

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2008-125390

(P2008-125390A)

(43) 公開日 平成20年6月5日(2008.6.5)

| (51) Int.Cl. | F I | テーマコード (参考) |
|--------------------------------|-----------------|-------------|
| C 1 2 N 15/09 (2006.01) | C 1 2 N 15/00 A | 4 B O 2 4 |
| C 1 2 Q 1/68 (2006.01) | C 1 2 Q 1/68 Z | 4 B O 6 3 |

審査請求 未請求 請求項の数 3 O L (全 20 頁)

| | | | |
|-----------|------------------------------|----------|---|
| (21) 出願番号 | 特願2006-311855 (P2006-311855) | (71) 出願人 | 504013775 |
| (22) 出願日 | 平成18年11月17日 (2006.11.17) | | 学校法人 埼玉医科大学 埼玉県入間郡毛呂山町毛呂本郷38 |
| | | (74) 代理人 | 100088155 弁理士 長谷川 芳樹 |
| | | (74) 代理人 | 100092657 弁理士 寺崎 史朗 |
| | | (74) 代理人 | 100126653 弁理士 木元 克輔 |
| | | (72) 発明者 | 神田 将和 埼玉県入間郡毛呂山町毛呂本郷38 学校 法人埼玉医科大学内 |
| | | (72) 発明者 | 岩佐 泰靖 埼玉県入間郡毛呂山町毛呂本郷38 学校 法人埼玉医科大学内 |

最終頁に続く

(54) 【発明の名称】 逆ホモ接合マッピング法による同祖領域の抽出方法および遺伝子スクリーニング方法

(57) 【要約】

【課題】多型マーカーを用いて、遺伝子疾患の原因遺伝子の候補領域を効率よく探索する方法を提供すること。

【解決手段】 対象とコントロール対象とに由来するDNA間の多型マーカーを比較し、対象とコントロール対象とに共通する非候補領域を抽出する、非候補領域抽出ステップと、2以上の対象に由来するDNA間の多型マーカーを比較し、対象間に共通する候補領域を抽出する、候補領域抽出ステップと、上記候補領域から上記非候補領域を除いて同祖領域を抽出する、同祖領域抽出ステップと、を含む、同祖領域の抽出方法。

【選択図】なし

【特許請求の範囲】

【請求項 1】

対象とコントロール対象とに由来するDNA間の多型マーカ-を比較し、対象とコントロール対象とに共通する非候補領域を抽出する、非候補領域抽出ステップと、

2以上の対象に由来するDNA間の多型マーカ-を比較し、対象間に共通する候補領域を抽出する、候補領域抽出ステップと、

前記候補領域から前記非候補領域を除いて同祖領域を抽出する、同祖領域抽出ステップと、

を含む、同祖領域の抽出方法。

【請求項 2】

対象とコントロール対象とに由来するDNA間の多型マーカ-を比較し、対象とコントロール対象とに共通する非候補領域を抽出する、非候補領域抽出ステップと、

2以上の対象に由来するDNA間の多型マーカ-を比較し、対象間に共通する候補領域を抽出する、候補領域抽出ステップと、

前記候補領域から前記非候補領域を除いて同祖領域を抽出する、同祖領域抽出ステップと、

前記同祖領域のDNA配列情報および遺伝子情報を基に、対象とコントロール対象との遺伝子を比較し、対象の原因遺伝子を検索するステップと、

を含む、遺伝子スクリーニング方法。

【請求項 3】

前記対象が、特定の遺伝子疾患を有する患者である、前記請求項 1 または 2 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、多型マーカ-を用いて、遺伝子疾患の原因遺伝子の候補領域を効率よく探索する方法に関する。

【背景技術】

【0002】

疾患の発症原因または関連する遺伝子を同定する疾患遺伝子解析研究は、疾患の治療法の開発に極めて重要であり、ポストゲノム研究の重要な課題の一つである。疾患感受性遺伝子の同定方法としては、疾患感受性遺伝子の領域を特定する連鎖解析や罹患同胞対解析、ホモ接合マッピング法などの解析方法が開発されている。

【0003】

連鎖解析とは、表現型に関連する遺伝子座と染色体上のマーカ-遺伝子座との連鎖がどの程度あるのかを手がかりにして、染色体上の原因遺伝子の存在領域を絞り込んでいく方法である。また、罹患同胞対解析とは、同じ疾患を有している兄弟姉妹を比較して、染色体上の原因遺伝子の存在領域を絞り込んでいく方法である。

【0004】

連鎖解析による劣性疾患遺伝子の同定法としては、多型を用いたホモ接合体マッピング (Homozygosity Mapping) 法が知られている。ホモ接合体マッピング法は、ある特定の1遺伝子の変異により起きる劣性かつ単一遺伝子疾患の原因遺伝子を取る方法として1987年にエリック・ランダー博士らにより提唱され、実際に成果を挙げている方法である (非特許文献1)。これらの解析において、多型マーカ-が用いられている (非特許文献2)。ホモ接合体マッピング法では、多型としてSNPの一つである制限酵素切断断片長多型 (Restriction Fragment Length Polymorphism、RFLP) を用いたもの (非特許文献1) や、マイクロサテライト多型を用いたもの (非特許文献3) がある。

【0005】

また、疾患と対立遺伝子 (もしくはマーカ-遺伝子座) が関連しているかを調べる方法

10

20

30

40

50

として、関連解析が知られている。関連解析は、コントロール集団と疾患を有している集団で特定の多型マーカーの出現頻度を比較することによって、疾患と特定のマーカーとの関連を生じる連鎖不平衡を見出し、疾患と密接に連鎖した遺伝子座を絞り込む方法である。

【0006】

このように、おおよその染色体位置に関する情報を得て、その情報に基づいて原因遺伝子を同定するポジショナルクローニングによって、疾患の原因遺伝子、例えば、X連鎖性慢性肉芽腫症の原因遺伝子（非特許文献4）などが同定されている。また、上記連鎖解析及び関連解析によって同定された疾患感受性遺伝子としては、ヒト2型糖尿病の原因遺伝子が知られている（特許文献1）。

【特許文献1】特開2002-339901

【非特許文献1】Lander, E. S. et al. Science 236:1567, 1987.

【非特許文献2】中村祐輔著「ゲノム医学からゲノム医療へ」、羊土社、2005年

【非特許文献3】Kobayashi, K. et al. Nature Genetics 22:159, 1997.

【非特許文献4】Royer-Pokora et al. Nature 322:32-38, 1985.

【非特許文献5】Dietz H. et al. Nature 352:337-339, 1991.

【非特許文献6】Sakai L. et al. Cell Biol. 103:2499-2509, 1986.

【発明の開示】

【発明が解決しようとする課題】

【0007】

しかしながら、連鎖解析や罹患同胞対解析は家系解析を基本とするものであり、その遺伝子解析を行う前段階となる対象を得る過程において困難性を有している。同一家系内の患者は、共通祖先に由来する大きな染色体断片を共有するため、上記解析により得られる候補領域は広い上に、原因遺伝子を含まない領域をも多く共有している可能性がある。また、ホモ接合体マッピング法も、おおよかに候補領域を抽出するにとどまり、数十以上の候補遺伝子が存在する候補領域内から真の原因遺伝子を探し出す工程が必要であるため、膨大な費用と時間がかかるという課題があった。また、ポジショナルクローニングの困難さは、染色体上の候補領域の広さによって決まるため、効率的に候補領域を絞り込む方法の開発が望まれていた。

【0008】

また、近年ヒトゲノムにおけるDNA多型のパターンが解析され、ハプロタイプ地図が完成したことにより、ハプロタイプと疾患感受性遺伝子の関係が研究されている。しかし、ヒトにおいて染色体は2組存在し、一人の遺伝子座（多型）がすべてわかったとしてもいずれのハプロタイプに由来するかを特定することはできない。そのため、関連解析等により疾患感受性遺伝子の推定を行うが、関連解析は、対立遺伝子間や表現型間の関係を統計学的に述べているにすぎず、関連を生じる要因は連鎖不平衡以外にも存在するため、偽陽性が多く再検定を行わなければならないという問題点が存在する。また、症例群とコントロール群との不一致によっても関連が認められる可能性がある。したがって、偽陽性の可能性、再現性の低さやコントロール群の選定における不確実性が欠点となっており、有意な解析結果を得るためには、膨大な費用と時間がかかるという課題があった。

【0009】

本発明は、上記従来技術の有する課題に鑑みてなされたものであり、多型マーカーを用いて、遺伝子疾患の原因遺伝子の候補領域を効率よく探索する方法を提供することを目的とする。

【課題を解決するための手段】

10

20

30

40

50

【 0 0 1 0 】

ある集団における疾患の発症原因は、同一遺伝子の突然変異が異なる個人に同時に起こり発症するのではなく、一人の先祖の単一遺伝子の変異であることが多いと考えられている。すなわち、同一疾患を有する患者は、同じ祖先に由来する疾患感受性遺伝子の存在する染色体領域を共有し、このような領域内の遺伝子変異、SNP、マイクロサテライト多型などのすべての塩基配列が、同一疾患を有する患者において保存されていることになる。そのため、ある集団の異なる個人間において、同一の塩基配列を有している領域内に疾患感受性遺伝子が存在していることになる。

【 0 0 1 1 】

本発明者らは、鋭意研究した結果、遺伝子疾患の原因遺伝子の候補領域を効率よく探索するために、患者と健常者との間で共有される染色体領域に着目する方法を見出した。これにより、同じ祖先に由来し、疾患感受性遺伝子と連鎖または関連することのない染色体領域（非候補領域）を検索することができる。また、患者と患者との間で共有され、疾患感受性遺伝子の存在が疑われる候補領域から上記非候補領域を差し引くことにより、原因遺伝子の候補領域を効率よく絞り込むことが可能であることを見出した。かかる知見に基づき、本発明者は、逆ホモ接合マッピング法（Reverse Homozygosity Mapping法）、すなわち、遺伝子疾患の原因遺伝子の候補領域を効率よく探索する方法を提供することを可能とし、本発明を完成した。

10

【 0 0 1 2 】

(1) すなわち、本発明は、対象とコントロール対象とに由来するDNA間の多型マーカーを比較し、対象とコントロール対象とに共通する非候補領域を抽出する、非候補領域抽出ステップと、2以上の対象に由来するDNA間の多型マーカーを比較し、対象間に共通する候補領域を抽出する、候補領域抽出ステップと、上記候補領域から上記非候補領域を除いて同祖領域を抽出する、同祖領域抽出ステップと、を含む、同祖領域の抽出方法を提供する。

20

【 0 0 1 3 】

上記方法における対象は、特定の遺伝子疾患を有する患者であることが好ましい。本発明の同祖領域の抽出方法によれば、遺伝子疾患の原因遺伝子の候補領域を効率よく探索することができる。特に、単一遺伝子疾患の原因遺伝子を含む候補領域、中でも、浸透率の高い単一遺伝子疾患の原因遺伝子を含む候補領域を効率よく探索することができる。例えば、家系解析によって得られた候補領域が広範であっても、本発明の方法を行うことにより、原因遺伝子を含まない領域を除くことができるため、候補領域を効率よく絞り込むことができる。

30

【 0 0 1 4 】

また本発明は、ある特定の疾患や遺伝子変異を含まない領域または含む可能性の低い領域を除くことによって、候補領域を絞り込むという考えに基づいている。そのため、関連解析とは異なり、連鎖不平衡以外の要因によって得られた偽陽性の領域が加えられることなく、関連解析におけるコントロール選定の不確実性を排除することができ、コントロールの選定が容易となる。

【 0 0 1 5 】

(2) 本発明は、また、対象とコントロール対象とに由来するDNA間の多型マーカーを比較し、対象とコントロール対象とに共通する非候補領域を抽出する、非候補領域抽出ステップと、2以上の対象に由来するDNA間の多型マーカーを比較し、対象間に共通する候補領域を抽出する、候補領域抽出ステップと、上記候補領域から上記非候補領域を除いて同祖領域を抽出する、同祖領域抽出ステップと、上記同祖領域のDNA配列情報および遺伝子情報を基に、対象とコントロール対象との遺伝子を比較し、対象の原因遺伝子を検索するステップと、を含む、遺伝子スクリーニング方法を提供する。

40

【 0 0 1 6 】

この構成により、原因遺伝子を含まない領域を除くことができるため、候補領域を効率よく絞り込むことができる。つまり、候補領域に存在する候補遺伝子の数を絞り込むこと

50

ができるため、その中から真の原因遺伝子を同定するための費用、労力および時間を低減することができる。したがって、従来と比較して、原因遺伝子の同定が容易となる。また、上記方法における対象は、特定の遺伝子疾患を有する患者であることが好ましい。

【発明の効果】

【0017】

本発明は、家系内および同一祖先に由来する集団内での解析に応用することができ、さらに、劣性遺伝疾患のみならず、優性遺伝疾患の原因遺伝子の探索にも適用することができる。また、従来の疾患感受性遺伝子の探索方法に限らず、2対象間の染色体領域を比較するあらゆる方法に適用することができる点において有用である。また、本発明者らが発明した *homozygosity fingerprint* 法および/またはホモ接合ハプロタイプ法と合わせて解析を行うことにより、劣性遺伝の疾患感受性遺伝子の同定の精度を向上させ、少数の対象数で判定することができる。

10

【0018】

また本発明は、家系内の患者に限らず、ある特定の遺伝子疾患を有する患者のデータを利用することができ、コントロールの選定も容易であることから、従来よりも広範囲な対象およびコントロール対象を利用することができる。つまり、他の解析によって得られたデータや公共のデータベースなどの公知のデータを活用することができる。したがって、原因遺伝子の検索のための費用、労力および時間を低減することができる点において優れている。

20

【0019】

さらに、本発明を植物や動物に対し行うことにより、疾患に関してはヒトと同様に原因遺伝子の探索が可能となるとともに、有用な機能を果たす遺伝子や、有用な形質を発現するような遺伝子の探索が可能であり、品種改良等の分野においても利用することができる。

【発明を実施するための最良の形態】

【0020】

以下に、各発明を実施するための最良の実施形態について詳細に説明する。本発明は、これらの実施形態に何ら限定されるものではなく、その要旨を逸脱しない範囲において、種々なる態様で実施し得る。

【0021】

まず、本実施形態の基本となる原因遺伝子の遺伝様式について、図1を用いて説明する。この図は、ある一族の家系図を示している。突然変異等により、Aが疾患の原因となる異常疾患原因遺伝子(黒色)を有しているものとする。この場合に、その子供であるB、CはAから1本の染色体を受け継ぐが、減数分裂時の交叉によりAの原因遺伝子を有している染色体との共通部分(グレー部分)は短くなる。この共通部分が「同祖領域」である。ここで、異常疾患遺伝子が劣性遺伝疾患である場合には、Fのように共通の祖先Aに由来する原因遺伝子がホモ接合となった場合に、疾患を発症する。また、優性遺伝疾患である場合には、原因遺伝子を有しているB、C、D、E、Fの全員に疾患発症の可能性がある。しかし、劣性遺伝、優性遺伝のどちらの場合においても、原因遺伝子とその近傍の領域が受け継がれており、この領域においてはAからFの全員が同じハプロタイプを有していることになる。

30

40

【0022】

一方で、同祖領域の中には、原因遺伝子および原因遺伝子座を含む領域とは異なる領域(ハッチング部分)も存在する。したがって、従来の解析方法によって得られた候補領域には、原因遺伝子とは関係のない領域(非候補領域)も含まれる。この事実に基づき、本発明は完成されたものである。

【0023】

したがって、前述の事実から、対象およびコントロール対象間で同じハプロタイプを有する領域を特定することができれば、原因遺伝子とは関係のない非候補領域を特定することができる。本発明は、2対象間に共通する染色体領域を抽出する方法であればあらゆる

50

方法を応用することができ、対象およびコントロール対象間で共有されている非候補領域を抽出し、候補領域を効率的に絞り込むことができる。

【0024】

次に、本実施形態の概要について、図2および図3を用いて説明する。

本実施形態の同祖領域の抽出方法は、対象とコントロール対象とに由来するDNA間の多型マーカーを比較し、対象とコントロール対象とに共通する非候補領域を抽出する、非候補領域抽出ステップ(ステップ1)と、2以上の対象に由来するDNA間の多型マーカーを比較し、対象間に共通する候補領域を抽出する、候補領域抽出ステップ(ステップ2)と、候補領域から非候補領域を除いて同祖領域を抽出する、同祖領域抽出ステップ(ステップ3)と、を含む。

10

【0025】

はじめに、候補領域抽出ステップ(ステップ2)を説明する。候補領域抽出ステップ(ステップ2)は、従来の疾患遺伝子解析に相当する。つまり、患者間に共通する同祖領域を抽出し、この領域を疾患の原因遺伝子を含む候補領域と見なす方法である。この場合、対象は特定の遺伝子疾患を有する患者である。以下に、図2を用いて説明する。図2は、従来のホモ接合マッピング法を模式的に示した図である。劣性遺伝子疾患の原因遺伝子を同定するための方法であるホモ接合マッピング法の原理は、原因遺伝子を含む染色体上の領域が、単一患者内でホモ接合しており、さらに、患者間で共通に保有されていることを前提としている。まず、それぞれの患者において染色体上に点在するマーカーを比較し、二本の染色体間でホモ接合している領域を抽出する(図2、黒色部分)。次に、複数の患者間でマーカーを比較し、患者間で共通する領域を抽出する(破線で囲んだ部分)。以上が、ステップ2である。さらに従来の疾患遺伝子解析では、得られた領域を疾患の原因遺伝子を含む候補領域と見なし、染色体の位置情報を基に、原因遺伝子を同定する。

20

【0026】

本実施形態においては、候補領域抽出ステップ(ステップ2)に加えて、非候補領域抽出ステップ(ステップ1)を行う。ステップ1は、患者と健常者とに共通する領域、つまり、原因遺伝子とは関係のない領域(非候補領域)を抽出するステップである。抽出方法は、比較対象を患者および健常者にする以外は、ステップ1と同様の方法で行うことができる。

【0027】

図3は、本実施形態の同祖領域の抽出方法における同祖領域抽出ステップ(ステップ3)を模式的に示した図である。ステップ3では、候補領域抽出ステップ(ステップ2)で得られた患者間で共通する領域(図3、黒色部分)から、非候補領域抽出ステップ(ステップ1)で得られた患者-健常者間で共通する領域(図3、グレー部分)を差し引くことによって、疾患の原因遺伝子に関する同祖領域を抽出し、候補領域を絞り込む。

30

【0028】

また、本実施形態の遺伝子スクリーニング方法は、対象とコントロール対象とに由来するDNA間の多型マーカーを比較し、対象とコントロール対象とに共通する非候補領域を抽出する、非候補領域抽出ステップ(ステップ1)と、2以上の対象に由来するDNA間の多型マーカーを比較し、対象間に共通する候補領域を抽出する、候補領域抽出ステップ(ステップ2)と、候補領域から非候補領域を除いて同祖領域を抽出する、同祖領域抽出ステップ(ステップ3)と、同祖領域のDNA配列情報および遺伝子情報を基に、対象とコントロール対象との遺伝子を比較し、対象の原因遺伝子を検索するステップ(ステップ4)と、を含む。

40

【0029】

本実施形態の遺伝子スクリーニング方法におけるステップ1~3は、本実施形態の同祖領域の抽出方法におけるステップ1~3と同様に行えばよい。また、本実施形態の遺伝子スクリーニング方法における原因遺伝子を検索するステップ(ステップ4)は、まず、ステップ1~3によって得られた同祖領域のDNA配列情報および遺伝子情報を、公知のデータベースや実験的手法により得る。次に、領域内のDNA配列情報および遺伝子情報が

50

ら、候補遺伝子を抽出し、それぞれの候補遺伝子の情報から、有力な候補遺伝子を選択する。候補遺伝子の情報としては、候補遺伝子が既知遺伝子の場合、遺伝子の機能などの公知の解析結果、未知遺伝子の場合には既知のタンパク質のドメインとの比較などによる機能の予測などが利用可能である。疾患の様態、発症組織、時期などの情報から、疾患に関与することが期待できる有力な候補遺伝子を選択すればよい。また併せて、各候補遺伝子の塩基配列を、対象とコントロール対象との間で比較し、遺伝子疾患に関与する可能性のある塩基置換を検索する。そして、有力な候補遺伝子を選択したら、その候補遺伝子の機能を、分子遺伝学的手法、例えば、実験動物を用いた遺伝子改変実験などによって調べる。このように、それぞれの候補領域および候補遺伝子に対して、原因遺伝子を同定するまでステップ4を繰り返す。本実施形態の方法によれば、従来法よりも、候補領域および候補遺伝子を絞り込めるため、従来法よりも容易に原因遺伝子を同定することができる。

10

【0030】

上述のように本発明は、ある特定の疾患や遺伝子変異を含まない領域または含む可能性の低い領域を除くことによって、候補領域を絞り込むという考えに基づいている。患者と患者との間で共有される候補領域から原因遺伝子とは関係のない領域（非候補領域）を差し引くことにより、原因遺伝子の候補領域を効率よく絞り込むことが可能である。そこで本発明者は、この多型マーカーを用いた遺伝子疾患の原因遺伝子の検索方法および遺伝子スクリーニング方法を、「逆ホモ接合マッピング法（Reverse Homozygosity Mapping法）」と名づけた。

【0031】

20

さらに、本実施形態についてより詳細に説明する。

本発明における「対象」とは、個体の染色体上に存在する特定の遺伝子変異によって、特定の特徴的な形質を有する個体を指し、また特に、特定の遺伝子疾患を有する個体を指す。ここで、対象はヒト、ヒト以外の動物、植物、その他の微生物などであってもよい。また対象は二倍体以上の個体が好ましい。多くの場合、対象はヒトである。ヒトである場合には、1つの拡張した家系をなす集団、つまり、祖先集団を共有する集団（例えば、日本人）から対象を選定することが好ましい。祖先集団を共有する集団から対象を選定することが好ましい理由は、共通の祖先に由来する同祖領域を有する確率が高いからである。それに対して、民族間の交わりが多く近交係数の小さな集団（例えばアメリカのような集団）では、交叉によって同祖領域が短くなっているため、同祖領域の判定が難しい。また、多型マーカーを構成している塩基を他検体と比較した場合に、偶然に多型マーカーが一致しているのか、あるいは同祖であるために一致しているのかの判断が困難となるためである。また、DNAは、染色体を構成するゲノムDNAであり、血液、唾液、組織、細胞などから採取することができる。ゲノムDNAの調製方法は、多型のタイピング方法に合わせた方法で行えば、特に限定されない。また、多型のタイピングには、後述のマーカー多型のデータベースなどの公知のデータを活用することができる。

30

【0032】

「コントロール対象」とは、解析対象とする単一遺伝子変異を有しない個体を指す。ヒトでは、コントロール対象は健常者とも言うが、この場合、解析対象とする単一遺伝子疾患の患者でないことが保証されていれば良いと考えられる。コントロール対象は、上記対象と同一の生物種から選定する。また、同一集団から選定することが好ましい。DNAの調製および多型のタイピングについては上述の方法を用いることができる。また、健常者のマーカー多型のデータは、国際HapMap計画が公開しているサイト（<http://www.hapmap.org/>）などから得ることができる。国際HapMap計画は、ヒトの病気や薬に対する反応性に関わる遺伝子を発見するための基盤を整備するプロジェクトである。

40

【0033】

「多型」とは、DNA上の塩基の違いをいい、ある塩基の変化が人口中1%以上の頻度で存在しているものと定義されている。しかし、実際には1%以下の頻度で存在している塩基の変化を多型と呼ぶこともある。本発明においては、塩基に多様性のあるものをすべ

50

て多型とする。「多型マーカー」とは、DNA上の特定の多型を、疾患感受性遺伝子を探索する際にマーカー（標識）として用いたものである。多型マーカーとして、マイクロサテライト多型、VNTR（Variable number of tandem repeat）多型、一塩基多型（Single Nucleotide Polymorphism：SNP）などが解析に用いられている。DNA上には2～数十塩基の繰り返し配列が存在する。その多くは遺伝情報を持たず、機能不明な領域にあり、生物の個体間で違いが生じやすい部分である。この繰り返し部分の回数が個人間で異なっており、多型となっている。この部分の多型のうち数塩基から数十塩基のものを「VNTR多型」、2から4塩基のものを「マイクロサテライト多型」という。また「SNP」とは、DNA中の一塩基の違いによる多型をいう。SNPには、RFLPも含まれる。SNPは、塩基配列中に高頻度に存在しており、ヒトにおいては300塩基あたりに1個程度存在し、染色体全体では300万から1000万個あるといわれている。近年、このSNPの違いを利用して疾患感受性遺伝子の探索が行われている。本発明において、多型マーカーとしてマイクロサテライト多型又はVNTR多型を用いることができる。また、多型の箇所の多さから、本発明においては、多型マーカーとしてSNPを用いることが好ましい。さらに、SNP、マイクロサテライト多型、VNTR多型のいずれか二以上を組合せて用いてもよい。また、多型のデータベースが公開されており、疾患感受性遺伝子の解析に利用されてきている（非特許文献2等参照）。データベースには、上述の国際HapMap計画に加えて、例えば、NCBIが公開しているdbSNPデータベース（<http://www.ncbi.nlm.nih.gov/SNP/index.html>）や、科学技術振興事業団と東京大学医科学研究所が共同で公開しているJSNP（日本人のSNP）データベース（<http://snp.ims.u-tokyo.ac.jp>）などがある。本発明は、これらのデータベースを適用することができる。また、DNAマイクロアレイなどの方法を用いて、SNPなどの多型を検出してもよい。

10

20

30

40

50

【0034】

「ホモ接合」とは、相同染色体の全部又は一部の領域が同一塩基であることをいう。父親由来と母親由来の対立する塩基（対立塩基対）がどちらも同じ塩基であることをいい、ホモ接合である塩基対をホモ接合体という。

【0035】

「疾患感受性遺伝子」とは、遺伝子疾患における原因遺伝子、または疾患に関係する遺伝子のことである。本発明においては、「疾患の原因遺伝子」、または単に「原因遺伝子」とも言う。疾患の原因となる染色体上の変異、つまり塩基置換は、タンパク質をコードするコード領域のみならず、遺伝子の発現を制御する領域などの非コード領域にも存在し得る。したがって、本発明における遺伝子とは、コード領域および非コード領域をも含む広義な意味を包含する。また、「候補遺伝子」とは、原因遺伝子であることが疑われる、または期待できる遺伝子であり、原因遺伝子の検索によって得られた候補領域に存在する遺伝子を指す。なお、本発明を品種改良等の分野において利用するばあいは、原因遺伝子は有用な機能を果たす遺伝子や有用な形質を発現するような遺伝子に置き換えてもよい。

【0036】

「同祖遺伝子」および「同祖領域」とは、それぞれ共通祖先に由来することが明らかである対立遺伝子および染色体上の領域を指す。ここで、領域とは染色体領域とも言い、染色体上の位置を意味する。ある祖先における突然変異等により生じた遺伝子疾患では、疾患の原因遺伝子は同祖領域に含まれる同祖遺伝子として子孫に受け継がれる。本発明における同祖領域の抽出とは、特に、目的とする原因遺伝子に関係する領域を抽出することを指す。

【0037】

「候補領域」とは、2以上の対象間に共通する領域、つまり、2以上の対象間で共通するそれぞれの領域およびその総和の領域を指す。特定の遺伝子疾患を有する患者間の場合、候補領域は同祖領域を多く含むか、同祖領域にほぼ等しいことが期待でき、また、原因遺伝子を含むことが期待できる。なお、候補領域および同祖領域の中には、原因遺伝子お

よび原因遺伝子座を含む領域とは異なる領域も存在すると考えられる。また、本発明における「非候補領域」とは、対象およびコントロール対象間で共通するそれぞれの領域およびその総和の領域を指す。特定の遺伝子疾患を有する患者および健常者間の場合、非候補領域は、原因遺伝子と連鎖または関連することのない染色体領域に相当する。

【0038】

共通領域の抽出方法は、2つの対象間、ならびに、対象およびコントロール対象間のいずれの場合も、以下に述べる2個体間に共通する領域の抽出方法を用いることができる。2個体間に共通する領域の抽出方法は、2個体それぞれから得た多型データを比較することで、2個体ペア間で共通する領域をそれぞれ抽出し、得られたすべての領域を足し合わせて、共通領域とすることによって行う。具体的には、多型のデータの比較および共通領域の抽出は、染色体上に隣り合って配置されている多型が連続かつある一定数(N)以上続くことを条件とする。ここで、多型は、上述のSNP、マイクロサテライト多型、VNTR多型を用いてもよく、いずれか二以上を組合せて用いてもよい。特に、SNPが好ましい。

10

【0039】

定数(N)は、それぞれの多型、設定された多型間の平均距離および扱う対象に応じて、当業者が適切に設定することができる。多型として、染色体上に平均5.8kb間隔で設定されているSNPを用いたときの定数(N)は、非候補領域を抽出する場合には、70が好ましく、70以上がより好ましく、例えば、75, 80, 85, 90または70以上の任意の数値を設定することができる。定数(N)が70未満の場合は、同一祖先に由来しない領域を含む可能性が高くなり、また、短い(170~400kb以下)共通領域が数多く得られるため、効率的な原因遺伝子の検索が困難となる。この場合、定数(N)は70以上のより大きな数値が、得られる共通領域の信頼性が高くなるため好ましいが、定数(N)が大きすぎると、例えば200以上であると、170~1200kb程度の共通領域を非候補領域から除くこととなるため、2個体間に共通する領域の抽出が困難となる。また、N数が多くなるにつれ、偶然に2個体ペアで一致する領域の出現頻度は低くなる。したがって、実際の運用においては、N数が多い領域を優先的に非候補領域と見なすことが好ましい。なお、特定の遺伝子疾患を有する患者間の場合、患者-健常者間に比べて、共通領域は長くなる傾向がある。

20

【実施例】

30

【0040】

以下、実施例に基づいて本発明をより具体的に説明するが、本発明は以下の実施例に制限するものではない。

【0041】

本発明における逆ホモ接合マッピング法による同祖領域の抽出方法および遺伝子スクリーニング方法を、以下に述べる通り、擬似疾患モデルに適用し、本発明の方法が有効に機能することを確認した。

【0042】

<マルファン症候群>

マルファン症候群(Marfan syndrome: OMIM #154700参照)は、骨格、肺、目、心臓や血管といった多くの器官に症状がでる結合組織の疾患である。その75%が遺伝により、25%は自発的な(新たな)突然変異の結果として発症することが知られている。発生頻度は人口の約5000人に一人程度である。マルファン症候群を持つほとんどの患者に大動脈の拡大が認められ、これに伴い、動脈解離が起きて胸の中心、腹部、背中などに激しい痛み等の症状を訴える。大動脈が裂けた場合は、死亡に至ることもあり、緊急手術や薬物治療が必要となる。

40

【0043】

マルファン症候群の原因遺伝子は、連鎖解析および遺伝子クローニング解析から、第15番染色体上に存在するフィブリリン-1遺伝子(fibrillin-1、FBN1)であることが明らかとなった(非特許文献5)。フィブリリン-1は、線維組織の形成に

50

関わる細胞間接着因子であり、分子量およそ350kDaの結合組織タンパク質として単離され、皮膚、肺などの結合組織に広く発現していることが知られている（非特許文献6）。フィブリリン-1遺伝子の変異により、正常なフィブリリンタンパク質が生産されなくなり、結合組織が脆弱となり細胞に弾力性を減少させ大動脈や網膜、硬膜、骨の形成等に異常をもたらす。

【0044】

<対象>

本実施例では、原因遺伝子の知られている遺伝子疾患を擬似疾患モデルとして、本発明の同祖領域の抽出方法を適用した。上述のマルファン症候群を擬似疾患モデルとして扱った。対象は、マルファン症候群の患者9名である。また、コントロール対象は、上述の国際HapMap計画のサイトにて利用可能である健常者（45名）とし、そのデータを本実施例の解析に用いた。

10

【0045】

<多型マーカーの選択>

多型マーカーは、染色体の全範囲に亘って満遍なく配置されているAffimetrix社のGeneChip（登録商標）Human Mapping 500k Array Setを用いて行った。GeneChip Human Mapping 500k Array Setは、テロメアとセントロメアを除く領域を幅広くカバーし、約50万個のSNPを一度に検出することができる。10kb以内に少なくとも1個のSNPが含まれる領域は、全DNAの85%に相当し、SNP間の物理的距離の中央値は2.5kb、平均距離は5.8kbであるため、疾患の原因が分かっていない場合の同祖領域の判定において好ましい。なお、上述の国際HapMap計画のサイトにて利用可能な多型マーカーのデータは、GeneChip Human Mapping 500k Array Setを用いて得られたデータである。

20

【0046】

<候補領域の抽出>

対象であるマルファン症候群の患者9名から得られたDNAサンプルを用いて、ホモ接合マッピング法の類似法により、患者9名に共通する領域を抽出した。具体的には、Human Mapping 500k Array Setを用いて、1サンプルあたり約50万個のSNPデータを得た。そして、各SNPデータを比較して、染色体上に隣り合って配置されているSNPが連続かつある一定数（N）以上続くことを条件として、サンプル間で共通な領域を抽出した。定数（N）には、80を用いた。その結果を、図4に示す。図4は、マルファン症候群の患者9名に共通する染色体上の領域、つまり、候補領域を示している。黒色の領域が、候補領域を示している。既に特定されているマルファン症候群の原因遺伝子（FBN1、クロスハッチング部分）は、本実施例で同定された複数の候補領域の中で、一番広い領域（候補領域A）に存在していた。候補領域Aは、染色体15番の45,660,841番目から49,978,435番目の塩基に相当する領域であった。図5は、候補領域Aに含まれる遺伝子を模式的に示した図である。現在の時点でのこの領域には、31個の遺伝子の存在が遺伝子データベースに登録されている。

30

【0047】

<非候補領域の抽出>

非候補領域の抽出は、以下の手順で行った。まず、それぞれ1名の患者および健常者からなる患者-健常者ペアを設定した。次に、各患者-健常者ペア間のSNPデータを比較して、ペア間の共通領域を抽出した。複数の患者-健常者ペアから得られた共通領域を足し合わせ、非候補領域とした。健常者のSNPデータ（45名分）は、国際HapMap計画から得た。患者-健常者ペア間のSNPデータを、染色体上に隣り合って配置されているSNPが連続かつある一定数（N）以上続くことを条件として比較し、共通領域の抽出を行った。定数（N）として、60~100を用いて検討した。

40

【0048】

<同祖領域の抽出>

50

上述の手順により得られた候補領域 A から非候補領域を差し引いて得られた領域を、図 6 および図 7 に示す。図 6 は、候補領域 A およびその周辺領域のみを示しており、それぞれの定数 (N) を定めた際に得られる非候補領域 (グレー部分) を候補領域 A (黒色部分) から差し引いて得られた同祖領域を示すものである。定数 (N) = 60 の場合は、原因遺伝子を含む形で誤検出している。これは、設定された N 数が小さいことが原因と考えられた。N 数が多くなるにつれ、偶然に患者-健常人ペアで一致する領域の出現頻度は低くなる。このことから、実際の運用においては、N 数が多い領域を優先的に非候補領域と見なすことが好ましい。図 7 は、定数 (N) = 80 の非候補領域 (グレー部分) を候補領域 A (黒色部分) から差し引いて得られた候補領域 A およびその周辺の同祖領域を示すものである。定数 (N) = 80 に設定した場合、非候補領域として、染色体第 15 番の 47, 074, 166 番目から 47, 839, 497 番目まで、および、48, 408, 175 番目から 49, 422, 538 番目の塩基までの 2 箇所の領域が得られた。また、定数 (N) = 100 に設定した場合は、染色体第 15 番の 48, 408, 175 番目から 49, 422, 538 番目の塩基までの非候補領域が得られた。

【0049】

<考察>

以上の結果から、従来の手法で得られた候補領域から非候補領域を差し引くことによって、原因遺伝子に関する同祖領域を絞り込むことができることが確認できた。本実施例においては、候補領域 A に含まれる 31 の候補遺伝子から 18 の候補遺伝子に絞り込むことができることが確認できた。したがって、逆ホモ接合マッピング法が有用であることが証明された。なお、今回の実施例では、優性遺伝形式と分類される Marfan 症候群を扱っているが、本発明は劣性遺伝形式の単一遺伝子疾患にも適用可能であると考えられる。したがって、本発明の同祖領域の抽出方法および遺伝子のスクリーニング方法は、疾患遺伝子の同定において非常に有効な解析方法を提供するものであることが判明した。

【産業上の利用可能性】

【0050】

遺伝子疾患の原因遺伝子を同定する疾患遺伝子解析研究において、本発明の逆ホモ接合マッピング法による同祖領域の抽出方法および遺伝子のスクリーニング方法を適用することによって、遺伝子疾患の原因遺伝子の候補領域を効率よく探索することができる。特に、単一遺伝子疾患の原因遺伝子を含む候補領域、中でも、浸透率の高い単一遺伝子疾患の原因遺伝子を含む候補領域を効率よく探索することができる。本発明は、家系内および同一祖先に由来する集団内での解析に応用することができる。また、劣性遺伝疾患のみならず、優性遺伝疾患の原因遺伝子の探索にも適用することができる。さらに、従来の疾患感受性遺伝子の探索方法に限らず、2 対象間の染色体領域を比較するあらゆる方法に適用することができる点において有用である。

【0051】

また本発明は、従来よりも広範囲な対象およびコントロール対象を利用することができ、他の解析によって得られたデータや公共のデータベースなどの公知のデータを活用することができる。したがって、効率的に原因遺伝子の検索を行うことができる。また、同定された遺伝子は創薬分野においての利用性が非常に高い。

【0052】

さらに、本発明を植物や動物に対し行うことにより、疾患に関してはヒトと同様に原因遺伝子の探索が可能となるとともに、有用な機能を果たす遺伝子や有用な形質を発現するような遺伝子を同定することも可能であるため、品種改良等の分野においても利用可能であり、畜産・農業上の利用性も非常に高い。

【図面の簡単な説明】

【0053】

【図 1】 遺伝子疾患の原因遺伝子の遺伝様式について、ある一族の家系図を模式的に示す図である。

【図 2】 従来のホモ接合マッピング法を模式的に示す図である。

10

20

30

40

50

【図3】本実施形態の同祖領域の抽出方法における同祖領域抽出ステップを模式的に示す図である。

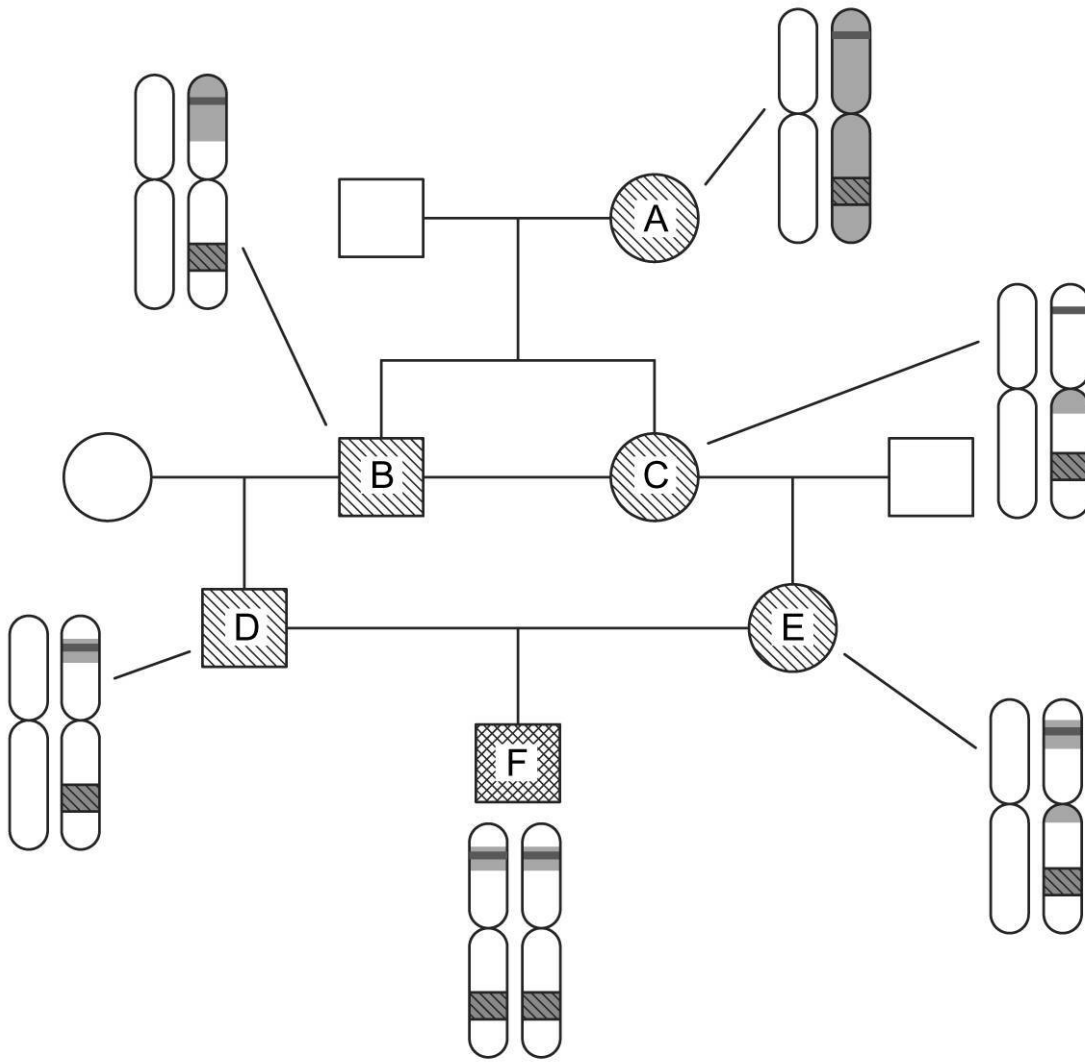
【図4】本実施例におけるマルファン症候群患者に共通する染色体上の領域を模式的に示す図である。



【図5】本実施例におけるマルファン症候群患者に共通する染色体上の領域（候補領域A）および領域内に含まれる遺伝子を模式的に示す図である。

【図6】本実施例におけるマルファン症候群の原因遺伝子に関する同祖領域を模式的に示す図である。定数（N）を、それぞれ（a）60、（b）80および（c）100に設定した。

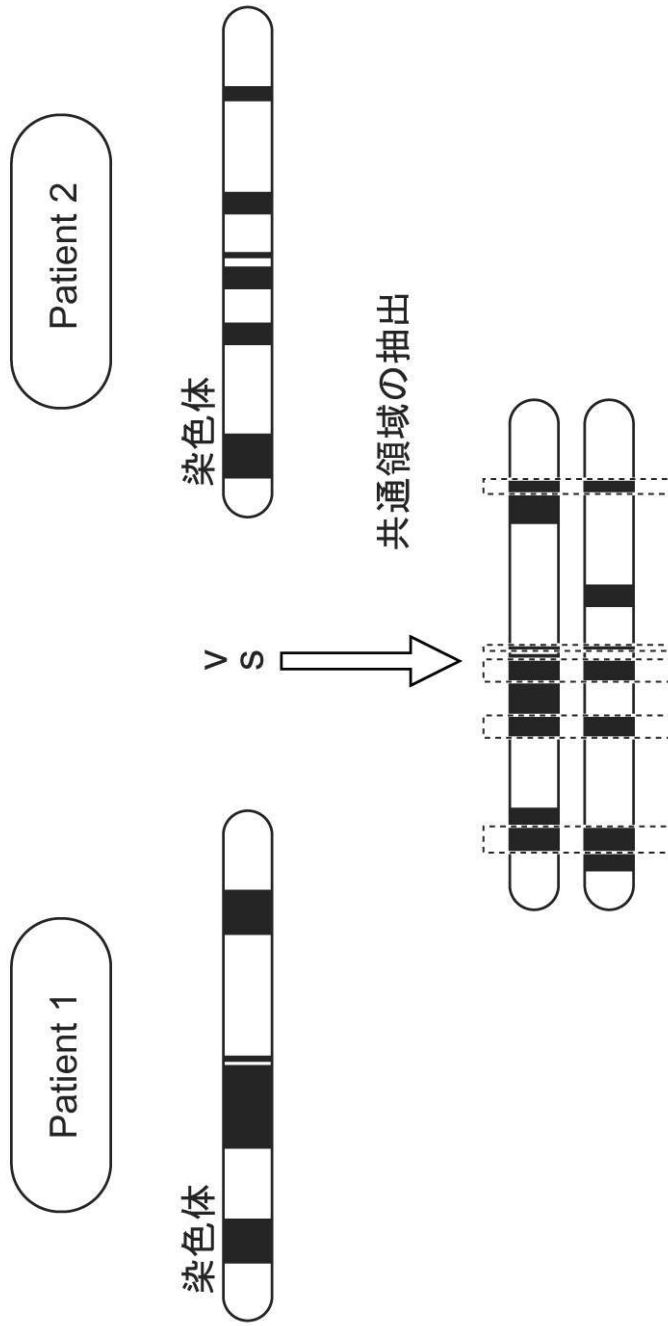
【図7】本実施例におけるマルファン症候群の原因遺伝子に関する同祖領域および領域内に含まれる遺伝子を模式的に示す図である。

【 図 1 】

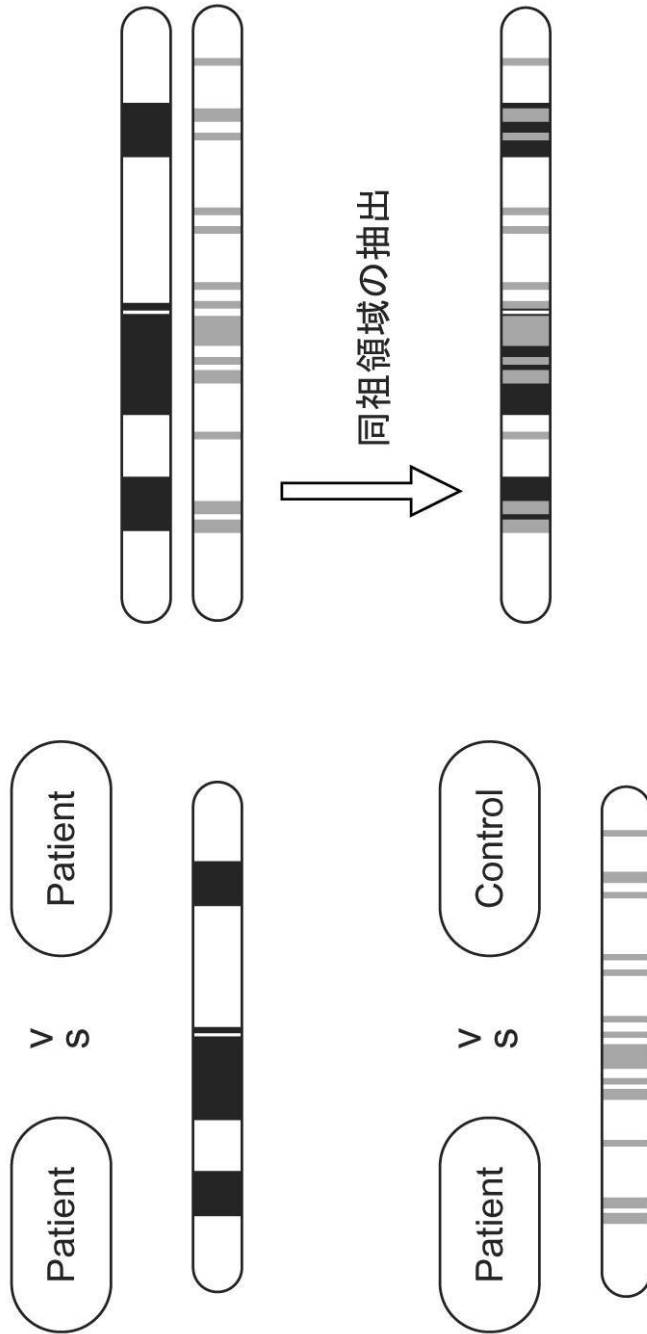


| | |
|---|---------------------------|
|  | 異常疾患原因遺伝子について ヘテロ接合の女性 |
|  | 異常疾患原因遺伝子について ホモ接合の男性 |

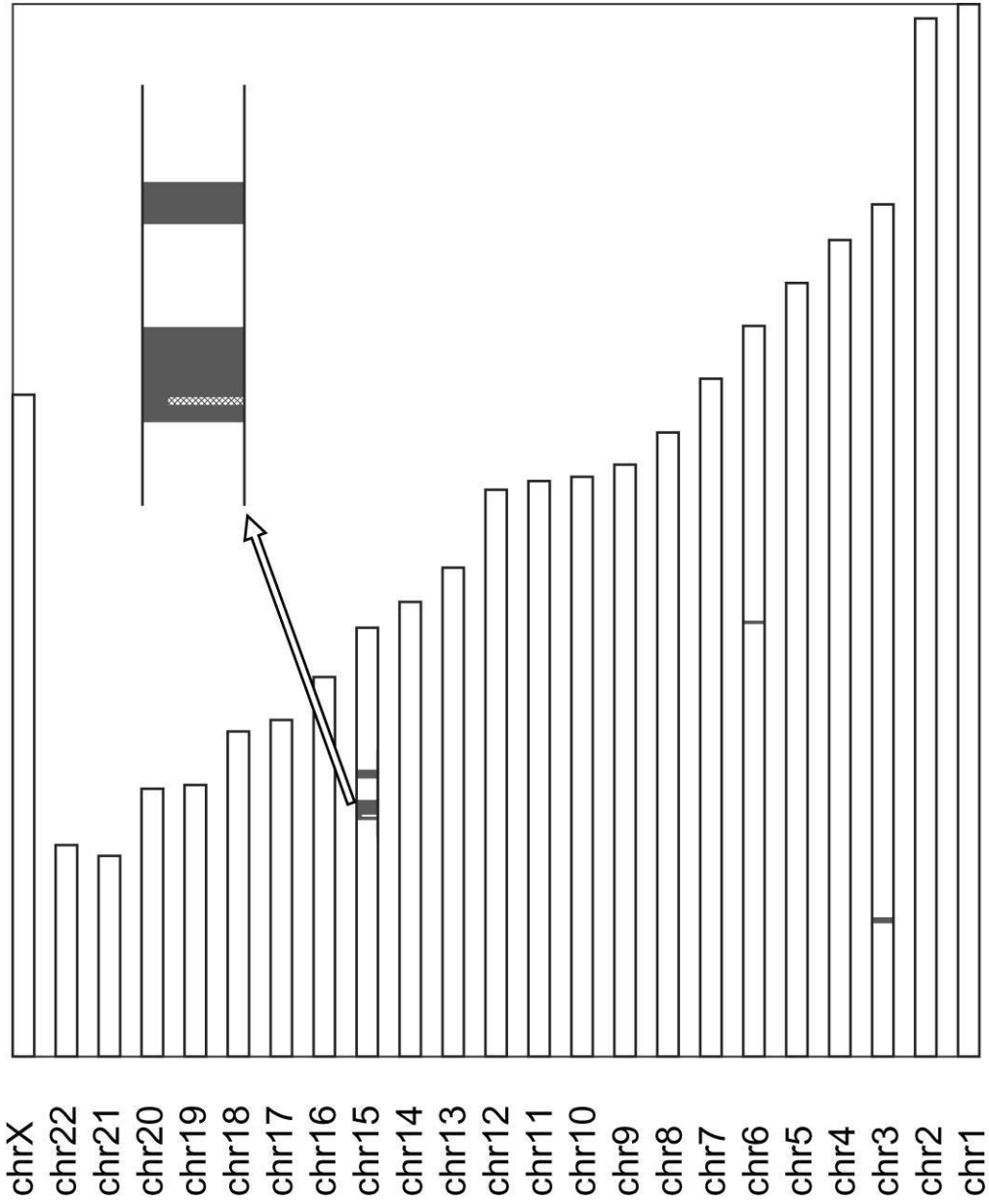
【 図 2 】



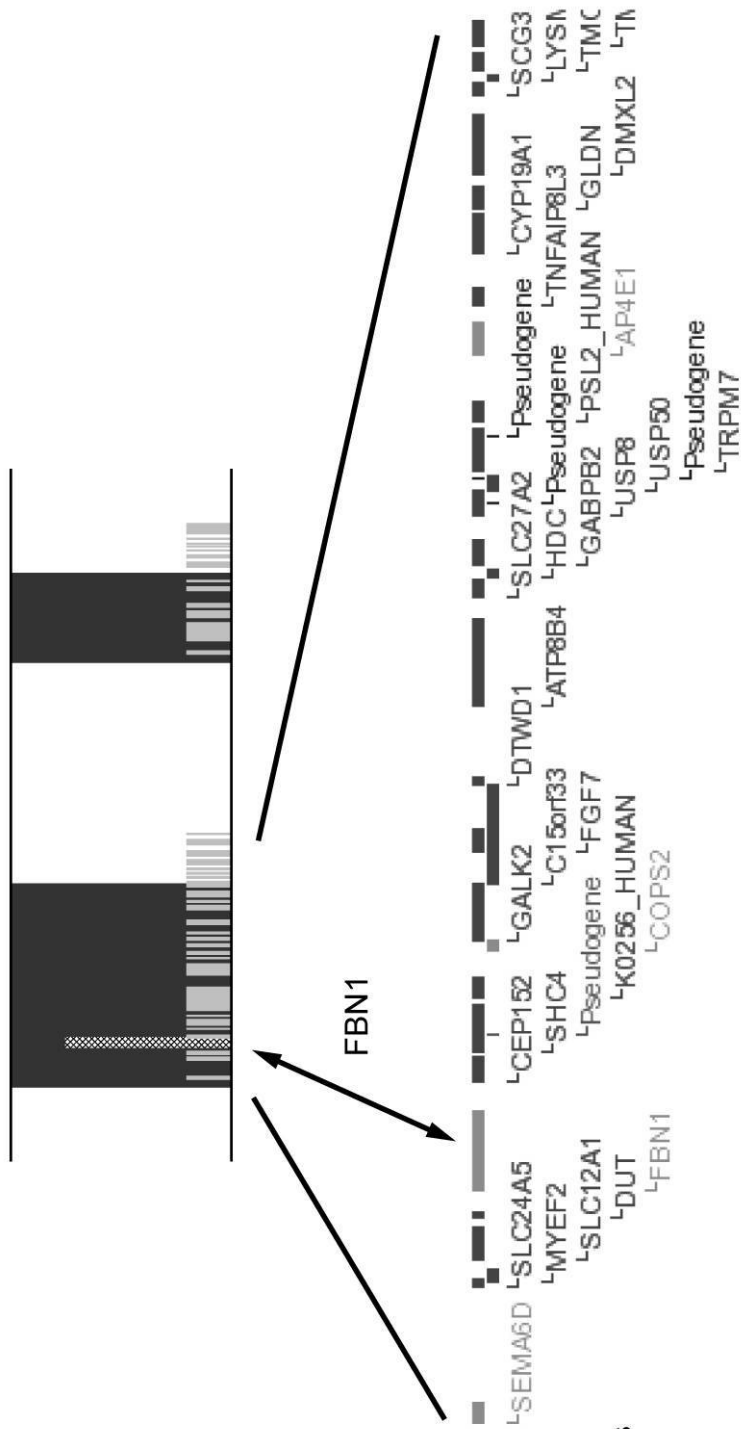
【 図 3 】



【 図 4 】



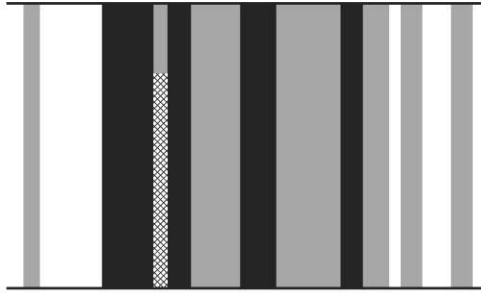
【 図 5 】



【 図 6 】

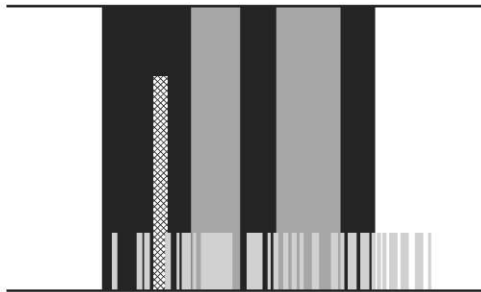
(a)

#ofSNP=60



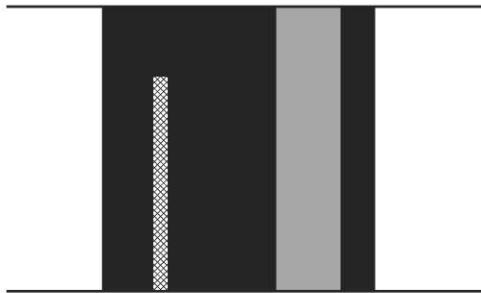
(b)

#ofSNP=80

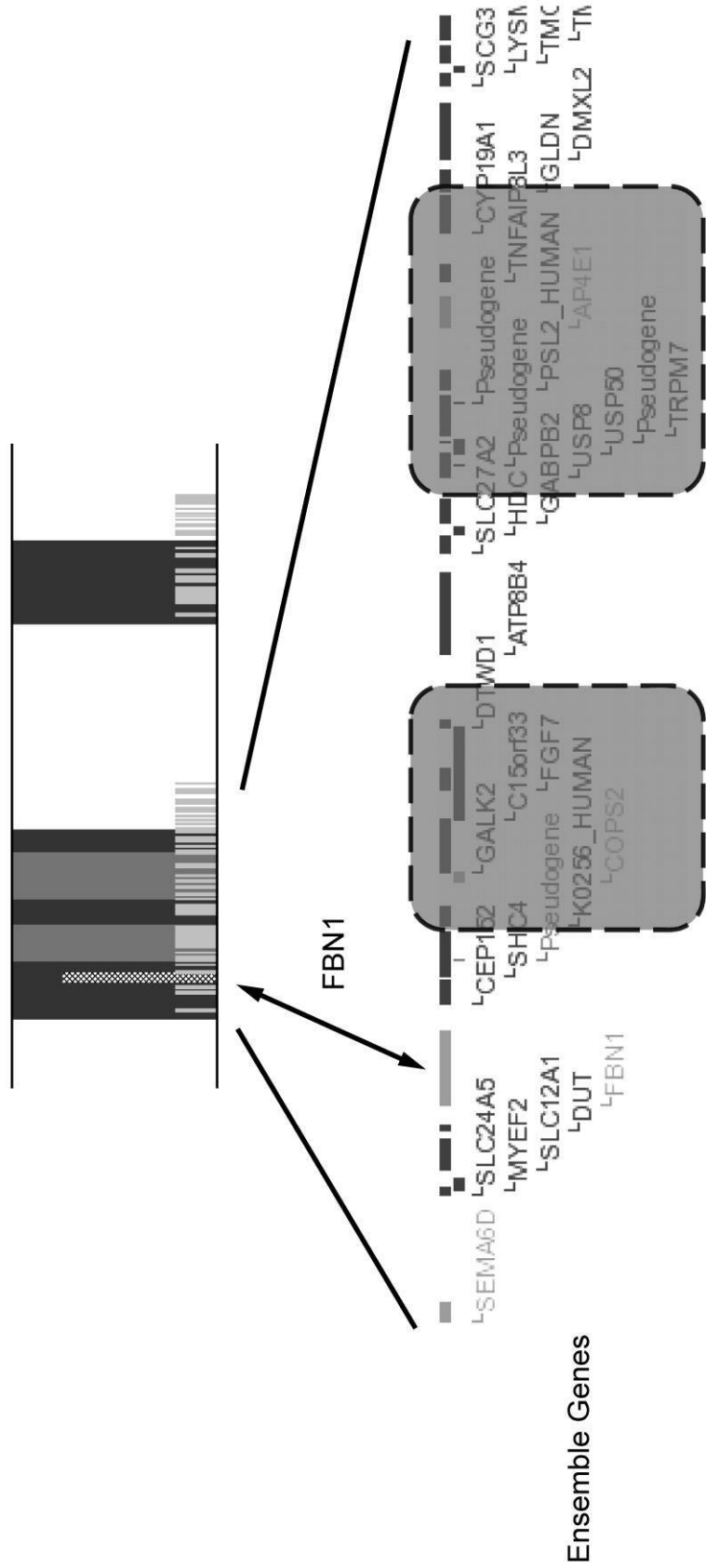


(c)

#ofSNP=100



【 図 7 】



フロントページの続き

(72)発明者 萩原 弘一

埼玉県入間郡毛呂山町毛呂本郷 3 8 学校法人埼玉医科大学内

(72)発明者 岡 崎 康司

埼玉県入間郡毛呂山町毛呂本郷 3 8 学校法人埼玉医科大学内

Fターム(参考) 4B024 AA01 AA11 AA20 CA02 HA11

4B063 QA13 QA18 QQ43 QR32