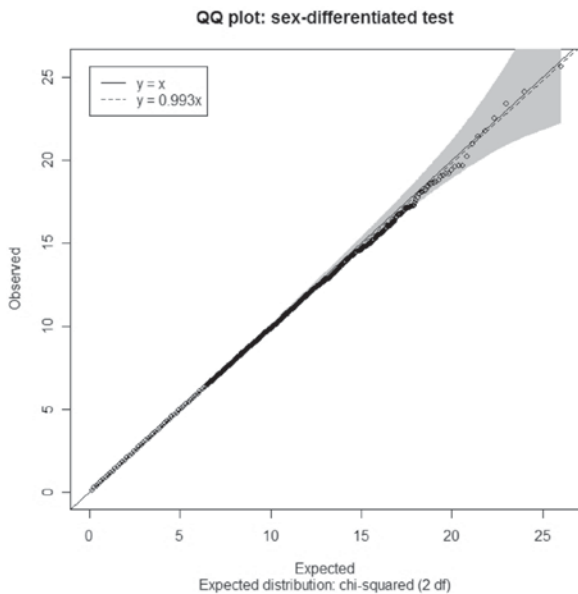
Illumina HumanHap300 BeadChipを用いた計500例のジェノタイプングが完了した。BeadChipのバージョン更新に伴い、300例はHumanHap300 v1、残りの200例はHumanHap300-Duo (v2-2)を用いた(2つのチップに共通のSNPは312,712個)。全体のコール率は99.5%を超えるものであった。

タイプングデータに関して、サンプル除外基準を次のように設定し、サンプルのクオリティコントロール (QC) を行った (括弧内は除外対象となったサンプル数): (i) コール率97%未満 (2例)、(ii) サンプルの重複 (2例)、(iii) 未報告の血縁者対 (2例)、(iv) 異なる祖先集団由来 (12例)。項目 (iv) の結果は、日本人サンプルにおいても階層化が生じうることを考慮する必要性を示唆している。実際、祖先集団の明らかなHapMap CHBや韓国人サンプルのなすクラスターとは異なるルートでの日本列島への流入の可能性を示唆している。QC基準を満たしたサンプルは症例288例、対照194例であった。また、SNPのQCについては、実験的な

タイプングエラーの他にコピー数変異やプローブ上に局在するSNPが影響を及ぼすため、その効果を調査した。その結果、コール率の低下、Hardy-Weinberg平衡 (HWE) からホモ接合過剰な配位への有意なシフトが生じる傾向を見出した。その上で除外基準をコール率、対照サンプルにおけるHWEからのずれの割合 (正確検定の $p$ -値) およびその向き (ホモ接合超過度) という3つの指標を用いて設定し、計512個のSNPを関連解析から除外した。関連検出の検出力を考慮し、サンプル全体でのマイナーアレル頻度の下限を5%に設定した。最終的に250,570個の常染色体上のSNPを関連解析に供した。X染色体上のSNPに関しては独立な解析が必要であり、X染色体の不活化をモデルに適切に取り込む必要があるが、これに対する共通認識が得られていない状況である。そのため、常染色体上のSNPに限定した。

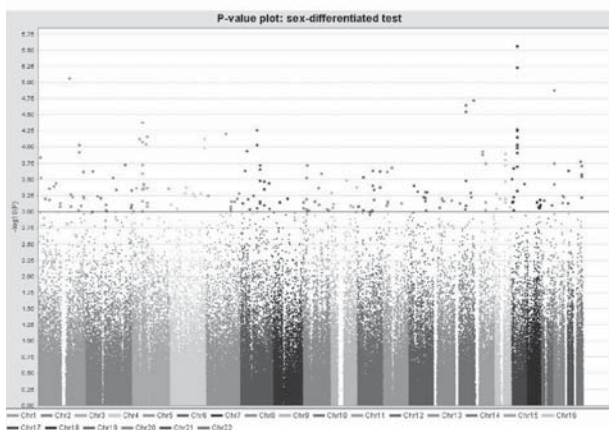
関連検定に際して、脳動脈瘤の既知の危険因子の一つである性別 (女性が男性の1.5から2倍のリスクをもつ) による脳動脈瘤とSNPの関連への効果の修飾および交絡の可能性が考えられた。そこで、サンプルを性別で層別化し、各層でのCochran-Armitage傾向性検定 (C-A検定) のカイ2乗値の和を基本統計量とした (性別層化検定と呼ぶ)。効果の修飾が有意でないSNPに関しては、男女間の効果の均質性を仮定できるMantel拡張検定も考慮することとした (交絡の可能性も示唆されない場合には単純なC-A検定も参照した)。性別層化検定の分位点-分位点 (QQ) プロットを図1に示す。Bonferroniの補正を考慮した有意水準 $2.0 \times 10^{-7}$ を下回るSNPは存在しなかった。本研究の規模では、多因子性疾患で一番尤もらしいと予想されるオッズ比の範囲 (1.3-1.6) ではSNPの検定 $p$ -値は0.001以上になることが期待される。したがって、図1は (1) データQCの結果、階層化は除去された; (2) HumanHap300チップにアッセイされたSNPでタグづけされている範囲には、非常に強い効果を示す脳動脈瘤感受性変異は存在しない、ということを示唆している。全染色体にわたる性別層化検定 $p$ -値の分布を図2に示す。

疾患感受性変異ないしそのtagSNPを直接タイプングしていない場合、間接的関連は連鎖不平衡にある複数のSNPをセットとして検出する。GWAS規模のハプロタイプ解析は実質困難であるため、相を推定しない解析法を採用した。単純に領域内の複数SNPに対するHotellingの $T^2$ 統計量を用いた関連検定はあまり良い性能を示さなかった。この理由としては、多自由度の負荷が大きく、かつ、強いSNP間連鎖不平衡による冗長性を調整できないことが考えられる。そこで、Gauderman et al. (2007) による、初めに主成分分析を行って重要な主成分だけを抽出して自由度を落とした上でロジスティック回帰分析を行うという2段階アプローチを採用した。検定は提案された尤度比統計量ではなく、スコア統計量を用いた。これは実践的な計算効率の問題の改善だけではなく、尤度比統計量では反保守的な結果を生み出す傾向が見られたためである。性別などの共変量の取り扱いやデータのsparsenessの影響の分析が完全ではないが、下記の2次スクリーニング候補SNPを用いた解析では、3番染色体と17番染色体に $10^{-6}$ のオーダー (単一座位検定よりも $p$ -値が1/100から1/1000倍) の比較的強い関連が見出された。



1次スクリーニングの解析プロセスにおいて、C-A 検定で最小の p-値を示した 21 個の SNP に対して TaqMan アッセイを用い、脳動脈瘤患者と対照各 196 例に対して追加タイピングを行った。これは、Parkinson 病の GWAS を一例として、一般の多因子性疾患では数 100 サンプル規模の 1 次スクリーニングでは「トップ 10」戦略がうまく機能しない可能性が高いことが示唆されているが、脳動脈瘤の遺伝様式が不明確であるため検証実験を試みたものである。この結果は、示唆されている通りのものであった。対照サンプルを RIKEN JSNP550 のサンプル (934 例) と拡大して得られた有意性の上位 21 個の SNP に対しても同様の結果が得られた。したがって、脳動脈瘤感受性変異同定のためには、包括的な 2 次スクリーニングが必要不可欠であることが明確になった。

包括的な 2 次スクリーニングのための SNP 選抜には、単一座位関連検定の結果を用いた。まず、p-値が 0.008 以下の SNP を抽出し、この集合から SNP 対の遺伝子型の相関係数を用いたクラスター分析 (最長距離法) により tagSNP を選定し、GoldenGate アッセイ可能な SNP 2,304 個を 2 次スクリーニングに供した。



#### <国内外での成果の位置づけ>

未だ我々の脳動脈瘤研究が世界のフロントランナーとはいえるが、deCODE Genetics による脳動脈瘤解析が報告され、Yale 大学では 1500 例の GWAS 完了の話の聞くと、我々もかなりスピードアップする必要がある。

#### <達成できなかったこと、予想外の困難、その理由>

Affymetrix 500K のジェノタイピング結果の品質評価を行った結果、標準的な品質基準を満たすサンプルが半分未満となり、関連を適切に評価することができなかった。(Affymetrix Genome-Wide Human SNP Array 5.0 を用いて再実験を行った。)

従来、日本人集団は遺伝的に均質であるという説もあったが、必ずしもそうではないことを示唆する結果を得た。我々のデータからは結論付けるまでには至らないものの、日本列島への異なる流入経路が示唆された。このことから、ゲノムワイドな関連解析に際し、階層化の補正やデータクリーニングによる均質化というプロセスが非常に重要であるという認識を得るに至った。

コピー数変異およびプローブ上に存在する SNP により、遺伝子型判定に用いられる蛍光強度のクラスター図に乱れが生じる。このことに起因する遺伝子型判定の誤りの可能性はコール率や HWE からのずれの評価といった統計処理だけでは除外できなかった。2 段階スクリーニングデザインでは数千から数万の規模で候補 SNP が追跡されるが、現状ではタイピングエラーによる artifact を目視で除外していく以外にない。しかし、特定のパターンが出現するとは限らないので、一定の基準を設定することが非常に難しい。クラスター分析の方法を応用してそのような artifact を除外する方法を検討したが、計算負荷の問題もあり、現実的に利用可能な形までには至らなかった。

#### <今後の課題>

今回の GWAS 解析に関する検体、SNP のフィルタリング等、クオリティコントロールは脳動脈瘤のみでなく、すべての疾患に普遍性のある解析法である。脳動脈瘤感受性遺伝子同定のための研究であるが、広く応用ゲノムで計画されている GWAS 研究の解析基盤となるよう準備しており研究協力も可能である。

#### <成果公表リスト>

1. Uno Y, Suzuki Y, Wakaguri H, Sakamoto Y, Sano H, Osada N, Hashimoto K, Sugano S, Inoue I. Analysis of expressed sequence tags from liver in cynomolgus monkey (*Macaca fascicularis*): A systematic identification of drug-metabolizing enzyme genes. *FEBS Let* in press.
2. Saigo K, Yoshida K, Ikeda R, Sakamoto Y, Murakami Y, Urashima T, Asano T, Kenmochi T, Inoue I. Integration of hepatitis B virus DNA into the MPP (mixed lineage leukemia) 4 gene and rearrangements of MLL4 in human hepatocellular carcinoma cells. *Hum Mutation* in press.
3. Tomoyasu Y, Yamaguchi T, Tajima A, Nakajima T, Inoue I, Maki K. Further evidence for an association between mandibular height and the growth hormone receptor gene in the Japanese population. *Am J Ortho and Dento Ortho* in press.
4. Sasahara A, Kasuya H, Akagawa H, Ujiie H, Kubo O, Sasaki T, Onda H, Sakamoto Y, Kriscsek B, Hori T, Inoue I. Increased expression of ephrin A1 in brain arteriovenous malformation. *Neurosurg Rev* 30, 299-305, 2007.
5. Akagawa H, Narita A, Yamada H, Tajima A, Kriscsek B, Kasuya H, Hori T, Kubota M, Saeki N, Hata A, Mizutani M, Inoue I. Systematic screening of lysyl oxidase-like (LOXL) family genes demonstrates that *LOXL2* is a susceptibility gene to intracranial aneurysms. *Hum Genet* 121, 377-387, 2007.
6. Ikeda R, Yoshida K, Inoue I. Identification of FAZF as a novel BMP2-induced transcription factor during osteoblastic differentiation. *J Cell Biochem* 101,147-154 2007.