

20030368

厚生労働科学研究

ヒトゲノム・再生医療等研究事業

びまん性汎細気管支炎等遺伝要因を有する
慢性呼吸器疾患の疾患感受性遺伝子の研究

平成 15 年度 研究報告書

主任研究者 慶長直人

平成 16 年 3 月

目 次

I. 総括研究報告書

- びまん性汎細気管支炎等，遺伝素因を有する慢性呼吸器疾患の疾患感受性遺伝子の研究 …………… 3

II. 分担研究報告

1. びまん性汎細気管支炎における疾患感受性候補領域と候補遺伝子 …………… 9
土方美奈子 他
2. 非定型抗酸菌症の疾患感受性遺伝子同定へ向けて—マイクロサテライトマーカーを用いたゲノムワイド相関解析の第一次スクリーニングの実施 …………… 12
田中 剛 他
3. マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発と連鎖不平衡ブロックの生成に高頻度組換え点を与える影響の理論的検討 …………… 15
大橋 順 他
4. びまん性汎細気管支炎（DPB）におけるムチン遺伝子群プロモーター多型の解析 …………… 24
神尾孝一郎 他

III. 研究成果の刊行に関する一覧表 …………… 33

IV. 研究成果の刊行物・別刷

1. Diffuse Panbronchiolitis …………… 43
K.Kudoh 他
2. Direct determination of MUC5B promoter haplotypes based on the method of singlestrand conformation polymorphism and their statistical estimation. …………… 54
K.Kamio 他
3. Power of genome-wide linkage disequilibrium testing by using microsatellite markers. …………… 87
J. Ohashi 他
4. Power of association test for detecting minor histocompatibility gene causing graft-versushost disease following bone marrow transplantation …………… 92
J. Ohashi 他
5. CD36 polymorphism is associated with protection from cerebral malaria …………… 98
K. Omi 他
6. Identification of IkbL as the second major histocompatibility complex-linked susceptibility locus for rheumatoid arthritis. …………… 109
K. Okamoto 他
7. Diabetes mellitus may increase risk for idiopathic pulmonary fibrosis …………… 119
T.Enomoto 他
8. 呼吸器疾患の分子疫学—びまん性汎細気管支炎 …………… 124
慶長 直人 他

I . 総括研究報告書

びまん性汎細気管支炎等、遺伝素因を有する慢性呼吸器疾患の 疾患感受性遺伝子の研究

主任研究者 慶長直人 国立国際医療センター 研究所・呼吸器疾患研究部

研究要旨 我々はびまん性汎細気管支炎の主要疾患感受性遺伝子の候補領域を HLA クラス I 領域 HLA-B 遺伝子座より、300 kb ほど A 座側の約 200 kb と推定した。平成 13 年度はその 200 kb の候補領域について、遺伝子予測部位を中心に単一塩基置換（SNPs）のスクリーニングを行い、平成 14 年度は、さらに候補領域内の SNP を探索し、のべ 100 個以上の SNPs を同定した。症例対照でタイピングを行い、連鎖不平衡解析とハプロタイプ分析を加え、さらなる候補領域の絞り込みを行った結果、200 kb の候補領域の中には、きわめて連鎖不平衡の強い 80 kb の領域が判明し、その領域内の複数の SNP が疾患と有意な関連を示した。さらに、領域内に、構造上、膜型ムチンと推測される新たな遺伝子をクローニングした。平成 15 年度は、膜型ムチン様遺伝子の同定を完了し、80 kb の候補領域のさらに詳細な変異解析を行った結果、従来のいずれの遺伝マーカーよりも疾患との関連性が強い変異が見いだされた。それらの変異は、主要な疾患感受性ハプロタイプが複数あることをよく説明し、韓国人患者でも共有されていること、欧米人にはまれであるという条件を満たし、分子遺伝学的に有力な疾患感受性遺伝子と考えられた。また、ムチン遺伝子群のプロモーター多型を解析した結果、機能的にも、MUC5B 遺伝子の転写活性を低下させるハプロタイプが同定された。このハプロタイプは、気道過分泌を特徴とするびまん性汎細気管支炎患者には有意に頻度が少なく、疾患抵抗性ハプロタイプとして病態を修飾するものと推測された。

非定型抗酸菌症については、多施設共同研究により、200 例の血液試料を集積し、2 万 マイクロサテライトマーカーによる全ゲノム関連分析の 1 次スクリーニングを行った。現在、統計解析を行い、有意差のあるマーカーを解析中である。

A. 研究目的

びまん性汎細気管支炎は、厚生労働省特定疾患で、難治性稀少疾患として知られ、一連の研究により、ヒト白血球抗原（HLA）と疾患と強い関連が報告されている。我々は未知の DPB 感受性遺伝子が HLA-A 座と B（C）座との間に位置し、DPB に感受性を示す遺伝子の変異が HLA B54-Cw1-A11 を保有する祖先染色体に生じたという仮説の下に、疾患感受性ハプロタイプの推定、連鎖不平衡分析の手法を駆使して、HLA 関連感受性遺伝子座の候補領域を推定し、最終的に疾患感受性遺伝子を同定することを目的と

した。非定型抗酸菌症は、慢性、難治性の非定型抗酸菌感染により生じる疾患であり、特別な免疫力の低下徴候のない、中年以降の女性に発症することが多く、近年、わが国で増加傾向にあることが報告されている。このような疾患の宿主側の感受性遺伝子をスクリーニングすることは、ゲノム創薬上も、意義深いものと思われる。これらの疾患は、いずれも、慢性気道炎症性疾患に属し、共通の分子が発症に関与している可能性も示唆される。

B. 研究方法

厚生省班診断基準に基づき、診断された症例 100 例前後をびまん性汎細気管支炎の疾患群とした。ほかに、慢性閉塞性肺疾患、サルコイドーシス、非定型抗酸菌症についても、副次的に検討を行った。びまん性汎細気管支炎の感受性遺伝子の候補領域となる 200 kb の領域について、コンピュータ予測されるエクソン部位、プロモーター部位を PCR 増幅し、直接シーケンス法によりその領域の全 SNP を同定し、遺伝子型タイピングを行った。さらに新規にマイクロサテライトを領域内に同定し、連鎖不平衡の強い 80 kb の領域については、20 検体で全領域の塩基配列を決定し、網羅的に遺伝子変異を同定し、タイピングを行った。

候補領域内に既知遺伝子は存在せず、新規遺伝子のクローニングには、遺伝子予測プログラムで推測されるエクソン位置に基づき、3'-, 5'-RACE 法を施行、RT-PCR により、組織発現を検討した。

ハプロタイプの推定には、EM アルゴリズムを用いた方法とベイズ推定を応用した方法を用い、SSCP 法を利用したハプロタイプの直接決定法との対比を行い、推定法の信頼性と限界について検討し、permutation 法により、症例と対照の推定ハプロタイプ頻度差に関する検定を行った。

連鎖不平衡の検討には、ハプロタイプ頻度を EM アルゴリズムにより推定し、D' 値と r² 値を解析した。遺伝子座周辺のマイクロサテライトマーカーと SNP の間の連鎖不平衡の及ぶ範囲をこれらのパラメーターおよび尤度比検定により、検討した。また、多因子疾患の関連解析にとって問題となる、マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発と連鎖不平衡ブロックの生成に高頻度組換え点を与える影響の理論的検討を行った。

(倫理面への配慮)

本研究における遺伝子解析に関しては、いずれも三省合同の「ヒトゲノム・遺伝子解析研究に関する倫理指針」に準拠した当センターの遺伝子解析に

係わる倫理委員会の承認を受けている。

C. 研究結果

1) 候補領域における SNPs の同定とタイピング、関連解析

予測されたエクソン部位を中心にプライマーをデザインし、PCR 増幅し、塩基配列を決定し、100 以上の SNPs を同定した。関連解析の結果、200 kb の候補領域のうち、前半の 80 kb の複数の SNPs が疾患と関連性を示した。

2) 候補領域における SNPs の連鎖不平衡構造と遺伝子変異の同定

疾患と関連する 80 kb の領域は強い連鎖不平衡の島状構造を呈していたため、20 個体について、全領域の SNPs 検索を行い、その結果、さらに 200 以上の遺伝子変異を同定した。

3) 候補領域における新規遺伝子の同定

この 80 kb の領域に、新規遺伝子をクローニングした。構造上、新規の膜型ムチン様遺伝子と推測され、気管支上皮細胞に発現していた。

4) 新規遺伝子に見られる遺伝子変異の同定

新規遺伝子のコード領域に疾患と強く関連する遺伝子変異を同定した。

5) 既知のムチン遺伝子のプロモーターハプロタイプの直接決定法の開発と関連解析

SSCP を利用した方法により、プロモーター領域にある遺伝子変異からなるハプロタイプを直接決定した。その特定のハプロタイプとびまん性汎細気管支炎との関連が推測された。in vitro でそのプロモーター活性は、変異の有無により、増減が認められた。

6) 候補遺伝子周辺のマイクロサテライトマ

カーを用いた候補遺伝子のスクリーニング法

サルコイドーシスの疾患感受性遺伝子検索のため、Th1 免疫関連遺伝子座より、100 kb 以内の領域にあるマイクロサテライトマーカートを同定し、疾患との関連性を検討したところ、STAT4 のマーカーとサルコイドーシスとの関連を示した。連鎖不平衡の及ぶ範囲内に、疾患関連 SNP を同定した。

7) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発と連鎖不平衡ブロックの生成に高頻度組換え点を与える影響の理論的検討

これらの検討は、びまん性汎細気管支炎などの多因子疾患の感受性変異探索研究で利用可能であり、疾患感受性変異探索研究の基礎情報として利用できる。

8) 非定型抗酸菌症の疾患感受性遺伝子同定のための多施設研究の進行

マイクロサテライトマーカーを用いたゲノムワイド相関解析を行うための、多施設共同研究組織を立ち上げ、すでに 200 例以上の症例サンプルを蓄積し、2 万マイクロサテライトマーカーによる第一次スクリーニングを実施した。

D. 考案

慢性、難治性呼吸器疾患の多くは、多因子疾患であるが、主要疾患感受性遺伝子が限られ、相対危険度も高いのではないかと推測される。たとえば、本研究の主要テーマのひとつであるびまん性汎細気管支炎は、欧米人に同様の疾患が見られず、欧米においてはアジア系の移民が罹患することが多いという事実から、人類創世後、アジア系集団の形成時に生じた遺伝子変異の一つが、疾患感受性を規定する有力な因子ではないかと考えられる。本年度は、膜型ムチン様遺伝子の同定を完了し、80 kb の候補領域のさらに詳細な変異解析を行った。この遺伝子は気

道上皮の分化に伴い、alternatively spliced form が出現することから、気道の傷害、修復に密接な関連を持つ分子である可能性が示唆される。この遺伝子内の新規遺伝子変異は、従来のいずれの遺伝マーカーよりも疾患との関連性が強く、主要な疾患感受性ハプロタイプをよく説明し、韓国人患者でも共有されていること、欧米人にはまれであるという条件を満たし、分子遺伝学的に有力な疾患感受性遺伝子と考えられた。さらに既知のムチン遺伝子群と疾患との関連についても検討した結果、MUC5B のプロモーターの特定のハプロタイプが、気道過分泌を特徴とするびまん性汎細気管支炎患者に有意に頻度が少なく、機能的にも、MUC5B 遺伝子の転写活性を低下させるものであった。したがって、MUC5B の特定のプロモータハプロタイプは、疾患抵抗性を示し、病態を修飾するものと推測された。

本研究では、このようにこれまで取り残されていた、呼吸器病の臨床に直結した疾患感受性遺伝子探索というテーマを掲げ、初年度、呼吸器病学の研究者と第一線のゲノム研究者が共同戦線をひき、次年度は、各疾患について、データを蓄積し、ここに有意な結果を報告することができた。

E. 結論

80 kb の候補領域内に新規膜型ムチン様遺伝子を同定し、詳細な変異解析を行った結果、従来のいずれの遺伝マーカーよりも疾患との関連性が強い変異が見いだされた。それらの変異は、主要な疾患感受性ハプロタイプが複数あることをよく説明し、韓国人患者でも共有されていること、欧米人にはまれであるという条件を満たし、分子遺伝学的に有力な疾患感受性遺伝子と考えられる。また、ムチン遺伝子群のプロモーター多型を解析した結果、機能的にも、MUC5B 遺伝子の転写活性を低下させるハプロタイプが同定された。このハプロタイプは、気道過分泌を特徴とするびまん性汎細気管支炎患者には有意に頻度が少なく、疾患抵抗性ハプロタイプとして

病態を修飾するものと推測された。非定型抗酸菌症については、多施設共同研究により、200 例の血液試料を集積し、2 万 マイクロサテライトマーカーによる全ゲノム関連分析の 1 次スクリーニングを行った。現在、統計解析を行い、有意差のあるマーカーを解析中である。

本研究では、このようにこれまで取り残されていた、呼吸器疾患の臨床に直結した感受性遺伝子検索というテーマを掲げ、呼吸器病学の研究者と第一線のゲノム研究者が共同戦線をひき、各疾患についてデータを蓄積し、最終年度、ついに分子遺伝学に有意な疾患感受性遺伝子を同定することができた。

参考文献

- 1) Risch N, Merikangas K : The future of genetic studies of complex human diseases. *Science* 273 : 1516-1517, 1996
- 2) Koch HG, McClay J, Loh EW et al : Allele association studies with SSR and SNP markers at known physical distances within a 1 Mb region embracing the ALDH2 locus in the Japanese, demonstrates linkage disequilibrium extending up to 400 kb. *Hum Mol Genet* 9 : 2993-2999, 2000
- 3) Barcellos LF, Klitz W, Field LL et al : Association mapping of disease loci, by use of a pooled DNA genomic screen. *Am J Hum Genet* 61 : 734-747, 1997
- 4) Daniels J, Holmans P, Williams N et al : A simple method for analyzing microsatellite allele image patterns generated from DNA pools and its application to allelic association studies. *Am J Hum Genet* 62: 1189-1197, 1998
- 5) Collins HE, Li H, Inda SE et al : A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. *Hum Genet* 106 : 218-226, 2000

II . 分担研究報告

びまん性汎細気管支炎における疾患感受性候補領域と候補遺伝子

主任研究者	土方美奈子 慶長 直人	国立国際医療センター 研究所・呼吸器疾患研究部 国立国際医療センター 研究所・呼吸器疾患研究部
分担研究者	松下 育美	国立国際医療センター 研究所・呼吸器疾患研究部
分担研究者	中田 光	国立国際医療センター 研究所・呼吸器疾患研究部
分担研究者	徳永 勝士	東京大学大学院医学系研究科人類遺伝学教室
	大橋 順	東京大学大学院医学系研究科人類遺伝学教室
分担研究者	吾妻安良太	日本医科大学第四内科
	工藤 翔二	日本医科大学第四内科

研究要旨 びまん性汎細気管支炎の疾患感受性候補遺伝領域として、HLA-A, B 遺伝子間の 200kb の領域が推定されている。昨年度までの検討で、200kb の候補領域の中に特に連鎖不平衡の強い 80kb の領域が存在すること、その連鎖不平衡の島状構造の中にある複数の一塩基多型(SNPs)が疾患と有意な関連を示すこと、同部位に新しい膜型ムチン様遺伝子が存在することが判明していた。このムチン様遺伝子には疾患との関連を示す多型も見出されていたが、連鎖不平衡の結果である可能性も否定できなかった。本年度の検討では、連鎖不平衡の強いこの 80kb の領域の中の遺伝子多型の網羅的探索および遺伝子予測プログラムで予測された遺伝子のクローニングを試みた。その結果、80kb の領域の中に疾患と強く関連する新たなマイクロサテライトマーカーが見出された他、新規膜型ムチン様遺伝子に疾患と関連する遺伝子変異が新たに見出された。網羅的探索の結果より、80kb の領域からクローニングされたもうひとつの遺伝子には疾患により強く関連する多型がないこと、この膜型ムチンの遺伝子の中には、昨年度報告したようにコントロールと比較して疾患群で有意に頻度の低い多型もあることなどから、このムチン様遺伝子がびまん性汎細気管支炎の発症に関連する可能性が高いと考えられた。

A. 研究目的

びまん性汎細気管支炎(diffuse panbronchiolitis, DPB)は日本でその疾患概念が確立した疾患であるが、日本だけでなく韓国や中国などの東アジア地域で発症がみられ、アジア人特有の遺伝要因が DPB の発症に関与している可能性が示唆されている。日本人の DPB 患者集団では HLA-B54 の保有率が有意に高いのに対し¹⁾、韓国では HLA-B54 とは明らかな関連を示さず、かわって HLA-A11 が DPB と強い関連を示すことが報告された²⁾。このことから、特定の HLA 分子そのものが疾患感受性を規定するのではなく、HLA-A, B 遺伝子座間に DPB の発症に関連する遺伝子が存在し、HLA-B54 と A11 をあわせもつ日韓共通の祖先染色体上にその遺伝子の疾患感受性変異が生じ、その後歴史的組み換えを経て、日本では B54 との連鎖不平衡が保たれ、韓国では A11 との連鎖不平衡が保たれたことにより、現在それぞれの HLA 型と関連

性を示すに至ったと考えられた。この仮説に基づいた検討により、HLA-A 座より 300kb ほど HLA-A 座よりの 200kb の領域が疾患感受性遺伝子の候補領域として推定されていた³⁾。

昨年までの検討で、200kb の候補領域の中を、遺伝子予測プログラムで予測したエクソン部位を中心に検討した結果、特に連鎖不平衡の強い 80kb の領域が存在すること、その連鎖不平衡の島状構造の中にある複数の一塩基多型(SNPs)が疾患と有意な関連を示すこと、新しい膜型ムチン様遺伝子が同部位に存在することが判明していた。しかし、疾患と有意な関連を示す遺伝子多型が連鎖不平衡の結果である可能性も否定できなかったため、本年度は、この 80kb の領域の遺伝子多型の網羅的探索および 80kb の領域内に予測されたすべての遺伝子のクローニングを試みた。

B. 研究方法

昨年度までの検討により、疾患と有意な関連を示す複数の SNPs を含む 80kb に渡る連鎖不平衡の強い領域を見出していたが、これは遺伝子予測プログラム GENSCAN でエクソンと予測された部位を中心に解析した結果であるため、今年度は、さらにこの領域について、エクソン予測部位以外の部分でも網羅的に SNPs を探索した。材料としては、疾患に関連するマイクロサテライトのアレル (C2_2_2 が 251 かつ C2_4_4 が 231) をホモ接合で持つ DPB 患者 12 例と健常コントロール 12 例、計 24 例を対象とし、この 80kb を約 20 領域に分割して平均約 4kb を増幅する PCR を行い、シーケンスプライマーを約 600 塩基程度ごとに設定して、PCR 産物を ABI PRISM3100 DNA シーケンサー (Applied Biosystems) を用いた直接シーケンス法により塩基配列を決定し、SNP を同定した。得られた SNPs を用いて、ハプロタイプ構造の詳細な解析を行った。また、疾患に関連するハプロタイプのみが持つと考えられる変異が見出された場合には、DPB92 例と健常コントロール 100 例を対象として関連解析を行った。

同時に、この 80kb の領域内で GENSCAN により予測された 4 つの遺伝子のクローニングを試みた。ヒト気管支上皮細胞株の NCI-H292 細胞・A549 細胞・BET1A 細胞、EBV 不死化ヒト B 細胞株を材料にし、total RNA を抽出し、ランダムプライマーを用いて cDNA 化した。予測された遺伝子エクソン部のうち、上流エクソンにセンスプライマーを設定し、下流エクソンにアンチセンスプライマーを設定し、cDNA を PCR で増幅した。増幅産物は直接シーケンス法で塩基配列を決定した。この 4 つの材料からは増幅産物が得られなかった予測遺伝子については、Human MTC パネル I および II (CLONTECH) を用いて、16 種類のヒト組織由来の cDNA を材料にして PCR 増幅を試みた。

図 1 に本年度の研究方法の概要を示す。

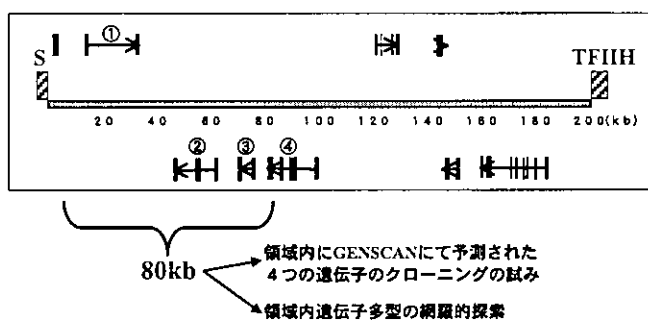


図 1 S 遺伝子から TFIIH 遺伝子までの 200kb の疾患感受性遺伝子の候補領域のうち、連鎖不平衡の強い 80kb の候補領域に対する今年度の研究方法

(倫理面への配慮) 本研究における遺伝子解析に関しては、三省合同の「ヒトゲノム・遺伝子解析研究に関する倫理指針」に準拠した当センターの遺伝子解析に係わる倫理委員会の承認を受けている。

C. 研究成果

疾患に関連するマイクロサテライトのアレル (C2_2_2 が 251 かつ C2_4_4 が 231) をホモ接合で持つ 24 例を対象にした 80kb の網羅的な SNPs 探索については、リピート配列構造を持つ約 10kb の領域が解析不能であったが、解析できた約 70kb の領域から 100 個以上の SNPs が同定された。また、新たに疾患と強く関連するマイクロサテライトマーカーもこの領域内に同定された。この 80kb の領域はもともと強い連鎖不平衡にある領域であるが、ハプロタイプ構造を詳細に検討した結果、疾患に関連するハプロタイプが存在することが判明した。また、疾患に関連するハプロタイプのみが持つ遺伝子多型が少数ながら存在し、これらの遺伝子多型については、DPB92 例と健常コントロール 100 例を対象として関連解析を行い、疾患と関連することが明らかになった。

遺伝子クローニングの試みについては、GENSCAN による 4 つの予測遺伝子のうち 2 つの遺伝子 (図 1 の②と④) が RT/PCR で実際に増幅された。また、この 2 つともヒト肺組織で陽性であった。このうち 1 つ、図 1 の④は昨年度の研究結果である膜型ムチン遺伝子である。残りの 2 つの遺伝子 (図 1 の①と③) は、16 臓器由来の cDNA パネルを用いた検討まで行ったが、いずれの臓器からも増幅されなかった。

今年度新たにクローニングされた遺伝子 (図 1 の②) には、膜型ムチンの遺伝子と比較して、疾患とより強く関連する変異は見出されなかった。一方、膜型ムチン遺伝子 (図 1 の④) は、80kb の領域の辺縁部にその一部が位置し、残りの部分は 80kb の領域外にあるため、更に範囲を広げて遺伝子変異を網羅的に探索した。その結果、疾患と関連する遺伝子変異がこの遺伝子のエクソン 2 の内部に新たに見出された。

D. 考案

マイクロサテライトマーカーを用いた検討により、HLA-A 座より 300kb ほど HLA-A 座よりの 200kb の領域が疾患感受性遺伝子の候補領域として推定されていたが、予測されたエクソン部位を中心に検討して得られた SNPs を用いた昨年度までの結果により、更に 80kb まで候補領域が絞り込まれていた。しかし、この 80kb の領域は非常に連鎖不平衡の強い領域であったため、見出された疾患と関連する遺伝子多型が真に疾患と関連する変異なのか、それとも真に疾患と関連する変異は別に存在し、連鎖不平衡の結果として見かけ上疾患と関連しているのかの判断は困難な状況であった。そのため、今年度の検討は、遺伝子クローニングの面でも、多型の探索の面でも連鎖不平衡のブロック構造内について網羅的に行うことを目標とした。

その結果、80kb の領域に新しい遺伝子がひとつクローニングされたが、その遺伝子には疾患と特に強く関連する遺伝子多型は見出されなかった。それに対し、膜型ムチン遺伝子について、更に範囲を広げて遺伝子変異を網羅的に探索した結果、疾患と関連する遺伝子変異がエクソン 2 に新たに見出された。80kb の領域のうち、解析不能であったリピート配列構造を持つ約 10kb の領域を除いて残りの領域は網羅的に探索しており、遺伝子が存在する領域以外に、疾患と強く関連する DNA 多型がある可能性はほぼ排除されている。また、昨年度の検討で、膜型ムチン遺伝子のエクソン 2 の繰り返し配列について、最も短いアリルが症例・対照集団において出現頻度に偏りが認められ、症例での頻度が有意に低いことも判明しており、このエクソン 2 が短いものは疾患抵抗性の性質を有すると考えられた。以上の状況より、この膜型ムチン遺伝子が DPB の発症に関与する可能性が高いと判断された。

E. 結論

びまん性汎細気管支炎の HLA 関連疾患感受性遺伝子候補領域の中に、疾患と関連する遺伝子多型を有する膜型ムチン遺伝子がクローニングされた。

「謝 辞」

虎の門病院、天理よろず相談所病院、日本医科大学第 4 内科の諸先生のご協力をいただきましたことに感謝いたします。

参考文献

1) Keicho N et al. Contribution of HLA genes to

genetic predisposition in diffuse panbronchiolitis. *Am J Respir Crit Care Med* 158:846-850, 1998.

2) Park MH et al. Association of class I antigens with diffuse panbronchiolitis in Korean patients. *Am J Respir Crit Care Med* 159:526-529, 1999.

3) Keicho N et al. Fine localization of a major disease-susceptibility locus for diffuse panbronchiolitis. *Am J Hum Genet* 66:501-507, 2000.

F. 健康危険情報 (略)

G. 研究発表

1. 論文発表

2. 学会発表

H. 知的財産権の出願 (略)

非定型抗酸菌症の疾患感受性遺伝子同定へ向けて—マイクロサテライトマーカーを用いたゲノムワイド相関解析の第一次スクリーニングの実施

研究協力者 田中 剛、庄嶋淳子、慶長直人 国立国際医療センター 研究所・呼吸器疾患研究部
分担研究者 田宮 元 猪子英俊 東海大学医学部基礎医学系分子生命科学

研究要旨 非定型抗酸菌症（MAC症）の疾患感受性遺伝子のゲノムワイドなスクリーニングのために、連鎖不平衡のおよぶと考えられる 100kb に 1 個の間隔で計 3 万個のマイクロサテライトマーカーを配置すると、ゲノムワイドな連鎖不平衡マッピングを行うことが可能となると考えられる。すでに、ヒトゲノムドラフト配列より、2~6 塩基繰り返しの新規マイクロサテライト配列を抽出し、プライマー設計を行い、日本人一般健常者集団 100~200 人を対象として多型検索を行い、約 3 万マーカーの設定が終了しており、100~200 人程度のゲノム DNA を混合（Pool）した DNA を用いたマイクロサテライトタイピング法（Pooled DNA 法）を利用し、効率的にタイピングを行うことが可能となったため、この方法を用いて、多施設共同研究による、非定型抗酸菌症の一次スクリーニングを開始した。

A. 研究目的

ヒトゲノムの全塩基配列がほぼ決定した現在、ヒトゲノムの全 DNA 配列を利用したヒトゲノム多様性解析は、解読された DNA 配列全体に多型性情報を蓄積し、遺伝的マーカーとして利用することによって、ゲノム全体にわたって高解像度でかつ特異的に感受性遺伝子同定を行えるアプローチとして注目されている 1)。

このような状況の中、100kb 以上にわたって連鎖不平衡を維持できることが知られているマイクロサテライトマーカー 2) を、100kb に 1 個の間隔で設定・配置すれば、ゲノムワイドに連鎖不平衡マッピングを行うことが可能となると試算された。また、マイクロサテライトは、塩基配列をコンピュータ解析することにより、容易にそのゲノム上の位置を推測できる。

非定型抗酸菌症は呼吸器疾患の中でも、慢性、難治性の非定型抗酸菌感染により生じる疾患であり、特別な免疫力の低下徴候のない、中年以降の女性に発症することが多く、近年、わが国で増加傾向にあ

ることが報告されている。本症は、緩徐に進行することが多く、有効な治療法が乏しく、新たな特効薬の開発が急がれる。このような疾患の宿主側の感受性遺伝子をゲノムワイド相関解析により、スクリーニングすることは、ゲノム創薬上も、意義深いものと思われる。

B. 研究方法

1. 対象

非定型抗酸菌症については、対象を MAC 症に限定し、ATS1997 基準（Am J Respir Crit Care Med Vol. 156, pp. S1-S25, 1997）に基づいて診断された症例（治療中、治療後を問わない）で、HIV 症例、血液疾患症例（白血病、悪性リンパ腫など）、免疫抑制剤（ステロイドを含む）使用例は対象から除く。既存の肺病変の有無は問わない。説明と同意の可能な症例を対象とする。国立療養所近畿中央病院（井上義一、鈴木克洋、坂谷光則、岡田全司、木村謙太郎）を中心に、当センター、国立療養所東京病院（永井英明、倉島篤行）の協力を得、さらに国立療養所

大牟田病院（原田 進）、国立療養所西新潟中央病院（桶谷典弘）の参加により、1次、2次スクリーニングに必要な臨床検体を確保した。試料提供施設を追加して、最終的に300例を集積する。

2. マイクロサテライトマーカーを用いたゲノムワイド相関解析

ゲノムワイドな遺伝的相関解析の進め方として、理想的には、対象集団を以下の三段階に分けて解析を行う。すなわち、①非血縁患者集団と非血縁健常者集団を用いた検出力の高い相関解析を補正無しで行うことにより、有意な相関を示すマーカーの一次スクリーニングを行い、②相関を示したマーカーに関して別の集団を用いて補正有りで二次スクリーニングを行い、③最終的には患者の両親を内部コントロールにする TDT (transmission disequilibrium test) 法を採用して相関を追試する。現時点では、非定型抗酸菌症 300 例と、健常 300 例を、100 対 100 の三段階に分け、アリル頻度で、 $p < 0.05$ をカットオフとして、スクリーニングを行う (表)。

マイクロサテライトのタイピングと統計解析 (1stスクリーニング)

作業の流れ

- ・PCR (蛍光プライマー, 96マーカープレート単位)
- ・PCRプロダクトの調整 (希釈, dry up...)
- ・ABI 3700での泳動
- ・サイジング
- ・ピークの取得
- ・統計解析
- ・データベース登録
- ・エラーの処理

表 作業工程

C. 研究結果

本研究は、マイクロサテライトゲノムワイド相関解析の総責任者である、東海大学医学部分子生命科学2の猪子英俊教授に、平成13年12月17日、本研究班の主任研究者（慶長）が問い合わせた結果、非定型抗酸菌症のマイクロサテライトゲノムワイド相関解析のための血液検体を集めること、集まった

ものから、実際に相関解析を開始してよいとの許可を得たため、まず当センター倫理委員会で、平成14年3月8日、当センター単独の検体収集に関する承認を受け、ここから検体の蓄積を開始し、さらに国立療養所近畿中央病院を中心に、話し合いを進め、300検体収集のための多施設研究を組織し、平成15年3月14日に多施設研究に関わる承認を受け、国立療養所近畿中央病院でも平成14年6月に承認を得た。そこで、これを受けて、平成15年8月18日付で、主任研究者、猪子英俊教授、分担研究者、田宮 元、岡 晃、慶長直人、中田 光、田中 剛、ほか三名で、「非定型抗酸菌症感受性領域の全ゲノム高解像度マッピング」の申請がなされ、同10月15日に東海大学の医の倫理委員会にて承認されたため、同11月より、実際に、田中 剛を中心とした、一次スクリーニングが開始された。

D. 考案

非定型抗酸菌症は、たとえば、肺結核に見られるような、貧困、低栄養など、あきらかな環境要因がなく、また集団感染が見られないことから、菌側の要素として、強い感染性が発症の原因になっていると考えられない。個体のもつ何らかの内的な素因が発症に重要であることが、臨床的に推測されている。対照集団を3段階に分けて解析を行うことにより、サンプル構造に依存した偽陽性を示すマーカーを、無理な補正無しに大幅に減少させることができるものと期待される。また、Pooled DNA 法は、Individual タイピングに比べて時間とコストの大幅な軽減が期待できるため、効率的にタイピング・多型検索を行う方法として注目されている (3-5)。本法では、増幅した PCR 産物の電気泳動上の波形パターンをもとにして推定対立遺伝子頻度を求める。すなわち、波形パターンから全てのピークの高さの総和を求め、その値に対してそれぞれのピークの高さの割合を求めることで推定上の頻度を算出する。

このように、マイクロサテライトマーカーを用

いたゲノムワイド相関解析により感受性候補領域を100kb程度まで絞り込んだ後には、各感受性候補領域内に存在するSNPマーカーを用いて相関解析・ハプロタイプ解析を行うことにより、ゲノム中に存在する全ての疾患感受性遺伝子および疾患感受性遺伝子座を、高い検出感度を維持したまま同定できると予想される。

E. 結論

われわれは、これまで遺伝学的解析に広く利用されてきたマイクロサテライトマーカーの有用性に着目し、約3万遺伝子座という前例のない規模で遺伝的相関解析を行っている。本解析により、未だ確実な方法論のない疾患感受性候補領域の絞り込みから感受性遺伝子の同定に至る戦略を確立し、それをシステム化することで、今後迅速に非定型抗酸菌症の疾患感受性遺伝子の同定が可能となると期待される。

参考文献

- 1) Risch N, Merikangas K : The future of genetic studies of complex human diseases. *Science* 273 : 1516-1517, 1996
- 2) Koch HG, McClay J, Loh EW et al : Allele association studies with SSR and SNP markers at known physical distances within a 1 Mb region embracing the ALDH2 locus in the Japanese, demonstrates linkage disequilibrium extending up to 400 kb. *Hum Mol Genet* 9 : 2993-2999, 2000
- 3) Barcellos LF, Klitz W, Field LL et al : Association mapping of disease loci, by use of a pooled DNA genomic screen. *Am J Hum Genet* 61 : 734-747, 1997
- 4) Daniels J, Holmans P, Williams N et al : A simple method for analyzing microsatellite allele image patterns generated from DNA pools and its application to allelic association studies. *Am J Hum Genet* 62: 1189-1197, 1998

- 5) Collins HE, Li H, Inda SE et al : A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. *Hum Genet* 106 : 218-226, 2000

マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発と連鎖不平衡ブロックの生成に高頻度組換え点を与える影響の理論的検討

分担研究者 大橋 順 東京大学大学院医学系研究科人類遺伝学教室

分担研究者 徳永 勝士 東京大学大学院医学系研究科人類遺伝学教室

研究要旨 本研究では、びまん性汎細気管支炎などの多因子疾患の感受性変異探索研究で利用可能な「(1) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発」と、疾患感受性変異探索研究の基礎情報として有用な、「(2) 連鎖不平衡ブロックの生成に高頻度組換え点を与える影響の理論的検討」を行った。これらの研究内容の要旨は以下の通りである。

(1) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発

マイクロサテライトマーカーでの突然変異は再帰的であり、かつリピート数の近いアリルへ突然変異を起こす確率が高いことが分かっている。本研究では、そのようなマイクロサテライトマーカーの特徴を考慮して、マイクロサテライトマーカーを用いる連鎖不平衡マッピングで利用可能な、並べ替え検定に基づいた新規統計検定法を開発した。また、マイクロサテライトマーカーでの突然変異をコンピュータシミュレーションにより再現すべく、K 対立遺伝子単一ステップ突然変異モデルを構築した。

(2) 連鎖不平衡ブロックの生成に高頻度組換え点を与える影響の理論的検討

未知の疾患感受性変異をマーカーとの間の連鎖不平衡によって同定しようとする連鎖不平衡マッピングにおいては、調べるゲノム領域の連鎖不平衡の程度（連鎖不平衡ブロックの有無）に応じてマーカーを設定する必要がある。連鎖不平衡マッピングでのマーカーの配置を考える際に、連鎖不平衡ブロックの生成メカニズムを理解しておくことは重要であり、本研究では、連鎖不平衡ブロック間の境界が高頻度組換え点によって生成される条件をコンピュータシミュレーションに基づいて理論的に検討した。

A. 研究目的

(1) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発

多因子疾患の感受性遺伝子変異の同定を目的とする連鎖不平衡マッピングでは、単一塩基多型 (SNP) がマイクロサテライトマーカーのいずれかが遺伝マーカーとして利用されるが、ゲノム中での密度の高さ（約 300-500bp に 1 個）とタイピングの容易さから、SNP が連鎖不平衡マッピングのためのマーカーとして注目されている。また、SNP を単独で解析す

るだけでなく、SNP 同士の組合せ (SNP haplotype や tagged SNP) をマーカーとして利用する手法の有効性も報告されている。しかし、SNP はアリルがただか 2 個しかなく、研究対象集団中において、未知の疾患感受性変異（多くは 2 対立遺伝子からなる多型）と SNP マーカーとの間の相関が高く無い（両多型間での連鎖不平衡係数 r^2 が小さい）場合には、その検出力が著しく低下することが知られており、同様の条件下において、マイクロサテライトマーカーを用いた方がより検出力が高くなる可能性が指摘

されている。特に、浸透率は高いが、疾患感受性変異の集団頻度が低いような場合には、マイナーアレル頻度が高い SNP のみを選んで行う連鎖不平衡マッピングでは、感受性変異を見逃す（偽陰性が起こる）可能性もあると予想される。連鎖不平衡マッピングにおいて、マイクロサテライトマーカーを利用することの最大の利点は、マイクロサテライトは複数のアレルを有することから、そのうちの一つのアレルと疾患感受性変異との間の相関が高いことが期待され、その結果として検出力が高くなることが挙げられる。その一方、SNP に比べて突然変異率が高いことから、感受性変異との間の連鎖不平衡が崩れやすいというマーカーとしての不利な一面もある。しかし、マイクロサテライトマーカーで起こる突然変異は、現在のアレルのリピート数より 1 個多い、または 1 個少ないアレルへの突然変異が大部分と考えられ、組換えによる連鎖不平衡の崩れ方とは異なった連鎖不平衡の崩れ方を示す。すなわち、たとえマイクロサテライトマーカーで突然変異が起こったとしても、もとのアレルとリピート数が近いアレルへの突然変異であることから、感受性変異を有する染色体のマイクロサテライトマーカーアレルは互いにリピート数が近いことが期待される。このことは、感受性変異を有する個体が多い case 集団では、その感受性変異の近傍のマイクロサテライトマーカーでアレルのリピート数のペアワイズな差が、control 集団でのそれに比較して小さくなることを示唆する。そこで、本研究では、そのようなマイクロサテライトマーカーの特徴に鑑み、マイクロサテライトマーカーを用いる連鎖不平衡マッピングで利用可能な、新規統計解析手法を開発することを主要な目的とした。なお、マイクロサテライトマーカーでの突然変異モデルの検討と 2 染色体間で期待されるリピート数の差の理論的

考察も併せて行った。

(2) 連鎖不平衡ブロックの生成に高頻度組換え点が与える影響の理論的検討

未知の疾患感受性変異をマーカーとの間の連鎖不平衡によって同定しようとする連鎖不平衡マッピングにおいては、調べるゲノム領域の連鎖不平衡の程度（連鎖不平衡ブロックの有無）に応じてマーカーを設定する必要がある。もし連鎖不平衡が強い領域であれば、少数のマーカーによって当該領域をカバーすることで（連鎖不平衡マッピングにおける当該領域の代表マーカーになる）、タイピングコストを軽減することが可能となるからである。また、一般的には対象集団ごとに連鎖不平衡ブロックの位置は異なるため、集団ごとに連鎖不平衡ブロックの位置を同定する必要がある。しかし、連鎖不平衡ブロック間の境界が高頻度組換え点によって生成されているとすると、そのようなゲノム領域では対象集団に関わらず同一の連鎖不平衡パターンをとると予想される。そのため、連鎖不平衡マッピングでのマーカーの配置を考える際に、連鎖不平衡ブロックの生成メカニズムを理解しておくことは重要である。そこで、本研究では、連鎖不平衡マッピングのための基礎情報として、連鎖不平衡ブロックの生成に高頻度組換え点が与える影響の理論的検討を行うことを第二の目的とした。

B. 研究方法

(1) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発
(i) マイクロサテライトマーカーでの突然変異モデルの検討と 2 染色体間で期待されるリピート数の差
マイクロサテライトマーカーでの突然変異が起こる主要なメカニズムはスリップページである（不等交差による突然変異は稀と考えられている）。ヒトのマイクロサテライトマーカーでの突然変異率は、マーカーごとに異なるが、減数分裂当り 10^{-5} ~ 10^{-3} と推定されている。本研究では、最初にマイクロサテライトマーカーでの突然変異モデルの検討を行った。マイクロサテライトマーカーでの突然変異モデルとして、これまでに次の 4 つの代表的なモデルが提唱

されている。なお、以下では 2 倍体集団の有効集団サイズを N_e 、突然変異率を v とおき、 $M = 4N_e v$ 、平均ホモ接合度（集団集からランダムに選んだ二つの染色体が同じアレル（リピート数）を有する確率）を H とする。

(a) 無限対立遺伝子モデル (infinite allele model, Kimura and Crow 1964)

このモデルでは、突然変異が起こると、それまでに集団中には存在しなかった新たなリピート数をもつアレルが誕生すると仮定する。この場合には、 $H = (1+M)^{-1}$ である。しかし、このモデルは、マイクロサテライトマーカーでの突然変異を考えるためのモデルとして明らかに不適切と思われる。

(b) K 対立遺伝子モデル (K -allele model, Kimura and Crow 1970)

このモデルでは、あるマイクロサテライトマーカーでは K 種類のアレル (K 個の異なるリピート数) のみが存在し、突然変異によって、残りの $K-1$ 個のアレルの何れかに $v/(K-1)$ の確率で突然変異を起こすと仮定するものである。この場合には、 $H = [1 + M/(K-1)] / [1 + MK/(K-1)]$ である。

(c) 単一ステップ突然変異モデル (single-step mutation model [SSMM], Ohta and Kimura 1973, Kimura and Ohta 1975, Kimura and Ohta 1978)

このモデルでは、あるリピート数もつアレルは、等確率 ($v/2$) で 1 個リピート数の多い、または 1 個リピート数の少ないアレルに突然変異を起こすと仮定する。なお、リピート数の上限や下限については考慮しない。この場合には、 $H = (1+2M)^{-0.5}$ である。

(d) 多重ステップ突然変異モデル (multi-step mutation model [MSMM])

これまでに複数の MSMM が提案されているが、代表的なものに一般化突然変異モデル (generalized mutation model [GMM]) がある。これは、あるリピート数をもつアレルが突然変異を起こし、 X リピート数増えるか、 X リピート数減るかする確率が、 $P(X$

$= x) = a(1-a)x^{-1}$ という幾何分布に従うと仮定するモデルである。なお、リピート数の上限や下限については考慮しない。

以上の 4 つのモデルのうち、リピート数には明らかに上限と下限が存在することを考えると (b) は妥当であり、突然変異はリピート数が一つ増えるか一つ減るかするケースが多いという観察結果をふまれば、(c) や (d) は妥当と思われる。そこで、(b) と (c) を組み合わせた K 対立遺伝子単一ステップモデル (K -allele SSMM) を仮定して [図 1]、モンテカルロシミュレーションを行った。なお、(d) を仮定しなかったのは、パラメタ数の軽減を図るためである。

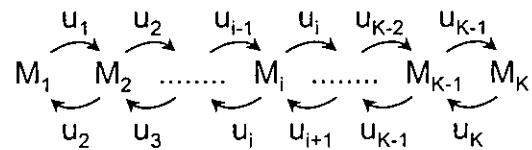


図1 K 対立遺伝子単一ステップ突然変異モデル
 M_i はアレルを示し、コンピュータシミュレーションでは $u_i = v/2$ とした。

[図 1]

次に、集団中から二つの染色体をランダムに抽出した際の、マイクロサテライトマーカーでのリピート数の差の期待値を計算するアルゴリズムを考察した。集団中から 2 つの染色体を無作為に選んだ場合に、それらは必ず共通祖先をもつ。すなわち、着目するマイクロサテライトマーカーでの祖先遺伝子は共通である。二つの染色体が分岐後、突然変異はそれぞれの経路で独立に起こり、両染色体のマイクロサテライトマーカーでのリピート数には差が生じるようになる [図 2]。

[図 2]

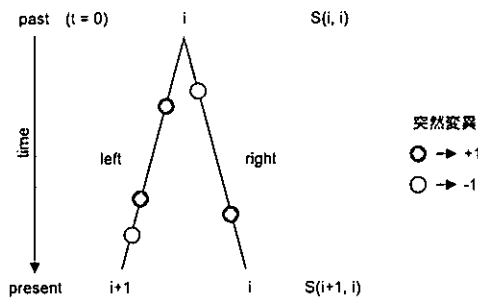


図2 二つの染色体間でのリピート数の差
 二つの染色体は必ず共通祖先（リピート数は*i*）を有しており、その時点ではリピート数の差*d*は0である。時間が経過するとともに、各染色体で突然変異が起これ、左側の染色体では時間*t*後にリピート数は*i+1*、右側の染色体では時間*t*後にリピート数は*i*となっている。結果として、これら2本の染色体では、時間*t*後のリピート数の差は1である。

ここで、 $S(i, j)$ は、便宜上 i が左側の染色体上でのリピート数、 j は右側の染色体上のリピート数を示している。すなわち、以下では染色体は区別して考える。共通祖先から分岐後の時間 t に左側の染色体が i リピート、右側の染色体が j リピートもっている確率を $P(i, j)_t$ とおく。図 1 の K 対立遺伝子単一ステップ突然変異モデルを仮定すると、その状態変化は $P(i, j)_t$ の漸化式を利用して記述することが出来る。例えば、 $K = 3$ であれば、

$$P(1, 1)_{t+1} = (1-v)P(1, 1)_t + (v/2) \{P(1, 2)_t + P(2, 1)_t\}$$

$$P(1, 2)_{t+1} = (1-3v/2)P(1, 2)_t + (v/2) \{P(1, 1)_t + P(2, 2)_t + P(1, 3)_t\}$$

$$P(1, 3)_{t+1} = (1-v)P(1, 3)_t + (v/2) \{P(1, 2)_t + P(2, 3)_t\}$$

$$P(2, 1)_{t+1} = (1-3v/2)P(2, 1)_t + (v/2) \{P(1, 1)_t + P(2, 2)_t + P(3, 1)_t\}$$

$$P(2, 2)_{t+1} = (1-2v)P(2, 2)_t + (v/2) \{P(1, 2)_t + P(2, 1)_t + P(2, 3)_t + P(3, 2)_t\}$$

$$P(2, 3)_{t+1} = (1-3v/2)P(2, 3)_t + (v/2) \{P(2, 2)_t + P(3, 3)_t + P(1, 3)_t\}$$

$$P(3, 1)_{t+1} = (1-v)P(3, 1)_t + (v/2) \{P(2, 1)_t + P(3, 2)_t\}$$

$$P(3, 2)_{t+1} = (1-3v/2)P(3, 2)_t + (v/2) \{P(2, 2)_t + P(3, 3)_t + P(3, 1)_t\}$$

$$P(3, 3)_{t+1} = (1-v)P(3, 3)_t + (v/2) \{P(2, 3)_t + P(3, 2)_t\}$$

なる漸化式が成立する。ここで、 t を無限大にすると、 $P(i, j)_t$ は $1/K^2$ に近づく。ある時間 t に二つの染色体間でのリピート数の差が d である確率を $Q(d)_t$ とおくと、 $|i-j| = d$ となる全ての組合せに対して $P(i, j)_t$ を足し合わせたものが $Q(d)_t$ となる。個体数が N の二倍体集団で 2 染色体が時間 t で合体する確率は $C(t) = (1/2N) (1 - (1/2N))^{t-1}$ で与えられることから、となり、期待されるリピート数の差は d で表される。 $P(i, i)_0 = 1/K$ 、 $P(i, j)_0 = 0$ for $i < j$ or $i > j$ とすれば、上の漸化式を解くことができ、 $Q(d)_t$ を計算することも可能である。なお、 $K = 3$ に対して上の漸化式を解くと、

$$P(1, 1)_t = (1/9) + (1/6)(1-v)^t + (1/18)(1-3v)^t$$

$$P(1, 2)_t = (1/9) - (1/9)(1-3v)^t$$

$$P(1, 3)_t = (1/9) - (1/6)(1-v)^t + (1/18)(1-3v)^t$$

$$P(2, 2)_t = (1/9) + (2/9)(1-3v)^t$$

$$P(2, 3)_t = (1/9) - (1/9)(1-3v)^t$$

$$P(3, 1)_t = (1/9) - (1/6)(1-v)^t + (1/18)(1-3v)^t$$

$$P(3, 2)_t = (1/9) - (1/9)(1-3v)^t$$

$$P(3, 3)_t = (1/9) + (1/6)(1-v)^t + (1/18)(1-3v)^t$$

(ii) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計検定法の開発

ただ一度の突然変異によって生じた疾患感受性変異が集団中に広がったと仮定する。突然変異が生じた染色体の感受性変異近傍のマイクロサテライトマーカーのリピート数が i だったとすると、感受性変異と当該アレル i との間に強い連鎖不平衡が存在すると期待される。マイクロサテライトマーカーで突然変異が起こったとしても、もとのアレルとリピート数が近いアレル ($i+1$ や $i-1$ など) に変異する可能性が高いことから、感受性変異を有する染色体のマイクロサテライトマーカーアレルは互いにリピート数が近いことが期待される。このことは、感受性変異を有する個体が多い case 集団では、その感受性変異の近傍のマイクロサテライトマーカーでアレル

のレポート数のペアワイズな差が、control 集団でのそれに比較して小さくなることを示唆する。いま、case を $N1$ 個体、control を $N2$ 個体サンプルしたとする。二つのアليل i と j のレポート数の差を評価する関数を $D_{i,j} = |i-j|e$ とする。なお、 e には 1、2、0.5 などの値を用いる。case 群で観察された $2N1$ 個の遺伝子の全ペアワイズに対して $D_{i,j}$ を計算しそれらを平均したものを $x1$ 、control でのそれを $x2$ とする。これらをサンプルサイズに応じて重み付けし、足し合わせた検定統計量を $d0$ とすると、

$$d0 = x1N1/(N1+N2) + x2N2/(N1+N2)$$

となる。この値を観察値とし、 $N1 + N2$ 個体からなるサンプルを $N1$ 個体と $N2$ 個体とに分ける全組合せ（観察値を含む）に対して d の値を計算する。組合せ総数は $(N1 + N2)! / (N1!N2!)$ である。その中で、観察値と同じかそれ以下の d を示した回数を y とすると、並べ替え検定に基づく正確な P 値は $(N1!N2!)y / (N1 + N2)!$ で与えられる。組合せ総数が多い場合には、モンテカルロ再抽出法やピアソンのタイプ III 近似によって P 値を求めることも可能である。

(2) 連鎖不平衡ブロックの生成に高頻度組換え点が与える影響の理論的検討

次のようなコンピュータシミュレーションプログラムを構築した。集団中には $2N$ 本の染色体があると仮定する（集団は任意交配を行い、そのサイズは一定）。染色体上には等間隔（遺伝距離）に連鎖して存在する n 個の二対立遺伝子サイトを仮定する。各サイトにおける突然変異率を u とし、多型的なサイトでは突然変異は起こらないと仮定する。この仮定により、シミュレーション中の全ての変異は単一起源であることが保証される。隣り合う多型間の組換え率を r とし、交差は 2 箇所以上で起きない（二重組換え体は生じない）と仮定する。 $n/3$ 番目のサイトと $n/3+1$ 番目のサイトの間が高頻度組換え点と仮定し、そこでの組換え率は rh とする。 n 個のサイトを図 3 のように、それぞれ $n/3$ 個のサイトを有す

る 3 つの領域に分け、各領域内、および領域間でのペアワイズな連鎖不平衡係数の平均値を計算した。シミュレーションでは各サイトの対立遺伝子頻度は機会的遺伝子浮動によって変化するが、連鎖不平衡係数 (r^2) の計算には、マイナー対立頻度が 15% 以上であるサイトのみを利用した。

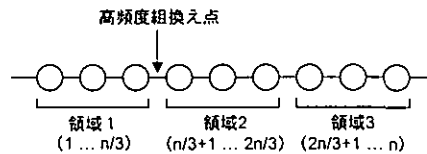


図3 連鎖不平衡係数を計算するためのモデル
 染色体上には n 個の二対立遺伝子サイトがあり、隣り合うサイト間の遺伝距離は等しいと仮定した。ただし、 $n/3$ 番目と $n/3+1$ 番目のサイトとの間に高頻度組換え点を仮定した（したがって、この間の遺伝距離は相対的に長いことになる）。 $n/3$ 個のサイトをそれぞれ含む 3 つの領域にわけ、領域内および領域間でペアワイズ連鎖不平衡係数を計算した。

[図 3]

C. 研究結果

(1) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発

(i) マイクロサテライトマーカーでの突然変異モデルの検討と 2 染色体間で期待されるレポート数の差

K 対立遺伝子単一ステップ突然変異モデルを仮定して行ったシミュレーション結果の一部を図 4 に示す。

[図 4]

これらの分布は、平衡状態に達してから、ある一定期間の平均頻度分布である。ここでは、 $K = 20$ を仮定し、 $4Nv = 40$ （上の図）と $4Nv = 4$ （下の図）とパラメタを設定した。数回のシミュレーションしか行っていないが、 $4Nv = 40$ を仮定した場合（突然変異率が高い場合）には多数のアليلが観察され、 $4Nv = 4$ を仮定した場合（突然変異率が低い場合）には少数のアिलが観察される傾向は一貫していた。

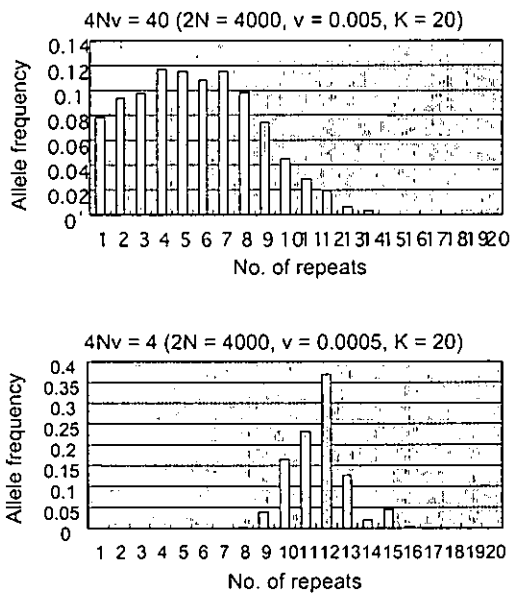


図4 K対立遺伝子単一ステップ突然変異モデルを仮定したコンピュータシミュレーションでの平衡状態におけるアレル頻度分布
上の図は $4Nv = 40$ を仮定した場合、下の図は $4Nv = 4$ を仮定した場合の平衡状態に達してから一定期間での平均頻度分布を示す。

(ii) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計検定法の開発
一般的な並べ替え検定の枠組みの中で、新規統計検定法を開発することが出来た。

(2) 連鎖不平衡ブロックの生成に高頻度組換え点を与える影響の理論的検討

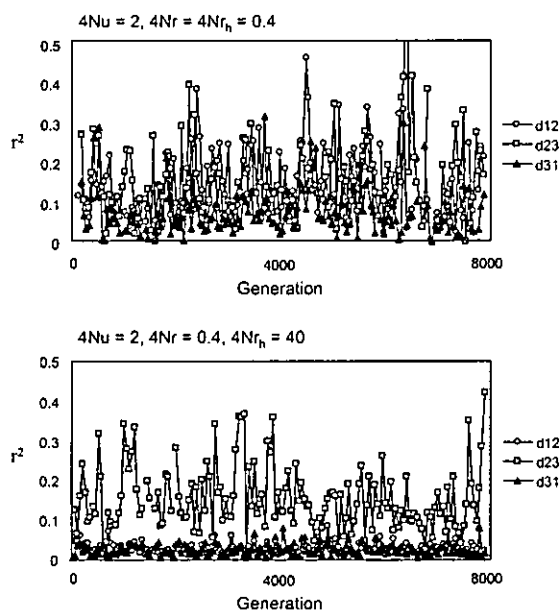


図5 領域間で観察された連鎖不平衡係数 (r^2)
上の図は $4Nu = 2, 4Nr = 4Nr_h = 0.4$ を仮定した場合、下の図は $4Nu = 2, 4Nr = 0.4, 4Nr_h = 40$ を仮定した場合の結果を示す。 d_{ij} は、図3の領域のサイトと領域のサイトとの間で観察された連鎖不平衡係数 r^2 の平均値。

シミュレーションでは、 $n = 60$ (高頻度組換え点は20番目と21番目のサイトの間に存在)、 $4Nu = 2$ 、 $4Nr = 0.4$ とパラメタを設定した。なお、シミュレーションは全サイトが単型的な状態から開始した。8000世代経過時までの、図3で定義する領域をまたいで存在する2つのサイト間での r^2 の平均値を図5に示す。 r^2 の計算には、マイナーアレル頻度が15%以上のサイトのみを使用した。

[図5]

上の図は $4Nr_h = 0.4$ を仮定した(高頻度組換え点を仮定しない)場合の結果であり、下の図は $4Nr_h = 40$ を仮定した(高頻度組換え点で100倍の組換え率を仮定した)場合の結果である。領域1と領域2との間に高頻度組換え点が存在すると、高頻度組換え点をまたいで観察される連鎖不平衡係数 r^2 は明らかに小さくなった。すなわち、領域1と領域2との間で観察される値は、領域2と領域3との間で観察される値よりも常に小さかった。なお、このシミュレーションプログラムの一部は、平成13年度および平成14年度厚生科学研究費補助金(ヒトゲノム・再生医療等研究事業)によって開発されたものであり、それぞれの報告書内に記載してある。

D. 考案

(1) マイクロサテライトマーカーを用いた連鎖不平衡マッピングにおける新規統計解析手法の開発
(i) マイクロサテライトマーカーでの突然変異モデルの検討と2染色体間で期待されるリピート数の差

K対立遺伝子単一ステップ突然変異モデルを構築できたことは、マイクロサテライトマーカーの理論的な研究を行う上で極めて有用である。今後は、実データとシミュレーション結果とを比較することで、現実のデータをうまく説明できるようなパラメタの組合せを求める必要があろう。