

## SNP 解析支援体制の確立と疾患関連遺伝子の同定

●徳永 勝士<sup>1)</sup> ◆大橋 順<sup>1)</sup> ◆宮寺 浩子<sup>1)</sup> ◆本多 真<sup>2)</sup> ◆慶長 直人<sup>3)</sup>

1) 東京大学医学系研究科 2) 東京都精神医学総合研究所 3) 国立国際医療センター研究所

### <研究の目的と進め方>

本研究は2つの主要な目的を持つ。第一に、「基盤ゲノム・ヒト SNP タイピングセンター」として、ハイスループットかつ低コストの SNP 解析システムを確立し、「応用ゲノム」班員が目指す各種疾患の発症や病態に関わる遺伝子多型の探索研究を支援する。このため、数十万種以上の SNP を用いるゲノムワイド関連分析システムを導入し、多数の疾患遺伝子候補領域を検出する。また、新たな高効率・低コストの SNP 解析システムを開発して候補領域の高密度関連分析を効率化する。

第二は個別的研究であり、免疫学的機序が発症に関与すると考えられる睡眠障害や感染症の感受性遺伝子を同定し、病態機序への理解を深め、新しい治療法や予防法の開発に貢献する。このためヒトナルコレプシーについてゲノムワイド関連分析および網羅的遺伝子発現解析を実施し、過眠症関連遺伝子を同定してその病態機序を探る。またマラリアの臨床亜型に関して、候補遺伝子アプローチおよび候補領域アプローチを行い、重症化と関連する感受性遺伝子多型を検出する。結核についても、ゲノムワイド連鎖分析、伝達不平衡検定などを用いて感受性遺伝子を同定し、その機能を解明する。

### <2007年度の研究の当初計画>

**SNP解析支援：**「応用ゲノム」班員によるゲノムワイド関連分析(GWAS)を支援するため、500K SNP Array (50万種のSNPsを同時解析)による健常者群試料のタイピングを行い、信頼度の高い解析結果を得るための基準や健常者データを作成する。さらに多系統萎縮症およびパニック障害については患者群試料をタイピング後、関連分析を実施する。また、前年度に開発した新しいマルチプレックスSNPタイピングシステム (DigiTag2) を用いて、「応用ゲノム」班員による疾患関連の再現性検証や候補領域の高密度関連分析 (絞り込み) のためのSNPタイピングを行う。特に2型糖尿病については、これまでに検出された候補領域・遺伝子について、再現性確認や絞り込みのためのSNPタイピングを継続する。

**個別的研究：**ナルコレプシーについて、500K SNP Arrayを用いたGWASを完了し、さらに検出された候補領域の再現性検証や絞り込みを実施する。また網羅的な遺伝子発現解析結果およびGWAS結果に基づいて同定されたナルコレプシー関連遺伝子について、血中の遺伝子発現を比較定量するとともに機能的検討を進める。マラリアについては、染色体5q33および5q31領域において、関連解析により有意差を示すマイクロサテライトマーカーを見出した。そこで、その近傍のSNPマーカーを高密度に解析し、当該領域に存在する関連変異の同定を目指す。結核については、タイ症例について、伝達不平衡検定、罹患同胞対法によるゲノムワイド連鎖分析をさらに進めるとともに、結核等抗酸菌感染症に関連する遺伝子多型の機能解析を実施する。

### <2007年度の成果>

**SNP解析支援：**500K SNP Arrayにより450名の健常者試料のタイピングを完了し、一定の品質基準を満たした試料とSNPsに基づいて健常対照群データを作成した。結果の一部はデータベースに格納後公開、または班員共有する予定である。また、多系統萎縮症患者230名およびパニック障害患者200名の試料についても500K SNP Arrayによるタイピングを完了し、品質管理基準

を満たしたデータについて疾患関連分析を行った。さらにパニック障害については、検出された候補SNPsの中から30種類を選択し、独立した患者・健常者試料セットを用いてDigiTag2法による再現性検証を開始した。これらに加えて、最も新しい900K SNPsと900K CNV probesを搭載したarrayについても健常者200名の試料のタイピングを実施した。500K Arrayに比較して信頼度が高く、統計解析可能SNP数も多いことを確認したうえで、健常対照群データを作成した。

これまでに検出された候補遺伝子・領域の絞り込みに関しては、2型糖尿病については、先の特定領域「ゲノム医科学」で実施されたゲノムワイド関連分析および候補遺伝子の大規模関連分析によって検出された複数の有望な遺伝子について、絞り込みのためのSNPタイピングを実施した。この結果、ヨーロッパ人で発見されたTCF7L2が頻度は低いながら日本人でも有意な関連を示すことが確認されたのに加えて、ヨーロッパ人で報告のない新規感受性遺伝子も見出された。また、岩崎 (東京女子医大) らがゲノムワイド連鎖分析によって検出した候補領域のひとつから有意な関連を示す遺伝子を2個見出し、再現性確認、絞り込みを終え、現在機能解析を行っている。

**個別的研究：**ナルコレプシーに関しては、220名の患者および約400名の健常者 (共有データ) を対象として500K SNP Arrayを用いたGWASを完了し、数十個の候補領域を検出した。このうち少なくとも1個のSNPについて、独立な試料セットを用いて再現性が確認され、さらに絞り込みも完了した。他集団における関連分析も行った結果、韓国人で全く同様な関連を認めた。しかも、このSNPが、近傍に位置する2個の遺伝子の発現レベルに有意に影響することもわかった。

一方、網羅的遺伝子発現解析から同定された新規ナルコレプシー関連遺伝子IGFBP3の機能を *in vitro* のレポーターアッセイ系および *in vivo* のIGFBP3過剰発現マウスで検討した。IGFBP3がオレキシン遺伝子の転写抑制を行い、行動面でも覚醒を減少させる作用をもつことが示されたことから、ナルコレプシーの病態に関連することが証明された。また、定量的RT-PCR法による血中遺伝子発現解析により、MX2遺伝子がHLA遺伝子多型と関連して発現を変化させることを見出した。

マラリアについては、染色体5q33領域の解析から、TIM1遺伝子プロモーター領域多型と脳性マラリアとの統計学的関連を見出した。また、この多型がTIM1遺伝子の発現量と関連することも見出した。

結核については、タイ罹患同胞対連鎖解析によりGM-CSF、IL-4などサイトカインクラスターが存在する第5染色体長腕上に連鎖を示唆する領域を検出した。また若年発症例では、第17番と20番に新たな連鎖領域を見出した。さらにヒト気道上皮のパネルにより、抗酸菌感染に関連するCFTR遺伝子の多型が、エクソン9のスプライシング効率と関連することを明らかにした。

### <国内外での成果の位置づけ>

**SNP解析支援：**数十万種のSNPsを用いたゲノムワイド関連分析法は最も先進的な疾患遺伝子探索法である。本年度の対象となった多系統萎縮症グループおよびパニック障害グループは国内、国外で最大規模の試料収集を実現しており、その成果は世界の先頭を切る。また2型糖尿病について、先の特定領域「ゲノム医科学」で実施されたGWASによって検出された候補領域の再現

性検証と絞り込みが進展し、アジア系集団の新規感受性遺伝子が特定された意義は大きい。さらに、連鎖分析によって検出された候補領域から、やせ型2型糖尿病の新規疾患感受性遺伝子が同定された。

**個別的研究：** ナルコレプシーについては、国内最大、世界でも有数の試料を収集し、世界に先駆けてマイクロサテライト多型およびSNPsを用いたGWASを実施して新しい疾患感受性・抵抗性遺伝子を同定した。また死後脳を用いた遺伝子発現解析から同定されたIGFBP3は、オレキシン遺伝子の発現制御を通じて個体レベルでの睡眠覚醒も変化させた。網羅的遺伝子発現解析が神経疾患の病態解明に有用であることのモデルとなる。さらに、HLAアリルと関連した血中遺伝子発現変化の発見は、HLAアリルがMX2など他の免疫関連遺伝子発現変化を介して病態に関わる可能性を示す。マラリアについて、これまでヒトの遺伝子多型とマラリア重症化との関連研究はアフリカ集団を対象に行われてきた。本研究の対象はタイ人マラリア患者であり、アジア系集団に特異的な新規重症化関連多型が見出されることが期待されることから、本研究の成果は高い注目を受けている。結核についても、ベトナム人、タイ人などからアジアで最大規模の試料収集（家系試料を含む）を行い、まずゲノムワイド連鎖解析を実施して、これまでに報告のない候補領域を検出した。結核の感受性遺伝子研究は、まだ報告も少なく、その結果が一定していないため、また将来我が国の結核患者にも裨益するため、多くのアジア系集団における知見を蓄積する必要がある。

#### <達成できなかったこと、予想外の困難、その理由>

**SNP解析支援：** ほぼ計画に添って進行しており、予想外の困難もない。ただし、数十万種のSNPsを用いるGWASは本特定ゲノム領域が発足した当時に存在しなかったため、当初予想できなかった技術的検証作業やタイピング作業のための人員と予算を必要としている。

**個別的研究：** ナルコレプシーについて、死後脳を試料とした遺伝子発現解析で同定された疾患関連遺伝子の中で、特に血中で発現定量ができないものについては、脳組織学的な発現局在により機能の検討を行う予定であったが、抗体の特異性や良質のヒト脳組織の入手が困難であることから十分な検討が行えていない。

#### <今後の課題>

**SNP解析支援：** 最も新しい900K SNPsと900K CNV probesを搭載したarrayを用いて、より広範かつ大規模なGWASを実施できる体制を整える。またこれまでにタイピングしたデータに基づいて、日本人におけるCNV（copy number variation）の種類と頻度分布を明らかにし、疾患との関連についても解析を開始する。さらに、GWAS後の再現性検証、絞り込みの効率化を目指し、DigiTag2法によるタイピングから統計解析に至るシステムを改善する。

**個別的研究：** ナルコレプシーについては、GWASによって検出された多数の候補領域の再現性検証を進め、順次絞り込み作業を行う。また、同定された過眠症関連遺伝子について転写レベルと病態との関連について検討するとともに、血中遺伝子発現定量をすすめ、診断指標・重症度指標(主観的・客観的眠気尺度との関連)などとの相関を調べ過眠症の分子診断指標として臨床応用試行につなげる。マラリアについては、その臨床亜型に関して可能であればGWASを実施する。結核については、連鎖分析で見出した候補領域の絞り込みを行う。また可能であればタイ人、ベトナム人などの試料を用いたGWASにより、アジア系集団における結核感受性遺伝子の同定を進める。

#### <成果公表リスト>

(おもな発表論文のみ示す)

1.0702141227

Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K: Further development of multiplex SNP typing method, DigiTag2 assay. *Anal. Biochem.* 364: 78-85, 2007.

2.0702141146

Hananantachai H, Patarapotikul J, Ohashi J, Naka I, Krudsood S, Looareesuwan S, Tokunaga K: Significant association between TNF- $\alpha$  (TNF) promoter allele (-1031C, -863C, and -857C) and cerebral malaria in Thailand. *Tissue Ant.* 69: 277-280, 2007.

3.0705071534

Tanaka S, Honda Y, and Honda M: Identification of differentially expressed genes in narcolepsy blood cells. *Sleep* 30:974-979, 2007.

4.0801281514

Mai HN, Hijikata M, Inoue Y, Suzuki K, Sakatani M, Okada M, Kimura K, Kobayashi N, Toyota E, Kudo K, Nagai H, Kurashima A, Kajiki A, Oketani N, Hayakawa H, Tanaka G, Shojima J, Matsushita I, Sakurada S, Tokunaga K, Keicho N: Pulmonary Mycobacterium avium complex infection associated with the IVS8-T5 of the CFTR gene. *Int. J. Tuberc. Lung D* 11: 808-13, 2007.

5. 0801281526

Tanaka G, Shojima J, Matsushita I, Nagai H, Kurashima A, Nakata K, Toyota E, Kobayashi N, Kudo K, Keicho N. Pulmonary Mycobacterium avium complex infection: Association with NRAMP1 polymorphisms. *Eur. Respir. J.* 30: 1376-82, 2007.

6. 0801281545

Hoa BK, Hang NTL, Kashiwase K, Ohashi J, Lien LT, Horie T, Shojima J, Hijikata M, Sakurada S, Satake M, Tokunaga K, Sasazuki T, Keicho N: HLA-A, -B, -C, -DRB1 and -DQB1 alleles and haplotypes in the Kinh population in Vietnam. *Tissue Antigens* 71: 127-34, 2008.

7. 0801281607

Miyake K, Horikawa Y, Hara K, Yasuda K, Osawa H, Furuta H, Hirota Y, Yamagata K, Hinokio Y, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Yamamoto K, Tokunaga K, Takeda J, Makino H, Nanjo K, Kadowaki T, Kasuga M: Association of TCF7L2 polymorphisms with susceptibility to type 2 diabetes in 4,087 Japanese subjects. *J. Hum. Genet.* (in press)

8. 0801281624

Kohara K, Tabara Y, Nakura J, Imai Y, Ohkubo T, Hata A, Soma M, Nakayama T, Umemura S, Hirawa N, Ueshima H, Kita Y, Ogiwara T, Katsuya T, Takahashi N, Tokunaga K, Miki T: Systemic multiple candidate gene approach for identification of susceptible genes and susceptible pathways for hypertension: a millennium genome project for hypertension. *Hypertension Res.* (in press)

9. 0801280630

Tanaka S, Honda Y, Honda M: MX2 gene expression tends to be down-regulated in subjects with HLA-DQB1\*0602. *Sleep* (in press)

10. 801092012

Nuchnoi P, Ohashi J, Kimura R, Hananantachai H, Naka I, S Krudsood S, Looareesuwan S, Tokunaga K, Patarapotikul J: Significant association between TIM1 promoter polymorphisms and protection against cerebral malaria in Thailand. *Ann. Hum. Genet.* (in press)

11. 0801281640

Fujimoto A, Kimura R, Ohashi J, Omi K, Yuliwulandari R, Batubara L, Mustofa MS, Samakkarn U, Settheetham-Ishida W, Ishida T, Morishita Y, Furusawa T, Nakazawa M, Ohtsuka R, Tokunaga K: A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. *Hum. Mol. Genet.* (in press)