5. 多重検定についての考え方および解決策

5.1 はじめに

疾患原因遺伝子同定のための関連解析と言えば、複数の遺伝子多型を同時に分析の対象とするのが一般的であろう。近年は特に、数十万単位の SNP の情報を用いた大規模なゲノムワイド関連解析が注目を集めていることもあり、多重検定の考え方に基づいて、適正な有意水準を定めることが、ますます重要な課題の一つとなりつつある。

なぜ有意水準の補正が必要なのか、Šidák（1967）の手法に基づいて、有意水準を 5%とした場合を考えてみよう。例えば、互いに独立である二つの SNP について関連解析を行なうと、少なくともどちらか一方の SNP でタイプ I エラー（真である帰無仮説を誤って棄却してしまうこと、ここでは、実際には疾患と関連のない SNP を、関連ありと判定してしまうこと）をおかす確率は $1 - (1 - 0.05)^2 = 0.0975$ となり、有意水準を 5%と定めているにもかかわらず、実際のタイプ I エラーの確率は 10%近い値となってしまう。10 セットの SNP ともなれば、$1 - (1 - 0.05)^{10} \fallingdotseq 0.4013$ で、実に 4 割の確率で一つ以上のタイプ I エラーをおかしてしまうことになり、こうなればもはや有意性検定そのものが無意味と言える。

つまり、同時に複数の検定を行なう場合、通常の有意水準では不十分であり、すべての仮説のうち、少なくともいずれか一つでタイプ I エラーが生じる確率（family-wise error rate; FWER）を一定の水準に抑えるべく、単一検定の場合よりも、有意水準を高く設定することが必要である。本章では、多重検定における有意水準の設定について、現在までに提唱されている手法やその特徴などを概説する。

5.2 Bonferroni の補正法

Bonferroni の補正法は、多重検定で最も一般的に用いられる手法である。これは、有意水準 α を検定の数 N で除した値$\frac{\alpha}{N}$を、新たな有意水準 α'とするものである。例えば、10 の SNP を解析対象とし、有意水準を 5%と設定した場合、$\alpha' = \frac{\alpha}{N} = \frac{0.05}{10} = 0.005$ となり、個々の SNP が疾患と有意に関連していると判断されるためには、0.5%よりも高い有意性を示す必要がある。個々の検定における有意確率（帰無仮説を真としたときに、その検定統計量がデータから得られる確率。以下、p 値とする）と N の積が、有意水準を下回るかというように、逆に考えてもよい。なお、Bonferroni の補正法に近似的に一致するものとして、前述の Šidák（1967）の手法が挙げられ、有意水準は以下のような式（式 1）により求められる。

$$\alpha' = 1 - (1 - \alpha)^{\frac{1}{N}} \qquad \text{（式 1）}$$

上記の例では、$\alpha' = 1 - (1 - 0.05)^{\frac{1}{10}} \fallingdotseq 0.0051$ となる。

ただし、この補正法の大きな問題点は、一般に保守的な（＝過度に高い）有意水準を与えることであり、その理由は大別して二点挙げられる。一つは、実際のデータの多くで、仮説間に相関が認められるためである。SNP に限らず、遺伝子多型のタイピングデータで、互いに近傍に位置するもの同士であれば、連鎖不平衡の影響により、独立でないケースがほとんどである。極端な例ではあるが、10 の SNP がすべて完全な連鎖不平衡の関係にあり、どれか一つで、それらの SNP 全体が有する情報をカバーし得るとする。この場合、疾患との関連解析における検定統計量はどれもまったく同じ値を示すため、SNP の数がいくつであっても、実質的には単一検定であり、有意水準の補正の必要はない。しかし Bonferroni の補正法では、このような非独立性は考慮されないため、10 という数によって、本来必要のない補正がなされ、補正前とは逆に、タイプ II エラー（棄却されるべき帰無仮説を誤って採択してしまうこと。ここでは、疾患と有意に関連する SNP を、

関連なしと判定してしまうこと）が多数生じることになる。Moran（2003）は、以上の理由などから、Bonferroni の補正法を用いるべきではないとする見解を、かなり明確に示している。

二点目は、帰無仮説では説明できない、真の関連の存在である。仮説検定における p 値は本来、0 から 1 の範囲で一様分布を示す、すなわち、0 から 1 までの任意の値を等確率で取る指標と定義されている。ただし、その定義が成り立つためには、すべての検定において帰無仮説が真であるという前提条件が満たされていなければならず、一般的な有意水準を超える結果が得られたとしても、それらはすべて、帰無仮説における確率分布に基づいて生じた偽陽性ということになる。しかし、実際のデータにおいて有意とされた結果の一部は、偽陽性でない、真に存在する何らかの関連を示すものと考えられ（その関連を探し当てるのが、研究者の務めであるのだが）、その場合の p 値は、一様分布から 0 の方向に頻度が偏った分布を示す。したがって、有意水準の保守化を回避するため、厳密には N からその分を差し引いた数を補正に用いなければならない。この点に関しては、次項で詳細に述べる。

Rice（1989）は、N に代わり、$N - i + 1$（i は当該 SNP の有意性の強さの順位）で α を除す sequential Bonferroni 補正法を提唱している。しかし、Bonferroni の補正法ほどではないにしろ、その強い保守性には大差ない。

なお、Moran（2003）のコメントを受けて、Neuhäuser（2004）は、truncated product method（TPM）を妥協案として用いることを提唱している。これは Zaykin ら（2002）の手法をベースとしており、個々の p 値よりも、データ全体の p 値の生起確率に着目したアプローチである。TPM における p 値は、式（式 2）で求められる。

$$\Pr(W_\tau \le w) = \sum_{i=1}^{N} \binom{N}{i} (1-\tau)^{N-i} \times \left(w \sum_{j=0}^{i-1} \frac{(i \ln \tau - \ln w)^j}{j!} I(w \le \tau^i) + \tau^i I(w \ge \tau^i) \right) \quad \text{（式 2）}$$

ここで、$I(.)$は、カッコ内が真のとき 1、それ以外では 0 となる指標関数、N は

検定の数、そして τ は有意水準（0.05 など）である。また W_τ は $\prod_{i=1}^{N} P_i^{I(P_i \le \tau)}$、すなわち、1 から N までの p 値のうち、有意なもののみの積であり、式（式 2）を用いた計算の際には、$w = W_\tau$ として行なう。なお、N が大きい場合には、モンテカルロ法を用いて近似解を得ることが可能である。

例えば、10 の検定のうち 4 つが有意水準 0.05 で有意であり、p 値がそれぞれ 0.01、0.02、0.03、および 0.03 であるとすると、TPM での p 値は 0.0011 となる。すなわち、個々の有意性はそれほど強くないものの、10 検定すべてにおいて帰無仮説が成立するとした場合に、上記のような結果が得られる確率は非常に低く、したがってこの 4 つのうち、少なくともいずれかは真に注目すべき結果であると解釈される。これらをすべて有意でないと判定する Bonferroni の補正法とは対照的である。

また、10 の検定のうち一つだけが有意で、p 値が 0.03 であるとすると、TPM での p 値は 0.2752 となり、Bonferroni の補正法での結果に近い値となる。この場合は、すべての検定において帰無仮説が成立するとした場合でも十分考えられる結果であり、この有意性は特に注目には値しないと解釈される。

以上のように、TPM を適用することで、合理的な有意水準が得られるように思われるが、Bonferroni の補正法同様、それには前提条件として、すべての仮説が互いに独立でなければならない。互いに独立でない帰無仮説が複数棄却されるような場合、p 値が過小評価され、逆に反保守的な有意水準を与える傾向にあることは留意すべき点である。

その他、個々の p 値に着目しないなどの理由からか、あまり注目されていないが、参考までに押さえておきたい手法である。

TPM のツールは、Dr. Zaykin のサイト（http://statgen.ncsu.edu/zaykin/tpm/）で公開されており、Windows の場合、tpm.exe をダウンロードした後、コマンドプロンプトを起動し、

> tpm.exe 0.05 10000 pvalues_test.txt

と入力すればよい。なお、「0.05」は有意確率、「10000」はモンテカルロ法における計算の反復数、そして「pvalues_test.txt」は、各 SNP での p 値を一列に記したテキストファイルである。

5.3 False discovery rate（FDR）を制御する手法

False discovery rate（FDR）は、帰無仮説が棄却された検定のうち、実際には帰無仮説が真であるものの割合であり、この数値を一定の水準以下に制御する手法が提唱されている（Benjamini and Hochberg 1995; Storey et al.2003; Storey et al. 2004）。

これらの手法では、すべての検定を有意性の低い順に N、N - 1、...、i、...、2、1 と並べ替えた後、各検定での p 値をもとに q 値（各検定での p 値を有意水準としたときの FDR）を計算し、この値がある基準値 q^*（0.05 など）より小さいものを有意と見なす。

q 値は、式（式 3）で表される。

$$q_i = \begin{cases} \hat{\pi}_0 p_i & (i = N\text{の場合}) \\ \min\left(\dfrac{\hat{\pi}_0 N p_i}{i}, q_{i+1}\right) & (i \le N\text{の場合}) \end{cases} \qquad \text{（式 3）}$$

ここで、N は検定の数、また $\hat{\pi}_0 = \dfrac{\#\{p_j \ge \lambda\}}{N(1-\lambda)}$（ただし、$\lambda$ は $0 < \lambda < 1$ である定数、$\#\{p_j \ge \lambda\}$ は、λ より小さな p 値を示す検定の数）であり、帰無仮説が真である検定の、すべての検定に対する割合の推定値である。この π_0 が、前節で述べた、Bonferroni の補正法の保守化を招く原因の二点目に対する答えであると言える。

前節でも述べた通り、すべての帰無仮説が真であれば、図 5.1 で青色の破線で示されているように、p 値は 0 から 1 まで、高さ a の一様分布を示す。しかし、一部の検定における有意な結果が、帰無仮説では説明できない、明らかな関連を意味している場合には、赤色の曲線のように、0 に近い部分の頻度が高くなり、それに伴って、残りの一様分布を示す部分の頻度が相対的に低くなる。また、SNP を用いた関連解析の場合、一部の SNP において、Hardy-Weinberg 平衡から逸脱し、ホモ接合体の頻度が、遺伝子頻度から計算される期待値より高くなるケースがし

ばしば起こる。このとき、アレルモード（遺伝形式を仮定せず、症例・対照間のアレル頻度を検定する方法）での関連解析を行なうと、有意性が過大評価される傾向にあるため（Sasieni 1997）、やはり 0 に近い部分に p 値が偏在することが考えられる。

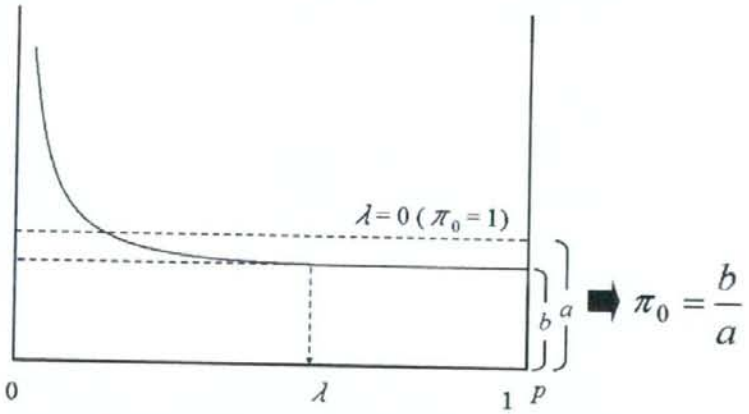


図 5.1　p 値の分布とパラメーターλおよび$\hat{\pi}_0$

すべての帰無仮説が真であれば、p 値は 0 から 1 まで、高さ a の一様分布（青色の破線）を示すのに対し、一部の帰無仮説が偽ならば、赤色の曲線のように、0 に近い部分の頻度が高くなり、それに伴って、残りの一様分布を示す部分の頻度が、赤色の破線で示したように、相対的に低くなる。この一様分布を示す部分の高さ b より下の部分すべてが、帰無仮説が真である SNP を表す。π_0 は、帰無仮説が真である（疾患との関連が有意でない）SNP の、すべての SNP に対する割合、すなわち $\pi_0 = \frac{b}{a}$、また、一様分布を示す部分とそうでない部分の境界点に相当する値が、$\hat{\pi}_0$ を与える λ となる。

この一様分布を示す部分の高さ b を赤色の破線で示しているが、これより下の部分すべてが、帰無仮説が真である検定であることから、$\pi_0 = \frac{b}{a}$、また、一様分布を示す部分とそうでない部分の境界点に相当する値が、$\hat{\pi}_0$ を与える λ となる。$\hat{\pi}_0$ の推定については、三次スプライン補間によるもの（Storey et al. 2003）と、ブートストラップを用いるもの（Storey et al. 2004）が提唱されている。詳細はこれらの文献を参照されたい。

なお、Benjamini と Hochberg（1995）はそれに先立って、$\hat{\pi}_0 = 1$（$\lambda = 0$）、すなわち、p 値の分布を一様分布と見なして計算する最も単純な方法（以下、BH 法とする）を提唱しているが、前記の理由から、その場合の有意水準は最も保守的と

なり、特に、有意性が最も高い（i = 1）もののq値は、Bonferroniの補正法による結果と一致する。またその結果が示す通り、保守化の原因の二点目については考慮されているものの、一点目、すなわち仮説間の相関については、Bonferroniの補正法同様、すべて独立と見なされるため、得られる有意水準は、依然として保守的な傾向が強いことに留意されたい。

以下に、BH法を用いた有意なSNPの抽出の具体的な例を示す。ある100のSNPについてのタイピングデータを用いて、疾患との関連解析を行なった結果、各SNPにおけるp値が、有意性の低い順に、次の通りであったとする。

$p_{100} = 0.8400$ （式4）

$p_{99} = 0.8170$

：

$p_4 = 0.0021$

$p_3 = 0.0009$

$p_2 = 0.0007$

$p_1 = 0.0002$

この場合、各SNPのq値（ただし$\hat{\pi}_0 = 1$とする）はそれぞれ以下のように計算される。

$$\hat{q}(p_{100}) = \frac{100 \times 0.8400}{100} = 0.8400 \qquad \text{（式5）}$$

$$\hat{q}(p_{99}) = \min\left\{\frac{100 \times 0.8170}{99}, \hat{q}(p_{100})\right\} = 0.8253 \qquad \text{（式6）}$$

：

$$\hat{q}(p_4) = \min\left\{\frac{100 \times 0.0021}{4}, \hat{q}(p_5)\right\} = 0.0525 \qquad \text{（式7）}$$

$$\hat{q}(p_3) = \min\left\{\frac{100 \times 0.0009}{3}, \hat{q}(p_4)\right\} = 0.0300 \qquad \text{（式8）}$$

$$\hat{q}(p_2) = \min\left\{\frac{100 \times 0.0007}{2}, \hat{q}(p_3)\right\} = \hat{q}(p_3) = 0.0300 \qquad \text{（式9）}$$

$$\hat{q}(p_1) = \min\left\{\frac{100 \times 0.0002}{1}, \hat{q}(p_2)\right\} = 0.0200 \qquad \text{（式 10）}$$

ここで、基準値 q^*を 0.05 とすると、p_3（$\hat{q}(p_3) = 0.0300 < 0.05$）までが採用される。ちなみに、Bonferroni の補正法を用いた場合は、p_1しか採用されない（$p_1 \times 100 = 0.0200 < 0.05$、$p_2 \times 100 = 0.0700 > 0.05$）。なお、この手法の改変版（Benjamini and Hochberg 2000; Benjamini and Yekutieli 2001）も提唱されているが、堀内と松田（2006）は、シミュレーション実験による検討を行ない、総合的に見れば、BH 法が最も安定した検定結果を示すとしている。

FDR の計算を行なうフリーのソフトウェアは、Dr. Storey のサイト（http://genomics.princeton.edu/storeylab/qvalue/）から、QVALUE が公開されており、いずれの OS にも対応している。ただし、使用には R のインストールが必要である。

実行にあたっては、各 SNP での p 値を一列に記したテキストファイルを用意するだけでよい。また、コマンドラインモード、ポイント・アンド・クリックモードのいずれでも操作が可能であり、前者の場合はコンソールウィンドウでコマンドを入力し、後者の場合はコンソールウィンドウとは別に表示される初期画面から操作を行なう。前者の場合、例えば C ドライブにあるフォルダー Program Files/R/QVALUE 内で作業を行ない、各 SNP での p 値を記したテキストファイルを test.txt、結果を出力するためのファイルを result.txt とすると、

```
> p <- scan("C:/Program Files/R/QVALUE/test.txt")
> qobj <- qvalue(p)
> qplot(qobj)
> qwrite(qobj, "C:/Program Files/R/QVALUE/result.txt")
```

と入力すればよい。

5.4 パーミュテーションテスト（並べ替え検定）

前節までに述べた手法は、p 値に直接手を加えるものであったのに対して、ここで紹介するパーミュテーションテスト（並べ替え検定）（Churchill and Doerge 1994）は、その理論的な背景や、得られる結果の特徴や傾向において、他とは一線を画す手法である。パーミュテーションテストではその名の通り、データセットのランダムな並べ替えと検定統計量の計算が多数繰り返され、得られた検定統計量の分布から、元のデータセットにおける検定統計量の有意性が判定される。すなわち、ランダムな並べ替えにより、すべての検定において帰無仮説を真と見なすことが可能なため、それらの検定統計量の分布から与えられる閾値を超える検定統計量は、偶然では説明できない、真の関連の存在を示唆しているという考え方に立脚している。なお、単一検定（$N = 1$）の場合、この手法で与えられる有意水準は、通常のカイ二乗分布から計算される値とほぼ一致する。

SNP を用いた関連解析における具体的な手順は以下の通りである。

1）あるデータにおける N セットの SNP すべてについて関連解析を行ない、それぞれの検定統計量（カイ二乗値を用いる場合が多い）を計算する。

2）データセットにおけるタイピングデータを固定しつつ、表現型のデータをランダムに並べ替え、オリジナルなものとは異なるデータセットを新たに作成する。

3）2）で作成された新たなデータセットを用いて、すべての SNP について関連解析を行ない、得られたカイ二乗値のうちの最大値 $\max_{1 \le i \le N} \chi^2_{i,perm_1}$ を記録しておく。

4）2）→3）を十分な回数（$T = 1{,}000 \sim 100{,}000$）繰り返すことで、カイ二乗値の最大値の分布 $\left\{\max_{1 \le i \le N} \chi^2_{i,perm_j}\right\}_{j=1}^{T}$ が得られる。この分布の 100*(1 − α)パーセンタイルを、有意水準 100*α%での検定閾値 τ とし、①で得られている検定統計量 $\left\{\chi_i^2\right\}_{i=1}^{N}$ との比較を行なう（$\chi_i^2 \le \tau$ ならば帰無仮説を採択、$\chi_i^2 \ge \tau$ ならば棄却）。

パーミュテーションテストの最大の特長は、仮説間の独立性を考慮した適正な有意水準を与えることである。すなわち、すべての仮説が互いに独立であれば、Bonferroni の補正法での結果にほぼ一致した有意水準を、逆にすべての仮説が強く相関し、実質的に単一検定であるようなデータであれば、通常のカイ二乗分布から得られる値とほぼ一致した有意水準を与える。また、ランダムな並べ替えにより、すべての検定において帰無仮説を真と見なすことが可能なため、p 値の一様分布からの歪みによる影響も受けない。一方で、多くの計算量を必要とするため、ゲノムワイドの超高密度データなど、大規模なデータセットには適さない。

フリーの関連解析ツールセットである PLINK（Purcell et al. 2007）には、パーミュテーションテストが実装されている。Windows の場合、ダウンロードした圧縮ファイルを解凍すれば、実行ファイル plink.exe をコマンドプロンプトで使用できる。詳細は PLINK のサイト（http://pngu.mgh.harvard.edu/~purcell/plink/）を参照されたい。

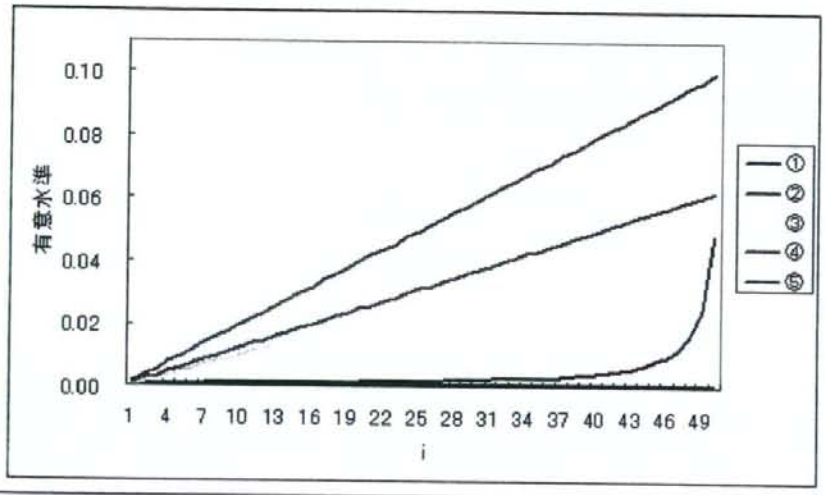
5.5 各手法のまとめ

以上、多重検定における有意水準の設定に用いられる手法について述べてきたが、どれを用いるべきかは、データセットの規模、連鎖不平衡、関連解析での p 値の分布などによって異なる。ここでは、用いる手法やパラメーターの選択のための参考となるよう、各手法の特徴や、用いる際の注意点について簡潔にまとめる。

まず、適正な有意水準を与えるという点では、仮説間の相関や、p 値の分布の歪みによる影響を受けないパーミュテーションテストが最も理想的な手法であり、極力これを用いるのが望ましい。しかし反面、多くの計算量を必要とするため、特に大規模なデータの場合には、Bonferroni の補正法など、パーミュテーションテスト以外の手法が推奨される。

これらの手法は、計算負荷の低さと引き換えに、仮説間の相関の影響を受けやすく、結果として保守的な有意水準を与える傾向が強い。したがって、連鎖不平衡の度合を示す指標である D' や r^2 に一定の基準を設けて、連鎖不平衡ブロックを代表する SNP 以外のものをデータセットから削除するなどの処理を、必要に応じて行なわなければならない。現時点での最良の方法は、すべての SNP のデータから、スペクトル分解によって、有効な（＝互いに独立な）検定の数を推定するソフトウェア SNPSpD（Nyholt 2004）を用いることであろう。

さらに、データ内の SNP が互いに独立の関係にあっても、関連解析での p 値が 0 付近に偏在し、一様分布からの歪みが生じている可能性は否定できない。この歪みもまた、有意水準の保守化を招く原因であるため、その傾向が認められる場合には、Storey ら（2003）の手法により、パラメーター π_0 として適切に考慮することが望ましい。図 5.2 に、これらの手法によって与えられる多重検定の有意水準を、互いに比較する形で示す。



p値の順位	①	②	③	④	⑤
1	0.00100	0.00100	0.00100	0.00125	0.00200
2	0.00100	0.00102	0.00200	0.00250	0.00400
3	0.00100	0.00104	0.00300	0.00375	0.00600
4	0.00100	0.00106	0.00400	0.00500	0.00800
			:		
i	α / N	$\alpha / (N - i + 1)$	$i\alpha / N$	$i\alpha / 0.8N$	$i\alpha / 0.5N$
			:		
50	α / N	α	α	$\alpha / 0.8$	$\alpha / 0.5$

図 5.2　各手法によって与えられる多重検定の有意水準

図中の①～⑤は、それぞれ Bonferroni の補正法、sequential Bonferroni 補正法（Rice、1989）、BH 法（Benjamini と Hochberg、1995）、Storey ら（2003）の手法（$\hat{\pi}_0 = 0.8$）、および Storey ら（2003）の手法（$\hat{\pi}_0 = 0.5$）を表す。番号が小さい手法ほど、保守性が強まる。

なお、上記の SNPSpD は、サーバーにファイルをアップロードして解析を行なうタイプのソフトウェアであり、結果はウェブブラウザで表示される。また入力ファイルは、Abecasis ら（2002）が開発した連鎖解析用ソフト MERLIN で指定されている形式（一部異なるが、基本的には LINKAGE 形式と同じである）に準じていなければならない。詳細は、SNPSpD のサイト（http://gump.qimr.edu.au/general/daleN/SNPSpD/）を参照されたい。

5.6 おわりに

多重検定における有意水準の設定について、従来は Bonferroni の補正法が一般的に用いられてきたが、その保守性がしばしば指摘されている。SNP を用いた関連解析の場合、特に連鎖不平衡の影響が強く、ほとんどのケースで、実際の SNP の数と比較して、有効な検定の数が小さいため、有意な結果を得るためのハードルが高く設定されてしまう。疾患と重要な関連を有する SNP を見出しながら、統計的に有意と見なされないばかりに、論文の作成や、その後の研究の進め方で頭を悩ませた研究者は少なくないであろう。

上述のように、この問題に対しては、FDR を制御する方法、およびパーミュテーションテストと、大別して二つの有効な解決策が提唱されている。計算量の問題などから、関連解析に関しては、前者の方がより頻繁に用いられているが、どの手法を用いるのが適切かは、データによって様々である。今後は、仮説間の独立性や p 値の分布といった、結果に与える影響が小さくないにもかかわらず、従来はあまり考慮されなかった点も踏まえつつ、手法やパラメーターをより慎重に選択することが望ましい。

また、適正な有意水準を考えることは確かに重要であるが、有意性検定の結果はあくまでも一つの目安である。Moran（2003）も述べているように、数字のみにとらわれて一喜一憂し、問題の本質を見失うことがあってはならず、研究におけるストーリーの説得性を高めるべく、むしろこれらの結果をうまく利用していくことが重要である。

6. 遺伝子間相互作用の検出法

6.1 はじめに

近年のゲノムサイエンスの目覚ましい発展に伴い、染色体上に存在するマイクロサテライトやSNPなどの遺伝子多型の情報を、データベースから取り出して利用することが可能となり、それらを用いた連鎖解析や関連解析などの遺伝統計学的アプローチを通じて、様々な疾患、特に単一遺伝性疾患に関与する原因遺伝子の同定が急速に進みつつある。特に単一遺伝病の遺伝子同定は手法がほぼ確立しており、一定以上の数の検体さえ収集できればほぼ可能である。一方、生活習慣病などのcommon disease（ありふれた病気）の発症リスクには、複数の遺伝子のほか、その名が示す通り、生活習慣や外的環境などの非遺伝性因子も深く関与しており、それらが相互作用を及ぼすことによって、非常に複雑なネットワークが形成されていると考えられる。多因子疾患の全容の解明には、この相互作用の理解が不可欠であることは言うまでもないが、その複雑さは、かつて糖尿病について Neel（1962）が用いた、「遺伝学者の悪夢」という言葉に象徴される通りであり、研究の進展を妨げている最大の要因であると言える。とは言え、研究者達の不断の努力により、新たに生物情報学のノウハウを取り入れるなどして開発された手法や、それらを実行するソフトウェアが数多く報告されてきており、この問題を取り巻く状況は、緩やかながらも着実に変わりつつあるように思われる。

ここでは、多因子疾患の発症リスクに関与する相互作用、特に遺伝子間相互作用の解析手法を紹介するとともに、今後の課題や、「悪夢」を良い夢に変えるための取り組みについて述べる。

6.2 遺伝子間相互作用とその検出法

ここで言う「遺伝子間相互作用」とは統計学的なものであり、図 6.1 に示すように、一方の遺伝子座の遺伝形式や効果の大きさが、他方の遺伝子座における遺伝子型の影響を受ける現象のことである。図 6.1 の右に示した例では、遺伝子座 B における遺伝子型が BB であるか、またはそれ以外かによって、A における遺伝形式と効果の正負が完全に逆転している。この場合、もし B の影響を考慮しなければ、これらが相殺することにより、破線で示したように、A の見かけ上の効果は 0 となり、疾患との関連はないものと見なされてしまう。したがって、このような関連構造を的確にとらえるためには、発症モデルに複数の遺伝子多型を同時に考慮しなくてはならない。

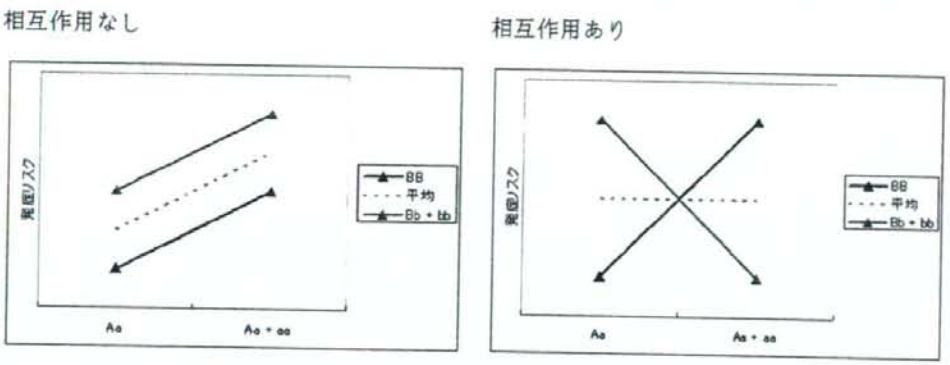


図 6.1　表現型への遺伝子間相互作用の影響

相互作用が存在しない場合、左図のように、遺伝子座 B においていかなる遺伝子型であろうと、A の遺伝形式、アレルの効果の大きさや向きは変わらず、逆もまた然りである。一方、相互作用が存在する場合には、B の遺伝子型によって、A の遺伝形式などに違いが生じる。右図に示した例では、A における遺伝形式と効果の正負は完全に逆転しており、もし B の影響を考慮しなければ、A の見かけ上の効果（＝破線の傾き）は 0、すなわち、疾患との関連はないものと見なされてしまう。

このように、遺伝子間相互作用の存在は、多因子疾患の感受性遺伝子の同定を困難なものにしている要因の一つである。

そこで最も一般的に用いられてきたのが重回帰分析法であるが、ここで研究者の行く手を阻むのが、「次元の呪い」（Bellman（1961）が用いた「curse of dimensionality」の訳語）、すなわち、解析に必要なデータの数は、モデルを構成するパラメーターの数に対して、指数関数的に増加するという問題である。例えば

通常の関連解析では、モデルに含まれるのは単一の SNP であるから、考えられる遺伝子型の数は 3 であり、現実的な数のデータで十分解析が可能である。しかし、10 もの SNP が関与する、非常に高次元な相互作用を考える場合、考えられる遺伝子型の数は $3^{10} = 59,049$ となり、一定の信頼性を有する結果を得るためには、少なくとも数十万単位のデータが必要となる。しかし実際に収集が可能な数はせいぜい数百から数千程度であり、そのようなデータに対して、無理に複雑なモデルを当てはめても、結果の信頼度を大きく損なう危険性が高い（Concato et al.1993）。特に近年では、ゲノムワイドでの関連解析が世界的な規模で進められており、数十万単位の SNP のデータをいかに取り扱うかが今後の課題となる。また、連鎖不平衡の影響により、変数としての SNP が互いに独立でない（多重共線性）ことも、「次元の呪い」と並んで、研究者を大いに苦しめている問題の一つである。さらに、重回帰分析では、それ自身が単独で有意な関連性を有している SNP のみ検出が可能であり、いかに他の SNP との間に強い相互作用を及ぼしていたとしても、単独で有意な効果を示さない SNP は、関連因子の候補として選択される可能性が低くなってしまう（Culverhouse et al. 2002）。

遺伝子間相互作用解析にまつわる以上のような問題に対処すべく、新たな理論に基づく手法がこれまでに数多く提唱されている。図 6.2 に示すように、これらはノンパラメトリック法、ニューラルネットワーク、そしてグラフィカルモデリングの三つに大別され、ノンパラメトリック法はさらに、組み合わせ法、再帰分割法、セット関連解析法の三つに分けられる。

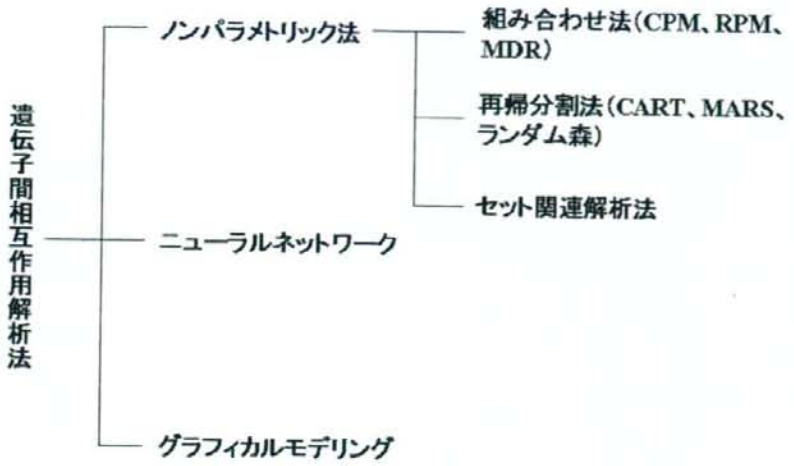


図 6.2　現在までに開発されている遺伝子間相互作用解析法

6.3 ノンパラメトリック法

6.3.1 組み合わせ法

組み合わせ法は、次元削減およびパターン認識をベースとした手法である。現在までに、CPM（combinatorial partitioning method）（Nelson et al. 2001）、RPM（restricted partitioning method）（Culverhouse et al. 2004）、そして MDR（multifactor dimensionality reduction）（Ritchie et al. 2001）の三つが提唱されている。

CPM は、量的形質における遺伝子間相互作用、特に 3 以上の因子に由来する高次元相互作用の解析を目的として開発された手法である。この手法により、アポリポタンパク質など複数の遺伝子における多型間の相互作用が、血漿トリグリセリド濃度に影響を与えていることを示唆している（Nelson et al. 2001）。なお RPM は、計算の効率化を図るべく、CPM に発見的アルゴリズムを導入したものである。

MDR は、CPM にヒントを得て、疾患など、質的形質における遺伝子間相互作用の解析を目的として開発されたものであり、現在最も注目を集めている手法の一つである。具体的なアルゴリズムは以下の通りである。

1）K セットの SNP 群から、i（$1 \leq i \leq K$）セットを選抜する（強い連鎖不平衡を示す SNP 同士は、後述の理由により、避けた方が望ましい）。

2）i セットの SNP から考えられるすべての遺伝子型の組み合わせについて、全データの一部（10 分割交差検定の場合は 9/10）を学習用データとして、罹患者および非罹患者の数をカウントする。

3）罹患者数と非罹患者数の比をもとに、各遺伝子型を「高頻度発症グループ」または「低頻度発症グループ」に二分する。

4）学習用データ以外のデータ（同 1/10）を検証用データとして、交差検定により、3）で高頻度発症型にカテゴライズされた遺伝子型を持つ被験者が実際に罹患しているか（感度）、低頻度発症型とされた遺伝子型を持つ被験者が実際に非罹患者であ

るか（特異度）を各々求め、両者の平均値（testing accuracy; TA）をもとに、各モデル（SNP の組み合わせ）の適合度を測る。

5）1）～4）を n（10 分割交差検定の場合は $n = 10$）回行ない、TA が最大となった回数（cross-validation consistency; CVC）が最も多い SNP の組み合わせを、i ごとに求める。

上述の通り、一般的には TA と CVC が最適なモデルを特定する指標となり、これらを最大化する SNP の組み合わせが、発症リスクに関わる相互作用をもたらす可能性が最も高いことを示す。なお、MDR を用いる際に注意すべき点は、強い連鎖不平衡を示す SNP の取り扱いである。データに強い連鎖不平衡を示す SNP が存在する場合、一方の SNP を含むモデルと、もう一方の SNP を含むモデルとの間で、CVC をランダムに二分してしまう傾向にあるため、結果として CVC の低下を招き、最適なモデルを誤って評価してしまう危険性を孕んでいる。したがって、互いに独立な SNP のみをあらかじめ選抜しておくことが望ましい。

現在までに、MDR により、孤発性乳ガン（Ritchie et al. 2001）、II 型糖尿病（Cho et al. 2004）、心房細動（Tsai et al. 2004）、および脳動脈瘤（Akagawa et al. 2007）などで、有意な遺伝子間相互作用の存在が報告されている。しかし、このような統計学上の結論を、より説得性の高い形で生物学的な見解と関連付けるためには、方法論も含め、さらなる検討が必要であろう。

MDR のソフトウェア（図 6.3 参照）は、アメリカ Dartmouth Medical School の Computational Genetics のサイト（http://www.epistasis.org/index.html）からダウンロードが可能であり、2008 年 5 月には、最新バージョンである 1.2.1 がリリースされている。使用の際には、表 6.1 に示すようなフォーマットで入力ファイルを作成するだけでよい。