◆総説◆

原因遺伝子の同定 ポジショナルクローニング

尾内 善広^{*1*2}, 倉橋 浩樹^{*1}

Identification of the Disease Causing Gene by Positional Cloning

Yoshihiro ONOUCHI^{*1*2}, Hiroki KURAHASHI^{*1}**Key words:** positional cloning, linkage analysis, translocation, shotgun sequencing

1. はじめに

近年シーケンス技術の向上やデータベースの充実により、ヒトの遺伝子単離は飛躍的に速度が増し、今日では週の単位で新たな遺伝子が報告されるまでになっている。ここでは遺伝子同定の戦略としてポジショナルクローニングに焦点を当て、その実際を具体例を用いて解説する。

2. ポジショナルクローニングとは

ポジショナルクローニングとは疾患責任遺伝子を既知の蛋白質の情報なしに染色体上の位置情報のみをたよりにしてクローニングしてくる手法であり、先天性酵素欠損症や、ホルモン欠損症のように予め疾患の原因となる蛋白のアミノ酸配列から遺伝子の塩基配列がある程度予測されるものでない場合に用いられることが多い。主な流れとしてはまず連鎖解析や、染色体異常をあしがかりにして遺伝子を染色体上のある区域に局限化し遺伝子座を決定する(マッピングするという)。次にマッピングされた領域の

マーカーを用いてさらに範囲をせばめたうえでその領域のゲノムコンティグを作製する。そしてその領域にエクソンを見つけだして原因遺伝子を単離し、患者で遺伝子変異を同定するというものである。

3. 連鎖解析

連鎖解析とは減数分裂時の組み換え(recombination)現象をDNA多型をマーカーとして用いて観察し、患者表現型と共分離(segregate)するハプロタイプを決定しそのマーカー近傍に存在する遺伝子を候補遺伝子とし、疾患との関連を証明していく手法である。図1に組み換えのモデルを示した。1回の減数分裂で染色体上のある2点間で組み換えが起こる確率(組み換え率)はその2点が近ければ近いほど小さく逆に遠ければ遠いほど大きくなるが、50%を越えることはない。この組み換え率を遺伝的距離と呼び、1%の組み換え率を有する2点間の遺伝的距離は1 cM(センチモルガン)と定義している。組み換え率の計算は減数分裂の数が多い

*1 大阪大学医学部 B9 臨床遺伝学研究室 [〒 565-0871 大阪府吹田市山田丘 2-2]

Division of Clinical Genetics, Department of Medical Genetics, B9 Biomedical Research Center, Osaka University Medical School, [2-2, Yamadaoka, Suita, Osaka, 565-0871, Japan.]

*2 大阪大学医学部小児科学教室 [〒 565-0871 大阪府吹田市山田丘 2-2]

Department of Pediatrics, D5, Osaka University Medical School, [2-2, Yamadaoka, Suita, Osaka, 565-0871, Japan.]

図1 Crossing over between homologous chromosomes in meiosis
減数分裂時に2本の相同染色体は交差し、ある確立で組み換えが起こる。ある位置で組み換えが起こった場合、(I) 2つの遺伝子座が離れて存在する場合にはこの組み換えによってこれらの遺伝子座に関する組み換え体(*)が生じるが、(II) 2つの遺伝子座が近接している場合にはこの組み換えによって組み換え体は生じない。

ば多いほど正確となるが実際のヒトの家系では調べる世代数に限りがあるなどして利用可能な情報に乏しいことが多い。その問題を解決するために導入されたのがロッドスコア (lod score) である。ロッドスコアは組み換え率の関数であり疾患の責任遺伝子とマーカーとが連鎖していない (組み換え率 0.5) と考えた場合と連鎖している (組み換え率 θ) と考えた場合とがそれぞれ起こりうる確率の比の対数として計算される。ロッドスコアは $Z_{\max} = \log_{10} \theta^r (1-\theta)^{n-r} / 0.5^n$ で得られる数値である。n は減数分裂の数、r は組み換え体の数、 $\theta = r/n$ であり $Z_{\max}$ は $\theta = r/n$ のときに得られるロッドスコ

図2 Lod score Analysis
a が疾患と連鎖していると考え、(*) が組み換え体となる。組み換え率 θ でこのような対応が生じる確立は $\theta (1-\theta)^5$ 、一方全くの偶然にこのような対応が生じる確立は $(0.5)^6$ である。これらの確率の比の対数がロッドスコアである。この場合 $\theta = 0.1 \sim 0.2$ でロッドスコアは最高値をとる。

アの最大値である¹⁾。図2に実際のロッドスコアの計算の例を示した。

ロッドスコアが 3.0 以上、すなわち連鎖している場合の確率が偶然の場合の確率に比べて 1,000 倍以上高い場合に連鎖ありと認められる。逆にロッドスコアが 2.0 未満の場合には連鎖を除外される。

連鎖解析に用いる DNA 多型マーカーとしてはかつては RFLP (制限酵素断片長多型) が用いられてきたがさまざまな欠点を有していた。現在ではマイクロサテライトマーカーを用いるのが一般的となっている。マイクロサテライトマーカーとはヒトゲノム上に分散して存在する短い繰り返し配列のことで、その繰り返し回数に多型が存在する。多型性が高い点や PCR により増幅して容易にハプロタイプを調べることができるので解析に要する DNA の量が少なく済むという点からも RFLP に勝っている。現在蛍光標識された primer のセットが市販されており複数のマイクロサテライトマーカーを同時に PCR にて増幅して調べられるようになっている。

実際の連鎖解析を Huntington 病の場合を例にとり解説する。Huntington 病は舞踏病運動、精神障害、痴呆を呈する常染色体優性の晩発性疾患で非常に高い浸透率を有することから連鎖解析に非常に適しており初めて連鎖解析により遺伝子座がマップされた疾患である。1983 年に Gusella らがベネズエラの大家系について RFLP の多型を調べたところ、G 8 という probe が *Hind*III の認識部位の多型を検出することが判明し、4 種類存在する多型の内 1 種類がこの大家系において疾患と共分離した。このことから Huntington 病の遺伝子座は G 8 と非常に近いことが分かった²⁾。G 8 は 4p16.3 にマップされ、D4S10 と登録された。その後の詳しい解析の結果 D4S10 より遠位の D4S180 と D4S182 の間の約 500 kb ではあるハプロタイプに疾患が強く連鎖していることから（連鎖不平衡）原因遺伝子の位置が同定され、同領域から単離された原因遺伝子は Huntingtin と命名された。その 5' 末端にある CAG の 3 塩基による繰り返し配列が正常では約 20 回であるところが Huntington 病患者では 40 回以上に伸張していることが判明したのである³⁾。

4. 染色体異常

疾患責任遺伝子をより直接的に染色体上にマップすることができるのが染色体異常の解析である。近年のゲノム解析の手法の進歩によって染色体転座や逆位の切断点の解析や微小欠失の検出が可能となり、そこに存在する遺伝子が同定されるようになった。染色体転座から疾患責任遺伝子が同定された代表的な例が Duchenne 型筋ジストロフィー (DMD) である。DMD はジストロフィン遺伝子の異常により進行性に骨格筋が萎縮する X 染色体劣性の疾患で、罹患した男児は幼児期より発症し 10 歳前後から車椅子が必要となりやがては呼吸さえも自力では不可能となり、20 歳前後で筋萎縮に起因する合併症により死亡する。DMD 遺伝子は連鎖解析に

よって 1982 年に Xp21 にマップされていたが、次に示す 2 つの染色体異常を用いた異なる手法によりさらに限局化され同定された。

そのひとつは X-常染色体間転座を有する DMD 女性患者の存在からのアプローチであった。この患者では X 染色体の短腕と 21 番染色体の短腕との転座があった。通常 X 染色体の不活化は発生のある時期に全ての細胞にランダムに起こるのであるが、X-常染色体間の転座がある場合、転座染色体に不活化が起こった細胞は同時に転座染色体上の常染色体由来の部分にも不活化が起こってしまいその部分にある遺伝子が発現しなくなる。そのような細胞は生きていくことができず淘汰され、結果として正常 X 染色体が不活化された細胞のみが生き残ることになる。すなわちこのような女性患者においては全ての細胞において転座切断点に存在する遺伝子産物が X 染色体からは全く作られないことになる。つまりこの切断点に原因遺伝子が存在する DMD 男性患者と同じ状況が起こっていることとなるわけである。21 番染色体の短腕には 13 番、14 番、15 番、22 番染色体の短腕と同じく、リボゾーム RNA をコードする繰り返し配列が存在することが分かっていたのでその切断点の解析から DMD 遺伝子の同定に成功したのである⁴⁾。

もうひとつは Xp21 に欠失が存在した DMD 男性患者にみられた隣接遺伝子症候群 (contiguous gene syndrome) からであった。この患者では DMD のほかに網膜色素変性症、慢性肉芽腫症、精神発達遅延を合併していた。この欠失領域をサブトラクションクローニングという手法を用いて単離したのである。得られた Xp21 領域の DNA クローンをプローブに DMD 患者でサザンハイブリダイゼーションを行なったところあるクローンが細胞遺伝学的には欠失のみられない DMD 患者の 6.5% において欠失を示し、この欠失が DMD と密接に連鎖していたのである。こうして同定された遺伝子はゲノム上の大きさが 2,400 kb、cDNA が 14 kb という巨

図3 Cloning of the breakpoint of chromosomes with translocation

(a) 染色体Gバンド法により切断点の大まかな位置を知る。(b) 染色体上のその部位にマップされたマーカーと切断点との関係をそのマーカーを含むクローンによるFISHで決定する。(c) 切断点を挟むマーカーから切断点に向けて染色体歩行を行う。

大なものであり427 kDの蛋白をコードしておりジストロフィンと命名された⁵⁾。

5. 現在行われているポジショナルクローニングの実際

転座や逆位の切断点の解析法は現在以下に示す方法が主流となっている(図3)。まず染色体Gバンド法により切断点は何番の染色体のどのバンド上にあるのかを調べる。次にその染色体上の部位にマップされているマーカーの位置をYAC(酵母人工染色体)、BAC(細菌人工染色体)、PAC(P1人工染色体)クローンをプローブとしてFISH(fluorescence in situ hybridization)法を行ない決定する。そして切断点をはさむマーカーの間にできたギャップをこれらのクローンによる染色体歩行を行って埋めてい

図4 Shotgun sequencing

(a) 目的のコスミドクローン(BAC, PAC)を超音波にて2~3 kbの長さに物理的に切断する。(b) プラスミドベクターにライゲーションする。(c) エレクトロポレーションにて大腸菌をトランスフォームしシングルコロニーを培養してプラスミドを抽出する。(d) シークエンス解析し整列化する。(e) コンピューターにてエクソンを予測。

く、プローブに用いたクローン上に切断点が存在すればシグナルは正常染色体では正常の位置に、転座染色体では転座点の近位と遠位の2カ所に認められるはずである。染色体歩行の際にすでにそのクローンを含むコンティグが報告されていればそれが利用可能となる。通常は末端部の塩基配列を決定してその部分に作製したプライマーによるPCRにてライブラリーのスクリーニングすることが多い。このようにして得られたクローンが切断点を含むことが判明したら、YACならばコスミドのサブライブラリーを作製し、FISH法にて切断点を含むコスミドクローンを単離する。また切断点を含むコスミドを患者のゲノムDNAのサザンプロットにハイブリダイズさせると患者においてあるバンドが薄くなるかあるいは消失して健常者にはない

表 1 エクソン予測プログラムの主なアドレス

program	address (E-mail server)
GRAIL	(GRAIL@ornl.gov)
HEXON	http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html (service@bioinformatics.weizmann.ac.il)
GENSCAN	http://gnomic.stanford.edu/GENSCANW.html (genscan@gnomic.stanford.edu)

現在われわれが主に利用しているエクソン予測プログラムのアドレスを示した。ここに挙げたもの以外にも多数存在する。

バンドが出現する。これらの切断点を含む BAC, PAC, コスミドの塩基配列の決定をショットガンシーケンスにて行う (図 4)。遺伝子が実際に転座によって切断を受けているとすればこのクローン上に遺伝子が存在する可能性が非常に高いこととなる。得られた塩基配列の中から遺伝子を探し出すにはコンピューターを用いる。エクソン予想プログラムを用いる方法と EST を用いる方法とがある。エクソン予測プログラムとして, Grail, Hexon, Genscan 等がインターネット上で利用可能である (表 1)。これらは, オープンリーディングフレーム, エクソン-イントロン境界部の保存された配列等の条件を満たす部位をエクソンとして予測してくるものである。EST (expressed sequence tag) とは cDNA の末端の 200~300 bp を塩基配列決定して得られたものであり, cDNA のどの末端から読んだものかによって 5'EST と 3'EST とがある。これらのうち染色体上にマップされているものは発現のある DNA マーカーとして利用価値が高い。切断点近傍から得られた塩基配列に対し相同性検索 (homology search) を行い, マッチする EST があるかを調べるのである。予測されたエクソンが本物であるか, 実際にどの臓器で発現しているかは RT-PCR で確認することができる。RT-PCR の産物でノーザンブロットや cDNA ライブラリーのスクリーニングを行い, cDNA の全長, 臓器間の発現量の違い等を知ることができる。最終的には患者でその遺伝子の遺伝子変異を同定することによって原因遺伝子であることを証明する。

6. ポジショナルクローニングから位置的候補遺伝子検索 (positional candidate gene approach) へ

近年ポジショナルクローニング以外の方法で疾患責任遺伝子が単離される例が増えてきている。それはあらかじめ機能や発現パターンの分かっている遺伝子が疾患と同じ染色体上の領域にマップされている場合に疾患の原因遺伝子の候補と考え患者でその遺伝子変異を検索するという手法である。このような手法を位置的候補遺伝子検索 (positional candidate gene approach) といい, アルツハイマー病, 網膜色素変性症などの遺伝子はこの手法で同定された。

ヒトの遺伝子の単離が進むにつれこのような候補遺伝子は今後ますます多く出現してくることが予想される。今後疾患責任遺伝子の同定法の主流は位置的候補遺伝子検索へと移行していき, 今以上の速度で疾患の原因そして病態の解明が進むと思われる。

7. おわりに

ポジショナルクローニングの手法として, 連鎖解析と染色体異常の解析を実例を交えて解説した。すでに単一遺伝子疾患の原因遺伝子の多くが単離され, 現在多因子遺伝疾患にも応用されている。今後さらに疾患責任遺伝子の同定が進み, そこから病態の解明, さらに新たな治療法の開発に結びつくであろう。また, 疾患の発症前診断, 保因者診断が可能となる一方でそ

これから新たな問題が生じてくる可能性があることをわれわれは忘れてはならない。

文 献

- 1) 中村祐輔: リンケージ解析法の原理とマッピングされた遺伝性疾患・実験医学 **12**: 660-664, 1994.
 - 2) Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, Watkins PC, Ottina K, Wallace MR, Sakaguchi AY, Young AB, Shuolson I, Bonilla E, Martin JB: A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**: 234-238, 1983.
 - 3) The Huntington's Disease Collaborative Research Group: A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**: 971-983, 1993.
 - 4) Worton RG, Thompson MW: Genetics of Duchenne muscular dystrophy. *Annu Rev Genet* **22**: 601-629, 1988.
 - 5) Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, Kunkel LM: Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50**: 509-517, 1987.
-