

2. RFLP による連鎖解析：

ロッド得点法ならびに RFLP マーカの選別

日下部 眞一* 服巻 保幸**

はじめに

DNA 診断の特徴の一つとして、病因遺伝子が不明な遺伝病さえも診断の対象となりえることが挙げられる。つまりゲノム上に存在する制限酵素切断片多型(RFLP)をマーカーとして連鎖解析を行うことにより、間接的に診断を行うことができるわけで、これは従来の遺伝子産物の異常に基づく診断法では不可能なことであった。この手法は上記のような遺伝病の診断以外にも、分子レベルでの病因が多岐にわたる遺伝病や、病因遺伝子が巨大なために変異の同定が困難な遺伝病の診断においても有効である。さらには病因遺伝子のクローニングや、最近注目を浴びつつあるヒト全ゲノム塩基配列決定においても、この連鎖解析による遺伝子地図作成は主要な位置を占めている。

ここでは、このような解析に用いられる統計的手法であるロッド得点法について述べ、さらに RFLP マーカークをきたすプローブの単離についても触れたい。しかし各遺伝病における連鎖解析の実際については、本紙の他の稿で詳細に触れられるため割愛し、ここでは連鎖解析の原理に重点を置きまとめることにする。

1. 連鎖解析とロッド得点法の原理

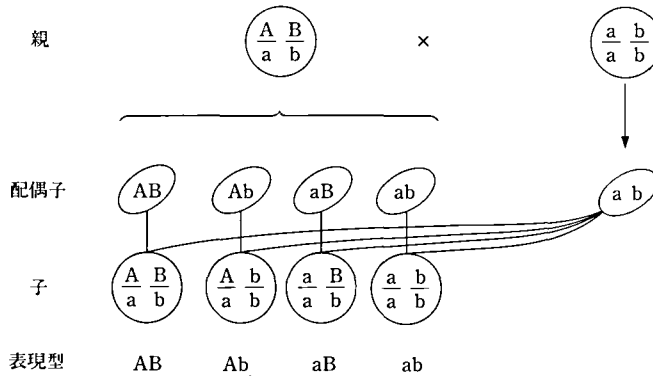
ロッド得点の計算は現在、LIPED や LINKAGE などの計算機プログラムを用いて手軽に行えるようになってきている。ここでは、現

在の連鎖解析法の基礎となった Morton(1955)の方法についてわかりやすく解説したい。このためには、まず第1に連鎖と組換え、および組換え頻度について、次にどのような遺伝子型の子供が生まれてくるかということは確率分布に従っているということ、したがって連鎖関係を判断するには定量的な方法が必要であるということを理解していただかねばならない。そこで、この節でこれらの考えの大筋を解説してロッド得点法の原理を理解していただき、次の節でいろいろな交配型についての具体的なロッド得点の計算法を詳述する。

a. 連鎖と組換え

2つの遺伝子座が異なる染色体上にあるときは、減数分裂の際それぞれの遺伝子座は独立に行動してメンデルの独立分離の法則に従うことになる(図1-1)。これに対し、同じ染色体上に位置するときには、それらの位置関係が非常に近ければ2つの遺伝子座のそれぞれの対立遺伝子対(たとえば図1-2のAとB、またはaとb)は行動をともにすることになるがこれは連鎖と呼ばれる。この位置関係が遠くなるにしたがいこれらの遺伝子座間で組換えがおこり組み合わせの異なる配偶子(図1-2のAbとaB)が作られてくる。2つの遺伝子座が同じ染色体上にあるときには、その位置が互いに遠くなるほどそれらの間で組換えがおこる確率が高くなり、これらの頻度(図1の θ)を測定することによって遺伝子座の位置関係を直線上に配置すること

1) 2つの遺伝子座が別の染色体上にあるとき(独立)



2) 2つの遺伝子座が同じ染色体上にあるとき(連鎖)

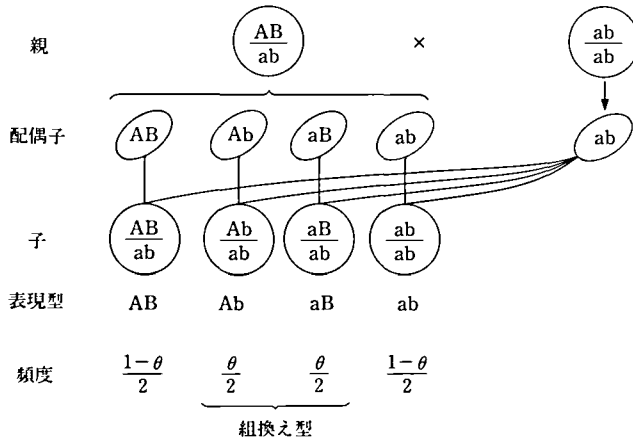


図1 連鎖と独立を示す図

ができ遺伝子地図が作成される。

ショウジョウバエやマウスなどの実験動物では図1に示したように二重劣性の個体に交配することによって二重ヘテロ接合体の親の配偶子の遺伝子型頻度を子供の表現型頻度からただちに測定できる(これを検定交配と呼ぶ)。そして、子供の中の組換え型の頻度(θ)をもって2つの遺伝子座の距離と定義することができる。 $\theta=0.01$ を1cM(センチモルガン)と呼び、これはほぼ 10^3 kbに相当する。

しかし、ヒトを対象とする場合には一家族の子供の数は多くても数人で、ある表現型をもった子供の出現は確率的事象となる。また、マウスやショウジョウバエの検定交配に相当する交配の組み合わせ(交配型)だけを集めることはほ

とんど不可能で、情報となりうる交配型すべてを集めて連鎖関係について考察することになる。結局、問題は形成された配偶子のうち組換え型か非組換え型かが容易に検出できればいいわけで、後にふれる、精子1個に対してPCRによってRFLPを検出できるようになってきたことは、連鎖解析を飛躍的に進歩させる可能性もっている。

b. 2項分布と多項分布

連鎖解析は二世資料などの家族構成をもとにして行われるが、どのような表現型の子供がうまれるかということは確率的事象である。これは、受精にあずかる配偶子は遺伝的には決定されているが何種類かの数多くの中のほんのわずかであるということを考えて理解されるで

あろう。

たとえば、男の子が生まれるか女の子が生まれるかということ、10 円玉を投げて表がでるか裏がでるかということと確率のモデルとしては同じと考えることができ、2 項分布のモデルがこれに相当する。

《2 項分布の法則》

ある事象 A がおこる確率が p で事象 B の確率が q ($p + q = 1$) であれば s 回の試行で A が a 回、B が b 回おこる確率は

$$P = \frac{s!}{a!b!} p^a q^b$$

である。

ある家族に 3 人子供がいるとして 3 人とも男の子かまたは 3 人とも女の子である確率は

$$\frac{3!}{3!0!} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^0 + \frac{3!}{0!3!} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

となる。2 項分布は事象(男か女, 表か裏など)が 2 つの要素についての法則であるが、多くの要素からなる確率事象に拡張できて多項分布として知られている。

《多項分布の法則》

ある事象 A がおこる確率が p , 事象 B の確率が q , 事象 C が r , ... であれば s 回の試行で A が a 回, B が b 回, C が c 回, ... おこる確率は

$$P = \frac{s!}{a!b!c!\dots} p^a q^b r^c \dots$$

である。ただし、 $p + q + r + \dots = 1$ である。

遺伝学の多くの問題でこの多項分布のモデルが使われる。たとえば先程の図 1-1 のような交配を行ったとする。総計 s 人の子供のうち表現型が AB, Ab, aB, ab の子供の数がそれぞれ a, b, c, d となる確率は

$$\begin{aligned} P &= \frac{s!}{a!b!c!d!} \left(\frac{1}{4}\right)^a \left(\frac{1}{4}\right)^b \left(\frac{1}{4}\right)^c \left(\frac{1}{4}\right)^d \\ &= \frac{s!}{a!b!c!d!} \left(\frac{1}{4}\right)^s \end{aligned}$$

で求められる。

たとえば 4 人の子供がすべて AB の表現型になる確率は

$$P = \frac{4!}{4!0!0!0!} \left(\frac{1}{4}\right)^4 = \left(\frac{1}{4}\right)^4$$

4 人の子供がそれぞれ AB, Ab, aB, ab の 4 種類の表現型になる確率は

$$P = \frac{4!}{1!1!1!1!} \left(\frac{1}{4}\right)^4 = 24 \times \left(\frac{1}{4}\right)^4$$

となる。

さて、連鎖解析で問題となってくるのは図 1-2 のような場合で、これは表現型の出現期待頻度としてただ組換え率 θ が変数としてはいつてくるだけで同じようにして多項分布の式を用いて

$$P = \frac{s!}{a!b!c!d!} \left(\frac{1-\theta}{2}\right)^a \left(\frac{\theta}{2}\right)^b \left(\frac{\theta}{2}\right)^c \left(\frac{1-\theta}{2}\right)^d$$

として計算できる。

Morton 法の場合、次に述べるように確率の比(尤度比)を用いて検定を行うので、同じ多項係数 ($s! / a!b!c!d!$ のこと)が分子と分母にあらわれて、その結果、消えてしまうことになる。

c. ロッド得点法の原理

一般論は後にまとめるとして、まずはじめにわかりやすい具体例を考えながらロッド得点法の考え方を説明していこう。図 2 に 2 つの二世代家系図を示している。黒印は患者である。病因遺伝子 (G) は優性と仮定する。病因遺伝子座の遺伝子型は、患者は Gg , 健康な人は gg と判断できる。連鎖解析に役立つには親の一方が少なくとも二重ヘテロ接合体でないといけないので親の患者の遺伝子型は調査の結果 GgT_1T_2 であったとしておく。最も簡単な解析となるように他方の親は ggT_1T_1 とする。病因遺伝子座と RFLP 遺伝子座の連鎖関係としては、①強く連鎖(相引(例 1)または相反(例 2))している、②独立である(例 3)、③やや連鎖している(例 4)という 3 つの可能性が考えられるわけで、この状況を想定して、RFLP を示す遺伝子座については DNA 診断の結果図に書いたような遺伝子型が得られたとする。

まず例 1 を考えてみよう。組換えのところで説明したように、二重ヘテロ接合体の親 GgT_1T_2 の配偶子の遺伝子型は相引 (GT_1/gT_2) のときには非組換え型が GT_1 と gT_2 で組換え型が

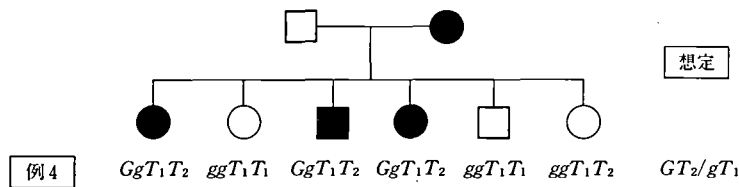
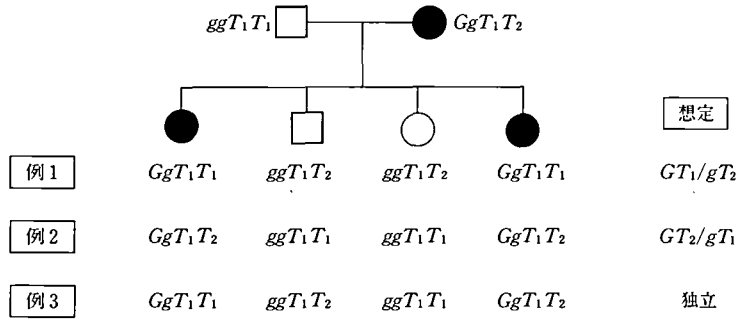


図2 優性遺伝病と RFLP を示す 2 世代家族資料の例

表1 図2の2世代家族資料の例を用いてロッド得点を求めるための表

		GgT_1T_2 の配偶子と頻度				子供の 総 数
		GT_1	GT_2	gT_1	gT_2	
		相 引	$\frac{1-\theta}{2}$	$\frac{\theta}{2}$	$\frac{\theta}{2}$	
		相 反	$\frac{\theta}{2}$	$\frac{1-\theta}{2}$	$\frac{1-\theta}{2}$	$\frac{\theta}{2}$
ggT_1T_1 の 配偶子	gT_1	GgT_1T_1	GgT_1T_2	ggT_1T_1	ggT_1T_2	
観 察 数	例 1	2	0	0	2	4
	例 2	0	2	2	0	4
	例 3	1	1	1	1	4
	例 4	0	3	2	1	6
	一 般	a	b	c	d	s

GT_2 と gT_1 である。組換え率を θ とすれば、 GT_1 , gT_2 , GT_2 , gT_1 の配偶子頻度は $(1-\theta)/2$, $(1-\theta)/2$, $\theta/2$, $\theta/2$ となることが期待される。親が相反のときには、非組換え型と組換え型とが入れかわるだけである。一方の親の配偶子は gT_1 だけだから、二重ヘテロ接合体の親との交配の結果、期待される子供の遺伝子型とその期待頻度は表1のように整理される。4種類

の表現型の出現期待頻度は θ がきまれば求まるので、実際に例1から例4のような資料が得られる確率を計算することが可能である。この計算に先程述べた多項分布が用いられることになる。

二重ヘテロ接合体の親はこの資料だけからは相引(GT_1/gT_2)か相反(GT_2/gT_1)かわからないので、それぞれの場合の確率を足して $1/2$ す

表 2 家族例 1 のロッド得点を求めた表(例 2 はこれと同じになる)

	θ					
	0	0.1	0.2	0.3	0.4	0.5
$f(y_1; \theta)$	0.188	0.123	0.077	0.047	0.029	0.023
$f(y_1; \theta)/f(y_1; \frac{1}{2})$	8.000	5.248	3.290	1.986	1.242	1.000
$Z(y_1; \theta)$	0.903	0.720	0.517	0.298	0.094	0.000

ればよい。相引のときに GT_1 が 2 人, gT_1T_2 が 2 人となる確率は多項分布の式を用いて

$$\frac{4!}{2!0!0!2!} \left(\frac{1-\theta}{2}\right)^2 \left(\frac{\theta}{2}\right)^0 \left(\frac{\theta}{2}\right)^0 \left(\frac{1-\theta}{2}\right)^2$$

となる。

また相反のときには

$$\frac{4!}{2!0!0!2!} \left(\frac{\theta}{2}\right)^2 \left(\frac{1-\theta}{2}\right)^0 \left(\frac{1-\theta}{2}\right)^0 \left(\frac{\theta}{2}\right)^2$$

である。したがって、組換え率が θ であると仮定すれば例 1 のような家族構成 (y_1) が得られる確率は

$$\begin{aligned} f(y_1; \theta) &= \frac{1}{2} \left\{ \frac{4!}{2!0!0!2!} \left(\frac{1-\theta}{2}\right)^2 \left(\frac{\theta}{2}\right)^0 \left(\frac{\theta}{2}\right)^0 \left(\frac{1-\theta}{2}\right)^2 \right. \\ &\quad \left. + \frac{4!}{2!0!0!2!} \left(\frac{\theta}{2}\right)^2 \left(\frac{1-\theta}{2}\right)^0 \left(\frac{1-\theta}{2}\right)^0 \left(\frac{\theta}{2}\right)^2 \right\} \\ &= \frac{6}{2^5} \{ \theta^4 + (1-\theta)^4 \} \end{aligned}$$

として求められる。

組換え率 θ が決まれば、たとえば $\theta=0.2$ のとき、例 1 のような家族が得られる確率が計算される。

ここで言葉を 1 つ定義しておこう。

今、ある事象(出来事、たとえば例 1 のような家族構成のこと)がおこったとする。このとき、あるモデル(仮説、たとえば 2 つの遺伝子座が連鎖していてその組換え率が θ であるということ)のもとでこのような事象が得られる確率を「尤度」という。

尤度という術語は何かしらかた苦しいひびきを感じさせるが、英語では 'likelihood' で、尤を訓読みしてもっともらしい度合いというわかりやすい。つまり現実におこった出来事を説

明するの、ある仮説のもとでどのくらいのもっともらしきで説明するかという度合いである。だから仮説がかわれば尤度はかわってくる。

さて例 1 にもどって、2 つの遺伝子座が独立なとき、すなわち別々の染色体上か、また同じ染色体上でも遠く離れていると 4 種の表現型の期待頻度はすべて等しく $1/4$ になる。したがって独立なときに例 1 のような家族構成を得る確率は

$$\frac{4!}{2!0!0!2!} \left(\frac{1}{4}\right)^2 \left(\frac{1}{4}\right)^0 \left(\frac{1}{4}\right)^0 \left(\frac{1}{4}\right)^2 = \frac{6}{2^8}$$

となる。

当然のことだが、これは先程の式 $f(y_1; \theta)$ で $\theta=1/2$ とおいた値 $f(y_1; 1/2)$ に等しいことに注意してほしい。ここで問題なのは、どのような数量を用いれば独立か連鎖かの判断の信頼性を表現できるかということである。このためには、このような家族構成が得られる尤度が、独立と仮定したときに比べて $\theta=\theta_0$ と仮定したときの方が大きくなるか小さくなるかということと判断するとよいことがわかる。つまり式を用いて表現すると

$$\frac{f(y_1; \theta_0)}{f(y_1; \frac{1}{2})}$$

が判断の基準となる。

これがオッズ(odds)で、統計学では尤度比といわれ、これを用いた検定は尤度比検定といわれている。このオッズは θ だけを変数とする関数となっていることに注意してほしい。 θ を 0 から 0.5 にわたって変化させることによって表 2 のような数値が得られる。予想どおりであるが $\theta=0$ 、つまり完全連鎖と仮定したときに最も見込み(オッズ)が大きくなっているのがわかる。ここで得られた結果は例 1 という 1 つの家系だ

表3 家族例3のロッド得点を求めた表

	θ					
	0	0.1	0.2	0.3	0.4	0.5
$f(y_3; \theta)$	0.000	0.012	0.038	0.066	0.086	0.094
$f(y_3; \theta)/f(y_3; \frac{1}{2})$	0.000	0.130	0.410	0.706	0.922	1.000
$Z(y_3; \theta)$	—	-0.887	-0.388	-0.151	-0.035	0.000

表4 家族例4のロッド得点を求めた表

	θ								
	0	0.02	0.05	0.1	0.17	0.2	0.3	0.4	0.5
$f(y_4; \theta)$	0.000	0.008	0.018	0.028	0.031	0.030	0.024	0.017	0.015
$f(y_4; \theta)/f(y_4; \frac{1}{2})$	0.000	0.579	1.238	1.890	2.147	2.105	1.668	1.192	1.000
$Z(y_4; \theta)$	—	-0.238	0.093	0.276	0.332	0.323	0.222	0.076	0.000

けの資料なのだが、実際の家系調査ではいろいろな交配型による2世代資料が集まる。それぞれの家族(y_1, y_2, \dots, y_n)で、その親の交配型にあうような方法(次節で述べる)に従ってオッズを計算してそれらの積をとればよい。積は繁雑で、積が和で表現される対数値に変換して和をとればよい。つまり、 θ の値を0から0.5まで変化させたときの

$$\sum_{i=1}^n \log \frac{f(y_i; \theta)}{f(y_i; \frac{1}{2})}$$

が計算されることになる。オッズ(odds)の対数(logarithm)というわけだからロッド得点(lod score)と呼ばれている。この例では表2でわかるようにロッド得点は $\theta=0$ で最も大きく約0.9である。ロッド得点の総和が3に近ければ連鎖ありと判断されるから、この例1のような家系資料が3つあつればロッド得点の総和は約3($0.9 \times 3 = 2.7$)になり連鎖ありと判断できることになる。

例2は、病因遺伝子が T_2 と連鎖していることを想定したものがロッド得点の計算結果は例1と全く同じになる。

例3は独立を想定したものである。例1にならってロッド得点を計算しよう。組換え率 θ を仮定すれば例3のような家族が得られる尤度は、

二重ヘテロ親の相引、相反を考慮して

$$\begin{aligned} f(y_3; \theta) &= \frac{1}{2} \left\{ \frac{4!}{1!1!1!1!} \left(\frac{1-\theta}{2}\right)^1 \left(\frac{\theta}{2}\right)^1 \left(\frac{\theta}{2}\right)^1 \left(\frac{1-\theta}{2}\right)^1 \right. \\ &\quad \left. + \frac{4!}{1!1!1!1!} \left(\frac{\theta}{2}\right)^1 \left(\frac{1-\theta}{2}\right)^1 \left(\frac{1-\theta}{2}\right)^1 \left(\frac{\theta}{2}\right)^1 \right\} \\ &= \frac{3}{2} \theta^2 (1-\theta)^2 \end{aligned}$$

となる。

例1と同様に θ が0から0.5にわたって0.1きざみにオッズとロッド得点を計算すると表3のような結果が得られる。ロッドの値が負になっているのは、独立と判断した方が可能性が高い(見込みが大きい)ということの意味しているわけである。

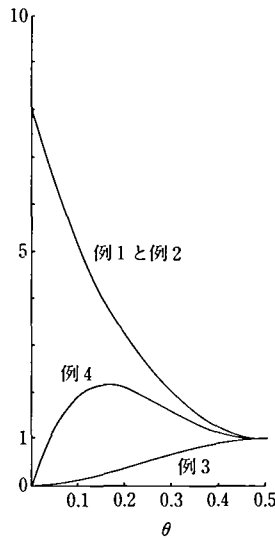
さて、以上は完全連鎖か独立の場合だったが、ゆるく連鎖している場合はどうなるであろうか。これを想定したのが例4である。今までと同じようにして尤度を計算すると

$$\begin{aligned} f(y_4; \theta) &= \frac{1}{2} \left\{ \frac{6!}{0!3!2!1!} \left(\frac{1-\theta}{2}\right)^0 \left(\frac{\theta}{2}\right)^3 \left(\frac{\theta}{2}\right)^2 \left(\frac{1-\theta}{2}\right)^1 \right. \\ &\quad \left. + \frac{6!}{0!3!2!1!} \left(\frac{\theta}{2}\right)^0 \left(\frac{1-\theta}{2}\right)^3 \left(\frac{1-\theta}{2}\right)^2 \left(\frac{\theta}{2}\right)^1 \right\} \\ &= \frac{15}{2} \theta (1-\theta) \{ \theta^4 + (1-\theta)^4 \} \end{aligned}$$

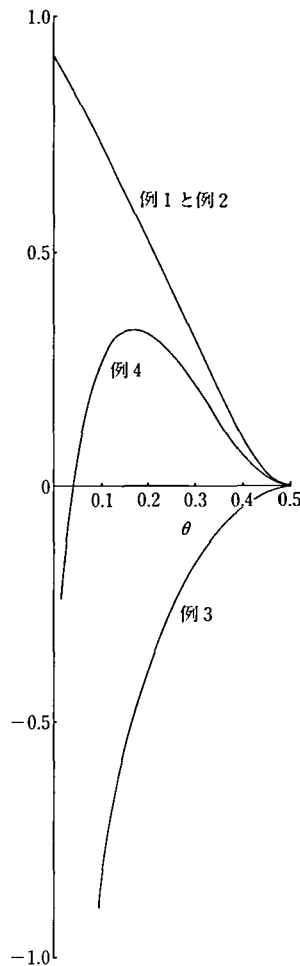
が得られる。

θ を0から0.5まで部分的には0.01きざみ

1) オッズ



2) ロッド得点



オッズについては
 例1と例2が $2^3 \{ (1-\theta)^4 + \theta^4 \}$
 例3が $2^3 \{ 2\theta^2(1-\theta)^2 \}$
 例4が $2^5 \{ (1-\theta)\theta^5 + (1-\theta)^5\theta \}$
 を θ の関数として図示したもので、
 ロッド得点はこれを対数変換した
 ものである

図3 オッズとロッド得点をあわす図

にしてオッズとロッドを計算すると表4の結果が得られる。ロッド値が最大となるのは $\theta=0.17$ 付近でロッド得点は0.33である。多くの家系を集めてロッド得点を集計して $\theta=0.17$ 付近で得点が最大となって3近くになればゆるく連鎖していて $\theta=0.17$ という判断が下せることになる。

以上の4例を用いた説明でロッド得点法の原理が理解されたと思われるのでここでまとめておく。

親の交配の組み合わせ(交配型)が定まればそ

の子供にどのような表現型が出現し、どれくらいの頻度で期待されるかということが計算できる(いろいろな交配型でのロッド得点の計算法は次の節で詳述する)。組換え率 θ を仮定したときにある家族 y_i が得られる確率は多項分布の式を用いて計算でき尤度 $f(y_i; \theta)$ が求まる。この式で $\theta=1/2$ とおいた値は独立と仮定したときの尤度である。したがって、

$$\frac{f(y_i; \theta)}{f(y_i; \frac{1}{2})}$$

はどれくらいの確からしきで独立にくらべて連鎖しているかどうかという見込みをあらわす量

表5 組換え率(θ)についての情報がなくなる交配型

		GT ₁ /gT ₂ 親の配偶子			
		GT ₁	GT ₂	gT ₁	gT ₂
		$\frac{1-\theta}{2}$	$\frac{\theta}{2}$	$\frac{\theta}{2}$	$\frac{1-\theta}{2}$
GT ₁ /GT ₁ 親の 配偶子	GT ₁	GT ₁	GT ₁ T ₂	GT ₁	GT ₁ T ₂
		$\frac{1-\theta}{2}$	$\frac{\theta}{2}$	$\frac{\theta}{2}$	$\frac{1-\theta}{2}$
GT ₂ /GT ₂ 親の 配偶子	GT ₂	GT ₁ T ₂	GT ₂	GT ₁ T ₂	GT ₂
		$\frac{1-\theta}{2}$	$\frac{\theta}{2}$	$\frac{\theta}{2}$	$\frac{1-\theta}{2}$
GT ₁ /GT ₂ 親の 配偶子	GT ₁ $\frac{1}{2}$	GT ₁	GT ₁ T ₂	GT ₁	GT ₁ T ₂
		$\frac{1-\theta}{4}$	$\frac{\theta}{4}$	$\frac{\theta}{4}$	$\frac{1-\theta}{4}$
	GT ₂ $\frac{1}{2}$	GT ₁ T ₂	GT ₂	GT ₁ T ₂	GT ₂
		$\frac{1-\theta}{4}$	$\frac{\theta}{4}$	$\frac{\theta}{4}$	$\frac{1-\theta}{4}$

でオッズ(odds)と呼ばれる。多くの家系について分析するからオッズの積をとればよいのだが対数値に変換すると和として計算できる。

$$Z(y_i; \theta) = \log \frac{f(y_i; \theta)}{f(y_i; \frac{1}{2})}$$

がロッド得点と呼ばれる。実際にはいくつかの家族の資料が集計されるが($\sum_{i=1}^n Z(y_i; \theta)$), たとえば組換え率 θ を 0.1 きざみにしてロッド得点を計算すると,

$$\sum_{i=1}^n Z(y_i; 0), \sum_{i=1}^n Z(y_i; 0.1),$$

$$\sum_{i=1}^n Z(y_i; 0.2), \dots, \sum_{i=1}^n Z(y_i; 0.4)$$

が得られる。必要であれば局所的にもっと小さいきざみ方で計算するとよい。ロッド得点が3以上になる θ のところで連鎖あり, -2以下であれば連鎖なし, -2から3の間であればもっと家族の資料を集めて検討する。ヒトの常染色体は22対だから偶然だけでも5%くらいの確率で同じ染色体上に位置している。だから5%の有意水準を考えると($\frac{1}{22} \times 0.05 \cong 0.0023$)の対数は約-2.6になる)ロッド得点が3くらいになれば連鎖ありと判断してもよいだろうという意味である。

例1から例4は, 一般式

$$Z = \log 2^{s-1} \{ (1-\theta)^{a+d} \theta^{b+c} + (1-\theta)^{b+c} \theta^{a+d} \}$$

において, s, a, b, c, d にそれぞれの数値を代入してただちに求められる。

$$\text{例1と例2は } Z = \log_{10} 2^3 \{ (1-\theta)^4 + \theta^4 \}$$

$$\text{例3は } Z = \log_{10} 2^3 \{ 2\theta^2(1-\theta)^2 \}$$

$$\text{例4は } Z = \log_{10} 2^5 \{ (1-\theta)\theta^5 + (1-\theta)^5\theta \}$$

実はこれらの関数に θ のそれぞれの値を代入して求めたのが表2~4であり図で示したのが図3である。

2. ロッド得点の計算法

連鎖解析は必ずしもすべての交配型で可能とはかぎらない。少なくとも一方の親は二重ヘテロ($G_1G_2T_1T_2, GgT_1T_2$ または $GgTt$) でないといけない。しかし, この場合でももう一方の親のどの遺伝子型との組み合わせで連鎖解析に情報となりうるかは対象となる遺伝子座の対立遺伝子の発現様式(優性, 共優性, 劣性)に依存してくる。

解析にあたって最も情報が多くなるのは2つの遺伝子座がともに共優性となる場合で, 図4に数字をつけて示した二重ヘテロ親の縦横のすべての36の組み合わせで情報が得られることになる。本稿で対象となる2つの遺伝子座は病因を支配する遺伝子座とRFLPを示す遺伝子

	GT_1/GT_1	GT_1/GT_2	GT_1/gT_1	GT_2/GT_2	GT_1/gT_2 (C)	GT_2/gT_1 (R)	gT_1/gT_1	gT_2/GT_2	gT_2/gT_1	gT_2/gT_2
GT_1/GT_1					1	2				
GT_1/GT_2					3	4				
GT_1/gT_1					5	6				
GT_2/GT_2					7	8				
GT_1/gT_2 (C)	9	11	13	15	17	19	21	23	25	27
GT_2/gT_1 (R)	10	12	14	16	18	20	22	24	26	28
gT_1/gT_1					29	30				
gT_2/GT_2					31	32				
gT_2/gT_1					33	34				
gT_2/gT_2					35	36				

図4 連鎖解析に情報を与える交配型

2つの遺伝子座の対立遺伝子($G, g; T_1, T_2$)がともに相互優性(共優性)の場合、数字をつけた36すべての型で情報が得られる。しかし、RFLPの連鎖解析では多くの病因遺伝子座は完全優性か完全劣性で模様をほどこした24の型だけが情報をもつ。模様(縦線, 横線, 斜線, 点)は、本文中のロッド得点の計算にあたっての分類をあらわしている。Cは相引(coupling), Rは相反(repulsion)をあらわす。

座であるが、一般にRFLPはホモ接合体($T_1 T_1, T_2 T_2$)とヘテロ接合体($T_1 T_2$)は明瞭に区別されうるからRFLP遺伝子座については共優性とと考えてよい。しかし、一方の病因遺伝子座は一般に優性か劣性の場合が多く、この場合、図4の中で模様をほどこした24の組み合わせだけしか情報となりえない。この理由は、たとえば $GgT_1T_2 \times GGT_1T_1$ の交配型を考えると、二重ヘテロ親が相引(GT_1/gT_2)のときは期待される子供の表現型とその頻度は表5のようになつて GT_1 と GT_1T_2 が $\frac{1}{2} : \frac{1}{2}$ で組み換え率(θ)に全く依存しない定数になってしまうからである。情報とはならない他の2つの交配型についても表5に示している。

以下、病因遺伝子座は優性か劣性とと考えて図4に示したおのおの場合についてロッド得点の具体的な計算法を考えてみよう。2つの遺伝子座がともに共優性となる場合については最後に簡単にふれる。あらためて用いる記号の定義を示すと、病因遺伝子座の対立遺伝子は G (優性遺伝子), g (劣性遺伝子) および RFLP 遺伝子座の対立遺伝子を T_1, T_2 とする。ここで注意

しておきたいことは、病因遺伝子が優性か劣性(このときは正常な遺伝子の方が優性である)かということはロッド得点の計算には無関係である。要は、(GG, Gg)と gg が表現型として明瞭に区別されうるということであり、病因遺伝子が優性のときは GG と Gg が患者であり、劣性の場合には gg が患者となる。

a. $GgT_1T_2 \times ggT_1T_1$ および $GgT_1T_2 \times ggT_2T_2$ (図4の21, 22, 27, 28, 29, 30, 35, 36)のとき(表6参照)

これは単純なもどし交配に相当する組み合わせである。二世代資料の場合、二重ヘテロの親は相引(GT_1/gT_2)か相反(GT_2/gT_1)か判断できないので、どちらかである確率を $1/2$ として計算することになる。出現が期待される子供の表現型4種はすべて区別することが可能で、その出現頻度は、メンデル遺伝の検定交配のときと同じように、二重ヘテロ親の配偶子頻度に一致することになる。二重ヘテロ親が相引のときと相反のときとのちがいは、二重ヘテロ親の配偶子の期待頻度 $\theta/2$ と $(1-\theta)/2$ を単純に入れかえるとよいだけである。

表6 $GgT_1T_2 \times ggT_1T_1$ および $GgT_1T_2 \times ggT_2T_2$ の交配型で期待される子供の表現型とその頻度

a. 相引(GT_1/gT_2)のとき

	GT_1/gT_2 親の配偶子			
	GT_1 $\frac{1-\theta}{2}$	GT_2 $\frac{\theta}{2}$	gT_1 $\frac{\theta}{2}$	gT_2 $\frac{1-\theta}{2}$
観察される子供の数	a	b	c	d
gT_1/gT_1 親の配偶子	GT_1 $\frac{1-\theta}{2}$	GT_1T_2 $\frac{\theta}{2}$	gT_1 $\frac{\theta}{2}$	gT_1T_2 $\frac{1-\theta}{2}$
gT_2/gT_2 親の配偶子	GT_1T_2 $\frac{1-\theta}{2}$	GT_2 $\frac{\theta}{2}$	gT_1T_2 $\frac{\theta}{2}$	gT_2 $\frac{1-\theta}{2}$

b. 相反(GT_2/gT_1)のとき

	GT_2/gT_1 親の配偶子			
	GT_1 $\frac{\theta}{2}$	GT_2 $\frac{1-\theta}{2}$	gT_1 $\frac{1-\theta}{2}$	gT_2 $\frac{\theta}{2}$
観察される子供の数	a	b	c	d
gT_1/gT_1 親の配偶子	GT_1 $\frac{\theta}{2}$	GT_1T_2 $\frac{1-\theta}{2}$	gT_1 $\frac{1-\theta}{2}$	gT_1T_2 $\frac{\theta}{2}$
gT_2/gT_2 親の配偶子	GT_1T_2 $\frac{\theta}{2}$	GT_2 $\frac{1-\theta}{2}$	gT_1T_2 $\frac{1-\theta}{2}$	gT_2 $\frac{\theta}{2}$

子供の4種類の表現型が出現する期待頻度が表6のようにして求められたので、実際に調査したある家族(y_i)でこれらの表現型を示す子供の数が a, b, c, d ($a+b+c+d=s$) となる確率は組換え率を θ とすると多項分布の式をそのまま適用して

$$\begin{aligned}
 f(y_i; \theta) &= \frac{1}{2} \left\{ \frac{s!}{a!b!c!d!} \left(\frac{1-\theta}{2}\right)^a \left(\frac{\theta}{2}\right)^b \left(\frac{\theta}{2}\right)^c \left(\frac{1-\theta}{2}\right)^d \right. \\
 &\quad \left. + \frac{s!}{a!b!c!d!} \left(\frac{\theta}{2}\right)^a \left(\frac{1-\theta}{2}\right)^b \left(\frac{1-\theta}{2}\right)^c \left(\frac{\theta}{2}\right)^d \right\} \\
 &= K \cdot 2^{-1} \cdot 2^{-s} \cdot \{(1-\theta)^{a+d} \theta^{b+c} + (1-\theta)^{b+c} \theta^{a+d}\}
 \end{aligned} \quad (1)$$

を得る。ここで K は多項係数 $s! / a!b!c!d!$ をあらわしている。式の前項は相引のときに、後項は相反のときに期待される確率であり相引か相反かはそれぞれ確率 $1/2$ だからこの係数がつい

ている。

2つの遺伝子座が連鎖していないときにこのような子供構成の実測値が得られる確率は上式(1)で $\theta=1/2$ とおいて得られる。

$$f(y_i; \frac{1}{2}) = K \cdot 2^{-1} \cdot 2^{-s} \cdot \left(\frac{1}{2}\right)^s = K \cdot 2^{-1} \cdot 2^{-2s}$$

したがってロッド得点(Z)の値はこれらの比の対数をとって

$$\begin{aligned}
 Z &= \log_{10} \frac{f(y_i; \theta)}{f(y_i; \frac{1}{2})} \\
 &= \log_{10} 2^{s-1} \cdot \{(1-\theta)^{a+d} \theta^{b+c} \\
 &\quad + (1-\theta)^{b+c} \theta^{a+d}\}
 \end{aligned} \quad (2)$$

として求まる。

b. $GgT_1T_2 \times ggT_1T_2$ (図4の25, 26, 33, 34)のとき(表7参照)

ggT_1T_2 の親で組換えがおこってもおこらな

表 7 $GgT_1T_2 \times ggT_1T_2$ の交配型で期待される子供の表現型とその頻度

a. 相引(GT_1/gT_2)のとき

		GT_1/gT_2 親の配偶子				
		GT_1 $\frac{1-\theta}{2}$	GT_2 $\frac{\theta}{2}$	gT_1 $\frac{\theta}{2}$	gT_2 $\frac{1-\theta}{2}$	
gT_1/gT_2 親の 配偶子	$gT_1 \frac{1}{2}$	GT_1 $\frac{1-\theta}{4}$	GT_1T_2 $\frac{\theta}{4}$	gT_1 $\frac{\theta}{4}$	gT_1T_2 $\frac{1-\theta}{4}$	
	$gT_2 \frac{1}{2}$	GT_1T_2 $\frac{1-\theta}{4}$	GT_2 $\frac{\theta}{4}$	gT_1T_2 $\frac{\theta}{4}$	gT_2 $\frac{1-\theta}{4}$	
合 計		GT_1 $\frac{1-\theta}{2}$	GT_2 $\frac{\theta}{2}$	gT_1 $\frac{\theta}{2}$	gT_2 $\frac{1-\theta}{2}$	T_1T_2 は 除く
観察される子供の数		a	b	c	d	

b. 相反(GT_2/gT_1)のとき

		GT_2/gT_1 親の配偶子				
		GT_1 $\frac{\theta}{2}$	GT_2 $\frac{1-\theta}{2}$	gT_1 $\frac{1-\theta}{2}$	gT_2 $\frac{\theta}{2}$	
gT_1/gT_2 親の 配偶子	$gT_1 \frac{1}{2}$	GT_1 $\frac{\theta}{4}$	GT_1T_2 $\frac{1-\theta}{4}$	gT_1 $\frac{1-\theta}{4}$	gT_1T_2 $\frac{\theta}{4}$	
	$gT_2 \frac{1}{2}$	GT_1T_2 $\frac{\theta}{4}$	GT_2 $\frac{1-\theta}{4}$	gT_1T_2 $\frac{1-\theta}{4}$	gT_2 $\frac{\theta}{4}$	
合 計		GT_1 $\frac{\theta}{2}$	GT_2 $\frac{1-\theta}{2}$	gT_1 $\frac{1-\theta}{2}$	gT_2 $\frac{\theta}{2}$	T_1T_2 は 除く
観察される子供の数		a	b	c	d	

くても生じる配偶子は gT_1 か gT_2 でその確率はともに $1/2$ である。二重ヘテロ親については前と同様に相引のときと相反のときにわけて考える。この交配で注意せねばならないことは表現型 GT_1T_2 と gT_1T_2 の出現期待頻度は θ の項が消えてともに $1/4$ となり、組換え率 θ には依存しない (θ の関数とはならない) 定数値で連鎖解析には何らの情報をも提供しないことになってしまう。したがって子供のうちのこれらの表現型 T_1T_2 をもつ個体は除いて解析せねばならない。これは相反のとき ($GT_2/gT_1 \times gT_1/gT_2$) も全く同じである。結局、出現する子供の表現型とその頻度は ggT_1T_2 、親の配偶子が

gT_1 か gT_2 である場合をあわせると表 7 に示すように a. の場合と全く同じような形になる。だから、ロッド得点の計算では、解析に際して表現型 T_1T_2 をもつ子供を除くこと以外、a. と同じで式(2)となる。

c. $GgT_1T_2 \times GgT_1T_1$ および $GgT_1T_2 \times GgT_2T_2$ (図 4 の 5, 6, 13, 14, 23, 24, 31, 32) のとき (表 8 参照)

b. の例と同じように GgT_1T_1 が親のときには GT_1 か gT_1 の配偶子が等確率 $1/2$ で、また GgT_2T_2 が親のときも GT_2 か gT_2 の配偶子が等確率 $1/2$ で生じる。したがって、二重ヘテロ親が相引のときには表 8 に示すように子供の表

表8 $GgT_1T_2 \times GgT_1T_1$ および $GgT_1T_2 \times GgT_2T_2$ の交配型で期待される子供の表現型とその頻度

		GT_1/gT_2 親の配偶子			
		GT_1 $\frac{1-\theta}{2}$	GT_2 $\frac{\theta}{2}$	gT_1 $\frac{\theta}{2}$	gT_2 $\frac{1-\theta}{2}$
GT_1/gT_1 親の配偶子	GT_1 $\frac{1}{2}$	GT_1 $\frac{1-\theta}{4}$	GT_1T_2 $\frac{\theta}{4}$	GT_1 $\frac{\theta}{4}$	GT_1T_2 $\frac{1-\theta}{4}$
	gT_1 $\frac{1}{2}$	GT_1 $\frac{1-\theta}{4}$	GT_1T_2 $\frac{\theta}{4}$	gT_1 $\frac{\theta}{4}$	gT_1T_2 $\frac{1-\theta}{4}$
合計		$\frac{GT_1}{2-\theta}$ $\frac{1-\theta}{4}$	$\frac{gT_1}{\theta}$ $\frac{\theta}{4}$	$\frac{GT_1T_2}{1+\theta}$ $\frac{1+\theta}{4}$	$\frac{gT_1T_2}{1-\theta}$ $\frac{1-\theta}{4}$
観察される子供の数		a	b	c	d
GT_2/gT_2 親の配偶子	GT_2 $\frac{1}{2}$	GT_1T_2 $\frac{1-\theta}{4}$	GT_2 $\frac{\theta}{4}$	GT_1T_2 $\frac{\theta}{4}$	GT_2 $\frac{1-\theta}{4}$
	gT_2 $\frac{1}{2}$	GT_1T_2 $\frac{1-\theta}{4}$	GT_2 $\frac{\theta}{4}$	gT_1T_2 $\frac{\theta}{4}$	gT_2 $\frac{1-\theta}{4}$
合計		$\frac{GT_1T_2}{2-\theta}$ $\frac{1-\theta}{4}$	$\frac{gT_1T_2}{\theta}$ $\frac{\theta}{4}$	$\frac{GT_2}{1+\theta}$ $\frac{1+\theta}{4}$	$\frac{gT_2}{1-\theta}$ $\frac{1-\theta}{4}$
観察される子供の数		a	b	c	d

		GT_2/gT_1 親の配偶子			
		GT_1 $\frac{\theta}{2}$	GT_2 $\frac{1-\theta}{2}$	gT_1 $\frac{1-\theta}{2}$	gT_2 $\frac{\theta}{2}$
GT_1/gT_1 親の配偶子	GT_1 $\frac{1}{2}$	GT_1 $\frac{\theta}{4}$	GT_1T_2 $\frac{1-\theta}{4}$	GT_1 $\frac{1-\theta}{4}$	GT_1T_2 $\frac{\theta}{4}$
	gT_1 $\frac{1}{2}$	GT_1 $\frac{\theta}{4}$	GT_1T_2 $\frac{1-\theta}{4}$	gT_1 $\frac{1-\theta}{4}$	gT_1T_2 $\frac{\theta}{4}$
合計		$\frac{GT_1}{1+\theta}$ $\frac{1+\theta}{4}$	$\frac{gT_1}{1-\theta}$ $\frac{1-\theta}{4}$	$\frac{GT_1T_2}{2-\theta}$ $\frac{2-\theta}{4}$	$\frac{gT_1T_2}{\theta}$ $\frac{\theta}{4}$
観察される子供の数		a	b	c	d
GT_2/gT_2 親の配偶子	GT_2 $\frac{1}{2}$	GT_1T_2 $\frac{\theta}{4}$	GT_2 $\frac{1-\theta}{4}$	GT_1T_2 $\frac{1-\theta}{4}$	GT_2 $\frac{\theta}{4}$
	gT_2 $\frac{1}{2}$	GT_1T_2 $\frac{\theta}{4}$	GT_2 $\frac{1-\theta}{4}$	gT_1T_2 $\frac{1-\theta}{4}$	gT_2 $\frac{\theta}{4}$
合計		$\frac{GT_1T_2}{1+\theta}$ $\frac{1+\theta}{4}$	$\frac{gT_1T_2}{1-\theta}$ $\frac{1-\theta}{4}$	$\frac{GT_2}{2-\theta}$ $\frac{2-\theta}{4}$	$\frac{gT_2}{\theta}$ $\frac{\theta}{4}$
観察される子供の数		a	b	c	d

現型としては GT_1 , gT_1 , GT_1T_2 , gT_1T_2 がそれぞれ $(2-\theta)/4$, $\theta/4$, $(1+\theta)/4$, $(1-\theta)/4$ の頻度で期待されることになる。相反のときも同じようにして計算される。子供の表現型と頻度が相引と相反のときとで計算できれば a の場合と同じようにして多項分布の式を適用し、ある家族 (i) でこのような子供構成が得られる確率は次式となる。

$$f(y_i; \theta) = \frac{1}{2} \left\{ \frac{s!}{ab!c!d!} \left(\frac{2-\theta}{4}\right)^a \left(\frac{\theta}{4}\right)^b \left(\frac{1+\theta}{4}\right)^c \left(\frac{1-\theta}{4}\right)^d + \frac{s!}{ab!c!d!} \left(\frac{1+\theta}{4}\right)^a \left(\frac{1-\theta}{4}\right)^b \left(\frac{2-\theta}{4}\right)^c \left(\frac{\theta}{4}\right)^d \right\}$$

$$= K \cdot 2^{-1} \cdot 4^{-s} \cdot \{ (2-\theta)^a \cdot \theta^b \cdot (1+\theta)^c \cdot (1-\theta)^d + (1+\theta)^a \cdot (1-\theta)^b \cdot (2-\theta)^c \cdot \theta^d \}$$

K は多項係数である。

連鎖してないときの確率は $\theta=1/2$ とおいて次のようになる。

$$f(y_i; \frac{1}{2}) = K \cdot 2^{-1} \cdot 4^{-s} \cdot \left(\frac{1}{2}\right)^{s-1} \cdot 3^{a+c}$$

したがってロッド得点 (Z) の値はこれらの比の対数をとって

$$Z = \log_{10} \frac{f(y_i; \theta)}{f(y_i; \frac{1}{2})}$$

$$= \log_{10} \frac{2^{s-1}}{3^{a+c}} \{ (2-\theta)^a \theta^b (1+\theta)^c (1-\theta)^d + (1+\theta)^a (1-\theta)^b (2-\theta)^c \theta^d \}$$

として求まる。

d. 二重ヘテロ親間の交配 $GgT_1T_2 \times GgT_1T_2$

T_2 (図 4 の 17, 18, 19, 20) のとき (表 9 参照)

二重ヘテロの親からは 4 種類の配偶子が出現するので、期待される表現型とその頻度の計算はやや複雑になるが今までと同じような方法で求めることができる。二重ヘテロ親は相引か相反なので可能な交配の組み合わせは相引×相引 ($GT_1/gT_2 \times GT_1/gT_2$), 相引×相反 ($GT_1/gT_2 \times GT_2/gT_1$), 相反×相引 ($GT_2/gT_1 \times GT_1/gT_2$), 相反×相反 ($GT_2/gT_1 \times GT_2/gT_1$) の 4 つの場合があって、その確率はすべて等しく $1/4$ である。この中で相引×相反, 相反×相引は同じであることに注意する。それぞれの組み合わせで期待される子供の表現型とその頻度を整理すると表 9 の **c** となる。ここまで計

算すればあとは多項分布の式を適用するだけで項の数が 6 つにふえるから計算は複雑になる。面当なので多項係数 $s!/a!b!c!d!e!f!$ は最初から K であらわすことにする。

$$f(y_i; \theta) = \frac{1}{4} \cdot K \cdot \left(\frac{1-\theta^2}{4}\right)^a \left(\frac{\theta^2}{4}\right)^b \left(\frac{\theta(2-\theta)}{4}\right)^c$$

$$\left(\frac{(1-\theta)^2}{4}\right)^d \left(\frac{2(1-\theta+\theta^2)}{4}\right)^e \left(\frac{2\theta(1-\theta)}{4}\right)^f$$

$$+ \frac{1}{2} \cdot K \cdot \left(\frac{1-\theta+\theta^2}{4}\right)^a \left(\frac{\theta(1-\theta)}{4}\right)^b \left(\frac{1-\theta+\theta^2}{4}\right)^c$$

$$\left(\frac{\theta(1-\theta)}{4}\right)^d \left(\frac{1+2\theta-2\theta^2}{4}\right)^e \left(\frac{1-2\theta+2\theta^2}{4}\right)^f$$

$$+ \frac{1}{4} \cdot K \cdot \left(\frac{\theta(2-\theta)}{4}\right)^a \left(\frac{(1-\theta)^2}{4}\right)^b \left(\frac{1-\theta^2}{4}\right)^c$$

$$\left(\frac{\theta^2}{4}\right)^d \left(\frac{2(1-\theta+\theta^2)}{4}\right)^e \left(\frac{2\theta(1-\theta)}{4}\right)^f$$

$$= \frac{1^{s+1}}{4} \{ 2^{e+f} \cdot \theta^{2b+c+f} \cdot (1-\theta)^{a+2d+f} \cdot (1+\theta)^a \cdot (2-\theta)^c \cdot (1-\theta-\theta^2)^e + 2\theta^{b+d} \cdot (1-\theta)^{b+d} \cdot (1-\theta+\theta^2)^{a+c} \cdot (1+2\theta-2\theta^2)^e \cdot (1-2\theta+2\theta^2)^f + 2^{e+f} \cdot \theta^{a+2d+f} \cdot (1-\theta)^{2b+c+f} \cdot (1+\theta)^c \cdot (2-\theta)^a \cdot (1-\theta+\theta^2)^e \} \cdot K$$

$$f(y_i; \theta) = \left(\frac{1}{4}\right)^{s+1} \cdot 3^{a+c+e} \cdot 2^{e+f} \cdot 2^{-2s} \cdot 4 \cdot K$$

ロッド得点 Z の値はこれらの比の対数をとって次のようになる。

$$Z = \log_{10} \frac{f(y_i; \theta)}{f(y_i; \frac{1}{2})}$$

$$= \log_{10} \frac{4^{s-1}}{3^{a+c+e} 2^{e+f}} \times \{ 2^{e+f} \cdot \theta^{2b+c+f} \cdot (1-\theta)^{a+2d+f} \cdot (1+\theta)^a \cdot (2-\theta)^c \cdot (1-\theta-\theta^2)^e + 2\theta^{b+d} \cdot (1-\theta)^{b+d} \cdot (1-\theta+\theta^2)^{a+c} \cdot (1+2\theta-2\theta^2)^e \cdot (1-2\theta+2\theta^2)^f + 2^{e+f} \cdot \theta^{a+2d+f} \cdot (1-\theta)^{2b+c+f} \cdot (1-\theta)^c \cdot (2-\theta)^a \cdot (1-\theta+\theta^2)^e \}$$

e. 2 つの遺伝子座がともに共優性を示すとき (表 10 参照)

詳しい説明は省略するが、図 4 に示した 36 の交配型のうち二重ヘテロ親間の交配を除く 32 のすべての組み合わせで、期待される子供の表現型の頻度は $(1-\theta)/2$, $\theta/2$, $\theta/2$, $(1-\theta)/2$ となり、ロッド得点の値は式(2)を用いて求めることができる。二重ヘテロ親間の交配については、上の **d.** に準じて複雑な計算を行うと子供の

表9 二重ヘテロ親間の交配($GgT_1T_2 \times GgT_1T_2$)で期待される子供の表現型とその頻度a. 相引×相引($GT_1/gT_2 \times GT_1/gT_2$)および相引×相反($GT_1/gT_2 \times GT_2/gT_1$)

		GT_1/gT_2 親の配偶子			
		GT_1 $\frac{1-\theta}{2}$	GT_2 $\frac{\theta}{2}$	gT_1 $\frac{\theta}{2}$	gT_2 $\frac{1-\theta}{2}$
GT_1/gT_2 親の 配偶子	GT_1 $\frac{1-\theta}{2}$	$\frac{GT_1}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1T_2}{4}$ $\frac{(1-\theta)^2}{4}$
	GT_2 $\frac{\theta}{2}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_2}{4}$ $\frac{\theta^2}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta^2}{4}$	$\frac{GT_2}{4}$ $\frac{\theta(1-\theta)}{4}$
	gT_1 $\frac{\theta}{2}$	$\frac{GT_1}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta^2}{4}$	$\frac{gT_1}{4}$ $\frac{\theta^2}{4}$	$\frac{gT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$
	gT_2 $\frac{1-\theta}{2}$	$\frac{GT_1T_2}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{GT_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{gT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{gT_2}{4}$ $\frac{(1-\theta)^2}{4}$
GT_2/gT_1 親の 配偶子	GT_1 $\frac{\theta}{2}$	$\frac{GT_1}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta^2}{4}$	$\frac{GT_1}{4}$ $\frac{\theta^2}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$
	GT_2 $\frac{1-\theta}{2}$	$\frac{GT_1T_2}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{GT_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_2}{4}$ $\frac{(1-\theta)^2}{4}$
	gT_1 $\frac{1-\theta}{2}$	$\frac{GT_1}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{gT_1}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{gT_1T_2}{4}$ $\frac{(1-\theta)^2}{4}$
	gT_2 $\frac{\theta}{2}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_2}{4}$ $\frac{\theta^2}{4}$	$\frac{gT_1T_2}{4}$ $\frac{\theta^2}{4}$	$\frac{gT_2}{4}$ $\frac{\theta(1-\theta)}{4}$

b. 相反×相反($GT_2/gT_1 \times GT_2/gT_1$)

		GT_2/gT_1 親の配偶子			
		GT_1 $\frac{\theta}{2}$	GT_2 $\frac{1-\theta}{2}$	gT_1 $\frac{1-\theta}{2}$	gT_2 $\frac{\theta}{2}$
GT_2/gT_1 親の 配偶子	GT_1 $\frac{\theta}{2}$	$\frac{GT_1}{4}$ $\frac{\theta^2}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1T_2}{4}$ $\frac{\theta^2}{4}$
	GT_2 $\frac{1-\theta}{2}$	$\frac{GT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_2}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{GT_1T_2}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{GT_2}{4}$ $\frac{\theta(1-\theta)}{4}$
	gT_1 $\frac{1-\theta}{2}$	$\frac{GT_1}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{GT_1T_2}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{gT_1}{4}$ $\frac{(1-\theta)^2}{4}$	$\frac{gT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$
	gT_2 $\frac{\theta}{2}$	$\frac{GT_1T_2}{4}$ $\frac{\theta^2}{4}$	$\frac{GT_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{gT_1T_2}{4}$ $\frac{\theta(1-\theta)}{4}$	$\frac{gT_2}{4}$ $\frac{\theta^2}{4}$

c. 合計

子供の表現型	GT_1	gT_1	GT_2	gT_2	GT_1T_2	gT_1T_2
子供の数	a	b	c	d	e	f
相引×相引	$\frac{1-\theta^2}{4}$	$\frac{\theta^2}{4}$	$\frac{\theta(2-\theta)}{4}$	$\frac{(1-\theta)^2}{4}$	$\frac{2(1-\theta+\theta^2)}{4}$	$\frac{2\theta(1-\theta)}{4}$
相引×相反 } 相反×相引 }	$\frac{1-\theta+\theta^2}{4}$	$\frac{\theta(1-\theta)}{4}$	$\frac{1-\theta+\theta^2}{4}$	$\frac{\theta(1-\theta)}{4}$	$\frac{1+2\theta-2\theta^2}{4}$	$\frac{1-2\theta+2\theta^2}{4}$
相反×相反	$\frac{\theta(2-\theta)}{4}$	$\frac{(1-\theta)^2}{4}$	$\frac{1-\theta^2}{4}$	$\frac{\theta^2}{4}$	$\frac{2(1-\theta+\theta^2)}{4}$	$\frac{2\theta(1-\theta)}{4}$

表 10 2つの遺伝子座が相互優性(共優性)のときに期待される
子供の表現型とその頻度

		$G_1 T_1 / G_2 T_2$ 親の配偶子				
		$G_1 T_1$ $\frac{1-\theta}{2}$	$G_1 T_2$ $\frac{\theta}{2}$	$G_2 T_1$ $\frac{\theta}{2}$	$G_2 T_2$ $\frac{1-\theta}{2}$	
観察される子供の数		a	b	c	d	
$G_1 T_1 / G_1 T_1$ 親の 配偶子	$G_1 T_1$	$G_1 T_1$ $\frac{1-\theta}{2}$	$G_1 T_1 T_2$ $\frac{\theta}{2}$	$G_1 G_2 T_1$ $\frac{\theta}{2}$	$G_1 G_2 T_1 T_2$ $\frac{1-\theta}{2}$	
	$G_1 T_2$	$G_1 T_1 T_2$ $\frac{1-\theta}{2}$	$G_1 T_2$ $\frac{\theta}{2}$	$G_1 G_2 T_1 T_2$ $\frac{\theta}{2}$	$G_1 G_2 T_2$ $\frac{1-\theta}{2}$	
$G_1 T_2 / G_1 T_2$ 親の 配偶子	$G_1 T_1$	$G_1 T_1 T_2$ $\frac{1-\theta}{2}$	$G_1 T_2$ $\frac{\theta}{2}$	$G_1 G_2 T_1$ $\frac{\theta}{2}$	$G_1 G_2 T_1 T_2$ $\frac{1-\theta}{2}$	
	$G_1 T_2$	$G_1 T_1 T_2$ $\frac{1-\theta}{2}$	$G_1 T_2$ $\frac{\theta}{2}$	$G_1 G_2 T_1 T_2$ $\frac{\theta}{2}$	$G_1 G_2 T_2$ $\frac{1-\theta}{2}$	
$G_2 T_1 / G_2 T_1$ 親の 配偶子	$G_2 T_1$	$G_1 G_2 T_1$ $\frac{1-\theta}{2}$	$G_1 G_2 T_1 T_2$ $\frac{\theta}{2}$	$G_2 T_1$ $\frac{\theta}{2}$	$G_2 T_1 T_2$ $\frac{1-\theta}{2}$	
	$G_2 T_2$	$G_1 G_2 T_1 T_2$ $\frac{1-\theta}{2}$	$G_1 G_2 T_2$ $\frac{\theta}{2}$	$G_2 T_1 T_2$ $\frac{\theta}{2}$	$G_2 T_2$ $\frac{1-\theta}{2}$	
$G_1 T_1 / G_1 T_2$ 親の 配偶子	$G_1 T_1 \frac{1}{2}$	$G_1 T_1$ $\frac{1-\theta}{4}$	$G_1 T_1 T_2$ $\frac{\theta}{4}$	$G_1 G_2 T_1$ $\frac{\theta}{4}$	$G_1 G_2 T_1 T_2$ $\frac{1-\theta}{4}$	$T_1 T_2$ は 除く
	$G_1 T_2 \frac{1}{2}$	$G_1 T_1 T_2$ $\frac{1-\theta}{4}$	$G_1 T_2$ $\frac{\theta}{4}$	$G_1 G_2 T_1 T_2$ $\frac{\theta}{4}$	$G_1 G_2 T_2$ $\frac{1-\theta}{4}$	
$G_1 T_1 / G_2 T_1$ 親の 配偶子	$G_1 T_1 \frac{1}{2}$	$G_1 T_1$ $\frac{1-\theta}{4}$	$G_1 T_1 T_2$ $\frac{\theta}{4}$	$G_1 G_2 T_1$ $\frac{\theta}{4}$	$G_1 G_2 T_1 T_2$ $\frac{1-\theta}{4}$	$G_1 G_2$ は 除く
	$G_2 T_1 \frac{1}{2}$	$G_1 G_2 T_1$ $\frac{1-\theta}{4}$	$G_1 G_2 T_1 T_2$ $\frac{\theta}{4}$	$G_2 T_1$ $\frac{\theta}{4}$	$G_2 T_1 T_2$ $\frac{1-\theta}{4}$	
$G_1 T_2 / G_2 T_2$ 親の 配偶子	$G_1 T_2 \frac{1}{2}$	$G_1 T_1 T_2$ $\frac{1-\theta}{4}$	$G_1 T_2$ $\frac{\theta}{4}$	$G_1 G_2 T_1 T_2$ $\frac{\theta}{4}$	$G_1 G_2 T_2$ $\frac{1-\theta}{4}$	$G_1 G_2$ は 除く
	$G_2 T_2 \frac{1}{2}$	$G_1 G_2 T_1 T_2$ $\frac{1-\theta}{4}$	$G_1 G_2 T_2$ $\frac{\theta}{4}$	$G_2 T_1 T_2$ $\frac{\theta}{4}$	$G_2 T_2$ $\frac{1-\theta}{4}$	
$G_2 T_1 / G_2 T_2$ 親の 配偶子	$G_2 T_1 \frac{1}{2}$	$G_1 G_2 T_1$ $\frac{1-\theta}{4}$	$G_1 G_2 T_1 T_2$ $\frac{\theta}{4}$	$G_2 T_1$ $\frac{\theta}{4}$	$G_2 T_1 T_2$ $\frac{1-\theta}{4}$	$T_1 T_2$ は 除く
	$G_2 T_2 \frac{1}{2}$	$G_1 G_2 T_1 T_2$ $\frac{1-\theta}{4}$	$G_1 G_2 T_2$ $\frac{\theta}{4}$	$G_2 T_1 T_2$ $\frac{\theta}{4}$	$G_2 T_2$ $\frac{1-\theta}{4}$	

表現型は 9 つ出てくる。興味ある方は Morton (1955) の論文¹⁾ の表 8 を参照していただきたい。

3. RFLP マーカーの選択

a. RFLP

ゲノムを構成する塩基 270 当りに 1 個は多型がみられるといわれている²⁾。さらにこの内

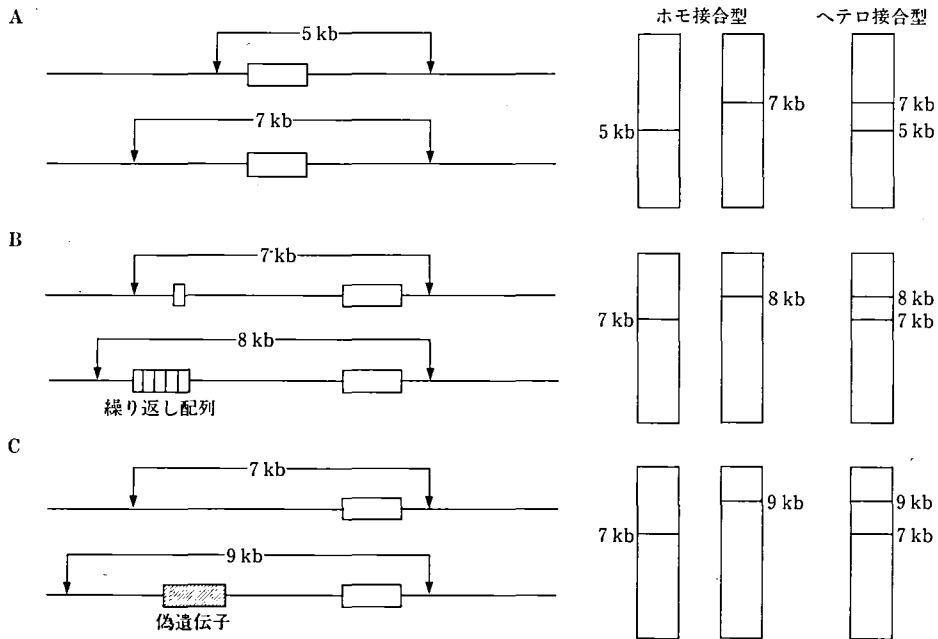


図5 RFLP をきたすゲノムの構造変化

A: 制限酵素切断パターンに変化をきたす塩基置換, B: 繰り返し配列のコピー数の相違, C: 偽遺伝子の存在の有無.

のいくつかにおいては、制限酵素の認識配列に変化をきたす。したがってこれらの塩基配列の多型は、ゲノム DNA を適当な制限酵素で切断後、Southern ブロットニングを行うことにより容易に検出でき、このような多型は制限酵素切断片多型(restriction fragment length polymorphisms; RFLPs)と呼ばれている。これは塩基の置換だけでなく、あるまとまった領域の付加や欠失によってもおこる。

図5にRFLPをきたしうるゲノムの構造上の変化につき模式的に示した。一つは制限酵素の認識部位に変化をきたす塩基置換の多型に基づくものである(A)。もう一つは繰り返し配列のコピー数の多型に基づくものである(B)³⁾。そしてまれであるが、個体によって偽遺伝子が存在する場合としない場合とがあり、これに基づくRFLPも報告されている(C)⁴⁾。これらRFLPはメンデルの遺伝様式に従って子孫に伝えられること、遺伝子産物のマーカーに比べ高頻度であること、検出が簡単であることなどの利点があり、1980年にBotsteinらによつ

てRFLPをマーカーとして用いればヒト染色体の遺伝子地図作成が可能であることが指摘され⁵⁾、この手法は現在では病因遺伝子が同定されていない遺伝病の連鎖解析の主役の座を占めている。この際問題となるのは、適当なRFLPをきたすプローブの選択であり、これについて以下記載することにする。

b. プローブ単離の概略

プローブの単離に際しまず考慮しなければならないことは、そのプローブが繰り返し配列をもっていないかどうかということである。ヒトゲノム DNA には約5 kbごとに *Alu* ファミリーや、約300 kbごとに *L1* ファミリーなどの繰り返し配列が存在することが知られており、これらがプローブに含まれていると多くの場合Southern ブロットニングはスミアーになってしまう。したがってまず制限酵素で切断したゲノム DNA を適当なファージベクターやプラスミドベクターにライゲーションしライブラリーを作製する。次にラベルした全ゲノム DNA を用いてプラークもしくはコロニーハイブリダイ

表 11 種々の遺伝子型をもつ両親から生まれた子供の遺伝子型の情報としての有効性

患者の遺伝子型	健康な配偶者の遺伝子型	情報となる子の確率	交配の頻度
$T_i T_i$	$T_i T_i$	1.0	$2p_i p_j (1 - p_i - p_j)^2$
$T_i T_i$	$T_i T_i$ もしくは $T_i T_j$	1.0	$2p_i p_j (1 - p_i - p_j) (2p_i + 2p_j)$
$T_i T_i$	全てホモ接合型	1.0	$2p_i p_j \sum_i p_i^2$
$T_i T_j$	$T_i T_j$	0.5	$2p_i p_j \times 2p_i p_j$
全てホモ接合型	全てホモ接合型	0	p_i^2

ゼーションを行う。もしクローン化されたゲノム DNA が繰り返し配列を有する場合、対応するプラークもしくはコロニーが X 線フィルム上、シグナルとして確認できる。このようにして選択したクローンから DNA を単離し、ライブラリーの作製に用いた制限酵素により切断後、ラベルした全ゲノム DNA をプローブとして Southern ブロッキングにより再び繰り返し配列の有無を確認する。以上の手順を踏んで単離したクローンをプローブとすることにより、正常者のゲノム DNA を各種の制限酵素を用いて切断後、Southern ブロッキングにより RFLP の有無を検討する。

以上がプローブ単離の大まかな手順であるが、以下いくつかのポイントについて述べることにする。

c. DNA ソース

DNA ソースとしては、末梢血白血球や胎盤などの比較的ゲノム DNA の単離が容易な組織が用いられる。これ以外でも目的とする遺伝子がどの染色体に位置しているかが知られている場合には、その染色体をもつヒトとマウスの雑種細胞や、セルソーターで分離した染色体 DNA が用いられることもある。

d. ライブラリーの作製とプローブの単離

DNA のサイズが大きいかほど RFLP の数も多くなるが、それにともない繰り返し配列が出現する頻度もあがることになる。一般にプラスミドベクターによるライブラリーでは 5 kb 前後の、ファージベクターの場合は 10~20 kb のインサートをもつクローンが得られる。しかしプローブと高濃度のゲノム DNA とをプレハイブリダイゼーションすることにより、繰り返し配

列がマスクされたプローブ(repeat-blocked probes)を得ることができることがわかり⁶⁾、最近では 35~40 kb のフラグメントがクローンできるコスミドをベクターとしたライブラリーが使用されることが多くなった。また *recB*, *recC*, *sbcB* または *recD* などの宿主を用いると多型性の高い領域が比較的効率よく得られることが知られており、これを宿主としてファージやコスミドライブラリーを作製しこれからクローンを得ることも行われている⁷⁾。しかしこの場合クローン化される領域にバイアスがかかることがあり注意を要する。

e. RFLP の検討

さてこのようにして単離したプローブが、多型性を示すかどうかを検討するには、どのような制限酵素を用いて、しかも何人程度を対象にすべきであろうか。これまでの報告によると RFLP をきたす制限酵素としては、*MspI* と *TaqI* がとりわけ多く、それぞれ認識部位は CCGG と TCGA といずれも CG 配列をもっている。これは動物細胞では CG 配列のシトシンの 5 位の炭素がメチル化を受けやすく、その後脱アミノを受けチミンに変化し、CG-TG もしくは CG-CA への置換をきたすためと考えられている⁸⁾。したがってこれらの酵素によるチェックはまず試みるべきである。そのほかの酵素としては、頻度や経済性から考えて、*HindIII*, *EcoRI*, *BamHI*, *PstI*, *PvuII*, *BglIII*, *RsaI* などがよく用いられている。次に DNA 提供者の数であるが、DNA 連鎖解析に有用な RFLP を得るためにはある程度頻度の高いものが必要であり、これを考慮すると 10 人程度でよいと考えられる。

次にこのようにして見いだされた RFLP について、少なくとも一つの家系においてこれがメンデルの遺伝様式にのっとって遺伝するかどうかをチェックする。そしてなるべく多くの検体について、それぞれの遺伝子頻度を明らかにする。最後にこのようにして単離された RFLP をきたすクローンについて、ヒトマウスの雑種細胞ならびに *in situ* ハイブリッド形成法を用いて染色体上の位置を決定する。

f. Polymorphism information content (PIC)

PIC はこのようにして得られた RFLP マーカーの連鎖解析への有用性を示す指標として、Botstein らにより定義されたものである⁹⁾。言い換えると、これはそのマーカーを用いた場合、病因遺伝子をもった親から生まれた子供のマーカーに関する情報をもとに、両親のマーカーの遺伝子型が推定できる確率と考えることができる。ここであるマーカーについて n 個のアレルがあり ($T_1, T_2, \dots, T_i, T_j, T_k, T_l, \dots, T_n$), i, j, k, l 番目のアレルの生じる頻度をそれぞれ p_i, p_j, p_k, p_l とすると、両親の遺伝子型とその子供のマーカーについての情報の有効性、およびそのような交配(結婚)が生じる頻度は表 11 のようにまとめることができる。つまり両親のマーカーについての遺伝子型がいずれも T_i, T_j の場合、その半数の子供の情報が無効になり、もし病因遺伝子をもった親がマーカーについてホモ接合型であれば、その配偶者の遺伝子型によらずその子供の情報はすべて無効である。一方これら 2 つの場合以外では、子供の情報はすべて有効である。以上から PIC は次のようにして計算できる。

$$PIC = 1 - \left(\sum_{i=1}^n p_i^2 \right) - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

PIC は 1 に近いほど informative であり、PIC > 0.5 の場合を highly informative, 0.5 > PIC > 0.25 の場合を reasonably informative そして PIC < 0.25 の場合を slightly informative と考えることができる。つまりマーカーとしてはより多くのアレルを有し、さらにそ

れらの頻度にかたよりのないものが多い訳で、このような PIC が高くなるようなプローブの選択が連鎖解析においては重要である。

g. PIC の高いプローブの単離

以上述べた操作で得られるプローブは、多くの場合制限酵素の認識部位の塩基置換に基づくものであり、認識部位の有無という dimorphic なマーカーに過ぎず、ヘテロ接合体である頻度 (heterozygosity) は 50% を超えることはなく、実際はそれよりはるかに低いことが多い。

そこでより高い確率で heterozygosity が得られるプローブが望まれるわけであるが、これに関しては 2 つの場合が考えられる。ゲノム中には比較的短い配列 (11~60 bp) が繰り返して存在することが知られミニサテライトと呼ばれている。Jeffreys らはヒトミオグロビン遺伝子のイントロンにみられる 33 bp の繰り返し配列がヒトゲノムの他のミニサテライト配列に共通なコア配列を有していることを明らかにして、これらの繰り返しを 600~800 bp ほど接続してプローブとして用いると高い多型性を示すことを見いだした⁹⁾。これは特定の部位にみられるこのような配列の繰り返し回数が、個体間で異なっていることが多いためである。このような繰り返し配列の数についての多型を VNTR (variation in the number of tandem repeats) と呼んでいるが、VNTR を示すプローブは先に述べたランダムプローブよりもより系統的に調製できる。つまり人工的に合成した繰り返し配列をプローブとしてゲノム DNA をスクリーニングすることにより、ランダムに単離したクローンを検討する場合にくらべ、約 5 倍から 10 倍の頻度で多型性を示すクローンを得ることができるからである¹⁰⁾。

もう 1 種類の高頻度に heterozygosity を示すプローブであるが、複数の制限酵素に対して RFLP が認められる場合であり、たとえば 2 種類の異なる制限酵素に対する多型を示すプローブでは、一つの制限酵素に対して 4 つの異なる長さをきたすプローブと同じ情報量を有しているといえる。このことはより長いサイズのプローブを用いることにより得られるはずで、コス

ミドクローンがプローブとして有利である。この際繰り返し配列の頻度も上がるが、これをクリアーする方法(repeat-blocked probe)については、d. で述べた。

h. 連鎖解析に必要なプローブの数およびプローブリスト

ゲノムの 20 cM ごとにマーカーがあれば、つまり RFLP をきたす 150 個程度のプローブがあれば如何なる遺伝子を考えても、少なくとも 10 cM 以内にマーカーが存在することになり連鎖解析が可能である⁹⁾。しかし実際はプローブの分布に片寄りがあるわけで、このことを考慮に入れ、Lange らはヒトゲノムの連鎖地図作成には約 1500 個のプローブが必要であると推定している¹¹⁾。実際 Donis-Keller らは独自に単離した PIC の高いプローブに既知のプローブを加えた合計 403 個のプローブを用いて 3 世代にわたる 21 家系を解析することにより、ヒトゲノム DNA の少なくとも 95 % についての連鎖地図を作成した¹²⁾。1987 年にパリで行われた Human Gene Mapping に関する会議によると、これまですでに 610 個の遺伝子がクローンされており、そのうち 216 個に RFLP が見いだされている。また既知の遺伝子以外の領域については 2057 個のプローブが単離されており、このうち 977 個が RFLP を示す。したがって、RFLP を示すクローンとしては 1193 個が知られていることになる。これらのプローブについては文献 13 に記載されており、おのおのについての availability が 6 段階で示されている。

4. RFLP マーカーを用いた連鎖解析

いくつかの家系において、以上のようにして得られた RFLP マーカーと病因遺伝子との連鎖解析を、先に述べたロッド得点を指標にして行うわけであるが、この際どれくらいの数の検体が必要であろうか。これは利用できる家族の数や構成、マーカーの PIC、さらにはマーカーと病因遺伝子との距離によっても異なってくる。Botstein らはロッド得点として 3.0 を得るには 300 検体が必要であるが、この際もしマカ

ーと病因遺伝子との高い連鎖が存在する 5 人の子供からなる理想的な家族が 3 つあれば、このような家族全員のつまり計 21 個の検体を解析すればよいとしている⁹⁾。検体の数を増やせない場合は、PIC を高めるためにマーカーの数を増やせば必要な検体数を減らすことができる。具体的には 6000 回の解析(検体数×マーカー数)を行えばロッド得点 3.0 を得ることができるといふ⁹⁾。

これまでこの手法を用いて病因遺伝子が不明な多くの遺伝病の解析が行われている。そのような疾患のリストや実例については特集号の他の稿で詳しく述べられるために、それに譲ることにする。ここでは最近のトピックについて触れておきたい。

従来の家系分析で解析可能な 2 つの座位間の距離は、組換え頻度の点からいってせいぜい 1 cM (1000 kb) 程度といわれている。しかし近年これを超えるような方法が開発された。それは微量の DNA を酵素的に増幅するいわゆる polymerase chain reaction (PCR) 法を用いたもので(本紙別稿参照)、この方法によると精子 1 個から単離した DNA について遺伝子型を決定できる。つまり精子 1 個を顕微鏡下で細いプラスチックの針を用いてチューブに取り、プロテアーゼ K や SDS を用いて DNA を抽出する。そして連鎖解析の対象となる 2 つの座位をはさむ 2 個ずつのオリゴヌクレオチドプライマーを用意し、DNA をまず 10 分間 95°C で 10 分間加熱して変性後これらのプライマーをアニールして *Taq* ポリメラーゼを用いて DNA 合成を行う。以後 95°C, 15 秒, 54°C, 15 秒そして 72°C, 1 分と変性、アニールおよび DNA 合成からなるサイクルを 50 回繰り返すことにより目的とする領域を 7.6×10^{10} 倍に増幅できる。この DNA を用いて 2 つの座位に特異的なオリゴヌクレオチドプローブ(allele-specific oligonucleotide probe)とのドットハイブリダイゼーションにより連鎖を調べる。

モデル実験として、第 6 染色体にある HLA の DQ- α と第 19 染色体にある LDL レセプターとの多型がそれぞれヘテロ接合型になってい

る一個人の精子 150 個についてこの解析が行われた¹⁴⁾。その結果、両座位についてシグナルが得られた 70 個の精子 DNA において、各多型の可能な 4 つの組み合わせがほぼ同頻度でみられた。これは両座位が異なる染色体上に位置することから予想される結果であり、この方法が信頼性のあるものであることを示している。1 個の精子を単離しかつ各試料間で PCR による増幅に差がでないような DNA 抽出の方法についてまだ改善の余地はあるものの、この方法によると 1 週間で約 500 回の減数分裂による産物を解析可能であり、比較的近い連鎖関係にある 2 つの座位間の距離を、ただし 2 つの遺伝子座が DNA レベルで検出できる RFLP などを示す場合に限られるが、家系分析よりも高い精度で推定できるものと期待される。

おわりに

RFLP を用いた連鎖解析が適用できる家系の条件として、少なくとも一方の親は二重ヘテロ接合体であること、また検討できる子供が 2 人以上であること、さらに 2 世代以上にわたる家系であることなどが挙げられる。またロード得点法による連鎖解析に影響を与える因子として、男女間での組換え率の相違、疾患の浸透率や、疾患のエンティティーの異質性などが考えられる。このような問題点はあるものの、すでに 10 をこえる病因遺伝子の不明な遺伝病に連鎖したマーカーが見いだされ、これによる診断が試みられている。今後さらに RFLP マーカーの数が増加し、多くの遺伝病の診断が可能になり、その病因遺伝子の単離、解析へと進み、最終的にはこれらの疾患の治療法や予防法の確立へとつながることが期待される。

文 献

- 1) Morton, N. E.: Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* 7: 277-318, 1955.
- 2) Cooper, D. N. et al.: An estimate of unique DNA sequence heterozygosity in the human genome. *Hum. Genet.* 69: 201-205, 1985.
- 3) Bell, G. I. et al.: The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 295: 31-35, 1982.
- 4) Anagnou, N. P. et al.: Chromosomal organization of the human dihydrofolate reductase genes: Dispersion, selective amplification, and a novel form of polymorphism. *Proc. Natl. Acad. Sci. USA* 81: 5170-5174, 1984.
- 5) Botstein, D. et al.: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331, 1980.
- 6) Schumm, J. W. et al.: Identification of more than 500 RFLPs by screening random genomic clones. *Am. J. Hum. Genet.* 42: 143-159, 1988.
- 7) Wyman, A. R. et al.: Factors which equalize the representation of genomic segments in recombinant libraries. *Gene* 49: 263-271, 1986.
- 8) Bird, A. P.: DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* 8: 1499-1504, 1980.
- 9) Jeffreys, A. J. et al.: Hypervariable 'minisatellite' regions in human DNA. *Nature* 314: 67-73, 1985.
- 10) Nakamura, Y. et al.: Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* 235: 1616-1622, 1987.
- 11) Lange, K. & Boehnke, M.: How many polymorphic genes will it take to span the human genome? *Am. J. Hum. Genet.* 34: 842-845, 1982.
- 12) Donis-Keller, H. et al.: A genetic linkage map of the human genome. *Cell* 51: 319-337, 1987.
- 13) Pearson, P. L. et al.: Human gene mapping by recombinant DNA techniques. *Cytogenet. Cell. Genet.* 46: 390-566, 1987.
- 14) Li, H. et al.: Amplification and analysis of DNA sequences in single human sperm and diploid cells. *Nature* 335: 414-417, 1988.