

## 論文の内容の要旨

論文題目：Genome-wide association study (GWAS) identified new candidate regions for narcolepsy

論文題目の和訳：ゲノムワイド関連解析によるナルコレプシーの  
新規疾患感受性領域の同定

指導教員名：徳永 勝士 教授

東京大学大学院医学系研究科

平成 17 年 4 月入学

博士後期課程

国際保健学専攻

氏名：宮川 卓

### 緒言

ナルコレプシーは、睡眠発作、情動脱力発作、入眠時幻覚及び睡眠麻痺を主症状とする異常な REM 睡眠を伴う代表的な過眠症である。一卵性双生児一致率は 20-30%、第一近親発症率は 1-2%と報告されていることから、ナルコレプシーは遺伝素因や環境素因といった複数の因子が複雑に作用し合って、発病に至る多因子疾患であると考えられている。ナルコレプシーの遺伝学的背景に関するこれまでの研究において注目すべき点は、ナルコレプシーとヒト白血球抗原 (human leukocyte antigen; HLA) との強い関連性である。HLA class II 領域に存在する DRB1\*1501-DQB1\*0602 ハプロタイプを日本人ナルコレプシー患者ではほぼ 100%、白人集団でも 85%の患者が保有している。しかしながら、DRB1\*1501-DQB1\*0602 ハプロタイプを持たないナルコレプシー患者がいること、日本人健常者の 10%以上が DRB1\*1501-DQB1\*0602 ハプロタイプを保有すること、及び遺伝統計学による計算の結果から、HLA 領域以外の遺伝子の関与が示唆される。そこで、HLA 領域以外のナルコレプシー疾患感受性遺伝子を同定することを目的とした研究を行った。

疾患感受性遺伝子を探索する方法は、候補遺伝子アプローチと位置的アプローチの二つに大別することができる。候補遺伝子アプローチは、遺伝子の機能や動物モデル等の情報から疾患に関与すると予想される遺伝子や、位置的アプローチにおいて同定された候補領域内に存在する遺伝子について、変異と疾患との関連を検討する方法である。候補遺伝子アプローチは簡便で直接的な方法であるが、疾患に関与すると予測できない遺伝子や未知の遺伝子に対しては適用できないアプローチ方法である。位置的アプローチでは、従来ノンパラメトリック連鎖解析が多用されてきたが、この解析方法は統計学的検出力が低いことから、比較的弱い疾患感受性領域を検出することが困難であるという弱点があった。近年、ノンパラメトリック連鎖解析より検出力が高く、ゲノム全域を探索できる方法であるゲノムワイド関連解析を実施するための実用的な技術が開発された。そこで、本研究では疾患感受性遺伝子同定のために最も効果的と考えられるゲノムワイド関連解析を行った。

また、ゲノムワイド関連解析では大規模 SNP タイピングキットを用いて遺伝子型のタイ

ピングを行う。しかし、最新の技術をもってしても全ての SNP が正確にタイピングされているとは限らないことが知られており、そのような不正確な SNP の情報を排除する必要がある（データクリーニング）。そこで、最適なデータクリーニング方法を同定することも目的として研究を行った。

## 対象と方法

### 最適なデータクリーニング方法の同定

Affymetrix 社の 50 万 SNP タイピング用のアレイでタイピングされた健常者群（389 例）を二等分して擬似の関連解析を行った。健常者群同士を比較していることから、得られた結果（観察値）が偶然において期待される値から逸脱しないという仮説が立てられる。擬似の関連解析の結果に対して、遺伝子型決定のための BRLMM アルゴリズムの設定値、各 SNP の call rate、ハーディ・ワインベルク平衡及びマイナーアレル頻度を用い、これらの条件を変更してデータクリーニングを行い、観察値と期待値を比較した。

### ナルコレプシーの疾患感受性遺伝子の探索

日本人集団におけるナルコレプシー患者（ケース）222 例と健常者（コントロール）389 例を用いて、ゲノムワイド関連解析を行った。タイピングキットとして Affymetrix 社の 50 万 SNP タイピング用のアレイを用いた。同定された最適なデータクリーニングを行い、約 25 万 SNP を選別した。次に、それら候補 SNP の再現性を確認するために、独立のサンプルセット（ケース 159 例、コントロール 190 例）を用意し、再度関連解析（Replication study）を行った。再現性の得られた SNP に関しては、真性過眠症候群 75 例を用いた関連解析も行った。

再現性の得られた SNP が、その近傍に存在する遺伝子の mRNA の発現量に影響を与えるか検討するために、白血球から抽出された RNA を用いて合成した cDNA を鋳型として、リアルタイム RT-PCR を行った。

## 結果

### 最適なデータクリーニング方法の同定

前述した各種条件でデータクリーニングを行った。その結果、BRLMM アルゴリズムの設定値 0.5、各 SNP の call rate（95%以上）、ハーディ・ワインベルク平衡（P 値 0.001 以上）及びマイナーアレル頻度（5%又は 1%以上）のデータクリーニング後の観察値において、期待値からの逸脱が小さいことを確認した。この方法の再現性を確認するために、健常者群（389 例）の二等分の組み合わせ方法を変えて、再度擬似の関連解析を行った。その結果に対し同様のデータクリーニングを行った後の観察値においても、期待値からの逸脱が小さいことを確認した。

### ナルコレプシーの疾患感受性遺伝子の探索

ナルコレプシーの疾患感受性遺伝子を探索するためにゲノムワイド関連解析を実施した。

タイピング後に、前述したデータクリーニングを行い、約 25 万 SNP を選別した。検定によって得られた各 SNP の P 値及び、それら SNP の周辺の遺伝子情報等を参考にし、疾患に関連し得ると予想される新規候補 SNP を 30 個選別した。それら候補 SNP の再現性を確認するために、独立のサンプルセットを用意し Replication study を実施した。その結果、一つの SNP（マーカー TM22.1）において再現性が確認された（表 1）。ゲノムワイド関連解析における TM22.1（T/C）の C アリルの頻度は、ケース 26%、コントロール 17%であり、ナルコレプシー患者において有意に増加していた（ $P=1.4 \times 10^{-4}$ 、OR=1.74）。Replication study においても、C アリルの頻度はケース 24%、コントロール 14%であり、ナルコレプシー患者において有意に増加していた（ $P=5.2 \times 10^{-4}$ 、OR=1.97）。両解析を統合した結果の P 値及び OR は  $4.4 \times 10^{-7}$  及び 1.79 となった。

表1. マーカーTM22.1とナルコレプシーとの関連

Genome-wide stage				Replication stage				Combined	
MAF (case)	MAF (control)	OR (95% c.i.)	$P_{allele}$	MAF (case)	MAF (control)	OR (95% c.i.)	$P_{allele}$	OR (95% c.i.)	$P_{allele}$
0.26	0.17	1.74 (1.31-2.31)	$1.4 \times 10^{-4}$	0.24	0.14	1.97 (1.34-2.90)	$5.2 \times 10^{-4}$	1.79 (1.43-2.25)	$4.4 \times 10^{-7}$

$P_{allele}$  値はアリル頻度に基づいて計算した（両側検定）。

（MAF：マイナーアリル頻度、OR:オッズ比、95% c.i.: 95%信頼区間）

TM22.1 に関して、真性過眠症候群における関連解析を行い、患者において有意に増加していることを確認した（片側検定  $P=0.04$ 、OR=1.44）。

TM22.1 の周辺の連鎖不平衡構造を明らかにするために、周辺に存在する多型（主に tagSNP）を用いて更なる解析を行った。その結果、TM22.1 を含む連鎖不平衡ブロック内には、Gene X 及び Gene Y の二つの遺伝子が存在することを確認した（図 1）。次に、連鎖不平衡ブロック内に存在する SNP を用いてハプロタイプ解析を行い、TM22.1 又は TM22.1 と強い連鎖不平衡にある多型が一義的な感受性変異であることを明らかにした。

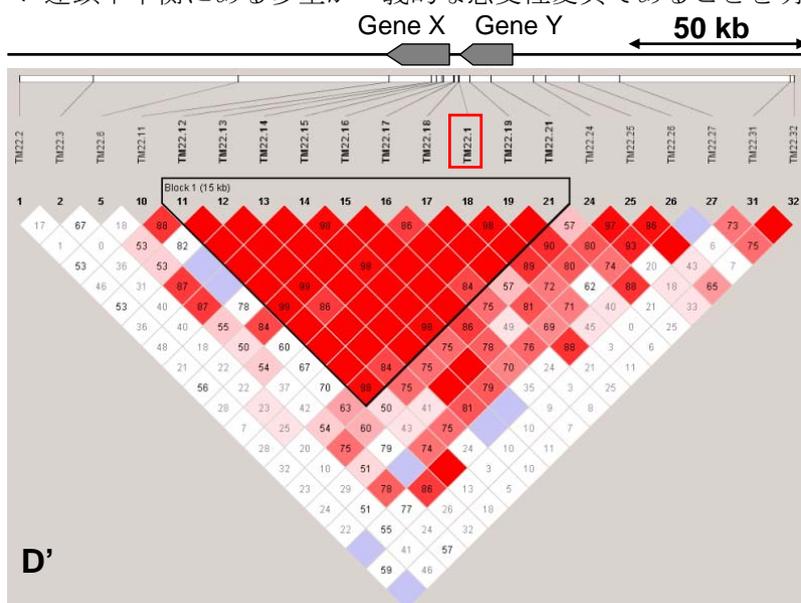


図1. マーカーTM22.1を含む領域の連鎖不平衡ブロック  
連鎖不平衡係数はD'を用いて計算した。

TM22.1 又は TM22.1 と強い連鎖不平衡にある多型が、*Gene X* 及び *Gene Y* の発現量に影響を与えるか検討した。両遺伝子の mRNA の発現量をリアルタイム RT-PCR を用いて定量し、TM22.1 の遺伝子型 (TT 及び TC) 間で比較した。その結果、図 2 のように両遺伝子において、TC 群で有意に mRNA の発現量が低下していることを確認した (*Gene X*:  $P=0.003$ ; *Gene Y*:  $P=0.01$ )。

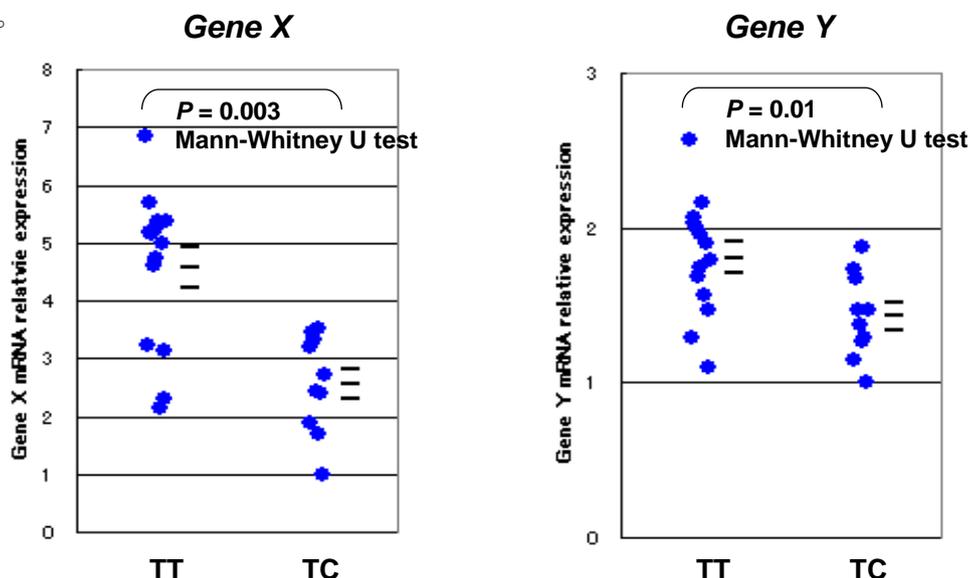


図2. *Gene X* 及び *Gene Y* の発現量と TM22.1 遺伝子型 (TT 及び TC) との関連 (中央の線: 平均値、上部及び下部の線: 土標準誤差)

## 考察

大規模 SNP タイピングキットでタイピングされたデータの中に存在する不正確な SNP のタイピング結果は、その後の統計解析や候補 SNP の選別に悪影響を与える可能性がある。本研究では、最適なデータクリーニング方法を同定することで、その後のナルコレプシーの関連解析を効率的に進めることを可能にした。また、今回のデータクリーニング方法は、他の研究者が行った大規模 SNP タイピングキットを用いた研究にも応用できると考えている。

本研究では、ナルコレプシーとの強い関連をマーカー TM22.1 に見出した。TM22.1 を含む連鎖不平衡ブロック内には、*Gene X* 及び *Gene Y* の二つの遺伝子が存在し、TM22.1 の C アリルを保有すると、両遺伝子の mRNA の発現量が低下することを明らかにした。この結果から *Gene X* 及び *Gene Y* とも疾患感受性遺伝子の候補となると考えられる。*Gene X* は  $\beta$  酸化に関わる遺伝子である。 $\beta$  酸化は  $\theta$  波周波数 (REM 睡眠の特徴) を調節する物質的基盤のひとつであることが判明しており、*Gene X* の発現量が低下することで REM 睡眠に影響を与える可能性が考えられる。*Gene Y* はコリンの代謝に関わる遺伝子である。コリンは、神経伝達物質のアセチルコリン、脳循環・代謝改善薬の一つである CDP コリンなどの合成材料である。*Gene Y* の発現量が低下することで、これら脳の活動に重要な働きを示す物質の合成量が低下する可能性が考えられる。今後、ナルコレプシーのモデルマウスに *Gene X* 又は *Gene Y* の働きを補う物質を投与することで、症状が改善するか確認したいと考えている。