

関西実験動物研究会会報

Kansai Journal of Laboratory Animals

平成 2 年10月 8 号

関西実験動物研究会

Kansai Laboratory Animal Research Association

目 次

< 第 2 1 回 研 究 会 >

講演会

毒性試験における統計解析

吉村 功 (名古屋大学工学部) 1

質疑応答 25

< 第 2 2 回 研 究 会 >

講演会

1. 霊長類における ATL-及び AIDS -関連ウィルス

速水 正憲 (京都大学ウィルス研究所) 29

2. 日本ザルの集団構造の遺伝学的解析

野澤 謙 (京都大学霊長類研究所) 36

< そ の 他 >

関西実験動物研究会だより 45

会員の動き 46

編集後記

〈第 2 1 回研究会〉

講演会

1. 毒性試験における統計解析

吉村 功 名古屋大学工学部

2. 癌原性試験における統計解析

菅野 純 東京医科歯科大学 (次号に掲載予定)

日時： 平成元年 3月 3日 (金)

場所： 三和化学研究所 大阪メディカルホール

「毒性試験における統計解析」

吉村功 (名古屋大学工学部助教授)

毒性試験の流れを見た場合それぞれの研究者の立場によって関心を払う対象は異なるであろう。生物系の研究者の場合、動物一匹一匹のあり方、動物の行動や摂食のあり方は大きな関心の対象となるであろうが、データ解析、統計解析の側面から見るとそちらの方はあまり気にならない。統計学者として毒性試験の流れを見た場合、以下に示す箇所が関心の対象となる。

表 1

私の理解している毒性試験の計画の流れ

- 1、薬物を定める (種類、数、形態)
- 2、動物を定める (種、性、齢)
- 3、測定変数を定める (反応変数、用量変数)
- 4、得たい結論の形を定める (記述、探索、確認*)
- 5、投与形態を定める (水準数*, 処置)
- 6、実験の大きさを定める (群の大きさ、期間*)

毒性試験を行う場合にはまず最初に薬物を定め、投与の形態 (内服、注射、塗布など) を決定する。

次に使用する動物を決定する。すなわち種類 (ほ乳類、細胞、細菌など) や性、齢を決

める。

測定変数としては用量変数（動物に与えた原因に関係する部分）と反応変数（現われる反応に関係する部分）を定める。

この後から統計学者の関心が深まってくる部分である。得たい結論がどのようなタイプのものか。何か分らないが、とにかく動物に投与して現われる反応を見よう、そして反応が出たらできるだけ丁寧に観察してできるだけ情報を手に入れようという立場が「記述」の立場である。つまり何が起こったかできるだけ詳しく知るという立場である。「探索」は単に記述する以上にある種の仮説のようなものを作りあげようとする発想が働く。つまりこのホルモンを与えればこの臓器に何らかの肥大が生じるに違いないといった仮説をその化学構造式から推定していく部分である。次にその仮説に対しその仮説が妥当であるかいなかを確かめることが「確認」になる。単なる「記述: descriptive」で終わるか、「探索: exploratory」の部分で攻め込むのか、それらを周知のこととして最終決定に結びつけるため「確認: confirmatory」するのかがデータ解析のうえで大きな分かれ目になる。これらの区別は実際のデータを取る時の目的として意識すべきである。

どのような目的であるかによって例えば動物の種類の定めかた、測定変数の決定の方法も変わってくるし、後のデータ解析の方法も変わってくる。ここで投与形態を定めるとするのは、薬物の形態を定めるのとは意味が異なる。群をいくつ作るのか、各群の用量をいくらにするのか、処置の種類をどうきめるのか、いつ、どのような時期に処置をし、いつ結果を見るのかを問題にする。また実験の大きさ、一群の匹数、投与期間、屠殺の時期を定めることである。なお、統計的用語で「群の大きさ（サイズ）」とは「一群の匹数」を意味し、それに対して何群あるかということは「群の数」と呼ぶ。ここで*印のついている部分は統計学者が特に関心を持つ部分である。

表 2

毒性試験実施の流れ

動物の群分け（ランダム化と層別*）
測定（誤差要因、欠損値、欠測値*）
データ解析（生データとクリーニング*、手法の選択*）
結果の表現・表示（誤解、再利用*）

*の部分は、統計の素養が役立つ、あるいはそれを役立たせるべき事項

実際に毒性試験を実施する場合には動物を群分けするが、これをランダムにする場合と層別して行う場合とがある。それぞれ一長一短あるがどのような理屈で一長一短あるか理解しておく必要がある。これは毒性試験のデザインの問題である。また測定すると誤差が生じるが誤差がどのような形ではいるかによってデータ解析の方法が変わってくる。頭を痛めるのは欠測値、欠損値である。欠測値、欠損値が原因と全く無関係の場合には単にバランスが崩れているだけですむが、原因と関係している欠測値、欠損値は困るわけである。つまり毒物の影響で死んでしまってデータが取れなくなって数が減ってしまったか、全く無関係な原因で死亡しデータが取れなくなってしまったかが問題になりこれによってデータの解析方法が変わってくる。ここにデータ解析のうえでの論争点が出てくる。実際にデータ解析をする場合にはいうまでもなく生データを必ずcleaningしないといけないという部分が出てくる。目的とは関係ないミスで思ったとおりの結果が出てこないことがあるがこのような場合には何らかの修正を施さなければならない。実際にデータをどのような手法で解析してどのような結論を導くかはさきほどのデータ解析の目的とのからみ合いでいろいろと分けがつく。結果の表現表示については従来、比較的軽視されてきていたが私は非常に重要と考えている。データをある程度整理し、最後に表現した時にそのpresentationの仕方によってmisleadingな結果となることが非常に多い。つまりもとのデータはごまかしも嘘もついていないのに提示方法によってとんでもない誤解を生むような提示の仕方がある。実際に一つのデータを出した時にそのデータが、後々最初に人が思ったのと違ったデータの使われ方をすることがよくある。自分としては毒性の有無の結論を出したつもりでいたら、後で他の人が用量やバラツキにまで及ぶ議論をする材料に使うような例もある。そのようなことを考えるなら単に自分の欲しい結論だけを提示するのではなく別の使い方があるのなら他の人が使える形でデータを整理して出しておく方がよい。そのような意味で結果の表現、表示はかなり重要であると私は考えるが、この問題は個人の趣味の問題としてややないがしろにされる傾向もあると思う。

ところで毒性試験について関心を持つ統計学者が増えてきており特にアメリカにおいてそれが顕著である。生物関係へのマンパワーの投入には著しいものがある。本邦ではアメリカに比べて一桁以上マンパワーが少ないのではないと思われる。とはいっても生物関係に興味を持つ統計学者が増えてきており、応用統計学会の最近の発表は半分ぐらいが生物関係の発表になっている。では統計学者がどのような部分に関心を持っているかというとそれは以下のとおりである。

表 3

<u>統計学者にとっての現在のトピックは？</u>	
1、手法の研究・開発面	2、利用面
効率のよい計画	標準化
水準数をいくつにすべきか？	手法採用の恣意性の排除
群の大きさは？	共同作業
多重性の処理	単なる計算業でなく
多項目、多時点、多群、多施設	コンサルテーション
多実験（種、性、齢）	現場の問題の実データによる
多重比較（多重仮説）	質疑応答
背景対照の利用法	
統計的管理状態？	
十分なデータ？	

この傾向は外国でも同様の状況のようでこの夏にはフランスでStatistical Methods in Biopharmacy というようなシンポジウムが開かれたりしている。この内容を見てみるとCross-over designsとかMulti-center studies and meta-analysisといった臨床試験に関係しているものや、動物実験に関係するものではMultiple comparisons（多重比較）、Analysis and description of laboratory data ,Philosophical and ethical problems of statistical methods in biopharmacy などがヨーロッパでは非常に大きな興味になっていると考えられる。

これらの問題が大きくとらえられるようになってきたのはなぜであろうか。動物実験は昔から行われており、なぜ最近になってきてこのような問題に関心が注がれるようになってきたのであろうか。

表 4

<u>これらのテーマが問題として大きくなったのは？</u>
<u>動物実験をめぐる最近の状況の変化</u>
規制・認可、監視（GLP）が社会の重要な手続き
そこで検定という統計手法が繁用
実験がルーチン化
検出したい毒性が（低レベル、特殊化）見にくくなった
損得共に、費用・時間の規模が大
動物愛護の強調

そのような状況の中で効率のよい計画というのが重要になる。つまりなるべく少数の動物で結論を出したいということである。昔薬害が問題になっていた時には検出力が重要視されていた。動物数を増やしどんな小さな変化も見逃さない試みがなされていた。私はこの当時から多くの動物を殺しすぎるのはおかしいと思っていたが最近では案の定動物を殺しすぎるのはよくないという意見が強まってきている。とするとぎりぎりの効率を追求する計画を真剣に考えることが一つのポイントとなってくる。

検定という手法は毒性があるかないかということをデータから断定する訳であるが、その場合、必ず第1種の過誤と第2種の過誤が生じる。第1種の過誤は実際には毒性がないにもかかわらず毒性があると判定してしまう、いわゆるfalse positiveである。それに対し第2種の過誤は毒性があるにも拘らず見逃してしまう、いわゆるfalse negativeである。検定という手法は第1種の過誤をある一定の確率、例えば5%に抑える、という前提でできている。ところが毒性試験での検査項目が600項目あったとすると1項目ずつのfalse positiveの確率が5%であっても、600項目の5%は何がなくても有意差ありという形で出てきてしまう。すると何のために統計処理をしているかわからなくなってしまう。この問題が多重性である。最近ではありとあらゆる毒性を調べよということになって項目数が増えてきて多重性の処理が問題になる次第である。

背景対照 (historical control) の利用法は何も処置しない動物が一試験に必ず一群出てくる。それを毎回やっているとやった実験の回数分だけ何も処置しなかった動物のデータが溜ってくる。そのようなデータが蓄積されてくれば毎回の実験での対照群の動物を省略して過去の対照群の動物のデータで代用できないか。そうすれば対照群の動物数だけ殺す動物が少なくてすむことになる。これが背景対照の利用ということである。

利用面からいうと標準化というのは、微妙なレベルでデータを扱っているとデータの扱い方次第で結論が変わるという現象が起きる。そうするとある種の意図を持って結論を出そうとしてそれに合わせて手法を選ぶという奇妙なことが起きる。統計手法選択の自由度が出てきて毒性を出したくない時には毒性の出ない統計手法を、毒性を出したい時には毒性の出る手法を選択することになりこれはある意味で非常に危険なことになる。そのような恣意的な手法の利用を排除したほうが客観的にはよいことになる。これが標準化である。これは新しい手法を選ぶのではなくどのような手法をどのように使い分けるかという問題である。これは重要なことであり私はいろいろな所で手法の標準化の重要性を強調している。

共同作業についていうと、統計学者は動物実験の苦勞を知らない。苦勞を知らないと実際のデータ解析のしかたを間違える。ロッシュの研究所を見学した時に気付いたことだが例えば 300匹のラットを解剖するには大変な時間がかかる。ある一定の時間が掛かっている間に最初の動物は食後に解剖しているのに最後の動物は空腹状態での解剖となる。臓器重量にも影響するようなことになる。また臓器の摘出切断部位によっても臓器重量は微妙に変わってくる。ところが統計学者は現場のこのような状況を知らない。ただ出てきたデータを全て同じだと思い込んでいる。どうしても一時間に 600匹を剖検しないとイケないとする人数を増やす以外にはない。すると作業者の偏りを考慮しなければならなくなる。そのような構造はわれわれ統計学者には分からないのである。一方、作業者の面からいうとそのような作業の違いがデータの統計解析をどのように狂わせるか分からないであろう。したがってデータ解析での結論の導き方には統計学者と現場の作業者の密接な共同作業が必要である。

consultationの問題であるが、統計学の研究者達はアカデミックな世界に浸ってしまい、論文などを書いている。ところが現場の人間は意外な所で意外な問題を抱えその解決法を見出せないでいる。しょうがないので適当な本を捜し、似たような事例への適用例を利用する。その結果、時に不適当な解析法を用いることになったりする。

ところで具体的な例をいくつか提示したい。第一の例は多重性についてのものである。ここでは検定手法の多重性を一つの例として提示する。同じデータに対していろいろな検定手法がある。これを勝手に選んだらどうなるかという例で図1は東洋醸造の松本一彦氏らが昨年の秋、自社のデータを細工して提出したものである。ある化学物質の対照群、32、96、320、640mg/kgを1群4匹のイヌに投与した薬物投与群4群の実験でのコリンエステラーゼのデータである。1群4匹であるから各群毎に4個のデータがあるが、群のデータを取り変えるとどうなるかをみたのが次の結果である。対照群、第2群～第4群のデータを入れ替えて、それぞれの場合に対してある種の統計手法を変えた場合に結論がどのように変わるか検討したわけである。

試してみた検定手法は以下のとおりであり、これらの検定手法はどれも同じデータに適用可能で、有意差の有無で結論が出る。いい換えれば毒性の有無という形で結論が出る。

表5

Control	Ch-E	Ch-E	Ch-E	Ch-E
N	4	4	4	4
Mean	393.37	393.37	393.37	362.15
S. D.	45.42	45.42	45.42	29.65
32mg/kg	Ch-E	Ch-E	Ch-E	Ch-E
N	4	4	4	4
Mean	397.00	397.00	362.15	393.37
S. D.	51.85	51.85	29.65	45.42
96mg/kg	Ch-E	Ch-E	Ch-E	Ch-E
N	4	4	4	4
Mean	477.02	362.15	397.00	397.00
S. D.	41.50	29.65	51.85	51.85
	t*			
	We*			
320mg/kg	Ch-E	Ch-E	Ch-E	Ch-E
N	4	4	4	4
Mean	362.15	477.02	477.02	477.02
S. D.	29.65	41.50	41.50	41.50
		t*	t*	t* W*
		We*	We*	We**
				Wc*
640mg/kg	Ch-E	Ch-E	Ch-E	Ch-E
N	4	4	4	4
Mean	480.25	480.25	480.25	480.25
S. D.	125.50	125.50	125.50	125.50
				W*

表6

1、対応のない2群の検定手法

t : 2群の平均の差のt検定

Wel : 2群の平均の差のWelch 検定

Wc : Wilcoxonの順位和検定

2、対応のない多群の検定手法

T : Tukey の多重比較

S : Scheffe の多重比較

D : Dunnett の多重比較

W : Williamsの多重比較

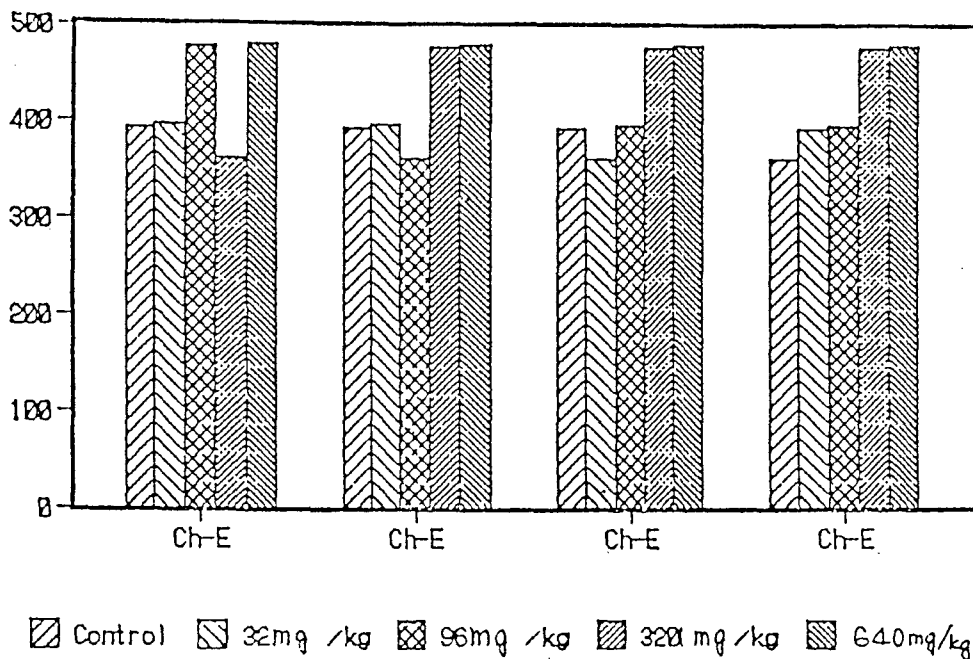
3、シグナル表示の意味

* : 両側5%有意 + : 片側5%有意

** : 両側1%有意 ++ : 片側1%有意

! : 両側1%有意 # : 片側1%有意

図1



このデータではどの程度の結論の食い違いが出るものであろうか。用量相関性のある綺麗なデータではどの統計手法を使っても同様の結論が出るが、きたないデータでは統計手法によって結論の出方が異なってくる可能性は十分にある。実際にやってみるとt検定、Welch 検定、Wilcoxonの順位和検定、Williamsの多重比較検定では1もしくは5%の危険率で有意のものが出てくるが、Tukey の多重比較検定、Dunnett の多重比較検定ではこのどの場合でも有意差なしという結論になる。これに対してt検定はペアごとの検定であるから、対照群と第3群（左の列）、対照群と第4群（左から2番目、3番目の列）との間に有意差が出てくるが、Williamsの検定ではこのような場合有意差なしという形で出てくる。そして一番右側の場合のみ第4群と第5群においてWilliamsの検定で有意差ありとなる。すなわち統計手法のどれを使うかによって結論が異なってくる。

そこでこのデータをみた時、どの統計手法が毒性の有無を直感的に判断したときの結論と一致しているかが重要となる。このようなときに使用する統計処理手法をあらかじめ決めておかず、出したい結論に添って統計手法を選択すれば、自分の好きな結論が得られることになってしまう。つまりもっとも合理的と思われる手法で処理することをあらかじめ決定しておかなければおそろしいことになるわけである。

そういうことを防ぐのに重要なことは統計手法の標準化である。皆さんであつたらこのような場合どの手法を標準として採用するであろうか。

種を明かすとこれらのデータは全て対照群のデータである。すなわち背景対照のデータを5つとってきて並べてみたのである。松本氏は対照群のデータを並べただけでこれだけ差がつくのだということを示したかったわけであるが、私が面白いと思ったのは対照群だけでもこれだけばらつくのに、その時ならべかたを変えると有意差がつくということである。データは5種類であるからその並べかたは5の階乗（120通り）ということになるが、そのうちの2通り（平均値477.02および480.25のデータが第4群と5群に来た時のみ）についてのみ有意差ありという結論を出してもおかしくないことになり、本来対照群データである2番目や3番目の列の並び方のような場合に有意差のつく統計手法はよくないという気がするのである。120個の中で2つだけ有意差ありと出すような結論というのは有意水準約2%であり、有意水準2%で結論を出す統計手法はいまの場合にはWilliamsの検定である。右の列で第2群と第3群が入れ替わっても有意差はつくかも知れないから、有意水準約4/120になるかも知れないが、この程度のオーダーで有意差ありと結論を出すような手法は合理的なのではないかと私は思う。だとすればこの種の問題に関して

はWilliamsの手法が合理的で、多分標準化にも向いていると思う。

先ほども述べたようにわれわれは統計学者であり、それに対して現場でいろいろ仕事をされている皆さんもいるわけである。そのような皆さんにconsultationをするとき単に聞いて指示することはあまりうまく行かないのではないかとと思っている。やはり共同作業が必須である。つまり私の方もいろいろ教えていただく、逆に私の知っていることもお教える。また一緒に考えることで何かが得られる。そのようなことが重要ではないかと思われる。

次に示す例は国立衛生試験所の林氏の示したものである。林氏はあるデータの処理についてある方法を提案した。しかしこの方法がどのようによいのかよくわからない。そこで私と一緒に種々検討、議論をし、やはりこれはこういう意味があるのではないかという結論を導いた例である。これはEnviromental Molecular Mutagenesis(1989)に採用されたものである。

このデータはどういう状況のもとで出てくるのかというと、先ほど述べたように日常的に同じやり方の実験が同じ動物種、性、年令で投与物質を変えて繰り返され、いわゆる背景対照データが十分に蓄えられている。そして処置群と対照群との差がかなり小さくなっているために微妙な問題が出てくるという場合である。

本当に実験が十分な感度で行われているかをチェックするには、陽性対照を入れる。ところが陽性対照というのは明らかに差があるため、実験をやっているうちにそれが陽性対照であることが分かってしまう。ブラインドで測っているはずなのに陽性対照であることが分かってしまい、多分これ位の値が出るはずだという測り方をしてしまったりする。これが陽性対照を入れた時の非常に困ったところである。陽性対照を入れないと不安だということで陽性対照を入れることが多いが、それがデータ解析に生かされていないのではないかという問題がある。皆さんの所ではいかがであろうか。実際に背景対照データがあるはずである。陽性対照も入れている。しかしそれをうまく使っているであろうか。

私がよく耳にするのは背景対照は得られたデータが自分に都合の悪い時だけ使うということである。用量相関性のあるデータが得られてもその最高用量のデータが背景対照に比較して差があるとは思えない。そういうときに有意差はなしにしよう、というようなことで背景対照データが利用されるのではなかろうか。

私はこのような背景対照データの使い方が 100%間違いというわけではないと思っている。かといって 100%正しいわけでもない。何が問題かということと都合の悪い時だけ使うの

が問題で、使うのなら都合のよい時も使うべきである。都合のよい悪いに拘らず背景対照データを使うべきなのである。そのためには背景対照データの使い方を恣意的ではなくそれをデータ解析に組み込んでおくべきである。もう一つは陽性対照を入れておきながら陽性対照をきちんとデータ解析に組み込んだ例があるのであろうか。だいたい飾り物になっていることが多い。それでは何のために入れているかわからないわけで、やはり陽性対照を入れたら入れたなりに使うべきだと思う。というような話をしていたら林氏がそれを自分の所でやってみましょうということになった。

林氏の直面した状況は以下に示すとおりである。

表 7

林氏の直面した状況

マウスを用いたin vivo 変異原性試験

骨髓液中の小核を有する多染性赤血球MNPCE 出現頻度

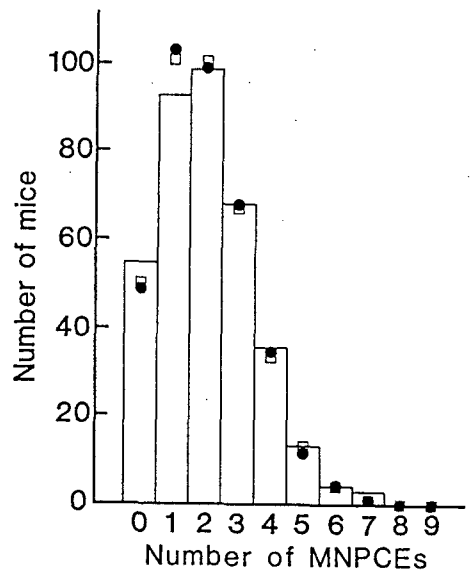
顕微鏡下で各動物から $m (=1000)$ 個の赤血球を調べその中のMNPCE 数 X_{ij} を調べる。

$X_{ij} \sim B(m, \pi_i)$ あるいは $Po(m\pi_i)$ すなわち $E\{X_{ij}\} = m\pi_i$

通常: $\pi_i \approx 0.002$ 群の数 $a = 2 \sim 4$ 群の大きさ $r = 4 \sim 6$

すなわちマウスを用いたin vivo 変異原性試験で骨髓液中の小核を有する多染性赤血球 (MNPCE) の出現頻度を問題にしている。MNPCE が高くなることが変異原性があることを意味する。実際の作業としては顕微鏡下で各動物から $m (=1000)$ 個の赤血球を調べその中のMNPCE の数 X_{ij} を調べる。 ij という添字をつけたのは第 i 群の第 j 番目の動物という意味である。実験は1週間ぐらいで片付くが顕微鏡を眺めるのが大変な作業であるということである。従って短期間で行われる実験である。ただし実際の動物は殺してしまうことになるから短期間ではあるけれど動物はちゃんと殺されるという仕組みになっている。その第 i 番目の群の第 j 番目の動物のMNPCE の数 X_{ij} が過去のデータによると二項分布あるいは非常に出現率が小さいのでポアソン分布で近似できるがそのような分布に従うということが分かっている。だいたいその出現率が対照群において 0.002ぐらいである。普通にやる実験では群の数 $a = 2 \sim 4$ 群である。群の大きさは $r = 4 \sim 6$ 匹/群である。それで過去の背景対照のデータを度数分布でとると図2のようになったという。

図2



つまり横軸が千個数えた時のMNPCE の数、出現率が0.002 というのは千個の中で2個ぐ
 らいあるのがコントロールの状態。これをみると平均が2ぐらいになっている。ヒストグ
 ラムのほうが実際に観測したデータで、●とか□は二項分布あるいはポアソン分布をあて
 はめたものである。この程度であるなら明らかにこのヒストグラムはあてはめた分布に非
 常によく合っている。だからこのデータは二項分布あるいはポアソン分布と看做して差し
 支えない。分布だけでなく、時間順での管理図をかいても図3のようによい管理状態にな
 っている。

図3

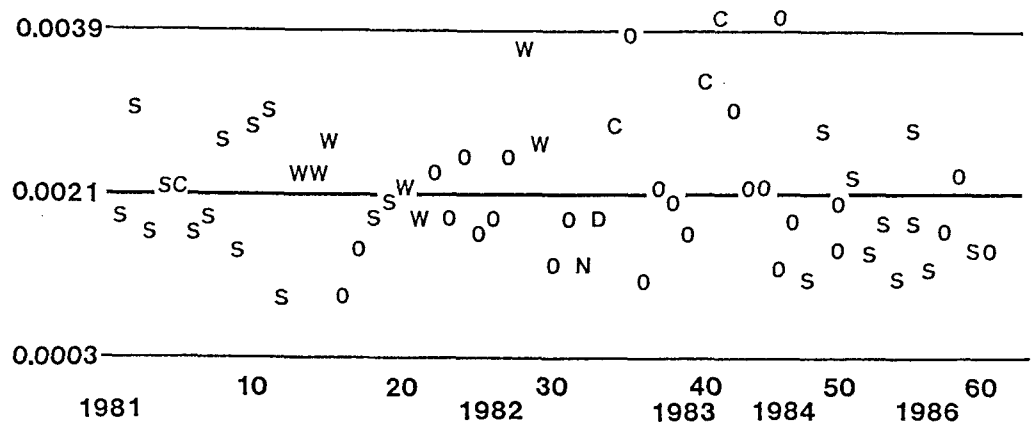
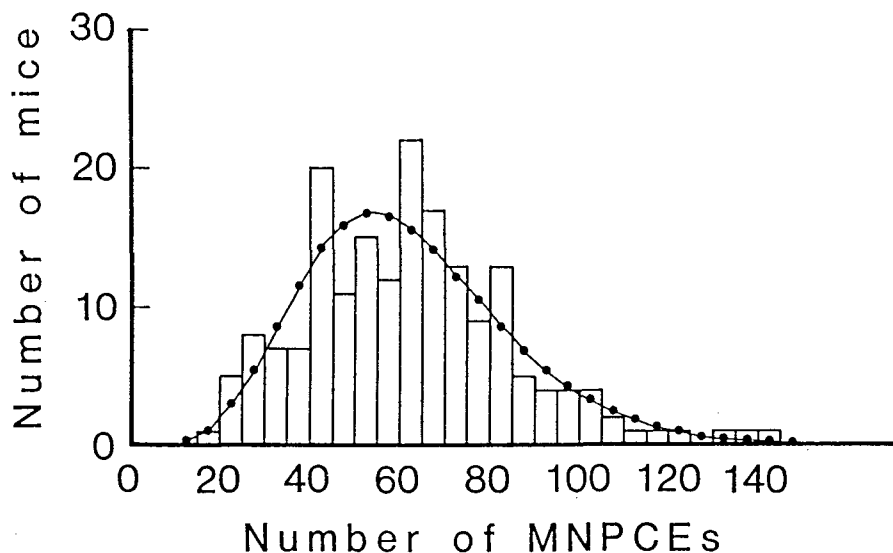
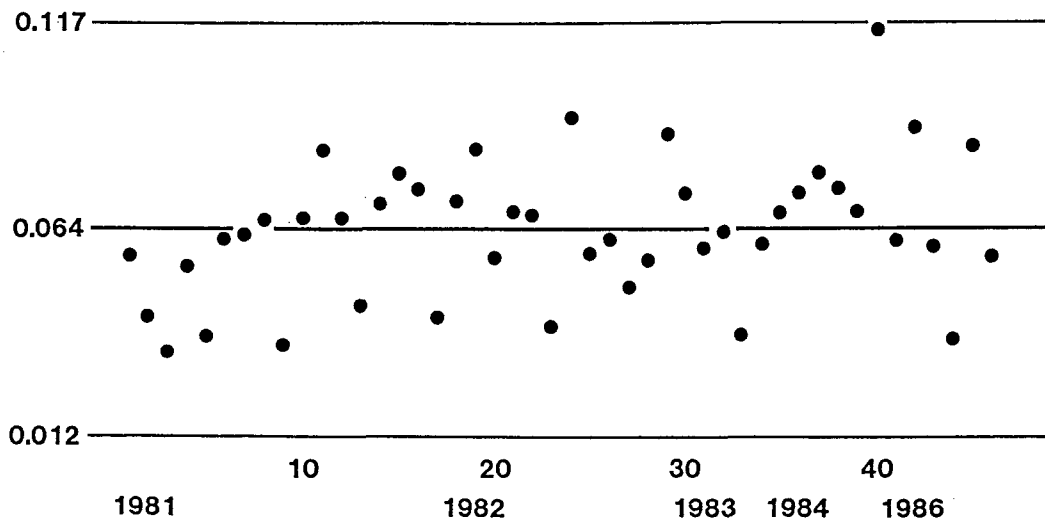


図4



ところが陽性対照は二項分布には従わない。図4のヒストグラムに示すようにばらつきがずっと大きい。動物というものはそれだけの処置をすると個体差が大きくなってくるのであろう。図中のなめらかな曲線は負の二項分布というものである。このヒストグラムは負の二項分布が非常によく適合する。だから陽性対照は陽性対照なりにきれいな分布に従う。さらに、時間的変化をみても、図5に示すようにこれ位だと管理図としては合格点である。

図5



林氏は一人でこの仕事をされているようであるので個人差は入らない。だから過去の背景対照群はうまく管理されている。だから対照群はこういう分布に従って変動すると看做して差し支えないであろう。だとするならこのデータをちゃんと使ってデータ解析をしたほうがいいに違いない。じゃあどうしたらよいか、林氏は次のように提案する。

表 8

林氏が提案する 3 ステップの判定法

陰性対照、陽性対照を入れて実験し、群単位で出現率をまとめ、次のルールで判定を行う。

- (1) p 管理図を用いて、当該実験の陰性および陽性対照群の結果が背景データの範囲内かどうか調べる。
上限、下限は背景データの平均値 $\pm 3 \cdot$ 標準偏差。
→ 範囲外なら実験のやり直し。
- (2) 各処理群の MNPCE 出現頻度を、陰性対照群の背景データから推定される二項分布と比較し、同一集団からの標本とみなせるかどうかを検討する。
→ 同一標本なら処理物質の変異原性が陰性。
- (3) 陰性対照群及び処理群のデータについて、Cochran-Armitage の傾向検定を行い、用量反応関係があるかどうか検定する。
→ 有意差なければ処理物質の変異原性が陰性。

なぜ管理図か？二項分布とどのように比較？有意水準は？

このようなデータの解析には 3 ステップの検定手法をうまく使い、総合的に判断するのがよいのではないかというのが林氏の主張である。第 1 ステップは何かというといわゆる p 管理図を用いて、当該実験の陰性および陽性対照群の結果が背景データの範囲内かどうか調べる。その時は 3 シグマの限界を使う。つまり陰性対照群の管理図、陽性対照群の管理図を作る。そして品質管理でいうところの管理限界を越えたならば実験は失敗とする。これはもちろん 1 週間ぐらいでやれる実験であるからやれるわけで、これが発癌実験のように 2 年たって実験はやり直しということになれば大変だが……。これはもちろん状況によってこういうステップを入れてよいかどうかということは異なってくる。こういう短期試験ではこのようなことが許されるわけである。このようにして陽性対照、陰性対照を

使おうというものである。

第2ステップとしては管理図をパスしたなら各処理群のMNPCE 出現頻度を、陰性対照群の背景対照データから推定される二項分布と比較し、同一集団からの標本とみなせるかどうかを検討する。つまり処置群はいくつかあるが、その処置群のすべてを背景対照と比べ平均値に有意差があるか否かを判断しよう。もちろん低いほうは問題なく、高いほうが問題となる。つまり処置をしてよくなることはないのであるから処置を受けて悪くなったものだけを問題にする。そうすると背景対照のほうを現対照よりも優先するということになる。

それだけではまずいので3番目のステップとして陰性対照群及び処理群のデータについて、Cochran-Armitageの傾向検定を行い、用量反応関係があるかどうか検定する。

Cochran-Armitageの傾向検定というのはご承知の方も多いと思われるが基本的には線形回帰である。つまり用量とともに直線的に出現率が増えているかという検定で直線的に出現率が増えていれば有意差あり、デコボコであれば有意差なしということである。だから要はここでやった実験が過去の陽性対照、陰性対照から大きくずれていたなら実験のやり直し。それから今度やった処置群が過去の背景対照群と明らかに差があって、しかもその傾向が単調増加であるときは変異原性ありと判定する。

そこで問題はそれぞれの検定の有意水準をどこに設けるかである。一回だけの検定では有意水準を1%か5%かにすれば有意水準はピタッと出るわけだが、今ここでは検定を3つやっているわけなので第1種の過誤をある一定の値以下に押えるためには1番目をどうするか、2番目をどうするか、3番目をどうするかということを全部評価しなければならない。それについて最初に投稿した時にやや曖昧だったためにレフェリーから苦情が出てこれについてはっきりさせるため林氏と去年の秋いろいろ工夫を試みた。そこではっきりさせたやり方が本当によいかどうかを調べた。調べるにはMonte-Carlo simulationというのを利用した。Monte-Carlo simulationというのは乱数実験。乱数を発生させて本当に出現率が増えていたらどうだろう、本当に出現率が増えていなかったらどうだろうということを実験を第1種の過誤と検出力 (true positive を発見できる確率) で調べてみた。調べるからには過去にいろいろ使われている手法と比較しなければ意味はない。ここでは4つの手法を比較した。

表 9

比較した 4 方法

□ Cochran-Armitage の検定
 $\Sigma (P_i, -p_{..}) (d_i - d_{..})$

 $\sqrt{p_{..} (1-p_{..}) \Sigma (d_i - d_{..})^2}$
 $> u(1\%)$

○ 陰性背景対照との二項検定
 $\text{Max}(y_i) > B(r_m, \Pi_-)$
 上側 1 / α % 点

△ 条件付き二項検定
 (Kasutenbaum-Bowman)
 $\text{Max}(y_i) > B(y_0 + y_i, 0.5)$
 上側 1 % 点

● 林 提案
 C - A 検定を $\alpha_1 = 5 / \alpha$ %
 二項検定を $\alpha_2 = 5$ %
 ---> 1% 有意水準

40000 回反復のモンテカルロ実験
 ---> 背景対照が信頼できるなら
 現対照を無視するのがよい。

一番上はCochran-Armitageの検定だけを使う。つまり単調に増加しているだけで判定する。これは有意水準を1%とした。これはルーチンワークなので実験回数は年間何十もするので有意水準5%などだとすると結構false positiveがでてしまうので1%にした。だからある意味で厳しい検定をやっているといえる。これはアメリカではCochran-Armitageの検定をやるべきだという論文が出ているから、比較の対象とした。

2番目は陰性背景対照との二項検定。つまり現在の対照群は気にしない。ひたすら過去の背景対照と差があるかどうかだけで判定しようというのが2番目の手法。なお4つの検定で第1ステップは全て共通にした。すなわち3シグマの限界を出たものはやり直すというのは全て共通にしている。

3番目は条件付き二項検定、Kastenbaum-Bowman の数表というのがあるが、要するに対照群と処置群との間の出現率が同じであるかどうかというのをいわゆる条件付き二項検定で調べる。これは変異原性関係では結構使われている手法である。

最後に林提案になるが検定が第2ステップと第3ステップで2つあるので第2ステップの方は5 / α %、第3ステップの方は5%で有意水準を決定した。つまりトータルとして

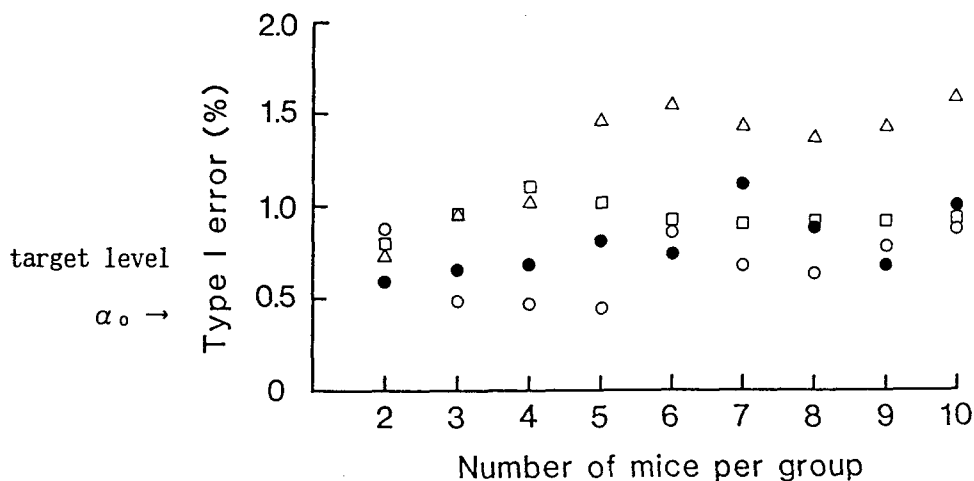
2つあるので1%とするためには両方を適当に調節するよりしようがない。ここではこのように調節した。なぜこのような調節のしかたをしたのかといえば、だいたいこのように調節すると不思議なことに有意水準1%になるのである。

そこで40000回反復のモンテカルロ実験をやった結果が図6である。全ての群が0.2%の出現率であるなら有意差なしであるべきなのに有意差ありという結論を出した確率は、調べると△印は有意水準1%を狙っているにも関わらず有意水準1.5%と割と高くなっている。つまり第1種の過誤が多くなっている。それに対して逆に○はやや小さくなる傾向にある。一番有意水準1%を守っているのは□である。□はCochran-Armitageのtrend testである。これはアメリカでは結構推奨されているものである。しかしながら第一種の過誤というのはある値を越えないというのがessentialであるからなるべく1%なら1%以下であることが重要である。で小さい方については小さいからけしからんということは余り言わない。問題はむしろ検出力である。本当にpositiveだった時にこれをちゃんと検出できることの方が重要である。

図6

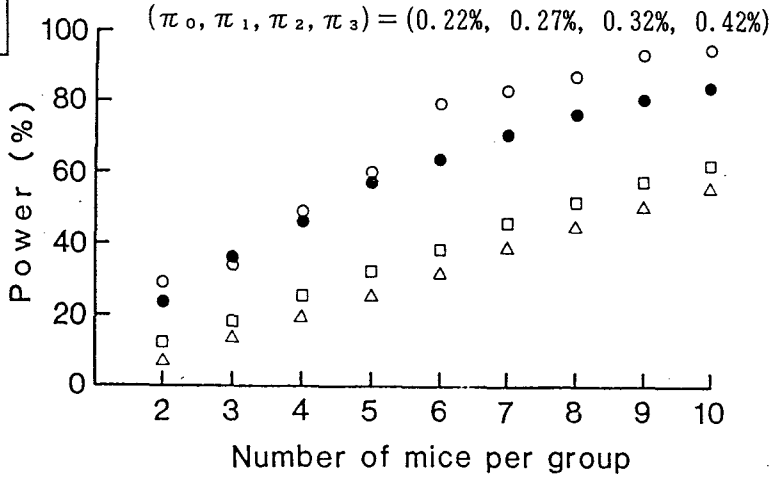
Result of Monte-Carlo simulation

$$(\pi_0, \pi_1, \pi_2, \pi_3) = (0.2\%, 0.2\%, 0.2\%, 0.2\%)$$



そこで調べるとたとえば図7の結果が得られる。△や□は検出力が悪い。Cochran-Armitageのtrend testや条件付き二項検定は検出力が悪いが背景対照との差をみる二項検定は○で検出力がいい。●は林氏の提案のものである。○と●は似たような傾向にあるが、他の場合では○がどうしても上にいく。したがって二項検定と林氏の検定は、ほぼcomparableであるけれどもほんのちょっと二項検定のほうがいい。

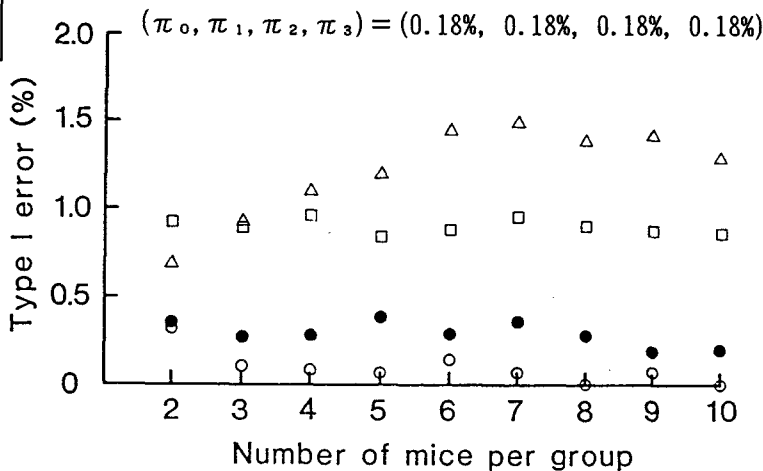
図7



これは考えてみればあたりまえのことである。なぜかというこのMonte-Carlo simulationは過去の分布が正しくて今回の実験でもちゃんと対照が出てくるとというのが前提である。このときは過去のデータの方がはるかに信用できるわけである。1回きりの実験よりも過去に延々とやってきたデータのほうが質的に差がなければこちらを使うほうがいいに決まっている。だとすれば二項検定が優れていることになり、林氏の提案は意味がなくなりこんな面倒くさいことをしなくても背景対照データを 100%信用してやればいいということになる。では今後の実験では陰性対照も陽性対照も入れないでおこう、データ解析に使わないでおこうと言ったとすると、きっと皆さんは反対するであろう。

そこで2人は考えた。つまり日常的に考えていることと結論が違っている。どこかでわれわれの考えの中におかしな所があるのではないかと約1カ月ほど考えた。そしてやっと解ったことが図8である。

図8



背景対照が安定しているときであっても、実際の実験に対照群を入れるのはなぜかという、いくら安定していても多少の変動はあり得るということにあると思われる。もちろん大きな変動であれば第1ステップで排除できるがバックグラウンドの変動の範囲内で出現率 0.2%のものがほんの少し上がったり下がったりすることがあるのではないか。このような場合どうなるのか。ここに対照群を入れる理由があると思われる。現在の実験に対照群を入れるのはこういう状況があるからではないであろうか。そうすると当然のことながらCochran-Armitageのtrend testは背景対照を考慮に入れていないから、有意水準は1%のままであるが、林氏の提案や二項検定は出現率が0.2%であることを前提にしているので、実際が0.18%であれば有意差が出ないことは当然であり、第1種の過誤が小さくなるのは当たり前である。問題は検出力である。検出力は図9である。

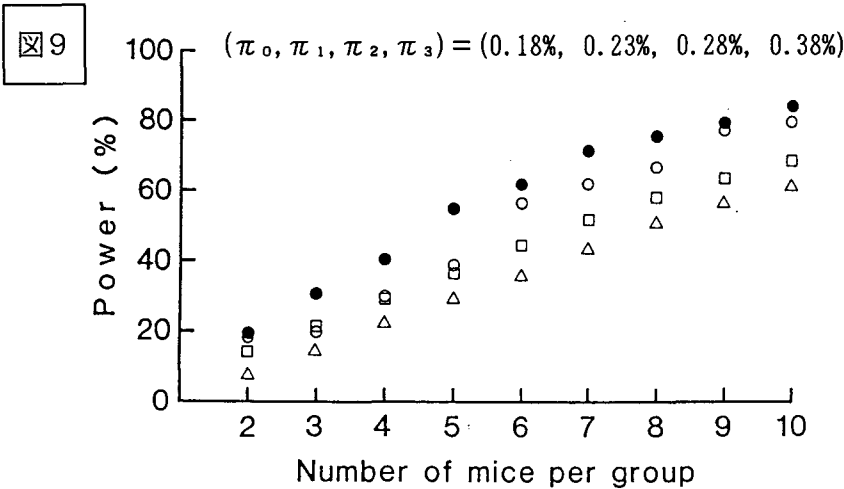
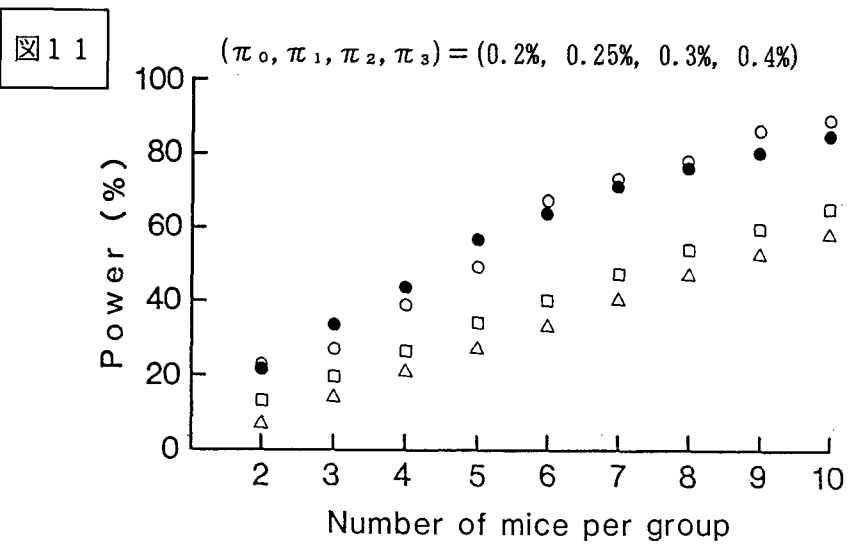
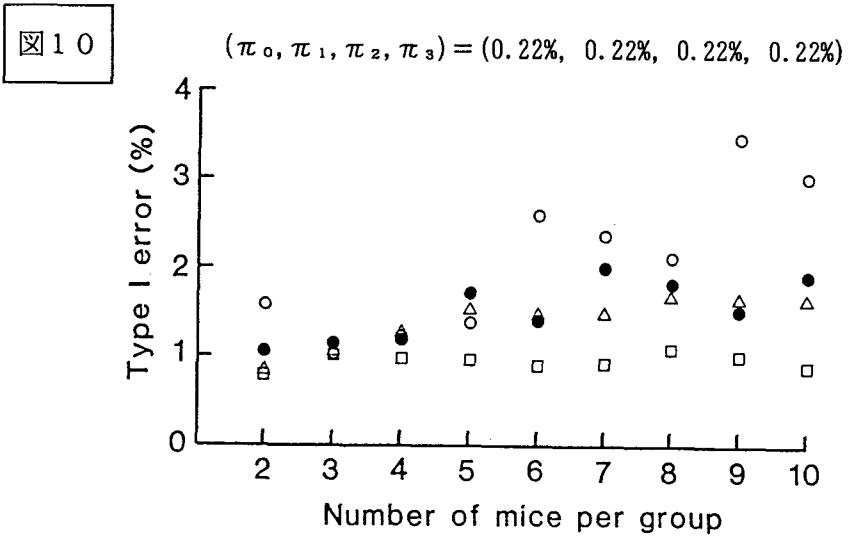


図8に示すように第1の過誤が小さくなったにも拘らず、提案した方法では検出力が保持されているのである。ここで面白いのはこの場合には●が○よりもあきらかに上にいくことである。ということは林氏の方法はこういう場合には強い、第1種の過誤が小さくなっているにも拘らず、検出力が他のものよりはるかによくなっているということである。これは実験全体にちょっとだけ差が生じるという要因が入っていた場合には現在の対照群を考慮に入れたデータ解析のほうが役に立つ。単に過去の対照群だけではなく・・・。一方現在の対照群だけに依拠するというのも弱い。だから過去のデータと現在のデータをほどほどに混ぜて使うのがよい。先ほどでも二項検定と比べるとよくないとはいえほとんどcomparableであった。この場合にはあきらかによい。ということで林氏の提案はなかなかいいということになる。

しかしながら逆の場合も調べてみる必要がある。つまりもともと 0.2%の出現率を0.18%にした場合にはうまくいったのだが、逆に0.22とか0.23%だったらどうだろうかというので調べたのが図10、11である。



逆の場合になると相変わらずCochran-Armitageのtrend testは1%ぴったりでおさまるが、他のものは第1種の過誤が少しふえる。少しふえるがそのふえ方が問題である。○の場合には大きくふえてしまう。それに対して●の場合には1%は最高2%弱が第1種の過誤となっておりそんなに大きくふえていない。そして検出力は○の場合には第1種の過誤が4%ぐらいまで非常に大きくふえてしまう。検出力が高くて第1種の過誤がこんなにふえてしまってはこまる。それに対して●では第1種の過誤はふえるがそのふえ方はそれ

ほどでもなく、検出力は依然として高い水準を保持している。だったらこの方法はいいぞ
 ということになる。それで投稿して採択されたわけであるが、結局これがたまたま林氏と
 私で議論したから検討したからこのような例になったが、この原理はかなり広い場合に使
 えるのではなかろうか。

表 10	結論：背景対照を 次のように使うのがよさそうだ 現対照が背景データから大きくはずれたら ー→実験のやりなおし 処置群と背景対照を 検定 T 1，有意水準 $\alpha 1$ で比較 処置群と現対照を 検定 T 2，有意水準 $\alpha 2$ で比較 (T 1， $\alpha 1$) と (T 2， $\alpha 2$) を 適当に決めるのがよいようだ どれだけ一般化できるか？ 目下検討中
------	---

いま二項分布ということにしたが、二項分布に限定することではなく、正規分布でもポア
 ソン分布でもガンマー分布でも何だっていいのだと思われる。要するに焦点は背景対照を
 次のように使うのがよいと思われる。まず現在の対照群が背景対照を大きくはずれたら実
 験のやり直しをする。処置群と背景対照をある種の検定のやり方である種の有意水準で比
 較する。処置群と現在の対照を別の検定で別の有意水準で比較する。それぞれの検定方法
 と有意水準を適当に決めるのがよい。そういう原理で他の場合についてやれば先ほどの場
 合と同様いいやり方が発見できるのではないだろうかとおもわれる。ではどういう場合に
 どのようなやり方をすればよいかというのは具体的にはやっていないが、たぶんこれでう
 まくいくのではないかと私は思っている。

もし皆さんで興味のある方がいらっしゃれば調べていただけると有難い。うちの研究所
 にはこういうデータがありこういう背景対照データがたまっている。だからいいではない
 か。ただし背景対照データの使い方のなかでいかなる統計的管理状態になっているかどう
 かということは essentialである。つまり分布はどんなに広がっていてもよいがこれが統

計的管理状態にないと困る。段々大きくなる傾向があるとか、逆に小さくなる傾向があるというのは背景対照データとして困る。多少ランダムであってもよいがばらつきの幅は安定していることが必要である。そのへんのチェックさえしっかりして背景対照を使えば、都合の悪い結論が出た時だけ使うというのよりはるかによい。

表 1 1

コンサルテーションの例

質問：急性毒性試験を行い
以下の結果を得ました

用量	雄	雌	(分子が死亡数)
2009	0 / 6	0 / 6	
2411	0 / 6	0 / 6	
2894	0 / 6	0 / 6	
3472	1 / 6	1 / 6	
4167	2 / 6	2 / 6	
5000	2 / 6	3 / 6	

このデータでLD50を計算しようと思います。そこでLitchfield-Wilcoxon 法だと低用量で死亡0の2群を除くことになりますがprobit法だと、それを除いても、除かなくても計算ができます。どのやり方でもLD50の推定値は同じくらいになりますが信頼区間に差がでてきます。

本来どちらで求めるべきでしょうか

次にconsultationについてであるが今の日本の場合にはこれが非常に必要だと思っている。ただ学会では余り評価されない。そんなものは金儲事業であるから商売人に任せておけ、アカデミックな人間がそういうことに手を出してはいけないという雰囲気である。銭金の問題ではなくそういうことに手を出すことを潔しとせずという学者の本質、習性がある。しかしそんなことをいってもはじまらないわけで私はconsultationが重要だと思うからいろいろな質問が来たら真面目に答える努力をしている。

これは1週間ほど前に来た質問で、やっと返事を出したものである。急性毒性試験を6用量で行った。雄6匹、雌6匹で実験をやり何匹か死んだ。低用量の3群に関しては6匹中1匹も死ななかった。上の方は6匹中1、2、3匹という具合に死亡した。つまり一番

高い用量で5割死んだということである。このデータでLD₅₀を計算しようという場合、種々の算出方法があるが、いいデータの場合どんな計算方法を使ってもほとんど同じ結論が出る。悪いデータの場合は計算方法によりまるっきり結論が変わってくる。悪いデータというのは用量に伴う変化が単調に変化せずジグザグのものである。このような場合にはどのような結果が出るかわからない。この質問をされた方はまず最初にLitchfield-Wilcoxon法でやったという。Litchfield-Wilcoxon 法というのはまずグラフを書いて直線を当てはめるというやり方である。そして50%死ぬのがどのへんだらうかというのをみる方法である。それに対してprobit法もやってみました。probit法というのはいわゆる正規分布曲線である。正規確率紙に直線が当てはまるかどうかというのを最尤法というのを使って計算するやり方である。ただしこれは普通逐次近似をやるので計算機を回すことになり、場合によっては時間のかかることになる。Litchfield-Wilcoxon 法では低用量で死亡のなかった2群は除くので実質、上の4群だけで直線を引くことになる。probit法の場合には下の死亡のない群を捨てるべきか残すべきかでこの人は悩んだ。probit法では下の2群を捨てても捨てなくても計算はできる。計算したところ、上のほうの4群使っても6群全部を使っても推定値そのものは同じになった。ところが信頼限界が違ってきた。どちらにすべきかという質問である。皆さんだったらどちらを使われるであろうか。

表 1 2

回答：6用量のデータを
全部用いるべきだと思います。

6用量を用いると、4用量を用いるより信頼区間が狭くなるでしょう。

なぜかというと低用量の3群の死亡数が0というのは低用量の1群だけで死亡数が0となるよりはるかに確実に死亡率が低いことを保証するデータだからです。

それが提供している情報を利用した方が結論が確実になるからです。

1989. 3. 2.
医薬安全性研究会

これは基本的にはconsultationの問題であるが、6用量全部用いるのがよいというのが私の回答である。6用量用いると4用量よりも信頼区間が狭くなるであろう。なぜならば低用量3群の死亡数ゼロというのは低用量の1群の死亡数ゼロよりもはるかに確実に低用量の死亡率が低いことを保証するデータだからである。つまりある用量で6匹やって死亡がゼロ匹ということは7匹やれば1匹ぐらい死ぬかも知れないということである。ところがそれより下の用量2群で死亡ゼロであったなら、これはかなりゼロに近いゼロであることを保証してくれているわけである。従って低用量での死亡数の曖昧さは減少し、信頼区間が狭まるという次第である。だからゼロというデータを入れることによりそれが提供している情報を入れるほうがいいのだということになる。これはある意味で当たり前のことであるが、このようなことで悩む人は以外に多い。統計学の原理的なものが分かっていると悩むことは少ないと思われるが専門が違ふとこういう現象が起こるのである。

表 1 3

ま と め

手法の研究・開発面

効率のよい計画

水準数をいくつにすべきか？

群の大きさは？

多重性の処理

多項目、多時点、多群、多施設

多実験（種、性、齢）

多重比較（多重仮説）

背景対照の利用法

統計的管理状態？

十分なデータ？

利用面

標準化

手法採用の恣意性の排除

共同作業

単なる計算業でなく

コンサルテーション

現場の問題の実データによる

質疑応答

これは「今何が問題になっているか」というシートの、「トピック」という所のラベルをはがしただけのものである。要するに今こういう毒性試験のデータ解析で重要になっている問題はこういう所にある。まず効率のよい計画を作ることにより殺す動物数を減らそう。多重性の処理というのは600項目とか700項目とか測った時には有意水準5%でやるとジャカジャカfalse positiveが出てしまうぞ、ということに対して何らかの手当をする必要がある。これについては結論が出ていないが何とかしないといけない。次に背景対照をどう使うかということの本気で考える必要がある。私が聞いている限り現在のやり方はちょっとでたらめすぎる。同じデータを処理するのにいろいろの手法がある。その手法を

恣意的に使うのはあぶない。手法の使い方は標準化すべきである。これは今のところ厚生省もあまり気にしていないようなのでどこかで頑張る必要があると思われる。先ほど述べたようにデータ解析も現場の皆さんとの共同作業で質のよいものができるのではないかと思われ、そういう意味で私も努力するし皆さんも共同作業をするという発想を持っていただきたい。consultationの問題はマンパワーの問題である。今こういう方面でconsultationができる統計学者にはマンパワーがない。なぜないのかというと、こういうデータ解析屋さんが単なる計算手段だと思われているからである。これは医学、薬学、生物学を通じてこういう傾向にある。データ解析なんて単なる電卓、コンピュータに過ぎない。だから必要な時だけちょっと電話して「ここどうすりゃいいの？」と聞いて教えてくれれば「ありがとう」で終らせればいい、という状況になっているがこれはよくない。やはりそれなりの役割を認識、評価してほしいというのが私の気持ちである。それなりに一つのパートを受け持つ存在なのだということを評価してほしいのである。

質疑応答

Q：統計についてはまったくの素人なので、基本的なことを確認させて頂きます。まず、standard deviation：標準偏差ですが、これは母集団のばらつきの程度を表わすもので、サンプリングの標本数とは本質的には関係のないひろがりを示すものですね。

吉村：はいそうです。

Q：それからstandard errorの方は母集団の平均値がどこにあるかを推定する範囲を示す方法でサンプリングの標本数が多くなればなるほど狭まってきて、標本数を無限にすればゼロになる性質のものでしたね。

吉村：はいそうです。

Q：実は先生のお話の中にもお答があったように思いますが、「背景対照の範囲を示せという」厚生省からの要求がありまして3シグマで表わすのか、standard errorで表わすのか、もめたことがありましたので確認させて頂いた次第です。

もうひとつ、ホルモンの産生を止めるような働きをする薬の生殖試験での経験ですが、このような作用を持つものでは当然流産が起こりますし、それに至る過程で胎児の致死も起こります。それでdoseを細かくとって検討しますと、致死ぎりぎりところで特定の奇形とは限らないのですが、いろいろな奇形がランダムに発生してきました。たとえばコルチコイドを投与致しますと口蓋裂が出ますが、そのような特定の奇形に限って発生するわけではないのです。しかしトータルを集計しますとdose dependentにふえる傾向がでました。その場合にこの薬は胎児に催奇形作用があると判断すべきかどうかということなのですが・・・。

吉村：種類を問わずトータルとして増える傾向がある、実際にひとつひとつの奇形について調べると数が少なくなるのでその傾向がわかりにくくなるということですね。その場合には基本的に催奇形性があると判断すべきだと思います。というのはもちろん奇形発生のメカニズムにもよりますが特殊な奇形だけに明らかに作用するのならともかく、そうでないときにdose dependentにふえることは、異常が生じるメカニズムになっているのではないかと思います。

一般的にいうと、いくつかの奇形が併存できるものであるなら、別々のものとして出てくる可能性があります、実際には発生分化のある段階で出てくるとほかのものが消える可能性があるのではないかと思います。トータルとして出るということは個々のものとしてはそうでなくても何か共通のメカニズムがある場合ではないかという気がします。

Q：個々の奇形は完全に背景対照データの範囲内にあるのですが・・・。

吉村：私だったらたぶんそのときには異常の発生がふえているという結論を出すだろうと思います。もちろんその時には発生学的な意味でどの異常がどの段階で出るかは専門家に聞くだらうと思いますが・・・。ただ、その結果がはっきりしなかったら、影響が出ているという判断を下すだろうと思います。

Q：次に低次元の話になりますが、毒性試験では測定項目がたくさんあります。これも先生がおっしゃった有意差検定の方法にもよると思いますが、たとえば5%の危険率で検定する場合、測定項目が20項目ありますと1項目は有意差ありという結論が出てくるのは当たり前なのでしょうか。？

吉村：平均的にはそうなるでしょう。その時に対応する処置として、皆さんはいろいろな理屈をつけて「これはたまたま出てきているがどうみても毒性あるとは思えない」というようなことをほかのデータを利用して作文しているのが現実ではないでしょうか。

それではやや恣意的になる危険があるので何とか標準化すべきではないかというのが私の意見なのです。具体的にどうすべきかということあまりよいアイデアはないのですが、ひとは項目ごとに重要な項目と重要でない項目クラス分けして重要な項目については有意水準を高く、そうでない項目については有意水準を1%とか0.5%に低くする。もうひとつ、同じタイプのものに系列化して分類することが可能ではないかと思うのです。たとえば臓器重量で特定の臓器だけに明らかに影響するのであれば注目することになると思いますが、そうでなければ臓器は成長の反映と考えられますので、たぶん同じ傾向になると思います。場合によってはこのようにまとめて判断することもひとつの手ではないかと思います。

Q：さきほどの先生のお話のように、その場合に背景対照の値と比べて問題な検定の仕方を応用すれば・・・。

吉村：その時にははっきり事前に一般的ルールを決めておいてそのルールのもとにこのような不合理なことは起こらないということを確認してからだと思います。ただ先程いいましたように都合の悪いときだけ使うということだけはすべきではないと思います。

Q：どうもありがとうございました。

Q：さきほどお話のありましたCochran-Armitageの検定ですが、これに関し2点ほど質問させて頂きたいと思います。

第1点は動物試験で傾向性検定をやるときに推奨されている方法と聞いていますが実際には癌原性とか毒性試験では出現率が非常に低くて結果として出現度数も非常に低い。たとえば5群の実験で4群までがすべてゼロで最高用量の1匹だけ死亡の発現があったというようなケースはしばしばあると思います。そのような場合、正規近似で検定するわけですが、有意水準が非常に変わって問題ではないかということです。

もう1点は傾向性を捉える場合、極端なケースではJ型というか、最高用量だけ多い場合にはCochran-Armitageで検定すると有意に出る場合があるのです。そういった意味で非線形性をまずチェックしてからCochran-Armitageを使用すべきではないでしょうか。つまりCochran-Armitageを使用する前にその妥当性を検討することが必要なのではないのでしょうか。

吉村：まず第一番の問題ですがCochran-Armitage検定は意外に近似がよいのです。だからかなり小さい場合にも有意と出ている場合にはそうでないもっと正確な検定、たとえばFisher流の条件付き検定によっても結論が同じになる場合が多いですね。もちろん微妙なところでは違いますが正規近似について私はあまり心配していません。

もうひとつの問題についてはたまたまデータがJ型になってしまったから線形性を検定するという議論ですが、そもそもその試験で用量をいくつか定めたのはなぜか、ということがあります。つまりあきらかに用量反応性があってどこかに閾値があってそこから上がるということがあるから、逆に用量をいくつか定めて、あるところからカットしてここから先は差なし、ここから先は差ありという議論をするならいわゆる多重比較型の検定をやるべきだと思うのです。それに対し用量をいくつか定めるのはrobustnessをみるのであって、結論としてこの物質はある性質をもっているか持っていないかだけを出したいというのであればCochran-Armitage検定でよいと考えられるのです。

多分、林さんがCochran-Armitage検定をやっているのは「変異原性ありやなしや」でだけで結論を出すのだ、もともとスクリーニングだから」ということで使うので、普通の薬物の毒性試験の場合だと、どの用量で差があり、どの用量で差がないということを議論したいわけですね。その場合にはCochran-Armitage検定は必ずしもよいものとはいえないと思います。そう丁寧に吟味したわけではありませんけれども、やはりnon-parametricな多重比較検定：Williams流、Shirley流のやりかたの方がよいのではないかなと思います。

Q：タイトルから2枚目くらいの表に「記述、探索、確認」というのがありましたが、
が、その中で「探索」というのがよくわかりませんでした。これは推計とか推定と考えてよいのでしょうか。

吉村：「探索」といったときにはいろいろな可能性をいろいろな側面から吟味する。
そうしますと検定という手法を使うわけですが、検定というのは仮説をデータを取る前、みる前に設定しないといわゆる「後知恵」の問題で第1の過誤が狂うんですね。つまりある現象があるということがわかってから、それを仮説にたてて検定するというと、それはいわゆる「後知恵検定」といって有意水準5%といっても、実は第一の過誤が何十%となる、そういった現象があるのです。そういうものを敢えてやってもとにかく目立つ現象だけは全部ひろいあげたいという発想でやることを「探索」というのです。たとえばいろいろな変数変換をやってみる。先程のJ型であるならスケールを横軸で変換してやれば直線的にすることだってできるのです。そういう形でデータを細工してある種の法則性が成り立つかどうかを吟味する。そういうやり方を普通「探索的」というのです。そうすると普通のデータをみる前にやり方を決めたのとは異なりどうしても「後知恵」になりますから結論がそれほど確実ではない。結果として出てきたものをまるまる受け入れる形となる。だからもしかしたらそれは誤差かもしれない、という感じのものができ上がる。したがって探索的な場合にはその次のステップとして確認的な実験データを取って確かめることが必要であるということになります。「推計」とは少し違うのです。

<第22回研究会>

学術講演会

1. 霊長類におけるATL—及びAIDS—関連ウィルス

速水正憲 京都大学ウィルス研究所

2. 日本ザルの集団構造の遺伝学的解析

野沢 謙 京都大学霊長類研究所

日 時 : 平成元年6月23日(金)

場 所 : 京大会館

霊長類におけるATL—及びAIDS—関連ウィルス

京都大学ウィルス研究所

速 水 正 憲

はじめに

1977年に成人T細胞白血病(ATL)が白血病の1つの病群として確立され、1981年にその原因ウイルスとしてヒトT細胞白血病ウイルスI型(HTLV-I)が発見された。またエイズは1980年に新たな疾患として認識され、1983年にその原因ウイルスであるヒト免疫不全ウイルス(HIV)が分離された。いずれの疾患もなぜ最近になって認識され、そのウイルスが分離されるようになったのかという点に関しては議論があり、元来ヒトに存在していたとする説と、他の動物、特にサルから伝播したのではないかとする説がある。

またこれらのサルウイルスはATLやエイズの感染様式や発症機序の解明、さらに治療薬やワクチン開発のための動物モデルとしても期待されている。

ここではこれらのウイルスの起源・進化に関連させながら、動物モデルとしての観点から、われわれのやってきた仕事を中心に紹介したい。

1. ATL関連ウイルス

1) 種々のサルにおける抗体保有状況

種々のサルのHTLV-Iに対する抗体保有状況を調査したところ、アジア・アフリカに生息する原猿類や中南米に生息するリスザルやマーモセットなどの広鼻類、いわゆる新

世界ザルには陽性例はみられなかった。アジア・アフリカに生息する狭鼻類いわゆる旧世界ザルでは、ニホンザル、アカゲザル、カニクイザルなどのアジア産マカカ属を始め、アフリカ産のセルコピテクス属のアフリカミドリザルや、ヒヒ属のマントヒヒに陽性例がみられている。また類人猿では、アフリカのチンパンジー、アジアのテナガザルなどに陽性例がみついている。このように霊長類においては、旧世界ザルと新世界ザルを境界にして、これより下等なサルはこのウイルスを持たず、高等なものが保有しているということになる。これはこのウイルスの起源と霊長類の進化を関連づけて考えるうえで興味がある。ヒトにおけるHTLV-Iの感染例は、日本西南部とカリブ海諸島、中央アフリカの3地域に限られ、世界におけるその局在は謎となっている。この究明のためには、ヒトの移動に伴うウイルスの移動だけではなく、サルを含めた霊長類の進化とそれに伴う大陸間の移動を考え合わせてみる必要があるのかも知れない。

2) ニホンザルにおける疫学調査

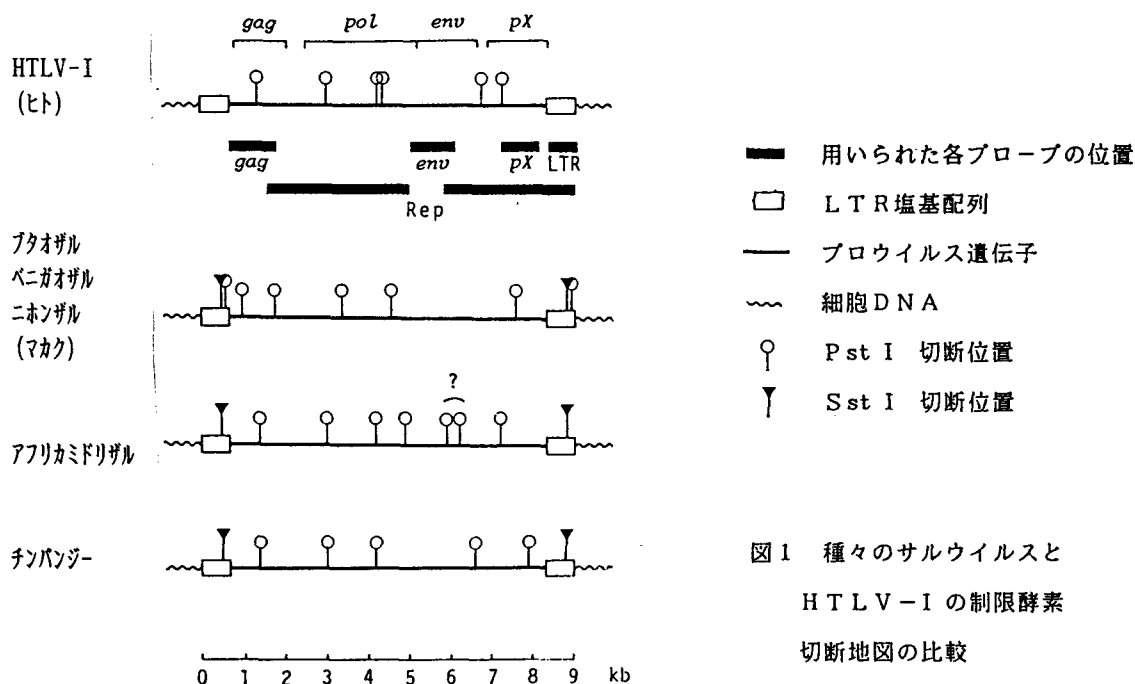
日本の各地に生息しているサルの抗体保有状況について疫学調査を行ったところ、ヒトの調査とは異なった結果を示した。すなわち、ニホンザルの抗体保有地域は、ヒトにおけるATL多発地域である九州地方のみならず、全国的に分布しており、しかも抗体保有率（全国平均28%）はヒトよりかなり高いことが明らかになった。成年ザルの抗体保有率が50%を超えている群は、全国の約半数に達し、なかには群れの80~90%が感染している例もみられた。

このようにヒトにおけるATL多発地域に限らず、非多発地域においても抗体保有ザルが高頻度に検出されたことから、現在ヒトとサル間でウイルスの直接的かつ容易な感染が起こっているとは考え難い。このことはサル飼育施設で感染サルと日常接している人々に感染例がないことから推察される。

3) サルT細胞白血病ウイルスI型の分離とその解析

まずウイルス分離のために、各種のウイルス保有ザルの末梢血リンパ球より、サル由来ウイルス産生細胞株を樹立し、現在までに11種のサル類から計18種類のウイルス産生細胞株を得ている。それらの細胞にはHTLV-Iと形態上区別できないC型ウイルスが観察され、これらの細胞培養上清中には、レトロウイルスの特徴である逆転写酵素活性が認められた。その細胞表面はヒトのATL細胞と同様にインターロイキン2受容体であるTac抗原が陽性である。ほとんどの細胞はT細胞マーカーを保有していた。これらのサルウイルスの主要蛋白質は、いずれもヒトと各種サル間で共通抗原性を示した。

次いで各細胞株中のプロウイルスについて、解析を行った。その結果、HTLV-Iと同様の遺伝子配列を持つ、プロウイルスが組み込まれていることが明らかになった。しかしこれらプロウイルスの制限酵素による切断地図は、ヒトとサルおよびサル種間で異なっていた。すなわち図1にみられるように同じマカカ属に属するブタオザル、ベニガオザル、ニホンザルは同一の切断地図を示したが、セルコピテクス属のアフリカミドリザルと類人猿のチンパンジーではおのおの異なるパターンを示した。



すなわち、これらサルウイルスはHTLV-Iと共通な祖先を持つHTLV-I群に属するが、明らかにHTLV-Iとは異なるウイルスであることが示された。このことから、ヒトおよび異なるサル種間における相互感染が起きたと仮定しても、それは最近ではなくかなり以前のできごとと考えられる。おそらくこれらのウイルスの“原ウイルス”が霊長類の進化のある過程で侵入し、その種属の進化に伴って、多少の変化はあったものの比較的安定な状況で、ヒトおよび各サル種属内に保持されてきたと推測される。

このように、HTLV-I類似サルウイルスはHTLV-Iと明確に区別されるウイルスであり、また次に述べるように、ATL類似のT細胞白血病を起こし得ることが判明したので、サルT細胞白血病ウイルスI型 simian T cell leukemia virus type I (STLV-I) と名づけた。

4) S T L V - I の白血病原性

このサルウイルスが原因ではないかと疑われる A T L 類似の T 細胞白血病がアフリカミドリザルにみられた。全身のリンパ節と脾臓の腫大がみられ、末梢血中には多数の異型リンパ球が観察されたが、それらは A T L 患者の分葉核リンパ球ときわめて類似していた。病理組織学的にはび慢性多形細胞型のリンパ系腫瘍であった。そのプロウイルスの解析を行ったところ、この腫瘍細胞はアフリカミドリザルに特有のウイルスに感染した細胞が、単クローン性に増殖した細胞集団であることが明らかになった。このことは、サルウイルスが原因となって A T L 類似 T 細胞白血病を起こしたことを示している。今後 H T L V - I による発症機序の解明において、このサルウイルスとサルの系が有用なモデルになるものと考えられる。

5) S T L V - I のサルにおける感染様式

ヒトにおいては抗体保有者の家族集積がみられることから、垂直感染および濃厚な接触による感染が疑われており、疫学的には夫婦間、特に夫から妻への水平感染が想定されている。

これまでにサルにおけるこのウイルスの感染経路の 1 つとして、垂直感染（おそらく経胎盤感染）の証拠を得ているが、今後乳汁感染や水平感染を含めた感染経路を明らかにしていく必要がある。

2. エイズ関連サルウイルス

1) 種々のサル類における H I V / S I V の抗体調査

類人猿から旧世界ザル、新世界ザル、原猿類に至るサル類 49 種由来、計 1875 例の血清について抗 H I V - I 抗体の検索を行ったところ、アフリカミドリザルのみが陽性反応を示した。そこで陽性のサルから分離したウイルス、S I V_{AGM} に対する抗体を再び各サル類血清について調べたところ、アフリカミドリザル、サイクスモンキー、ブラッザモンキー、マンドリルに陽性検体が認められた。これらはすべてアフリカ原産の旧世界ザルであり、アジア産のサルはすべて陰性であった。また、抗体陽性サルはすべて健康であり、これまでのところエイズ様症状は全く認められていない。

わが国の 4 箇所の実験用サル飼育施設のアフリカミドリザルはケニアもしくはエチオピアから輸入されたものであり、どのコロニーにおいても 20～50% が抗体陽性である。陽性の仔ザルはすべて陽性の母ザルから生まれており、陰性の母ザルからは陽性の仔ザルは 1 匹も生まれていないことから、母から仔への垂直感染が推測される。今までのところ

水平感染を疑わしめる例はみられていない。

2) アフリカミドリザルからの S I V の分離とその性状

抗 H I V - I 抗体陽性のアフリカミドリザルのリンパ球をヒトの T 細胞株と混合培養したところ、細胞変性効果 (C P E) が認められ、その培養上清中には逆転写酵素活性が検出された。電子顕微鏡では H I V と同様の形態であった。

このアフリカミドリザルからの分離株である S I V_{AGM} の gag 蛋白質はヒトのエイズ患者血清によって認識されたが、env 蛋白質は、H I V - I の env 蛋白質とはわずかな交差反応を示したのみであった。

S I V_{AGM} を種々のヒト C D 4 陽性の T 細胞株に接種したところ、感染・増殖し、融合細胞形成を中心とする C P E を引き起こした。しかしながら、このウイルスは Raji 細胞などの C D 4 抗原陰性の細胞には感染しなかった。

3) S I V_{AGM} の遺伝子配列の決定

S I V_{AGM} の遺伝子クローニングとその全塩基配列の決定を行ったところ、S I V_{AGM} は典型的な H I V / S I V レンチウイルスの構造をしていた。しかし今までに明らかになっているアカゲザル由来の S I V_{MAC}、H I V - 2 および H I V - 1 と相違がみられている。S I V_{AGM} の各遺伝子領域のアミノ酸配列の相同性を S I V_{MAC}、H I V - 2、H I V - I と比較したところ、おのおのに対して gag、pol 領域では約 50%、他の領域では 30~40% を示した。このことは S I V_{AGM} が明らかに H I V - 1、H I V - 2、S I V_{MAC} のウイルスとは異なるウイルスであることを示している。

4) マンドリルからの S I V の分離とその解析

西アフリカのガボンにある霊長類センターで飼育されているマンドリルの末梢血リンパ球をヒト T 細胞株と混合培養することによって、S I V_{MND} を分離することができた。このウイルスは S I V_{AGM} と同様の構成蛋白質からなっており、そのすべての構成蛋白質は S I V_{AGM} と交差反応性を示したが、H I V - 1 とは gag および pol 蛋白質が交差したのみであった。次いで S I V_{MND} の遺伝子クローニングとその全塩基配列の決定を行った。その結果、S I V_{MND} の遺伝子構造は S I V_{AGM} やその他の H I V / S I V ウイルスと基本的には同じであったが、若干の相違がみられている。S I V_{MND} の各遺伝子領域におけるアミノ酸配列の相同性を S I V_{AGM}、S I V_{MAC}、H I V - 2、H I V - 1 と比較したところ、S I V_{MND} は S I V_{MAC}、H I V - 2、H I V - 1 そして予期しなかった

したところ、 SIV_{MND} は SIV_{MAC} 、 $HIV-2$ 、 $HIV-1$ そして予期しなかった SIV_{AGM} の各ウイルスに対して、ほぼ等距離にあることが明らかになった。

5) HIV/SIV ウイルスの位置関係とその起源

今まで述べてきた SIV_{AGM} と SIV_{MND} との遺伝子解析をもとに、すでに知られている $HIV-1$ 、 $HIV-2$ 、 SIV_{MAC} との関係をシェーマ化したのが図2である。上段で示すように、各ウイルスはお互いにほぼ等距離にあり、正四面体を形成している。

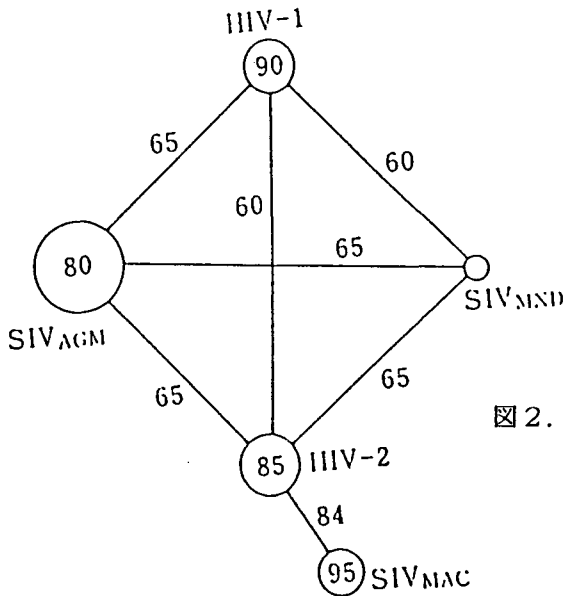


図2. HIV および SIV の相互関係。数値は pol 領域の核酸配列から計算した各グループ内あるいはグループ間の同一性(%)を表す。

例外は、 $HIV-2$ と SIV_{MAC} で、この2つは従来いわれていたように同一群と考えられる。米国の霊長類センターのアカゲザルから分離された SIV_{MAC} の由来については、現在、疑問視されている。すなわち、このウイルスが分離されたアカゲザルはニューイングランド霊長類センターで長い間、飼育されていたもので、野外例においてはわれわれおよび欧米の研究室の調査結果からも、アカゲザルのみならずアジア産のマカカ属のサルはすべて陰性であったからである。米国の霊長類センターの種々のサルコロニーは、かつてヒトの材料を接種されたこともあるので、ヒトの $HIV-2$ がその由来ではないかと考えられている一方、米国の別の霊長類センターのスーティマンガベイから分離された SIV_{SMM} の遺伝子が、 SIV_{MAC} とほぼ同じであったことから、 SIV_{MAC} はアフリカ産のスーティマンガベイ由来ではないかとも考えられている。

このように SIV_{MAC} 、 $HIV-2$ および SIV_{SMM} は、ヒトとサル間でウイルスが移行した可能性も存在するが、アフリカミドリザルとマンドリルは、ヒトにおけるエイズの

6) S I Vのサルにおける病原性

アカゲザルから分離されたS I V_{MAC}をアカゲザルに接種したところ、日和見感染やマクロファージ浸潤性脳炎、全身の消耗と免疫不全を呈し死亡した。

このようにS I V_{MAC}とS I V_{SMM}はアカゲザルにエイズ類似症状を起こすことから、病原性ウイルスと考えられるが、S I V_{AGM}とS I V_{MND}については現在のところ非病原性と考えられる。特にS I V_{AGM}については、実験感染例のみならず多くの自然感染アフリカミドリザルにおいてもエイズ症状が認められたことはなく、また免疫抑制も認められていない。in vitroにおいてもアフリカミドリザルの末梢血リンパ球によく増殖するものの、それを破壊することがないために、一般には非病原性であると考えられている。

もしこの病原性の相違がその遺伝子の相違によるならば、これら病原性と非病原性のウイルスを比較することは重要と思われる。現在、S I V_{AGM}の感染性クローンを用いて、病原性のH I V-1、H I V-2、S I V_{MAC}の種々の遺伝子を組み込んだ組換えウイルスを作製中であり、これらをサルに接種することにより病原性と関連する遺伝子を同定していくことが可能となろう。

おわりに

H T L V-I / S T L V-Iの起源と進化に関するわれわれの考え方は、現在、一般的に広く受け入れられている。しかし、H I V / S I Vについては依然としてサル由来説がくすぶっており、例えばS I V_{AGM}とS I V_{MND}についてはH I Vの起源とはなり得ないとしても、まだ発見されていないアフリカ産のサルが保有しているS I Vの可能性もあるのではないかとの議論もある。現在の状況ではそういった可能性も含め、いろいろな可能性が考えられ、われわれも決してそれらを否定するわけではない。これらのウイルスの起源を問うことは直接的にA T L・エイズの問題解決には役立たないが、これらのサルのウイルスをヒトウイルスと比較していくことは、ヒトにおける感染様式・病態解明のうえに、別の側面から光をあてることになる。特にA T L・エイズにおける動物モデルの欠如から、これらのサルウイルスは今後その感染様式・病態解明のみならず治療薬・ワクチン開発のうえで重要と考える。

ニホンザルの集団構造の遺伝学的解析

京都大学霊長類研究所

野 澤 謙

ニホンザル (*Macaca Fuscata*) は青森県下北半島を北限、屋久島を南限として、北海道を除く日本本土全域に自然生息地を有し、総個体数数万と推定されている。群れ (troop) を社会的単位として生存し、その社会構造や生態について豊かな知見の蓄積があるために、その集団構造を明らかにするために必要とする条件は整っている。筆者らは1971年以来、全国各地の群れから採血をおこなって、血液蛋白の構造を支配する遺伝子座30座位以上について変異を電気泳動法によって検索する作業を続行中である。ここにこれまでの結果を集計し、統計的分析を加えることによって、この種の繁殖構造が集団遺伝学の立場から見てどのような性格のものであるか考察したい。

調査対象となった群れの数33群、遺伝的変異が検索された血液蛋白は29種、遺伝子座の数は32座位である。電気泳動をおこなって、蛋白の泳動パターンから遺伝子型を判定し、それから遺伝子頻度を計算する。発見された変異のすべてが共優性形式の遺伝子支配を受けているので、単純カウント法で遺伝子頻度の計算をおこなうことができる。

群れ集団内の遺伝的変異性を次の2つの尺度で定量した。第1は検索された32遺伝子座のうち多型 (polymorphism) を示す座位の割合 (P poly) である。ここで多型とは、その座位における最高頻度をもつ対立遺伝子の頻度が0.99に満たぬ場合と定義する。第2は平均ヘテロ接合率 (average heterozygosity: $\bar{H}$) である。ある遺伝子座における第*i*対立遺伝子の頻度を q_i とすると、その座位におけるヘテロ接合率 (H) は $H = 1 - \sum_i q_i^2$ で計算され、この値を変異が見られない座位 (その場合、 $H = 0$) をも含め、全32座位にわたって算術平均を計算し $\bar{H}$ とする。

群れ集団間の遺伝的変異性は幾何学的遺伝距離 (Euclidean genetic distance: D) で測られる。第*j*および第*k*集団における、或る遺伝子座の第*i*対立遺伝子頻度をそれぞれ q_{ij} および q_{ik} とすると、両集団間の遺伝距離 (D) は $D = \sqrt{\sum_i (q_{ij} - q_{ik})^2}$ で計算される。ここで、右辺の横棒は変異が見られぬ座位をも含んだ全32座位にわたる算術平均を示すものとする。

(1) 個々の座位における遺伝的変異性

変異の検索がおこなわれた32遺伝子座のうち、1個またはそれ以上の群れで変異が見されたのは16座位(50%)であった。すなわちPi (Plasma protease inhibitor)、Alb (Plasma albumin)、Tf (Plasma transferrin)、Hb β (Hemoglobin β -chain)、PHI (Cell phosphohexose isomerase)、PGM-I (Cell phosphoglucomutase I)、PGM-II (同II)、Dia (Cell NADH-diaphorase)、CA-I (Cell carbonic anhydrase I)、Acp (Cell acid phosphatase)、MDH (Cell malate dehydrogenase)、LDH-A (Cell lactate dehydrogenase A)、LDH-B (同B)、Cell Es (Cell esterase)、AK (Cell adenylate kinase) およびEs-D (Cell esterase D)。

33群のすべてに変異が見られるような遺伝子座、すなわちユニバーサルな多型座位は存在しない。変異が見られた群れ数が最も多い座位はPHI座位とCA-I座位で、それぞれ14群に変異が見られ、Tf座位の13群がこれに次ぐ。Alb、PGM-IIの両座位には、ごく少数の変異個体がただ1群に見られるに過ぎない。

(2) 集団内遺伝的変異性

群れ集団内の遺伝的変異性を表1に示す。多型遺伝子座の割合(Ppoly)は3.1~21.8%の間に分布して平均9.1%、平均ヘテロ接合率($\bar{H}$)は0.1~3.3%の間に分布して平均1.3%となっている。NEVO et al. (1984)によると、全世界の哺乳類184種の平均Ppolyは19.1%、平均 $\bar{H}$ は4.1%となっているが、ニホンザルの集団内遺伝的変異性はこれよりも低いレベルにある。

(3) ニホンザルには前述のごとくユニバーサルな多型座位は見出されていないから、検索されたすべての遺伝子座において、遺伝学的意味での野生型遺伝子(wild-type allele)は容易に指示することができる。野生型遺伝子に対立する変異型遺伝子(mutant allele)の分布を、Tf、PHIおよびCA-Iの3座位について地図上に扇形グラフで示したものが図1である。Tf^G遺伝子が本州中部に、PHI²遺伝子が九州北部に、CA-I^{D2}遺伝子が小豆島、鹿島など四国周辺の島嶼に濃密に分布している。そのほか、Pi^D遺伝子は九州に、Hb- β ²遺伝子とCell Es²遺伝子とは伊豆地区に、LDH-A²遺伝子は九州北部に、AK²遺伝子は淡路島に集中して出現している。総じてニホンザルの遺伝的変異はこの種の全域に均等な分布を示すものではなく、地域的に集中して現われる傾向を有していると言うことができる。

表1. ニホンザルの群れ集団内遺伝的変異性

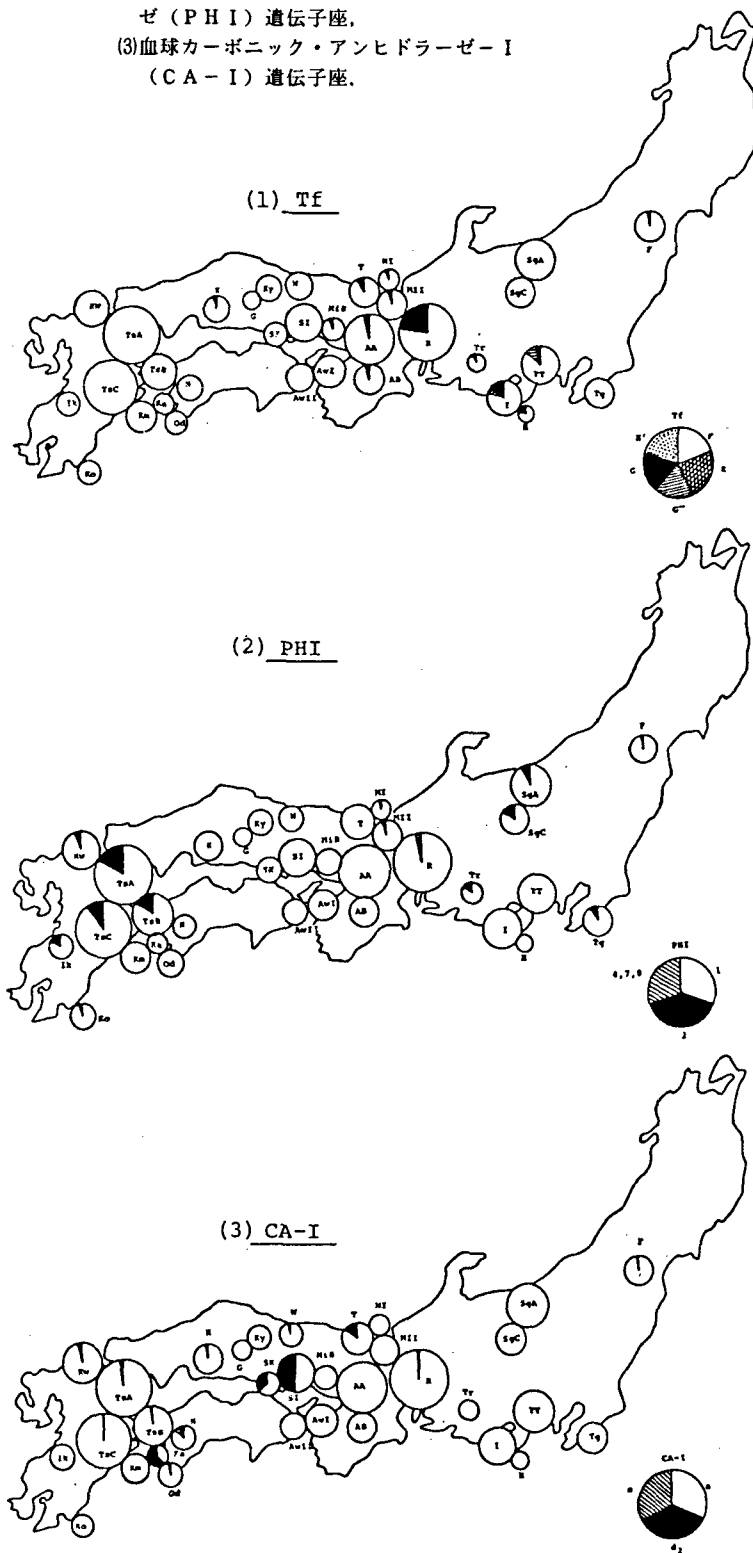
Troop	No. of poly-morphic loci	P _{poly}	$\bar{H}$
F	7	0.2187 $\pm$ 0.0730	0.0117 $\pm$ 0.0043
Tg	1	0.0312 $\pm$ 0.0307	0.0051 $\pm$ 0.0051
YT	6	0.1875 $\pm$ 0.0689	0.0277 $\pm$ 0.0114
I	5	0.1562 $\pm$ 0.0641	0.0333 $\pm$ 0.0168
H	3	0.0937 $\pm$ 0.0515	0.0279 $\pm$ 0.0173
Tr	4	0.1250 $\pm$ 0.0584	0.0253 $\pm$ 0.0125
SgA	3	0.0937 $\pm$ 0.0515	0.0121 $\pm$ 0.0068
SgC	3	0.0937 $\pm$ 0.0515	0.0172 $\pm$ 0.0111
R	6	0.1875 $\pm$ 0.0689	0.0204 $\pm$ 0.0118
MI	2	0.0625 $\pm$ 0.0427	0.0071 $\pm$ 0.0051
MII	4	0.1250 $\pm$ 0.0584	0.0112 $\pm$ 0.0058
T	4	0.1250 $\pm$ 0.0584	0.0236 $\pm$ 0.0117
AA	2	0.0625 $\pm$ 0.0427	0.0041 $\pm$ 0.0031
AB	2	0.0625 $\pm$ 0.0427	0.0037 $\pm$ 0.0029
MiB	3	0.0937 $\pm$ 0.0515	0.0141 $\pm$ 0.0097
W	4	0.1250 $\pm$ 0.0584	0.0082 $\pm$ 0.0040
G	1	0.0312 $\pm$ 0.0307	0.0012 $\pm$ 0.0012
Ky	3	0.0937 $\pm$ 0.0515	0.0053 $\pm$ 0.0034
K	4	0.1250 $\pm$ 0.0584	0.0141 $\pm$ 0.0081
AwI	1	0.0312 $\pm$ 0.0307	0.0085 $\pm$ 0.0085
AwII	1	0.0312 $\pm$ 0.0807	0.0052 $\pm$ 0.0052
SI	2	0.0625 $\pm$ 0.0427	0.0173 $\pm$ 0.0156
SK	1	0.0312 $\pm$ 0.0307	0.0146 $\pm$ 0.0146
Ka	1	0.0312 $\pm$ 0.0307	0.0151 $\pm$ 0.0151
N	2	0.0625 $\pm$ 0.0427	0.0106 $\pm$ 0.0079
Od	1	0.0312 $\pm$ 0.0307	0.0016 $\pm$ 0.0016
Kw	5	0.1562 $\pm$ 0.0641	0.0199 $\pm$ 0.0087
TsA	5	0.1562 $\pm$ 0.0641	0.0144 $\pm$ 0.0095
TsB	2	0.0625 $\pm$ 0.0427	0.0145 $\pm$ 0.0098
TsC	3	0.0937 $\pm$ 0.0515	0.0118 $\pm$ 0.0068
Km	2	0.0625 $\pm$ 0.0427	0.0040 $\pm$ 0.0027
Ko	3	0.0937 $\pm$ 0.0515	0.0095 $\pm$ 0.0050
Ik	1	0.0312 $\pm$ 0.0307	0.0086 $\pm$ 0.0086
Average	2.94	0.0918 $\pm$ 0.0094	0.0130 $\pm$ 0.0014

図1. ニホンザルの3つの多型遺伝子座における
対立遺伝子頻度の国内分布。

(1)血漿トランスフェリン (Tf) 遺伝子座,

(2)血球フォスフォヘキソース・イソメラー
ゼ (PHI) 遺伝子座,

(3)血球カーボニック・アンヒドラーゼ - I
(CA-I) 遺伝子座。



しかしながら注意すべきは、変異型が単一群に特異的に現われるのではなく、地域的に近接して位置する複数群に共通の遺伝的変異が見出されるということである。このことは、ニホンザルの群れという社会的単位が、遺伝学的な閉鎖系ではなく、近接群間になにほどこかの個体交換がおこなわれていることを示唆している。

(4) 個体交換率（移出入率）

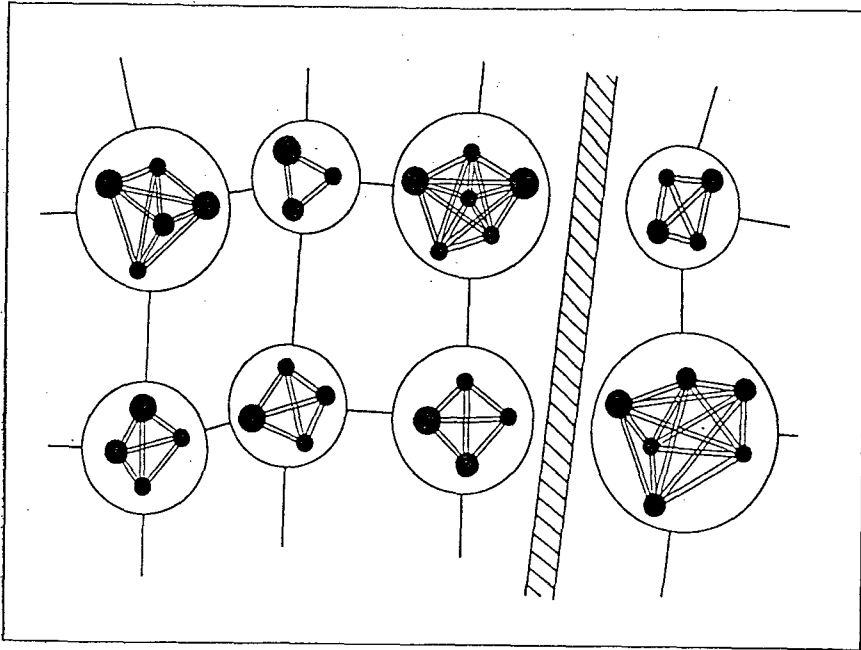
ニホンザルの集団構造はKIMURA and WEISS (1964) が提出した2次元飛石模型 (two-dimensional stepping-stone model) によって近似することができよう。このモデルは全集団が格子状に配列された同一有効サイズ (N) の副次集団 (subpopulations) の多数から構成され、各副次集団は隣接する4個の副次集団と毎代等しい率 (m) で個体交換をおこなうと仮定する。KIMURA and MARUYAMA (1971) のシミュレーション実験の結果によれば、淘汰的に中立な対立遺伝子がこのようなモデル集団中に多型的に維持される際、 $mN < 1$ であれば遺伝子頻度に顕著な地域的分化があらわれ、逆に $mN \geq 4$ であれば遺伝子頻度の分布はほぼ均等となり、全集団はあたかも単一の任意交配集団であるかのごとき様相を呈する。ニホンザルの血液蛋白変異は淘汰に対して実質的には中立で、変異の国内分布には著しい分化傾向があらわれている。すなわち、上記の $mN < 1$ の場合に相当する。よって $m < \frac{1}{N}$ 。即ち、個体交換率は副次集団の有効サイズの逆数以下と推定される。

もし、上記モデルにおける副次集団をニホンザルの群れと考えることができるとすれば、群れの遺伝学的有効サイズは全個体数の約 $1/3$ であり (NOZAWA, 1972)、群れの平均個体数は約66 (竹下、1964) であるから、 $m < \frac{1}{22}$ となり、隣接群間の個体交換率は世代当り5%以下と推定される。しかし、ニホンザルの生態観察をおこなうと、隣接群間の個体交換率はこれよりはるかに高い値を与える。

そこで生態観察によって個体交換をおこなっていることが明らかな隣接群間の個体交換率 (m) をWRIGHT (1965) の島模型 (island model) を使って推定してみた。その結果は志賀A、C間で世代当り36%、三方I、II間で53%、小豆K、I間で15%、高崎山A、B、C間で36%となり、決して5%などという低い値とはならない。

ニホンザル種内の遺伝的変異の分布が著しい地域的分化を示しているという事実と、隣接群間の世代当り個体交換率が著しく高い値となるという事実とを論理的に両立させようとするれば、群れと種との間にもう一つの構造が存在することを想定するほかはない。こうした考慮のもとにニホンザルの集団構造を図式化したものが図2である。大小の黒丸は個々の群れを表わす。

図2. ニホンザルの集団構造模式図



隣接して位置する数個の群れは相互に個体交換（主に雄の交換）によって密接に結合し、円で囲まれた地域集団とでも呼ぶべき一つの繁殖単位（deme）を構成する。それら地域集団相互は、距離が近い場合に限って稀に個体の出入りが可能であるという程度のゆるい結合を保っている。もちろん、中間に海とか高山とか都市とかがあれば、たがいの連絡は不可能となる。なお地域集団を構成する群れの数が平均5個であれば、群れの平均有効サイズを22として、近接地域集団間の世代当り個体交換率は $1 / (22 \times 5) < \frac{1}{100}$ 、平均10個であれば $1 / 200$ 以下、平均20個であれば $1 / 400$ 以下となろう。

(5) 変異遺伝子の有効分散距離

変異遺伝子の有効分散距離とは、ある群れにある変異遺伝子が存在するとき、群間の個体交換によって、その遺伝子が高群へ分散していく実際上の距離範囲である。33群のあらゆる対（その数は $33 \times 32 / 2 = 528$ 対）に対して群間の直線距離（km）と遺伝距離を計算し両距離の間の相関係数を求める（表2）。日本全体として見たとき、両距離間に有意な相関が認められる。しかし、それを2つにわけ、異なった島に生息する群間での相関は事実上0、同一の島、つまりたがいに陸続きの群間で相関は有意となる。さらにたがいに陸続きであっても100km以上離れて位置する2群間では無相関、100km以内であれば相関は有意である。これを50kmで調べてみると、50km以上でも相関係数は統計

的に有意となり、50 km以下では相関係数はかなり高い正の値を示すものの、対の数が少ないため統計的有意水準には達しない。こうした地理的距離と遺伝距離との相関関係の分析から、100 kmが遺伝的変異の有効分散距離のおよその推定値と考えられる。日本本州は北端から南西端まで1000 kmを越える長さを有しており、この島の中に多くのニホンザルの群れが生息しているけれども、100 km以上距って位置する2つの群れの遺伝子構成上の差異は、その間に海という越え難い障壁がある場合と事実上同程度のものであるとすることができる。

表2. ニホンザルの群れ間地理的距離 (km) と遺伝距離との間の相関係数

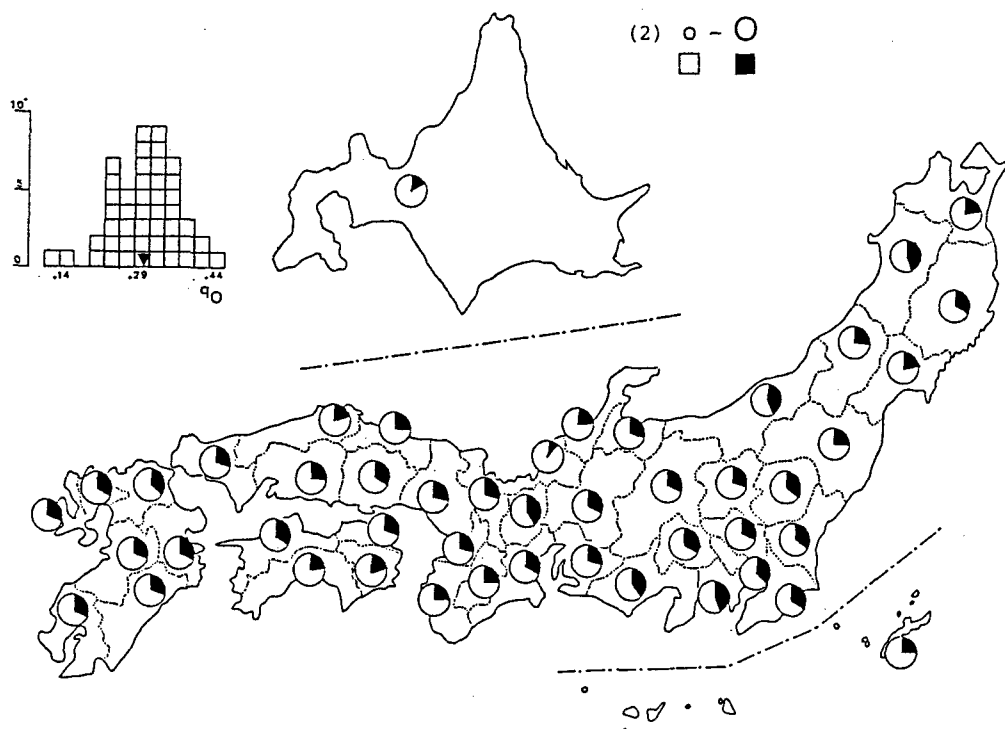
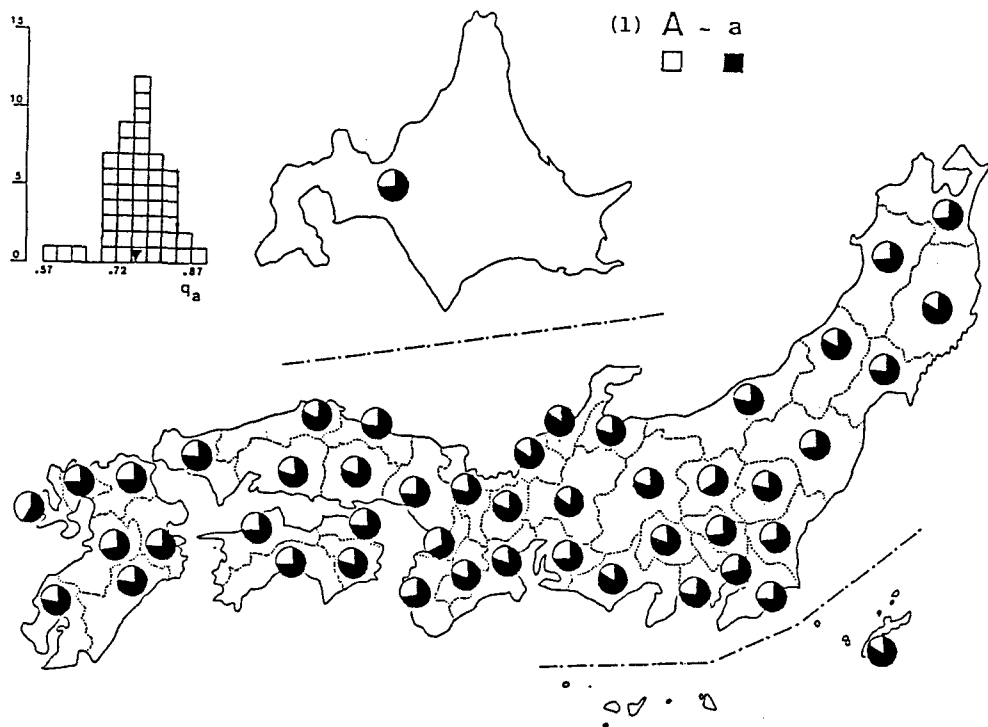
	No. of pairs	Correlation coefficient (r)	Test of significance (P)
Whole Japan	528	+ 0.1326	0.001 ~ 0.01
Different islands	339	+ 0.0503	0.3 ~ 0.4
Same island	189	+ 0.2339	0.001 ~ 0.01
Same island, distance > 100 km	145	+ 0.0275	0.7 ~ 0.8
Same island, distance < 100 km	44	+ 0.3987	0.001 ~ 0.01
Same island, distance > 50 km	174	+ 0.1489	0.02 ~ 0.05
Same island, distance < 50 km	15	+ 0.3252	0.2 ~ 0.3

(6) 種内の遺伝的分化

以上の遺伝学的分析結果から、ニホンザルという種の集団が、その内部に顕著な細分化傾向をもった構造をもっていることが知られる。総個体数として数万頭が数えられても、単位繁殖集団 (deme) のサイズは小さく、従って遺伝子頻度の機会的変動、つまり遺伝的浮動、(random genetic drift) が起りやすい性格のものである。このことは同じ日本列島内に生息している猫 (feral cat) の遺伝的多型形質の国内分布 (図3) と比較してみればよくわかる。日本人における多型形質の国内分布もまた図3と同様のものである。ニホンザルの個々の群れ集団内での遺伝的低変異性も、地域集団間での遺伝子構成の多様性も、共に上記のような構造的特質からくる帰結と考えられる。

図3. ニホンネコの2つの多型遺伝子座における
対立遺伝子頻度の国内分布.

- (1) Agouti (A-) ~ 黒色 (aa) 遺伝子座,
(2) 非オレンジ (o) ~ オレンジ (O) 伴性
遺伝子座.



文 献

KIMURA, M. and MARUYAMA, T. 1971. Pattern of neutral polymorphism in geographically structured population. Genet. Res. 18: 125-131.

KIMURA, M. and WEISS, G.H. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49: 561-576.

NEVO, E., BEILES, A. and BEN-SHLOMO, R. 1984. Evolutionary Dynamics of Genetic Diversity. pp. 13-213, Springer, Berlin, Heidelberg, New York & Tokyo.

NOZAWA, K. 1972. Population genetics of Japanese monkeys. 1. Estimation of the effective troop size. Primates 13: 381-393.

竹下 完 1964. 野生ニホンザルの分布及びポピュレーション.
野猿 19: 6-13, 20・21: 12-21.

WRIGHT, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19: 395-420.

＜その他＞

関西実験動物研究会だより

関西実験動物研究会会報7号に掲載した第24回研究会以後、以下の研究会が開催された。

1) 第25回研究会 (平成2年3月9日 国際交流会館 芝蘭会館)

講演会

1. ヨーロッパの実験動物施設を巡って 山中 久 (㈱ラビトン研究所)
2. 唾液腺涙腺炎ウイルス (SDAV) の繁殖障害
内海健二郎 (大日本製薬㈱総合研究所)

2) 第26回研究会 (平成2年6月8日 大阪科学技術センター)

講演会

1. ラットに自然発生した腎性骨異栄養症 (Renal osteodystrophy) の病態
飯田 晶敏 (大日本製薬㈱総合研究所)
2. 自己免疫疾患モデルマウス MRL に関する遺伝、育種学的研究
— 1pr 遺伝子発現に及ぼす Yaa 遺伝子の影響 —
宮脇 茂樹 (日本新薬㈱)

会員の動き

☆ 入会者 ☆

(個人会員)

藤井 良久
藤原 信之
宮本 政樹
江崎 孝三郎
西山 秀志
嵐田 好文
金田 平八郎
松田 庄司
都築 政起
加堂 洋一

ペーリンガー・マンハイム東宝(株)京都研究所
ベンリンガー・マンハイム東宝(株)京都研究所
日本チバガイギー(株)薬剤安全性研究室
大阪府立大学農学部
武田薬品工業(株)
日本クレア(株)石部生育場
(株)ラビトン研究所
(株)環境保健生物研究センター
大阪府立大学農学部
藤沢薬品工業(株)

(維持会員)

(株)ラビトン研究所
田辺製薬(株)研究統括センター

☆ 退会者 ☆

吉田 耕一
筒井 一男
塩本 泰久
三宅 可浩
高平 汎志
渡辺 嘉雄
木下 邦明
河野 一弥
八木 佐和子
高木 順彦
矢ヶ崎 修
山形 真三郎

日本商事(株)
(株)環境保健生物研究センター
大塚製薬工場(株)
国立循環器病センター研究所
協和醸酵工業(株)
神戸大学医学部
(株)ミドリ十字
大塚製薬(株)
藤沢薬品(株)中央研究所
朝日大学歯学部
大阪府立大学農学部
日本クレア(株)

平成2年8月31日

編集後記

今年の日本の夏は異常に暑かった。そんな中で、世界中の意表を突いてイラクのクウェート侵攻が8月2日に起こった。「ベルリンの壁」の崩壊や米ソ両国の接近により世界が東西対立の時代から新しい協調の時代に向かいつつある時にである。そして今、日本はいかに行動すべきかが問われている。これは日本が正確な国際情報とその分析能力を持っているかどうか、情報の語る流れや意味を読みとれるかどうかの問題であろう。

ところで、我々のような生物を扱う研究者にとっても正確なデータに加えて、それを判断する幅広い知識と適切な分析方法の重要性は言うまでもない。その意味で本号では吉村先生には統計解析の考え方とデータの生かし方を再めて教えて頂いた。また、速水、野澤両先生の記事は日頃サルやウィルスを扱わない我々のような門外漢にとっても興味深いもので、研究者としての視野の拡がりを与えてくれるものであろう。

- M. I. 記 -

平成2年10月5日 印刷
平成2年10月8日 発行

編集兼発行者 山田 淳 三
発行所 関西実験動物研究会
〒606 京都市左京区吉田近衛町
京都大学医学部附属動物実験施設
印刷所 関西ナショナル印刷株式会社
〒532 大阪市淀川区十三本町3-4-23