

# 疾患関連変異探索のための遺伝統計学的解析法の開発と Deng 出血熱関連解析への応用

●大橋 順

筑波大学大学院人間総合科学研究科

## <研究の目的と進め方>

本研究は、以下の5課題の解決を目的とする。(1) コピー数多型 (CNV; copy number variation) とは、約 1kb 以上の大きさをもつゲノム領域のコピー数が個人によって異なる多型のことをいい、ヒトゲノムの 10% 以上に当たる領域でコピー数多型が観察されている。その中には、多型領域が遺伝子全体を含んでいるものもある。コピー数多型を調べる方法の一つに、DNA チップによって、PCR 産物とプローブとのハイブリダイゼーションのシグナル強度をゲノム全域にわたって調べ、シグナル強度の相対値から、各コピー数多型領域におけるコピー数を推定する方法がある。しかし、このような方法では、各個体のコピー数の表現型 (2本の染色体のコピー数の和) を調べることはできても、コピー数多型の遺伝子型 (各染色体におけるコピー数) を知ることはできない。そのため、ある CNV アリルが疾患と関連していても (質的な関連、真の感受性変異との連鎖不平衡なども含む)、そのアリルを遺伝統計学的に検出することは困難である。そこで、各 CNV アリルの頻度を推定し、CNV アリルの頻度をケースとコントロールで比較する統計学的手法を開発する。(2) ゲノムワイド関連解析または候補遺伝子アプローチにより有意な関連が検出されたゲノム領域について、その連鎖不平衡情報を HapMap データベースから取得し、実際にはタイピングをしていない SNP についても、その統計学上の関連を評価し、その領域をさらに解析すべきか否かの判断を行うアルゴリズムを開発する。(3) SNP マーカーにおいて関連が見出され、その SNP がハプロタイプブロックに含まれていた場合に、系統樹理論に基づき、系統発生上可能な複数のハプロタイプの組合せを考えることで、検定の検出力を向上させるアルゴリズムの開発を目指す。(4) 関連解析の候補遺伝子が存在する場合、チンパンジーおよびゴリラとの配列比較からヒト系統における同義置換および非同義置換を同定するとともに、ヒト集団中での非同義座位の多型における頻度スペクトラムから、当該遺伝子に作用する負の自然選択強度を推定し、弱有害変異の数や対立遺伝子頻度を推測することで、re-sequencing に基づくその後の関連研究での最適戦略を提示するためのアルゴリズムを開発する。(5) 候補遺伝子アプローチに基づく Deng 出血熱感受性・抵抗性変異の探索を行うことも目標の一つとする。近年、地球温暖化の影響で Deng ウィルスを媒介するネッタイシマカやヒトシマカの生息域が拡大しており、つい先日 (平成 19 年 10 月) も台湾で集団感染が発生している。Deng 出血熱の死亡率は 5% 程度にも達するが、その発症機序および原因 (ウィルス側およびヒト側の因子を含めて) は未だ不明であり、治療としては対症治療しかない。治療上特に重要なことは、Deng 熱から Deng 出血熱への進行を予見し、重症化をくいとめることである。血漿漏出と出血傾向が Deng 出血熱の主症状であるため、ヒトの過剰な免疫応答がその原因の一つと考えられる。過剰な免疫応答がヒト側の遺伝因子によって規定されるとすると、Deng 出血熱と関連する遺伝子多型を同定すれば、原因となる遺伝子が明らかとなるだけでなく、その遺伝子産物をターゲットにした有

効な診断法・治療法の開発につながる可能性がある。免疫関連遺伝子を中心に候補遺伝子アプローチによって、Deng 出血熱と関連するヒト遺伝子多型を同定し、原因となる遺伝子を明らかにすることで、その遺伝子産物をターゲットにした有効な診断法・治療法の開発にむけた基礎情報を得ることを目指す。

## <2008 年度の研究の当初計画>

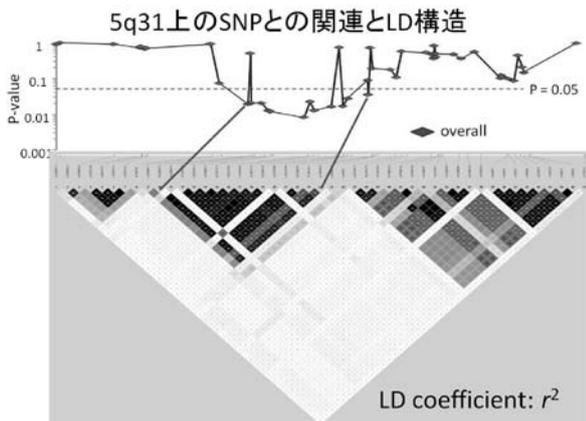
2008 年度は、課題 (1)、(2)、(5) の研究を重点的に行う。それら 3 課題の当初の研究計画は以下の通りである。(1) サンプルされた検体が属する集団におけるハーディ・ワインバーク平衡を仮定して、集団中のコピー数多型の遺伝子型頻度やコピー数アリル頻度 (または全検体中の各コピー数アリルの染色体数) を推定する EM (expectation-maximization) アルゴリズムを開発する。次に、EM アルゴリズムによるコピー数アリル頻度推定と並べ替え検定とを組み合わせた疾患関連解析手法を開発する。(2) HapMap のリファレンス集団 (日本人集団の関連解析であれば JPT 集団) における SNP-A のアリルと SNP-B のアリルとから構成されるハプロタイプの頻度情報を取得する (2つの SNP の連鎖不平衡の程度は高い方がよい)。ゲノムワイド関連解析または候補遺伝子解析によって、SNP-A のみが調べられたとする。ケース集団およびコントロール集団中での SNP-A の各アリル頻度から、SNP-B が強く疾患と関連すると仮定した条件下での、SNP-B の各集団中でのアリル頻度を推定し、サンプル数を乗じて擬似的なカイ 2 乗検定を行う (なお、HapMap データベースのハプロタイプ頻度は真の値であると仮定する)。カイ 2 乗検定の結果を参考に、SNP-B を実際に解析すべきかどうかを判断するアルゴリズムを開発する。(5) トール様受容体遺伝子 (TLR) ファミリー、各種サイトカインを候補遺伝子群として、Deng 出血熱関連解析を行う。有意な関連が見出された場合には、課題 (2) で開発した手法などを適用し、感受性変異の絞り込みを試みる。

## <2008 年度の成果>

課題 (1) EM アルゴリズムによって、集団中のコピー数多型の遺伝子型頻度やコピー数アリル頻度 (または全検体中の各コピー数アリルの染色体数) を推定することができる。そこで、コピー数多型の遺伝子型頻度やアリル頻度をケース群とコントロール群のそれぞれについて推定し、並べ替え検定によって各 CNV アリルに対してケース・コントロール関連検定を行う方法を確立した。さらに、本手法の検出力についても考察し、中間のコピー数をもつ CNV アリルを検出することが比較的困難であることを見出した。課題 (2) 連鎖する 2つの SNP-A と SNP-B があるとし、HapMap のリファレンス集団 (日本人集団の関連解析であれば JPT 集団) における SNP-A の  $i$  番目のアリルと SNP-B の  $j$  番目のアリルとから構成されるハプロタイプの頻度を  $p_{ij}$  とする。ここで、 $i$  も  $j$  も 1 か 2 である。ゲノムワイド関連解析または候補遺伝子解析によって、SNP-A のみが調べられたとする。N 人のケース集団中での 1 番目と 2 番目のアリル頻度がそれぞれ  $x_{A1}$  と  $x_{A2}$  であり、M 人のコント

ロール集団中での1番目と2番目のアレル頻度がそれぞれ $y_{A1}$ と $y_{A2}$ であったとする。このとき、リファレンス集団のハプロタイプ頻度が真の頻度だとすると、SNP-Bのケース集団中での1番目と2番目のアレル頻度は、 $x_{A1} (p_{11}/p_{11}+p_{12}) + x_{A2} (p_{21}/p_{21}+p_{22})$ と $x_{A1} (p_{12}/p_{11}+p_{12}) + x_{A2} (p_{22}/p_{21}+p_{22})$ と推定される。また、SNP-Bのコントロール集団中での1番目と2番目のアレル頻度は、 $y_{A1} (p_{11}/p_{11}+p_{12}) + y_{A2} (p_{21}/p_{21}+p_{22})$ と $y_{A1} (p_{12}/p_{11}+p_{12}) + y_{A2} (p_{22}/p_{21}+p_{22})$ と推定される。推定値に対して、2Nと2Mを乗じて、期待されるアレル数を求め、カイ2乗検定を行えば、SNP-Bにおける関連（またはallelic OR）を評価することが可能となる。本手法を用いれば、SNP-Bが疾患感受性変異であった場合に、SNP-Bを実際にタイピングした場合の検出力と、SNP-Aのみに対して検定を行った場合との、ほぼ中間程度の検出力を達成できることが確認された。これにより、一義的に関連するSNPの位置を予測可能となるだけでなく、replication studyでの費用軽減を図ることができる。

課題(5) インターロイキン遺伝子クラスターが存在する5q31を候補領域として、デング出血熱と47 SNPとの関連を評価した。二地域(LampangとRatchaburi)の結果を統合したところ、IL13領域のSNPとデング出血熱との有意な関連を見出した。さらに、関連するSNPまたはハプロタイプが存在しうるハプロタイプブロック領域を連鎖不平衡解析によって限定した。



<国内外での成果の位置づけ>

本研究で開発を目指す統計解析手法は対象疾患を選ばないため、国内外の関連研究をサポートすることができる。また、新規の統計解析手法の開発を行うため、国内外を問わず当該分野の進歩・発展に貢献しうるものである。デング出血熱関連解析については、本研究ほどのサンプルサイズ(約1000検体)をもって研究を行っているグループは稀であり、世界的にも信頼される成果をあげることができると期待している。

<達成できなかったこと、予想外の困難、その理由>

ほぼ計画通りに研究を遂行することができている。

<今後の課題>

2008年度は、デング出血熱と関連するヒト遺伝子多型の同定を目指し、5q31領域およびtoll like receptor遺伝子群を候補領域および候補遺伝子として関連解析を行う。5q31領域のSNPについては、目的(2)で開発した手法を用いて周辺の未タイピングSNPにおける関連P値の推定を行い、さらなる絞り込み(追加SNPタイピング)の必要性を検討する。目的(1)および目的(2)の研究成果については、論文として発表することを目指す。2009年度は、上述の関連解析を続けるとともに、目的(3)と目的(4)について重点的に研究を行う。

<成果公表リスト>

1)論文/プロシーディング(査読付きのものに限る)

1. 0901120251(論文)

Ohashi J. A practical case-control association test for detecting a susceptibility allele at a copy number variation locus. Journal of Human Genetics (in press)

2. 0811191110(論文)

Fujimoto A, Ohashi J, Nishida N, Miyagawa T, Morishita Y, Tsunoda T, Kimura R, Tokunaga K. A replication study confirmed the EDAR gene to be a major contributor to population differentiation regarding head hair thickness in Asia. Human Genetics 124: 179-185 (2008).

3. 0811191105(論文)

Teeranaipong P, Ohashi J, Patarapotikul J, Kimura R, Nuchnoi P, Hananantachai H, Naka I, Putaporntip C, Jongwutiwes S, Tokunaga K. A Functional SNP in the CR1 Promoter Region Contributes to Protection against Cerebral Malaria. The Journal of Infectious Diseases (in press).

4. 0811191115

Miyagawa T, Kawashima M, Nishida N, Ohashi J, Kimura R, Fujimoto A, Shimada M, Morishita S, Shigeta T, Lin L, Hong SC, Faraco J, Shin YK, Jeong JH, Okazaki Y, Tsuji S, Honda M, Honda Y, Mignot E, Tokunaga K. Variant between CPT1B and CHKB associated with susceptibility to narcolepsy. Nature Genetics 40: 1324-1328 (2008).

5. 801092012(論文)

Nuchnoi P, Ohashi J, Kimura R, Hananantachai H, Naka I, Krudsood S, Looareesuwan S, Tokunaga K, Patarapotikul J. Significant association between TIM1 promoter polymorphisms and protection against cerebral malaria in Thailand, Annals of Human Genetics, 72, 327-336 (2008).

6. 801271636(論文)

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, Human Molecular Genetics, 17, 835-843 (2008).

7. 806201536(論文)

Hirayasu K, Ohashi J, Tanaka H, Kashiwase K, Ogawa A, Takanashi M, Satake M, Jia GJ, Chimge NO, Sideltseva EW, Tokunaga K, Yabe T, Evidence for natural selection on leukocyte immunoglobulin-like receptors for HLA class I in Northeast Asians, American Journal of Human Genetics, 82: 1075-1083 (2008).

8. 806201546(論文)

Nuchnoi P, Ohashi J, Naka I, Nacapunchai D, Tokunaga K, Nishida N, Patarapotikul J, Linkage Disequilibrium Structure of the 5q31-33 Region in a Thai Population, Journal of Human Genetics 53: 850-856 (2008).

9. 0806201541(論文)

Naka I, Ohashi J, Nuchnoi P, Hananantachai H, Looareesuwan S, Tokunaga K, Patarapotikul J. Lack of association of the HbE variant with protection from cerebral malaria in Thailand. Biochemical Genetics (in press).

2)データベース/ソフトウェア

CNVデータを用いてケース・コントロール関連検定を行うためのwebツールを開発中である。