電子計算機による HLA 抗血清の特異性分析

田 村 眞 木 村 一 郎

(東京大学医学部血清学教室)

十 字 猛 夫

(東京大学附属病院輸血部)

萩 野 陽 一 郎

(東京大学附属病院産婦人科)

A NEW METHOD FOR THE STATISTICAL ANALYSIS OF HLA ANTIGEN SYSTEM BY THE LARGE-SCALED COMPUTER

by

M. TAMURA, T. JUJI, I. KIMURA and Y. HAGINO

(Department of Serology, Blood Transfusion Service and Department of Obstetrics & Gynecology.

Tokyo University School of Medicine)

Abstract

1) The processing of an enormous amount of serological data is indispensable for the statistical analysis of HLA antigen system. However, the suitable method for it is not yet developed even after the contemporary availability of the large-scaled electronic computer. The main purpose of this study is to develop the method for specificity analysis of antisera (lymphocytes) by the computer.

2) At first, the 2-dimensional chart of ρ , I' values and the serogram for each antiserum (lymphocyte) were drawn by the computer. Here, ρ and I' values mean the coefficients of correlation and inclusion between a given pair of reaction components. From the (ρ , I') chart of antiserum

the reactivities of antisera of unknown specificities were classified and clustered by statistical comparison with those of known specificities. Further analysis of selected antisera was carried out by referring to serograms which illustrate the reaction patterns the reaction patterns of lymphocyte in panel. An example of practical analysis of 400 antisera combined with 200 panel lymphocytes is shown in the text.

3) The method developed here can be said as one kind of cluster analysis using 2 similarity indices. The main merit of the use of it is that the classification of antisera is possible even when the antigen specificities of each panel lymphocyte are not known.

序言

HLA 抗原系の血清学的解析やタイピングは、移植のための組織適合性検査としてのみならず、輸血の副作用や、種々の疾病の発症等に関する遺

伝因子の検索にも重要な検査である。この検査には、各種の特異性を持つ抗血清を十分量確保することが要求されるが、それには新しく得た抗血清の特異性を敏速精密に同定する手順の確立が必要

となる。これには現在、個々の未知試料を特異性既知の抗血清やリンパ球と反応させ、その成績の主観的比較から同定する方法がとられているが、比較・同定の段階に電子計算機を用い試料集団を統計的に処理すれば、大量の試料を同時に取扱える利点がある。ところが、試料の反応性を見る方法はすでに臨床検査の手技としてほぼ確立している¹⁾のに対して、その統計的処理には多量の複雑な計算が含まれ、一貫した手順の確立にはまだいたっていない。また、特異性同定のためのデータ処理を困難にしている背景として、

1) パネル・リンパ球の抗原構造が複雑なうえ、使用抗血清中に含まれている抗体特異性は必ずしも単一ではない。

2) この両者の多数を組合せて得た成績を解析するので、処理の対象となるデータ量が膨大になる。

3) 日本人由来のパネル・リンパ球を用いる場合、これまで白人を中心に知られている HLA 抗原特異性のみならず、日本人に特有な未知の抗原系の存在の可能性も考慮する必要がある、等の要因が考えられる。

一方、HLA 抗原の特異性の同定に用いるデータ処理法には、血液型の研究に用いられた手法の応用が可能である。すなわち、血液型の分野で用いられていた 2×2 分割表を最初に白血球の分類に導入したのは、van Rood ら²⁾である。また、血液型に関する抗血清の特異性の同定に χ^2 検定もよく使われているが、Mi ら³⁾は χ^2 値から導き出される関連係数が特異性の分類には有効なことを述べている。

また、Mickey ら⁴⁾は白血球群の分類の方法の検討の際、同一特異性をもつ抗血清の集団 (cluster) を、Chi-square map を用いて図示したが、この際、包含係数 (inclusion index) の概念を導入して2種の抗血清がパネル・リンパ球に対して示す共通の特異性の包含の度合を表わしている。さらに、Albert ら⁵⁾は処理すべきデータの複雑化、大量化に伴い小型計算機を用いたプログラミングの方法を考案し実用化している。

このようにデータ処理に数学的基盤を与える報告はこれまでもあったが、現実の大量のデータについて、反応因子の特異性を敏速、的確に同定するための実用的手法に関する組織的な研究はまだ見当らない。

上述の如く、2種の抗血清の反応性から計算した関連係数、包含係数は抗血清相互が持つ同一、部分的に共通、または無関係の、反応特異性を測る重要な尺度となるから、両係数を2次的に組み合わせれば抗血清の特異性の相互関係をよりの確に把握できると思われる。そこで、この原理を応用して、大型計算機による HLA 系の特異性の解析を試み、その成績はこれまでも報告して来た^{6)~13)}。本報はその一連の実際的手法をまとめ、個人の HLA 系の決定というより集団としての抗血清特異性の分析方法として提案するものである。

この手法による HLA 抗原の特異性の解析は東京大学附属病院輸血部において実用に供されているが、これまで得られた成績は実状とよく合致し^{6)~8,13)}、HLA 型の抗血清の分類に有効であった^{10)~12)}ばかりでなく、日本人に特有と思われる新しい HLA 抗原等の発見につながった⁹⁾。

実験

I 材料と反応方法

(1) 抗血清

(a) 特異性未知の抗血清：抗白血球抗体陽性の妊婦の分娩時の出血血液から分離した血清をくり返し透析、濾過、限外濾過をして約400種の抗血清を得た。

(b) 特異性既知の抗血清：NIH Serum Bank, Dr. Terasaki, Dr. van Rood, Dr. Simons の好意で分与された約100余種の抗血清から50種の抗血清を選んで用いた。

(2) パネル・リンパ球

東京大学医学部および附属病院関係者で互いに血縁関係がなく、健康な約200名から定期的に提供を受けた。

(3) 血清学的反応：各抗血清とパネル・リンパ球のすべての組合せについて、Terasaki の方

法¹⁾¹⁴⁾による Microdroplet lymphocyte cytotoxicity test¹⁵⁾を行ない、死んだ細胞が40%以上のときを陽性とした。

以上の材料と反応に関する詳細は萩野ら⁹⁾により報告されている。

II データ処理の方法

上述の血清学的反応から得られたchecker-board型の反応成績は直接データ・シートに記入しパンチカード入力すると共に、図1のフローチャート

図1 HLA抗原の特異性分析のデータ処理の概要を示すフロー・チャート

に示した順序で抗血清、パネル・リンパ球の両面からデータ処理を行なった。以下に述べる計算と図の作製は、Förtran IVで書かれたプログラムによって電子計算機 Tösbac 5400で行なった。説明の繁雑さをさけるため、主として抗血清の面から述べる。

(a) 陽性率 (P_i) の算出: ある抗血清 (番号 i) が N_L 種類のパネル・リンパ球のうち、 n 種のものと陽性の反応を示すなら、この抗血清の陽性率 (P_i) は $P_i = \frac{n}{N_L}$ である。

(b) 二抗血清間の関連係数 (ρ_{ij}) と包含係数

表1 2×2分割表

抗血清(i番)

	+	-	
抗血清(j番)			
+	a	b	a+b
-	c	d	c+d
	a+c	b+d	a+b+c+d= N_L

全パネル・リンパ球を、 i と j 番の2抗血清に対する反応が陽性か陰性かにより分類した度数分布表
たとえば、 b は i 番抗血清に対して陰性、 j 番抗血清には陽性を示したパネル・リンパ球数を表わす。
(N_L は全パネル・リンパ球数)

(P_{ij}) の算出: 番号 i, j の2種の抗血清を、 N_L 個のパネル・リンパ球に反応させた場合の2×2分割表は表1に従って作製し、これをもとに、2種の抗血清間の関連係数または相関係数 (association coefficient, or coefficient of correlation¹⁷⁾) ρ_{ij} と包含係数 Γ_{ij} を次の式から求めた³⁾⁴⁾¹⁶⁾¹⁷⁾。

$$\rho_{ij} = \rho_{ji} = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} \\ = \pm \sqrt{\frac{\chi^2}{N_L}}$$

ここで

$$N_L = a + b + c + d, \text{ であり,}$$

$$\chi^2 = \frac{(ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

であるが、 ρ の符号を $(ad - bc)$ のそれに一致させれば、 ρ 値は $-1 \leq \rho \leq 1$ の範囲の値をとるが、その意義については後に詳述する。

次に Γ_{ij} 値は、

$$\Gamma_{ij} = \Gamma_{ji} = \frac{a}{\min\{(a+b), (a+c)\}}$$

であるが、この式は、比較すべき抗血清の両者に反応陽性だったリンパ球数を片方の抗血清のみに陽性を示したリンパ球数で除すことを示す。ただし、この際、2種の抗血清と反応したリンパ球数のうち値の小さい方を分母に用いる。 Γ 値の範囲は $0 \leq \Gamma \leq 1$ をとるが、 ρ, Γ 値のいずれも $(a+b)(a+c) = 0$ のとき計算できない。

2種の抗血清の特異性が全く同一のときの関連係数 ρ_{ij} の値は1.0となり、相互に類似するときにはこれに近い値になる。一方、2種の抗血清がパネル・リンパ球に対して全く逆の反応を示すときの ρ_{ij} 値は-1.0になる。実用的には100種以上の

パネル・リンパ球を用い関連係数の値が -0.20 付近より小であれば、2種の抗血清が示す特異性に対応するリンパ球上の抗原は対立遺伝子で支配されているとの考えも提出されている¹⁶⁾¹⁷⁾。

2種の抗血清のうち、一方の抗血清の抗体特異性が完全に他方の抗血清中に包含されていれば、包含係数 Γ_{ij} の値は 1.0 となり、一部を残してその大部分が含まれる場合はこれに近い値をとる。これに対して、一方の抗血清の示す抗体特異性が他方には全く含まれず無関係であれば Γ_{ij} 値は 0 になる。

次に、 i 番抗血清に対し比較に用いた抗血清の数だけの ρ_{ij} と Γ_{ij} ($j=1, \dots, N_s$; N_s は全抗血清数)値が計算できるが、それらの平均値、標準偏差を求め、 $\bar{\rho}_i, \sigma(\rho_i)$; $\bar{\Gamma}_i, \sigma(\Gamma_i)$ とした。

(c) $(\rho, \Gamma)_i$ 図の作製：抗血清が含む抗体の反応特異性の近縁関係を見るため、横軸に ρ 値、縦軸に Γ 値をとった座標上に、 i 番抗血清に対する他のすべての抗血清の ρ_{ij} と Γ_{ij} 値をプロットした図をつくる。これを $(\rho, \Gamma)_i$ 図と呼ぶが、ここには i 番抗血清に対する他の抗血清の近縁の度合いが点の位置として示されている。

(d) 血清反応図の作製： (ρ, Γ) 図だけでは抗血清の特異性をパネルとして用いた個々のリンパ球と対応させて解析することは難しい。そこで、特定の領域内に ρ, Γ 値を持つ抗血清の番号と、その反応性を検索できるようにした。こうして得た、 i 番抗血清と反応性が近縁な抗血清群の、個々のパネル・リンパ球に対する反応成績を書き出した図を、ここでは血清反応図と名付けた。この際、個々の抗血清とパネル・リンパ球の対応はかえらずに、比較すべき抗血清に対する反応の陽性率が高い順にリンパ球を並べかえると、反応のパターンは見やすく、詳細な分析が可能となる。この操作後、抗血清、パネル・リンパ球の順序は前述の一貫番号と異なってくるので、両者を対照して書き出せるようにした。

以上の諸操作から、 $(\rho, \Gamma)_i$ 図と共にその指定した特定の領域内にある抗血清群について血清反応図を画けば、特異性の比較が可能になる。

(e) chi-square map の作製： $(\rho, \Gamma)_i$ 図は i 番抗血清を標準とした場合、これと他の抗血清近縁関係をあらわすものであるから、全抗血清についてその相互関係を一覧することはできない。その目的には、従来から用いられている chi-square map が便利である。実験3, 4, 5で扱われた反応系について、Mickeyら⁴⁾とほぼ同じ方法で chi-square map を作製した。パネル・リンパ球数 N_L が200なので、関連係数 ρ 値の 0.50 は χ^2 値の 50.0 に対応するが、それ以上の値をもつ2種の抗血清の組合せを黒い小四角であらわした。この小四角が対角線に近くまとまりをつくって並ぶように抗血清の順序を並べかえてある。まとまりをつくる基準の χ^2 値を 50.0 とかなり高い値にとったことと、250種という多数の抗血清を対象とした点はMickeyら⁴⁾の場合と異なる。

成績

〔実験1〕特異性未知の抗血清とパネル・リンパ球間の反応

HLA系の特異性の統計学的検討には多数の抗血清とリンパ球を組合せて反応させる必要がある。従って、解析の対象となるデータはchecker-board型のものとなり、ほぼ同じ方法で抗血清とパネル・リンパ球の両側から検討できる。そこで特異性分析の基礎となるデータの性格を把握するため、特異性未知の抗血清とパネル・リンパ球の両者を反応させ、それらが示す陽性率の頻度分布、および (ρ, Γ) 図の比較から検討を始めた。

(1) 陽性率の頻度分布

特異性未同定の396種の抗血清と191種のパネル・リンパ球を反応させ、陽性率の頻度分布を求めた。図2(a)は抗血清についてのそれを示したもののだが、パネル・リンパ球に対して高い陽性率を示す抗血清ほどその頻度は低かった。検定の結果、分散が平均より大きいのでその分布は典型的ではないがPoisson分布に準ずるものと考えられる。

パネル・リンパ球側から見た陽性率の分布は図2(b)に示したが、この場合は12~14%の陽性率

図2(a) 全パネル・リンパ球に対する抗血清の陽性率の頻度分布

図2(b) 全抗血清に対するパネル・リンパ球の陽性率の頻度分布

図3(a) 抗血清 No. 12を標準として、全抗血清の近縁関係を示す $(\rho, \Gamma)_{12}$ 図 図3(b) パネル・リンパ球 No. 31を標準として、全パネル・リンパ球との近縁関係を示す $(\rho, \Gamma)_{31}$ 図

を示すパネル・リンパ球の頻度が最大となる一峰性の分布となった。これは正規分布検の結果、定ひずみがあるが、準正規分布に属すると考えられる。

(2) $(\rho, \Gamma)_{12}$ 図の示す分布

特異性未知の抗血清 No. 12を標準として画いた $(\rho, \Gamma)_{12}$ 図を、図3(a)に示した。図中の数

字はその領域に (ρ_{ij}, Γ_{ij}) 値を持つ個々の抗血清の個数をあらわす。これらの点の大部分は $\rho = 0(\bar{\rho}_{12}=0.05)$ 付近を中心として比較的平等の分布を示したが、抗血清の中には0.25~0.50、さらに0.50以上の ρ 値を持つものも認められた。 Γ 値はかなりばらついた分布を示しているが、 ρ 値が高いほど、 Γ 値も高くなる傾向があった。 ρ 値がー

0.25以下のものはごく少なく、 -0.40 以下のものは皆無であった。従って、図3(a)では ρ 値が -0.25 以下の領域は省略してある。この図では396種の抗血清のうち ρ 値が 0.50 以上のもの10種、 ρ 値が $0.25 \sim 0.50$ 、 Γ 値が 0.80 以上のもの7種が認められたが、残りはそれ以外の領域に分布していた。

図3(b)はパネル・リンパ球中のNo. 31を標準として得た $(\rho, \Gamma)_{31}$ 図である。ここでは、図3(a)に比し、比較的高い ρ, Γ 値でしかも狭い範囲に点が集中する傾向があり、抗血清の場合のようにいくつかの群に分けることは難しかった。ここには抗血清およびリンパ球の各1例について $(\rho, \Gamma)_{31}$ 図を示したにすぎないが、すべての抗血清とパネル・リンパ球について画いた $(\rho, \Gamma)_{31}$ 図を比較してもこの傾向は変わらず、リンパ球を群に分けることは難しいと思われる。

次に、各抗血清の ρ_{ij} および Γ_{ij} 値の平均値 $\bar{\rho}_i$ および $\bar{\Gamma}_i$ を求めてみた。 $\bar{\rho}_i$ 値が 0.05 から 0.20 の間にあるのに対し、 $\bar{\Gamma}_i$ 値は 0.15 から 1.0 の広い範囲を取った。一方、パネル・リンパ球ではほとんど全部の $\bar{\rho}_i$ 値は 0.25 から 0.40 の間にあり、 $\bar{\Gamma}_i$ 値も 0.50 から 0.80 の間の範囲にあった。これは1個のリンパ球上に多数の抗原があり、その一部が共通するため、リンパ球相互の特異性の相関が高いことによるのであろうか。

各抗血清やパネル・リンパ球の他に対する ρ_{ij} 、 Γ_{ij} の標準偏差 $\sigma(\rho_i)$ 、 $\sigma(\Gamma_i)$ を見ると、一般に抗血清の $\sigma(\Gamma_i)$ 値はパネル・リンパ球のそれより大きい値をとる傾向があったが、両者の $\sigma(\rho_i)$ 値はほぼ同じ値であった。

このように同一のデータを、抗血清とパネル・リンパ球の二方面から見たときの様相は、かなり異なることがわかった。特異性の分析という点に立つと、 (ρ, Γ) 図の分布のしかたから試料をいくつかの特異性類似の群と無関係な群に分けることは、パネル・リンパ球では難しいが、抗血清では可能であった。そこで、以後の実験では抗血清を群に分けることに主眼を置いた。

〔実験2〕特異性既知の抗血清の $(\rho, \Gamma)_{31}$ 図

実験1(図3)の特異性未知の抗血清とパネル・リンパ球の組合せから画いた抗血清の $(\rho, \Gamma)_i$ 図(図3)では、標準にした抗血清に対し高い ρ, Γ 値を持つ抗血清の存在が認められた。この事実の意義を明らかにするため、他より分与された特異性既知の50余種の抗血清に161種のパネル・リンパ球を反応させ、 $(\rho, \Gamma)_i$ 図を画いた。図4はHLA-B5に対する特異的抗体をもつ

図4 特異性既知の抗血清(No. 79)を標準として、特異性既知抗血清50種を、関連係数(ρ)と包含係数(Γ)についてプロットした $(\rho, \Gamma)_{79}$ 図

抗血清 No. 79 を標準として他の抗血清の ρ, Γ 値をプロットして画いた $(\rho, \Gamma)_{79}$ 図である。この図ではそれぞれの抗血清の ρ, Γ 値の位置にその抗血清が示す特異性を記入してあるが、 $\rho=1.0$ 、 $\Gamma=1.0$ の位置の点は抗血清 No. 79のものである。これを含めて検討した8種類の抗 B5 抗体特異性を持つ抗血清のうち7種類のものの ρ 値は 0.50 以上であった。 $\rho=0.34$ 、 $\Gamma=0.75$ の値を示した抗血清 No. 77 は、後述の血清反応図や他の血清学的検討の結果必ずしも典型的な抗 B5 抗体特異性をもつとはいえないことがわかっており、ここでの結果と一致する。抗 HLA-A9 抗体特異性をもった抗血清の多くが、 0.25 付近の比較的高い ρ 値を示したその理由については後に考察する。他の特異性を示す大部分の抗血清はそれぞれ $-0.25 \sim 0.25$ および $0.15 \sim 0.60$ の範囲の ρ, Γ 値を持っていたが、ここでは ρ 値が -0.25 以下の抗

血清は認められなかった。

他の特異性既知の抗血清を標準として (ρ, Γ) 図を画いても同様な所見が得られている。この事実は使用するリンパ球の抗原特異性が完全にわかっていなくとも、多数のパネル・リンパ球を用いれば、ある抗血清を標準として画いた $(\rho, \Gamma)_i$ 図から同じ特異性を持つ抗血清が高い ρ, Γ 値を持つものとして選り出せることを示している。従って、実験1で述べた特異性未知の抗血清による $(\rho, \Gamma)_{12}$ 図 (図3(a)) でも、高い ρ, Γ 値の領域にあった17種の抗血清は標準とした抗血清 No. 12と同一、またはほぼ同一の抗体特異性を持っていたことになろう。

〔実験3〕特異性既知と未知の抗血清集団とパネル・リンパ球間の反応

実験2 (図4) でわかった特性を利用くて、特異性既知の抗血清を標準とした $(\rho, \Gamma)_i$ 図を画けば、未知の抗血清から同じ特異性を持つものを選び出すことが可能と考えられる。そこで、実験2に用いた50種の特異性既知の抗血清と実験1に使用したものの中から選んだ200種の未知の抗血清と混ぜた集団を作り、200種のパネル・リンパ球に反応させた。この時の抗血清、リンパ球が示した反応の陽性率の頻度分布は実験1の場合と変わらず、上の操作をしても用いた抗血清、リンパ球の集団としての性質には大きな変化はなかったと思われる。ここでもパネル・リンパ球について画いた、 $(\rho, \Gamma)_i$ 図では、個々のリンパ球は狭い範囲に分布し、これからリンパ球の分類は困難であった。一方、図5の特異性既知の抗血清 No. 79を標準にした抗血清に関する $(\rho, \Gamma)_{79}$ 図では、特異性未知の抗血清の大部分および特異性は既知だがそれを異にする抗血清は何れも比較的低い ρ, Γ 値の領域に集まった。一方、特異性未知の抗血清の一部は、特異性が標準抗血清と同一の抗血清とともに高い ρ, Γ 値の領域に集まった。後者に属する抗血清群は標準とした抗血清と同一、またはほぼ同一の特異性を持つものと考えられるから、各種の特異性既知の抗血清を標準としてこの操作を繰り返せば、未知のものから特異性の同

図5 特異性既知の抗血清 (No. 79) を標準とした、50種の特異性既知抗血清と200種の未知抗血清の混合集団に対する $(\rho, \Gamma)_{79}$ 図

じものを選び出すことが可能となる。

図5中のI, II, III, IV, V群については実験4, 5で検討する。

〔実験4〕抗血清間の近縁関係の検討

このようにして $(\rho, \Gamma)_i$ 図から得て特定の抗血清 (i 番) と他の抗血清間の近縁関係の度合いについてももう少し詳しく検討する必要がある。 ρ 値はその定義から+1.0に近い程、標準にしたものに近く、-1.0に近い程対称的な性格を持つことを示している。同様に Γ 値が1.0に近い程その抗血清の特異性は標準にしたものの中に包含されている度合いが大きい。そこで、特異性未知の抗血清を用いた実験1およびその一部に既知のものを加えた実験3の組合せで全抗血清の持つ ρ_i 値の再平均値 $\bar{\rho}_i$ を求めると、その値はそれぞれ+0.03および+0.05であった。同様に、これらの場合の $\sigma(\Gamma_i)$ の中央値 m はそれぞれ0.15, 0.13であった。比較した大部分の抗血清の特異性は相互に無関係だから、これらの抗血清の Γ_i 値の多くは、 $\rho_i \pm 3m$ の間に分布すると思われる。そこでこれに上述の値を入れると+0.48, +0.44および-0.42, -0.34の値が得られた。したがって、標準抗血清と無関係な特異性を持つ抗血清の ρ 値の範

囲は $-0.40 \sim +0.50$ と考えられ、この範囲外の ρ 値を持つ抗血清はその範囲内にあるものに較べかなり異なった特徴を持つと云えよう。200種のパネル・リンパ球を用いた反応系では ρ 値が0.50、 -0.40 のときの χ^2 値はそれぞれ50.0、32.0となるが、これらの値は通常の有意差の検定に用いる限界値 $\chi^2=11.0$ に較べ十分高い値である。また、実験2で標準抗血清と特異性がほぼ同一と思われる抗血清の大部分は0.50以上の ρ 値を持ち、対称的な特異性を示す -0.40 以下の ρ 値を持つ抗血清は認められていない事実を考慮すると、‘ほとんど同一’または‘対称的’な抗血清を持つ ρ 値の下限と上限を $+0.50$ と -0.40 とするのが妥当であろう。しかし、抗血清の ρ 値が0.25 $\sim$ 0.50の場合でもこれを χ^2 値に換算すると13.5 $\sim$ 50.0になり、これらの抗血清もなお標準抗血清に対しかなりの類似性を持つと思われる。後述の血清反応図による検討から、 ρ 値がこの範囲内にあり Γ 値が0.80以上のものの反応特異性は標準抗血清のその中に含まれるのでこれを‘部分的に同一’の特異性を持つ抗血清と考えた¹²⁾。

以上の検討から、特定の抗血清を標準として画いた $(\rho, \Gamma)_i$ 図上に‘ほとんど同一’、‘部分的に同一’、‘無関係’、および‘対称的’な特異性を持

つ抗血清の分布する領域の設定が可能となつた。

〔実験5〕血清反応図

HLA-B5の特異性を持つ抗血清 No. 79を標準とした $(\rho, \Gamma)_{79}$ 図(図5)で、 ρ, Γ 値がそれぞれ0.50、0.80以上であった抗血清は8種類認められている。その中から、 ρ, Γ 値が0.75および0.90以上で抗 B5 抗血清として他から供与された2本の抗血清 No. 81, 83(第Ⅰ群)、および ρ, Γ 値がそれぞれ0.50 $\sim$ 0.75、0.80 $\sim$ 0.90の範囲内にある3本の妊婦抗血清 No. 382, 309, 223(第Ⅱ群)を選び、計算機に書き出させた血清反応図をもとにたれらの抗血清の200種のパネル・リンパ球に対する反応性を抗血清 No. 79のそれと比較した。このほかに ρ 値が0.25 $\sim$ 0.50の範囲にあるが、0.80 $\sim$ 1.00(第Ⅲ群)および0.60 $\sim$ 0.80(第Ⅳ群)の Γ 値を持つ抗血清が9種および11種類ある。これらの特異性も標準抗血清(No. 79)に近似する可能性が高いと思われるので、第Ⅲ群からは妊婦抗血清 No. 236, 221, 303の3本、第Ⅳ群から抗 B5 抗血清として供与されたが性質が多少異なると前に述べた抗血清 No. 77ほか2本の妊婦抗血清 No. 363, 378を選んだ。このほか、標準抗血清には特異性の関連性がない多数の抗血清(第Ⅴ群)から選んだ3本の妊婦抗血清 No. 216,

				パネル・リンパ球 NO.										(全パネル・リンパ球数200)										
				10	20	30	40	50	60	70	80	90	100											
I 群	血清番号	(ρ) 関連 係数	(Γ) 包含 係数																					
	NO. 79	1.00	1.00																					
	NO. 81	0.80	0.93																					
II 群	NO. 83	0.77	0.97																					
	NO. 382	0.73	0.85																					
	NO. 309	0.59	0.81																					
III 群	NO. 223	0.53	0.87																					
	NO. 236	0.36	1.00																					
	NO. 221	0.34	0.85																					
IV 群	NO. 303	0.27	0.90																					
	NO. 77	0.34	0.75																					
	NO. 363	0.30	0.70																					
IV 群	NO. 378	0.29	0.70																					
	NO. 216	0.00	0.34																					
	NO. 128	0.00	0.39																					
IV 群	NO. 87	0.00	0.55																					

1 抗血清とパネル・リンパ球との反応陽性は○、陰性は空白で示した。

2 抗血清No.79を標準とする $(\rho, \Gamma)_{79}$ 図(図5参照)のⅠ、Ⅱ、Ⅲ、Ⅳ、Ⅴ群の各領域から2,3種の抗血清について、パネル・リンパ球との反応を示した。

3 見やすくするためにパネル・リンパ球の順序を並べかえたが、パネル・リンパ球の一貫番号は繁雑をさけるため省略した。

4 並べかえられた順序のNo.105からNo.200までのパネル・リンパ球は、Ⅴ群の抗血清を除けば、他群のここにあげた抗血清に対してはすべて陰性なので省略した。

5 抗血清No.79を標準として算出した各抗血清の関連係数 (ρ) と包含係数 (Γ) は図の左に併記してある。

6 n はパネル・リンパ球No.106 $\sim$ 200の中で反応陽性を示した個数。

図6 血清反応図

128, 87 についても同様の検定を行なった。図 6 はこれらの抗血清のパネル・リンパ球に対する反応性を血清反応図として計算機で書き、これを見やすくまとめたものだが、反応陽性の組合せは 0 印で表わしてある。個々のパネル・リンパ球の一貫番号は別にあるが、ここでは抗血清 No. 79 および他の抗血清と反応性の高いリンパ球の順に並べかえてある。抗血清 No. 79 は 200 種のパネル・リンパ球中 68 種と反応陽性を示したが、 ρ , Γ 値が共に高い第 I 群の抗血清の反応性はこれに良く類似していた。すなわち、抗血清 No. 81, 83 は抗血清 No. 79 と反応したリンパ球の 57 および 65 種と反応陽性であると共に、標準抗血清と反応しなかったリンパ球の 3, 14 種とのみ反応陽性であった。次に、第 II 群に属する 3 本の妊婦抗血清 No. 382, 309 および 223 では標準抗血清に反応陽性のリンパ球の 57, 51 および 36 種と反応陽性であった。しかし、これらの抗血清は標準抗血清と反応しなかったリンパ球の 6, 13 および 4 種と反応していた。これらの抗血清は第 I 群のものに比べ、標準抗血清との類似性は多少低い。他の群のものと較べるとなおかなり高い類似性を保っており、この範疇に属する抗血清は第 I 群の抗血清と共に標準抗血清とほぼ同じ特異性を持つと思われる。

次に、第 III 群に属する抗血清は標準抗血清と反応するリンパ球の 10~19 種としか反応しなかったが、標準抗血清に反応陰性のリンパ球との反応性も低く、0~3 種のものと反応したに過ぎなかった。以上の事実は本群に属する抗血清の特異性は標準抗血清のそれの中に包含されていることを示し、‘部分的に同一’な共通部分に属するものと思われる。

ρ 値は第 III 群とほぼ同じだが Γ 値が低い第 IV 群の抗血清は、標準抗血清に反応性を示すリンパ球の 23~25 種と反応する反面、標準抗血清と反応しなかったリンパ球の比較的多数 8~10 種と反応した。これらの抗血清は ρ 値が 0.25 以下のものに比し、その特異性は標準抗血清に近いにしても、それが第 III 群に属する抗血清の如く部分的に同一と

いうほど近縁ではないから、一応、B5 特異性とは無関係の群に入れた。

第 V 群から選んだ 3 本の妊婦抗血清の ρ 値は何れも 0 であった。これらは標準抗血清と反応陽性のリンパ球のそれぞれ 23, 14, 6 種と反応し、反応陰性のものの 45, 25, 13 種と反応した。この群では図 6 内に示さなかったリンパ球とも反応陽性のものがあるので、その個数を右端に記した。これら反応陽性のリンパ球は全パネル・リンパ球中に無作為に分布するので、3 本の抗血清の特異性は標準抗血清のそれとは無関係と考えられる。

以上の検討から (ρ , Γ)_i 図の ρ , Γ 値が高い第 I, 第 II 群の抗血清は標準抗血清とほぼ同じ特異性を持ち、 ρ 値は 0.25~0.50 の間にあるが Γ 値の高い第 III 群の抗血清は標準抗血清と部分的に同一の特異性を持つことが明らかになった。この事実は個々のリンパ球の特異性が不明であってもパネルとして多数のリンパ球を用い、抗血清との反応を統計的に処理すればその特異性が同定可能な点で重要である。

〔実験 6〕抗血清の特異性の Chi-square map による検討

(ρ , Γ) 図による検討等からもわかるように、全抗血清の組合せで計算した ρ_{ij} 値が 0.50 (N_L が 200 なので χ^2 値は 50.0) 以上のものを集めると、抗血清は特異性のほぼ等しいいくつかの群または小集団 (cluster) に分けられる。このような小集団をくずすことなく同じ特異性を持つものの主要部分が対角線上の一定部位に来るよう抗血清の順序を並べかえると、抗血清間の相関関係を明瞭に表示できる、Chi-square map が画ける。図 7 は 250 種の抗血清について ρ_{ij} 値が 0.50 以上の場合を黒い小四角として表現した Chi-square map の一例である。繁雑をさけるため縦軸、横軸に同じ順序で並べた抗血清の番号は表示してない。

HLA-B5 系の例をみると特異性既知として分与された 8 種の抗血清と特異性未知として新たに得られた 10 種の妊婦抗血清が高い相関関係を持つ 1 集団を形成している。この対角線上にある B5 集団から上下、左右に離れて存在する黒い小四角は

小黒四角は $\chi^2 \geq 50.0$ すなわち相関係数の ≥ 0.50 の正の相関を示す。
パネル・リンパ球数は200である。
横軸・縦軸ともに同一の順序で抗血清を並べた。この順序は小黒四角が対角線の近くにまとまりとして集まるようにしたものである。
図中のA1, A2, …はHLA抗原系の特異性を示す。

図7 250抗血清間の相関を示す CHI-Square map

極めて少ない。HLA-A1, A2, ……等他の特異性を持った集団についても同様であった。この事実は $\rho = 0.50$ を下限としたとき、特異性を異にする2種以上の抗体を含む抗血清は稀になることを示し、リンパ球の場合と異なり chi-square map 法で抗血清がいくつかの集団に分類できた理由の1つでもあろう。

このようにして、特異性既知の抗血清と相関性の高い妊婦抗血清が多数得られたが、それらの中に集団内での相関性は極めて高いが、他の特異性既知の抗血清の何れとも有意の相関性を示さない抗血清集団が見出された。この集団が示す特異性はJ-1系と名付けられ、図7にも示されている。この系の表現型頻度、遺伝型頻度や他系との関連係数の計算、家系内の遺伝学的検討等を行なった結果、日本人に特有な HLA 抗原系であることがわかり、十字ら⁸⁾により既に報告されている。

考察

HLA 抗原系の本態の分析的解明には、家系内外の遺伝学的調査^{5)16~18)}や、抗血清の HLA 型既知のリンパ球による吸収など血清学的検討¹³⁾が広く行なわれている。しかし、特異性分析のための

情報処理の手法も、新しい HLA 抗原系の確立、新たに得た抗血清から特定の特異性のものの選出、さらに、リンパ球が持つ抗原の同定等に重要と思われる。ところが、HLA 系の研究の歴史が浅く、反応系が複雑でその全貌も不明なため、この方面の研究はあるが⁵⁾¹⁶⁾¹⁷⁾、まだ十分進んだとは云えない。また、ここでは HLA 型の検査には Terasaki の方法¹⁾¹⁴⁾を用いる。この方法は HLA 系の検査法のなかでは最もすぐれているが、反応成績の中に非特異的な陽性や陰性がまぎれこむ余地はなお多少残している。これは、抗血清特異性等の精密な決定に多数の試料について調べ、その結果をここで取扱った統計的方法で処理する必要性を生んだ理由の1つでもあろう。

そこで、本報では血液型の研究から、HLA 系に導入された方法²⁾⁴⁾⁵⁾をもとに、この面の解析法を主に検討したが、次のような特色を持つと思われる。従来の HLA 抗原系の特異性は特異性既知の抗血清やリンパ球を個々の未知試料に組合せ、その反応性を比較して同定するので、複雑なこの系の総括的配慮には多大の労力を要した。これを補うため、ここでは電子計算機を用いてすべての抗血清とパネル・リンパ球の反応の全体像を把握しつつ、HLA 特異性の分類、同定を試みた。その原理は、特異性未知として一旦電子計算機に入力したすべての試料の反応成績から類似度の高いものの集団 (cluster) を作り、それを特異性既知の集団と統計学的に比較し、同定することである。この操作には抗血清、パネル・リンパ球のすべての組合せについて反応を観る要があり、データは checker-board 型のものとなる。しかし、このようなデータを対象とした HLA 系の解析の例¹⁷⁾²⁰⁾は少なく、HLA 系の研究分野中の盲点でもある。

近年、多変量解析法¹⁹⁾がデータ解析法の分野で著しく進歩し、中でも主成分分析法 (principal component analysis) とクラスター分析法 (cluster analysis) が上述の目的に応用可能と思われる。主成分分析法は理論的に優れ、小集団の HLA 系を対象とした検討例²⁰⁾はあるが、大集団の取扱い

が困難なため、ここでは実用化の試みの対象としなかった。ここで狭義のクラスター分析法²¹⁾と呼んでいる方法は所定の式で分類すべき試料間の類似度を計算し、一定の手法で指定した数の小集団に順次まとめ、それらを統計学的に検討、比較するための基礎的理論を指すが、その一般論については電子計算機のプログラムもある。しかし、この理論の血清学、免疫学の領域への適用例はまだ見当たらない。ところが、HLA 系の研究の中に導入されている一定値以上の χ^2 値を持つ抗血清の類似度は高いとして順次クラスターを作っていく Chi-square map 法⁴⁾⁷⁾はこの種の分析法の初歩的なものだから広義のクラスター分析法に入る。

HLA 解析の統計学的処理法の理論的發展と実用化を計るため、上述の事情に次の諸点を考慮してそれに適したクラスター分析法の一種をわれわれは開発した。

- 1) χ^2 値は優れた統計数値だが、他にも優れた類似度の指標はないか。
- 2) これまでの分析法には類似度の指標は1種類しか使用されていない。性質の違う2種以上の指標を組合せれば、分析はより精密になる。
- 3) クラスター形成の段階で指標とした類似度の値が一定値以下でも、機械的に捨てずその意義も考える。
- 4) クラスター形成を電子計算機に自動的にまかせても、もとの反応データは常に参照できるようにしておく。

その結果、ここに報告した $(\rho, \Gamma)_i$ 図と血清反応図を併用する方法が開発された。まず、 $(\rho, \Gamma)_i$ 図では類似度の指標である関連係数 (ρ) と包含係数 (Γ) を2次元の座標上にとり、抗血清等の試料のクラスター化をした。ここで類似度の指標として χ^2 値の代りに $\rho \left(= \pm \sqrt{\frac{\chi^2}{NL}} \right)$ 値を用いたのは、比較すべき抗血清の χ^2 値は使用パネル・リンパ球数 (NL) によって変動するが、 ρ 値はこれに無関係だから、NL の異なる場合の比較が容易になるためである。また、 χ^2 値はその定義から負の値を取らないが、符号を適当にとれば ρ 値は $-1 \leq \rho \leq 1$ の範囲の値をとり、正負の

相関も表現できる。これは ρ 値を用いたもう1つの理由でもある。従って、この方法は広義のクラスター分析法に属するが、2種類の類似度を組合せたため、 (ρ, Γ) 図の標準にした特異性に対して‘ほとんど同一’‘部分的に同一’‘無関係’‘対称的’な特異性を示す試料の分布領域を区別できるようになった(実験4)。 (ρ, Γ) 図では一般に ρ 値が高いほど Γ 値は高くなる傾向があったが、これは相関性が高いほど包含率も高くなるのは当然として、両方の値は 2×2 分割表をもとに求められたことにも起因するのであろう。また、 (ρ, Γ) 図のみでは標準試料に対する未知試料の類似性はわかっても、個々のリンパ球に対する反応の状態を知ることはできない。

そこで、血清反応図でこの欠点を補うと共に (ρ, Γ) 図中の特定の ρ, Γ 値の領域を指定してそこにある試料の反応性を検討できるようにした(実験5, 図6)。その結果、従来の狭義や広義のクラスター分析法ではクラスター内にある試料相互の類似の程度を正確に把握できなかった欠点が補えたのは大きな進歩と云えよう。繁雑をさけるため、図6の血清反応図では並べかえた個々のリンパ球番号の細部および最初の一貫番号は示していないが、電子計算機からの出力にはそれらの表示、および両方の番号の対照を可能にしてあり、特異性の精密な検討に便である。

(ρ, Γ) 図は試料のクラスター化に役立つのみならず、図上における各試料の分布状態からその性質を知る手掛りにもなる。実験1の図3(a), (b)でも見られるように、 (ρ, Γ) 図上での抗血清とリンパ球の分布は異なる。抗血清の分布にはばらつきがあり、特異性が標準と無関係の大多数とこれが類似する少数に分けられる。これに対して、リンパ球は ρ, Γ 値の割合高い領域に一团としてかたまり、その分類は難しい。この分布状態の違いは、Chi-square map で抗血清の分類はできるが、リンパ球では難しく、その試みもない事実とも対応している。従って、現段階でのリンパ球の分類、同定には、従来の血清学的方法を用いるか、統計学的処理でこれを行なうためにはこ

れまでと発想の異なる方法の開発が必要であろう。この分布状態の違いは交叉反応性抗体の存在を考慮に入れても、抗血清は mono または oligo-specific な抗体よりなるものが多いのに対し、リンパ球はその表面に最大限6種の抗原を持つため、multi-specific なものとして働くためであろう。抗血清に関するこの推論は、図7の Chi-square map で特異性の相関を示す黒い小四角が対角線上に集まっており、その上下左右の離れた部位には極めて少ない事実からも裏付けられている。

以上の如く、ここで用いた方法は抗血清の分類、同定に有用なばかりでなく、実験2でも述べた如く、使用する個々のパネル・リンパ球の特異性は必ずしも既知でなくとも良いという優れた特質を持つ。しかし、この特質を十分発揮するには次の2条件を満たしたパネル・リンパ球集団を用いる必要がある。

1) 自明のことだが、パネル・リンパ球集団は抗血清から検出、分類すべき特異性に対応したすべての抗原を含んでいなければならない。また、特殊な場合を除き、それらの抗原特異性の頻度は母集団である民族のそれとほぼ同じであることが望ましい。

2) リンパ球は HLA 系のサブローカス A, B, C の各々に2種ずつまで抗原特異性を荷ないうるが、その組合せは母集団全体としてランダムでなければならない。Mittal ら¹⁸⁾は日本人では、A9 と B5 がリンクしたハプロタイプが非常に多いと報告している。実験2の反応に用いたパネル・リンパ球中にもこのようなリンパ球が多く含まれ、抗 B5 抗血清と反応したものの多くが抗 A9 抗血清とも反応するため、図4では抗 A9 抗血清が比較的高い ρ , Γ 値の領域に分布する原因になったのであろう。この時2種の抗原がリンクしたもののみを含むパネル・リンパ球が用いられていれば、それぞれの抗原に特異的な抗血清の反応態度は全く同じとなり、両者は区別できないことを考えれば、上述の条件は理解できよう。

この2条件を満たすにはパネル・リンパ球は可

能な限り多数の血縁関係のない者から無作為に蒐集する必要があり、クラスターを形成させる特異性既知、未知の抗血清についても同様である。従って、ここで述べた方法で良い結果を得るには、一定数以上の試料を組合せて得た多数のデータを処理するという問題が派生する。この点は中型、大型の計算機の使用と適したプログラムの作製・適用で解決できたが、さらに解析の能率を上げるには血清学的反応の実施、計算機への入力、出力の一貫したシステム化が必要となる。ここでは検査室で得られた反応成績を直接データシートに記入し、迅速、正確に入力できるよう配慮した¹¹⁾。

本報では HLA 特異性の分析のために新たに開発した高速情報処理法の基礎的検討について主に述べたが、この方法を多少修飾すれば、マウスの H-2 系をはじめ他動物の主要組織適合抗原系の解析や、一般の複雑な抗原抗体系の解析にも役立つと思われる。

むすび

1) HLA 抗原系の統計学的手法による解析には膨大なデータの処理を必要とするが、大型計算機の使用で初めてそれが可能になった。ここでは約400種の抗血清に約200種のパネル・リンパ球を組合せた大集団から抗血清特異性の統計学的解析を試みた。

2) 計算機による解析のため、類似度の指標として関連係数 (ρ)、包含係数 (Γ) を2次元的に組合せた (ρ , Γ) 図に血清反応図を併用した広義のクラスター分析法を新たに開発した。(ρ , Γ) 図は主に抗血清の特異性の解析に適するが、その原理は特異性未知の抗血清のパネル・リンパ球の集団に対する反応性を既知のもののそれと統計学的に比較し、分類、クラスター化することにある。この際、パネル・リンパ球個々の特異性は未知でも抗血清の分類は可能である。

3) クラスター形成の途上、実際の反応の様相を整理した形の血清反応図を計算機に画かせ、より精密な特異性の解析を可能とした。

謝辞 本研究に対して終始あたたかい御指導をいただいた東京大学医学部血清学教室鈴木 鑑名誉教授に

心から感謝いたします。また、多大の御助力をいただいた血清学教室教室員各位に深謝する。さらに、実験に関して御支援をいただいた東京大学医学部附属病院輸血部遠山博助教授をはじめとする関係各位および、計算機使用の便宜をいただいた東京大学附属病院電子計算機室開原成允助教授、桜井栄光掛長に深く謝意を表する。

また、貴重な抗血清を分与して下さった NIH 血液銀行, Dr. P. Terasaki, Dr. van Rood, Dr. Simons の御厚意に深謝する。

なお、本論文の一部は第21, 22, 23回日本輸血学会総会および、第3, 第4回日本免疫学会において発表した。

文 献

- 1) Terasaki, P., McClelland, J., Park, M.S. and McCurdy, B.: Microdroplet lymphocyte cytotoxicity test. Manual of Tissue Typing Techniques. National Institutes of Health, Bethesda, Maryland, 54—61, 1973.
- 2) van Rood, J.J. and Leeuwen, A. van: Leucocyte grouping, A method and its application. J. Clin. Invest., **42**: 1382—1390, 1963.
- 3) Mi, M.P. and Morton, N.E.: Blood factor association. Vox Sang., **11**: 434—449, 1966.
- 4) Mickey, M.R., Parmelee, L. and Terasaki, P.I.: Serotyping for homotransplantation. XV. Comparison for three methods of analysis in determining leukocyte groups. Histocompatibility Testing 1967. Munksgaard, Copenhagen, Denmark, 121—127, 1967.
- 5) Albert, E., Mickey, M.R. and Terasaki, P.I.: The use of computer programs for tissue typing and immunogenetic research. Tissue Antigens, **2**: 173—181, 1972.
- 6) 萩野陽一郎, 十字猛夫, 田村 真, 大河内一雄, 遠山 博: 抗白血球抗体陽性日本人妊婦血清における抗 HL-A 抗原の分析. 日本移植学会雑誌, **9**: 138—143, 1973.
- 7) 十字猛夫, 田村 真, 萩野陽一郎, 山田昭夫, 節原忠夫, 遠山 博, 松橋 直: 日本人妊婦血清における新しい HL-A 抗原の分析. 第3回日本免疫学会総会記録. 日本免疫学会, 京都, 1973, 161—163.
- 8) Juji, T., Hagino, Y., Tamura, M. and Okouchi, K.: A possibly new Japanese-specific HL-A antigen (J-1). Jap. J. Exp. Med., **43**: 447—449, 1973.
- 9) 田村 真, 鈴木 鑑, 十字猛夫, 萩野陽一郎, 大河内一雄: 抗 HL-A 血清の同定に関する統計的処理—リンパ球における問題. 日本移植学会雑誌, **9**: 60, 1974.
- 10) 田村 真, 十字猛夫, 萩野陽一郎, 大河内一雄: 抗 HL-A 血清の同定に関する統計学的処理における問題. 日本輸血学会雑誌, **20**: 91—92, 1974.
- 11) 田村 真, 鈴木 鑑, 十字猛夫, 遠山 博, 萩野陽一郎: 高速度情報処理による抗 HL-A 抗血清の解析法. 第4回日本免疫学会総会記録, 日本免疫学会, 京都, 1974, 348—350.
- 12) 田村 真, 鈴木 鑑, 十字猛夫, 萩野陽一郎: 抗 HL-A 血清の同定に関する統計的処理—(ρ , I') 図による検討. 日本輸血学会雑誌, **21**: 155—156, 1975.
- 13) Juji, T., Hagino, Y., Tamura, M., Ohtawa, T., Shibata, Y., Schreuder, I. and Rood, J.J. van: Absorption study on crossreactivity between W22 and J-1 antigens. Histocompatibility Testing 1975. Munksgaard, Copenhagen, Denmark, 283—287, 1975.
- 14) Terasaki, P.I. and McClelland, J.D.: Microdroplet assay of human serum cytotoxins. Nature(Lond.), **204**: 998—1000, 1964.
- 15) 十字猛夫, 奥山マチ子: Human leucocyte antigen の検査法, Micro lymphocyte cytotoxicity test. 臨床検査, **18**: 7—16, 1974.
- 16) Elandt-Johnson, R.C.: Survey of histocompatibility testing. Biometrics, **25**: 207—283, 1969.
- 17) Dausset, J., Ivanyi, P. and Feingold, N.: Tissue alloantigens present in human leukocytes. Ann. N.Y. Acad. Sci., **129**(I): 386—407, 1966.
- 18) Mittal, K.K., Hasegawa, T., Ting, A., Mickey, M.R. and Terasaki, P.I.: Genetic variation in the HL-A system between Ainu, Japanese, and Caucasians. Histocompatibility Testing 1972. Munksgaard, Copenhagen, Denmark, 187—195, 1973.
- 19) 奥野忠一, 久米 均, 芳賀敏郎, 吉沢 正: 多変量解析法. 日科技連出版社, 東京, 1971.
- 20) 安達 直, 林 喜男, 上田修一, 辻 公美: HL-A 抗血清の主成分分析による一考察. 日本移植学会雑誌, **10** (総会臨時号): 23, 1974.
- 21) 奥野忠一, 芳賀敏郎, 矢島敬二, 奥野千恵子, 橋本茂司, 古河陽子: 統多変量解析法. 日科技連出版社, 東京, 1976, 207—237.