

コンピュータ利用による植物病原細菌の細菌学的性質の データ集積と検索法

誌名	農業環境技術研究所報告
ISSN	09119450
著者	畔上, 耕児 西山, 幸司 渡辺, 康正 福田, 徳治
巻/号	2号
掲載ページ	p. 19-34
発行年月	1986年7月

コンピュータ利用による植物病原細菌の 細菌学的性質のデータ集積と検索法*

畔上耕児**・西山幸司**・渡辺康正**・福田徳治***

(1985年10月31日受理)

植物病原細菌の細菌学的性質を調べ、そのデータをコンピュータにファイルした。そして新たに検査した菌をファイル菌の中から検索し、類似菌をリストアップするプログラムを作成した。さらに同種(pathovar)とされている菌群の変異幅が小さく、別種(pathovar)とされている菌群間では違いが大きくなるような分類がなされているかどうかを調べるために、まずそれらの変異幅を明らかにした。その結果、一部の細菌は種としてよくまとまっており、検索をすると“ノイズ(別種であり、リストされることが期待されないにもかかわらず、類似度が高いためにリストされてしまう菌)”も少なく、期待される菌のみが拾い上げられ、検索は十分可能であることがわかった。*Pseudomonas syringae* 菌群では異なる pathovar も拾い上げられてしまうという難点があったが、それは当然のことといえる。また植物病原細菌のデータをコンピュータに集積する際の問題点を指摘した。

I 緒 言

コンピュータや通信システムの発達で大量の情報処理や伝達が可能となり情報化社会といわれる今日、細菌学においてもコンピュータは分類や同定に頻りに用いられるようになってきている。細菌の分類には様々な場所から分離された多数の菌株について多数の検査項目のデータが蓄積されている必要があり、その処理のために不可欠となっているからである。また同定においても、検査している菌の性質を既に報告されている菌の性質と照合して見る必要があり、その検索のためにデータをコンピュータに集積しておくことが望まれる。

植物病原細菌の場合には、従来、病原性に重きを置いた分類や同定が行われてきた傾向があり、細菌学的性質が十分に調査されていない菌株も少なくない。また調査されていても、その方法が異なるために結果の厳密な比較はできないことが多い。そこで著者らは手元にある植物病原細菌の細菌学的性質を調べ、それをコンピュータに集積し、各種細菌の細菌学的性質の変異幅を明らかに

するとともに、全国から他の研究者がその集積データにアクセスしてオンラインで細菌を検索する方法と問題点等について検討した。

II 植物病原細菌の分類・同定における コンピュータ利用について

細菌の分類を行った場合には、なぜそのように分類したか、あるいはなぜそのような分類体系を創ったかが明らかにされていなければならない。細菌は微小であり、多大な労力と時間を投じて調査しなければ情報が得られず、またそのようにして得られた情報も細菌の一側面をとらえているにすぎず、そのうちの何をよりどころとして分類するのかが常に大きな問題だからである。本実験は植物病原細菌各種の細菌学的性質の変異幅を調べて、その性質による細菌検索の可否を検討したものであり、分類体系そのものの妥当性にまで立ち入って論ずるわけではない。しかし、やはりまず植物病原細菌の分類・同定に対する現在の著者らの考え方を整理し、今回の実験にコンピュータを用いたことの原因について明らかにしておかなければならないと思われるので、やや長くなるが述べておきたい。

そもそも分類にコンピュータが用いられるようになったのは、細菌学者 SNEATH(1957a,b)や昆虫学者 SOKALらの論文が契機となったのであるが(COWAN, 1968)、その

* 本論文の概要は、日本植物病理学会夏季関東部会(1983年7月22日)で発表した(畔上ら, 1984a, b)。

** 環境生物部微生物管理科

*** 熱帯農業研究センター

分類は18世紀中葉のフランスの博物学者 ADANSON の考え方にもとづいている。ADANSON は植物の分類をしながら、共通の性質を最も多く有する生物を集めて分類群を構成するのが自然分類に最もかなったやり方であり、特別の性質のみを類別のカギとして重くみる正当な理由はなんら見出せない、という考え方をもつに至った。それは、観察しうるすべての性質を順位をつけなくて比較して、共通点の多い生物、すなわち形質全体にもとづいた相似度(overall similarity)の高いもの、を集めて分類群を作り、次に相似度の低い分類群へとまとめたらどうか、という考え方である(坂野, 1974)。それは当時の考え方からあまりにかけ離れすぎていたためと、処理に膨大な時間と労力がかかるためにかえりみられなかったが、1950年代に入りコンピュータの発達によってそれらを克服することが可能となり、それが SNEATH や SOKAL らによって受け継がれ、数値分類と呼ばれるようになった。

数値分類は、系統樹をうらづける化石が残らない細菌の場合には系統分類は意味をなすかどうか疑問があるとして、それとは全く別なものとして発達した。しかし、属よりも上級の分類群は別として、少なくとも種のレベルでは、数値分類も細菌の相互関係を主観的評価によって判断してきた古い分類学者の分類が妥当であることを認め、またはからずもそれまで見のがされてきた関係などを指摘したり(COWAN, 1968)して、必ずしも“従来の分類”と対立するものではないことがわかってきている(ときには従来の分類がはらんでいた矛盾を指摘することもある)。そして検査項目数が多ければ、常に安定した分類結果が得られることがわかってきているが、それは後述のように、混乱を招かないために重要なことである。ところで細菌の分類方法には、表現形質である細菌学的性質(形態的性質、培養的性質、生理・生化学的性質、など)にもとづいたいわゆる伝統的な分類、SNEATH や SOKAL らによって始められた数値分類、DNA 塩基組成や DNA 相同性あるいは遺伝子伝達などを調べる分子生物学的分類、細胞壁構成成分などを指標とする化学的分類、血清学的分類(後藤, 1980)、あるいは COWAN の異端分類(COWAN, 1965, 1970, 1971)と呼ばれるものがある。ただし、これらのうち COWAN 自身がそう呼んでいる異端分類は各分類群の細菌の性質を記号で表現、列挙し標識とするだけで、それ以上の属や科など上位の分類群を作ることは考えないので、分類といえるか議論のあるところである。細菌分類の窮極の目的が自然分類体系あるいは系統分類体系を作ることにあるとするならば、そして生物が生き残るために DNA を持ったのではなく、DNA

が残るために生物体という装置を作ったという最近の一般的な考え方によれば、分類のためには2次的情報というべき表現形質を調べるよりは1次的情報であり進化の跡が刻み込まれている DNA の塩基配列そのものを調べるべきである。しかしそれを取り入れることは技術的に可能となりつつあるが、労力的にはまだ不可能な状態である。DNA の塩基配列は細菌分類において、機器の発達や技術水準の向上によって今後ますます重要視されるようになるであろうが、それはやはり表現形質と結びつけられるべきものであるし、また後述の同定にはより簡単に結果が得られる性質が必要であり、表現形質を調べておくことは将来においても、細菌分類と同定を行うために現在と同じように必要なことである。

さてその表現形質にもとづいて細菌の同定を行おうとするならば、必ずいくつかの細菌学的性質、しかもそのうちの生理・生化学的性質を調査しなければならない。それは、細菌のように小さな体制で、形態的特徴が乏しい生物を形態的な性質のみによって分類することは非常に困難であり、多様な生理活性をもつ細菌が発見されるに従って、形態よりも生理、生化学的特徴を重要視するようになって、今日ではどの細菌分類体系においても、これらの特徴を分類の基準として取り入れている(坂野, 1974)からである。ここで別種あるいは同一種と分類されるゆえんが、特定の例外のない重要性のはっきりしている性質、すなわち常に他の多数の性質と符合する性質にあるのならば、細菌の同定はそれをメルクマールとして行える。ところがその分類は、特定の性質にもとづいてなされるのではない。結果的には、個々の菌株間や菌群間でいくつかの性質の違いが議論され、それが強調されて分類・同定に用いられることになるが、その性質は特定のものでなければならないということではない。それは次のような理由によると考えられる。すなわち、微小な細菌の生理・生化学的性質を調査してそれを特徴づけるためには、多大な時間と労力を投じなければならず、そのために変異を内包するであろう集団のごく一部しか認識できない。その集団の中からたまたま選ばれた菌株について、たまたま選ばれた検査項目の調査によって形成された部分集団、それから形成された概念が細菌の種であって、それと他の種を区別するのに有効であろうとされる性質の意味などについては明らかでない場合が多い。もちろんそのような性質でも分類学上重きを置くに値することが多かったが、誤って重きを置かれたものも少なくない。植物や動物の場合とは異なり、それまでにわざわざ時間と労力を投じてたまたま調査されたこと以

外には、分類学上役立つような情報はわずかしき得られない。これらのことから様々な地域から分離された多くの菌株について多くの調査が行われ、情報が蓄積するにつれて、分類・同定に有効とされる性質に例外ひいては分類体系の矛盾が出てくることは当然であり、分類に携わる研究者は日頃からそれを感じており、特定の性質が分類に重要あるいは有効であると強調することができないからであろう。

したがって細菌の分類は、新たに調査した菌株がある種のたまたま選ばれた代表というべき特定の菌株 (Type culture) に十分類似しているか否か、あるいは複数の菌株を別群として分けるほどに異なっているか否か、などの基準によって判断されるにすぎない。そしてそれ以上には、命名規約 (LAPAGEら、1975) はあっても、種についての厳格な統一の見解はなく、柳田 (1980) によれば各培養 (クローン) について、その表現型の諸性質を統計的に調べてみると、あるいくつかの培養集団同士では連続するが、ある培養集団と別の培養集団との間には断層が認められる。そこで連続する表現型を示す培養集団を“種”としてまとめるのが普通である。そのために、次のような問題が生ずる。すなわち、細菌の分類は表現型の断層がはっきりしている場合には種の断定は容易であるが、それが明確でない場合には研究者の主観が入る余地が十分にあり、それによって起こる混乱は避けられないのが現状である (柳田, 1980)。それゆえ両極端ははっきりしているが、その間には明確に区別できない2群を区別すべきであるか、あるいは区別するとしたら“どこに線を入れるか”は、結局は主観にゆだねられ、その点から“分類は芸術”と皮肉られることがある (そこには普遍的な美しい体系を創造すべきものという意味あいも含まれているであろう)。分類するということがそもそも人為的なことであるため、最終的には主観にゆだねられるのは避けられないとしても、より多くの研究者の同意が得られ、誰が行っても同じ結論に達するようなより客観性のある安定した分類体系が常に模索されねばならない。このような状況下で別種とするかどうかについての判断は、もし同一種とされている群の変異幅が明らかにされているならば、より客観的にそして容易に行えるであろう。坂野 (1974) によれば、一つの種について、多数の細菌培養の表現型のデータを集めることがその種内の遺伝子型変異の限界を知り、自然界における細菌種の境界を客観的に認知するために、まずなされなければならないことであるが、残念ながら細菌種を集団として取り扱った分類学的研究はあまり見当たらないのが現状である。

同定とは、新たに調査した細菌が既知の細菌のいずれであるかを、既知の細菌を識別するのに有効であろうと考えられている性質にもとづいて、判断することである。仮に同定しようとする菌株が、すべての性質で既知の種 (Type culture あるいはそれに十分類似しており同一種とされる菌株の集合) に一致するならば、同定は容易である。しかし検査項目数が多くなるとそのようなことはまれであり、一致しない場合、同定といえども特定の性質からだけではなく、総合的に判断されなければならない。その結果、あるいは分類が再検討されねばならないこともある。また、培養的性質や生理・生化学的性質などにもとづいて同一種とされている菌が、学問の進歩によって特定の性質の重要性が明らかになったり、機器の発達でそれまでは容易に行えなかった検査も行えるようになったり、それらによって違いが明らかとなり別種とされることがある。このように細菌の分類体系や分類方法は学問の進歩や技術の発達によって変わりうるものであり、実際が変わってきているが、分類には多数の菌株の様々な調査結果の蓄積が必要であることには変わりない。すなわち特定の性質にもとづいて人為分類を行うのでない限り、数値分類をするか否かには関係なく、またそれ以外のどのような考え方の分類方法をとるにしても、いつでも様々な性質をとり出してきてそれらを等価として評価しなおせるようにしておかなければならない。

細菌の分類にはその多量の情報が一つの表として表されなければならないが、この表は一方で様々な遺伝子の有無を表すものでもあり、微生物の様々な能力に期待が寄せられている今日、その要請に応えるためにも必要である。

このように細菌の多量の情報が盛りこまれた大きな表を作成するために、そしてその処理のために、コンピュータが用いられているのであるが、ここで植物病原細菌について見るならば、COLWELLら (1968) や SANDSら (1970) はおもに蛍光色素産生 *Pseudomonas* 属菌と *Xanthomonas* 属菌の分類に、また DYE (1981) は *Erwinia* 属菌の分類にコンピュータを用いている。そもそも分類とは、個々の菌株の性状は覚えきれないがゆえに行われるのである。そしてその分類群に与えられた名が呈示されるだけでその分類群に属する菌株の豊かで正確な有用情報が得られるためのものである。したがって農業上の必要性からすればおもに寄生性、農薬感受性、死滅温度などに関する情報が得られるような分類でもいいわけであるが、これらの例はやはり基礎に細菌学的性質などにもとづいた分類があることが求められていることを示すものと言

えるであろう。今日では、植物病原細菌の分類に用いられる細菌学的性質の検査項目数は属によっては150を越えており、コンピュータの使用は不可欠となっているが、その例は多いとは言えない。それは、様々な菌株を集めて大きな表を作成するのに多大な時間と労力を要することと、その割には、たとえそのように大きな表を埋めてたとえば数値分類を行ってみたところで、得られる結果はそれまでの分類体系を支持するのに役立つデータが得られるだけで、目新しい知見は多くないであろうと予想されること等が原因であると考えられる。

植物病原細菌を同定する場合は、それがどの種に近縁であるかは、コンピュータで検索するまでもなく、わずかな検査を行っただけで比較的容易に見当がつく。しかし調査結果がいずれの既知の細菌とも一致しない場合には、それをどの種と同定するかは、必ずしも容易ではない。それは、種の変異幅がはっきりしていないにもかかわらず新種が作られたために、2つあるいはそれ以上の種の間関係が判然としない場合が生じていることや、あるいは過去に報告された菌株の調査データが乏しかったりして比較できない場合があるためによる。また、調査時のデータは菌株単位であっても、文献では類似菌群単位に圧縮されてしまい、菌株間のわずかな違いは表記され得ず、個々の性質は同種のすべての菌株で陽性あるいは陰性であるのか否かが読みとれない場合があるためによる（正確に比較するには、データは菌株単位で記述されていなければならない。特に植物病原細菌の場合には分類・同定される菌株数が医学細菌と比べて限られており、それが望まれ、また可能でもある）。さらに、そのデータは個々の報告や研究者のノートにばらばらに記載されており、それらを探し出したり、比較検討したりするためにはかなりの時間と労力を費やさなければならないためにもよるの、述べるまでもないことである。分類のためばかりでなく、このような同定の際の多くの煩雑さや困難を回避するためにもデータを菌株単位でコンピュータに集積しておくことが望まれる。

植物病原細菌の分類・同定に係わる著者らは、以上に述べたような状況から、そのデータをコンピュータに集積し、様々な細菌学的性質による細菌検索の可能性について検討するの必要に迫られた。

III 材料および方法

1 供試細菌

表一1に示したとおり植物病原性の *Pseudomonas* 属菌 26種類135菌株、*Agrobacterium* 属菌 1種類 6菌株、

Xanthomonas 属菌 3種類 3菌株を用いた。

2 細菌学的性質

おもに西山(1978)の方法に従い、表一2に示した47項目について調査した(GC含量は一菌株のみ)。糖からの酸産生や有機酸の利用は、Ayers, Rupp and Johnson 培地(AYERSら, 1919)を基礎培地とし、それぞれ1, 0.15% 加用した。調査は移植2, 4, 7, 14日後に行った。

3 コンピュータとプログラム言語

農林水産研究計算センターの大型コンピュータを用い、検索用プログラムはFORTRANで作成した。検索用プログラムと次の細菌学的性質のデータはそれぞれ別にファイルした。

4 調査結果のコンピュータへの入力

細菌の性質は次のように3つに分けられる。すなわち、1. +かーかで表されるような2状態変量(0-1型データ)、2. 色やコロニーの性状のようにいくつかのカテゴリーに分けられる多状態変量、3. 大きさのような連続変量である。このうち多状態変量はいくつかのカテゴリーに分解し、たとえば黄色色素を産生するか否かのように2状態変量と同じように表すことができる。連続変量はいくつかの段階に分けて2状態変量に分解することも可能であるが、それではどこで分けるかについて問題があるし、なによりもせっかくのものとの情報が減ぜられることになる。したがって細菌の性質をコンピュータに入力するには、2状態変量と連続変量の2つを考える必要がある。細菌学的性質に代表される2状態変量は0~3で表した。すなわち、未調査の場合は0、陰性は1、弱陽性は2、陽性は3とした。2を設けたのは、データをとる際に弱陽性と記録しておきたい場合があるからであるが、次の類似度の計算では2と3は等価である。連続変量の一つGC含量はパーセンテージをそのまま入力した。

5 類似度の計算

コンピュータで類似する細菌を検索するためには、まず菌株間の類似度あるいは距離を何らかの“ものさし”で測る必要がある。その距離の計り方にはいろいろあるが(SNEATHら1973, 矢島1976, 吉沢1981, 鈴木1982, 金子1982)、今回はほとんどが2状態変量なので数値分類においてその場合に最も単純で頻繁に用いられている2つを用いることにした。それらには様々な呼び方があるが、混乱を避けるためにここではそれぞれ“単純一致率”および“陽性一致率”と呼ぶことにする。単純一致率は $SM = 100(A+C)/(A+B+C)$ で表される。ここでAは両細菌とも陽性となる性質の項目数、Bは両細菌の間で異なる項目数、Cは両細菌とも陰性となる性質の項目数

である。単純一致率はすなわち一致項目数の全調査項目数に対する割合である。陽性一致率は $PM=100 \cdot A / (A+B)$ で表される。陰性の性質はいくらでも考えられるということ、またたとえば代謝経路のどの個所の能力を欠いても陰性とするので、陰性の一致は見かけ上の一致にすぎないなどの理由から、単純一致率の算出においてCをはずしたのがこの値である。ただしグラム反

応とOF試験における一致は、CではなくAに含めた。なお両細菌のいずれかが0すなわち未調査の検査項目はこれらの計算には入ってこないようにした。

GC含量は、検索しようとする菌(以下ケンサ菌)のGC含量がファイルされている菌(ファイル菌)のGC含量の±5%以内に落ちていれば一致としたが、幅を広くとったのは、ファイル菌の中から似ている菌はもれなく拾い上

表-1 供試細菌

種 (pathovar)	菌 株 番 号*
<i>P. andropogonis</i>	NIAES 1005, 1007, 1116, 1154, KN17, 25
<i>P. avenae</i>	NIAES 1024, 1027, 1030, 1036, 1141, R24
<i>P. caryophylli</i>	NIAES 1060, 1100, 1192, 1195, 1406, 1410
<i>P. cichorii</i>	NIAES 1158, 1180, 1368, 1370, 1373, KN52
<i>P. fuscovaginae</i>	NIAES 1177, 1178, 1179
<i>P. gladioli</i> pv. <i>gladioli</i>	NIAES 1064, 1065, 1066
<i>P. glumae</i>	NIAES 1093, 1169, 1170, 1171, 1172, 1388
<i>P. marginalis</i> pv. <i>marginalis</i>	NIAES 1173, 1378, 1382, 1416, 1417, K04
<i>P. solanacearum</i>	NIAES 1067, 1069, 1189, 1418, KN117, R01, 14
<i>P. viridiflava</i>	NIAES 1130, 1157, 1293, 1343, 1347, KN01
<i>P. woodsii</i>	NIAES 1156, 1190, 1191, 1216, 1218, 1220
<i>P. syringae</i> pv. <i>aptata</i>	NIAES 1008, 1009, 1011, 1012, 1014, 1015
pv. <i>atropurpurea</i>	NIAES 1017, 1304, 1305, 1307, 1309, 1310
pv. <i>coronafaciens</i>	NIAES 1016, 1061, 1314
pv. <i>erobotryae</i>	NIAES 1062, 1063
pv. <i>glycinea</i>	KN 28, 31, 41, 44, 45, 168
pv. <i>japonica</i>	NIAES 1071, 1072, 1159, 1163, 1165, 1167
pv. <i>lachrymans</i>	NIAES 1290, 1291, 1316, 1318, 1321, 1415
pv. <i>maculicola</i>	NIAES 1174, 1175, 1419, KN83, 98, 110, 10
pv. <i>mori</i>	NIAES 1020, 1021, KN21, 22, 133
pv. <i>morsprunorum</i>	NCPFB 330, NIAES 1436, 1437
pv. <i>phaseolicola</i>	NIAES 1023, KN 79, 81, 85, 86, 132
pv. <i>pisii</i>	NIAES 1208, 1209, 1210, 1211, 1212, 1213
pv. <i>striaefaciens</i>	NIAES 1032, 1033, 1034, 1440
pv. <i>syringae</i>	NCPFB 524, NIAES 1429, 1430, 1431, 1434
pv. <i>tabaci</i>	NIAES 1073, 1074, 1075
<i>A. tumefaciens</i>	NIAES 1001, 1222, 1224, 1276, 1279, 1280
<i>X. campestris</i> pv. <i>campestris</i>	NIAES 1151
pv. <i>citri</i>	NIAES 1077
pv. <i>pruni</i>	NIAES 1421

* : NIAESは農業環境技術研究所、NCPFBはNational Collection of Plant Pathogenic Bacteria(英国)の保存菌株、それ以外は著者らの未発表の菌株である。

表-2 検査項目

GC 含量, グラム反応 : (-) = 1, (+) = 3. OF 試験 : O = 1, F = 3.

糖からの酸の産生および有機酸等の利用(グルコース, トレハロース, エリトリトール, マンニトール, ソルビトール, イノシトール, スクロース, ラムノース, ベタイン, アルギニン, β -アラニン, L-乳酸塩, L-酒石酸塩, レプリン酸塩, グリセリン, トリゴネリン, キシロース, セロビオース, ガラクチトール, L-アラビノース, D-アラビノース, フルクトース, D-酒石酸塩, 安息香酸塩, ガラクトース, マンノース).

アルギニン水解, セラチン液化, 蛍光色素産生, 非水溶性黄色色素産生, カタラーゼ, オキシダーゼ, 37°Cにおける発育, 40°Cにおける発育, エスクリン水解, NO₃還元, NO₃呼吸, ジャガイモ腐敗, H₂S産生, カゼイン水解, マーガリン水解, チロシナーゼ, ツイーン80水解, レシチナーゼ

げりストップしようとしたからである。GC 含量は今回は1菌株についてしか調べてないので、比較の計算には入っていない。

6 各細菌種(pathovar)の変異幅の表示

数値分類では、n 固体について P 個の性質が調べられているとき、各個体は P 次元空間内に散らばる n 個の点として表わされ(鈴木, 1982)、各点の位置によって分類群が作られる。その分類群は、n 次元空間内の 2 点間の距離として計られた 2 個体間の類似度あるいは距離のみをもとに構成される。各点は類似度の高い(あるいは距離の近い)ものから順次まとめられ、いくつかの分類群が形成される。そのまとめ方にはいろいろな方法が考えられ、どの方法をとるかによって形成される分類群も異なってくる。それゆえ、いくとおりにもできる分類あるいは方法が評価される。その評価の基準は、形成された分類群内では各点がよくまとまっており、異なる分類群間の距離が大きいかということである。

コンピュータで細菌が、いくつかの性質に重み付けするような操作を加えることなくうまく検索されるためにも、そのような分類がなされていることが望まれる。植物病原細菌の場合、従来細菌学的性質がなおざりにされ、病原性に重きを置いた同定や分類が行われてきた傾向があるので、細菌学的性質からも、その分類が妥当であり、検索が可能かどうかをまず調べなければならない。少なくとも検索の結果、類似菌がもれなくリストアップされるためには、同一種とされる菌株の細菌学的性質の変異幅が小さいことが望まれる。そこですべてのファイル菌のデータをケンサ菌のデータとして入力し、すべての組合せの菌株間で一致率を求めた。そして各細菌種(あるいは pathovar)ごとに、そのまとまり具合すなわち各細菌種の細菌学的性質の変異幅および他の細菌種との類似度を、その一致率をもとに図示して明らかにした。

IV 結果

1 プログラムの機能

- (1) 入力したケンサ菌の識別名, GC 含量, 細菌学的性質147項目が出力される(未調査でパンチカードにパンチしない場合には0と印刷される)(図-1)。なお検索できる項目は現在148であるが、増やすことも可能である。
- (2) 検索の際に、すなわち単純一致率あるいは陽性一

ホンキン・データ

ケンカブ	GC		*	*
TEST 1	67.8%	11311 33313 11131 11131	
		03333 23111 11111 31111	
TEST 2	70.0%	11333 33111 11111 11133	
		03331 31111 00000 00000	

図-1 出力結果(その1)。入力したケンサ菌の性質がGC 含量の場合は%で、それ以外の147項目は1~3(1:陰性, 2:弱陽性, 3:陽性)で印刷される。パンチカードにパンチしなければ0(未調査)と印刷される。

マスクド・データ

ケンカブ	GC		*	*
TEST 1	67.8%	11300 33313 11131 11131	
		03333 23111 11111 31111	
TEST 2	70.0%	11300 33111 11111 11133	
		03331 31111 00000 00000	

図-2 出力結果(その2)。図-1において4番目と5番目の検査項目をはずして(未調査として)検索したいときに出力される。

効率の計算の際に、すでにパンチカードにパンチしてしまった検査結果のいくつかをなんらかの理由で、たとえばその検査項目は同種であっても菌株によって結果が異なるので、あるいは判然としないので、採用しないで計算してみたいなどの理由からはずしたいとき、それらの検査項目をはずす（マスクする）ことができる。このマスク機能によって、自由に検査項目を選択して検索できる。マスクした場合、マスクド・データも印刷される（図-2）。

(3) ケンサ菌とファイル菌それぞれとの間の単純一致率または陽性一致率を計算し、ケンサ菌それぞれについてファイル菌のなかから、一致率の高いほうから指定したパーセンテージまで菌株をリストする。リストされた菌株については、1. ケンサ菌の菌株識別名、2. 単純一致率および陽性一致率、3. ファイル菌株番号、4. ファイル菌の種（pathovar）名、5. 文献、同定者あるいはコメント、などをプリントする（図-3）。

(4) ケンサ菌のデータの入力は、あらかじめ指定され

た書式でも、また検索者が指定した書式でも可能である。

(5) なお検索にあたりいくつかの制限を設けたので、それからはずれると図-3に示したものが印刷される。

2 細菌各種類の変異幅と検索結果

表-1に示した各細菌種(pathovar)の細菌学的性質の変異幅と検索可能性を調べた結果は次のようであった。

(1) 単純一致率による検索：図-4に示したように、まず *P. syringae* 群菌以外の *Pseudomonas* 属菌についてみるならば、*P. andropogonis*, *P. avenae*, *P. caryophylli*, *P. cichorii*, *P. fuscovaginae*, *P. gladioli* pv. *gladioli*, *P. glumae*, *P. viridiflava*, *P. woodsii* のいずれもが、それぞれ同種と同定されている菌株のグループ内によくまとまっていた。とくに *P. avenae*, *P. caryophylli*, *P. cichorii*, *P. fuscovaginae*, *P. gladioli* pv. *gladioli*, *P. glumae* では、その種と同定されておらず、リストアップされることが期待されないにもかかわらず、一致率が高いためにリストアップされてしまう菌株（いわば“ノイズ”）も少なく、よくまとまっており、検索は十分可能で

菌株	一致率（一致項目数）	ファイル菌	種(pathovar)名	文献・同定者等
ファイルキン (1985.08) トノヒカク ケッカ				
TEST 1	SM100.0% (46/46) PM100.0% (15/15)	N 1234	P. ANDROPOGONIS	
TEST 1	SM 97.8% (45/46) PM 93.3% (14/15)	N 1235	P. ANDROPOGONIS	
TEST 1	SM 97.8% (45/46) PM 93.8% (15/16)	N 2345	P. WOODSII	
TEST 1	80% ミマン ハ インサツ サレマセン			
TEST 1	PM100.0% (15/15) SM100.0% (46/46)	N 1234	P. ANDROPOGONIS	
TEST 1	PM 93.8% (15/16) SM 97.8% (45/46)	N 2345	P. WOODSII	
TEST 1	PM 93.3% (14/15) SM 97.8% (45/46)	N 1235	P. ANDROPOGONIS	
TEST 1	80% ミマン ハ インサツ サレマセン			

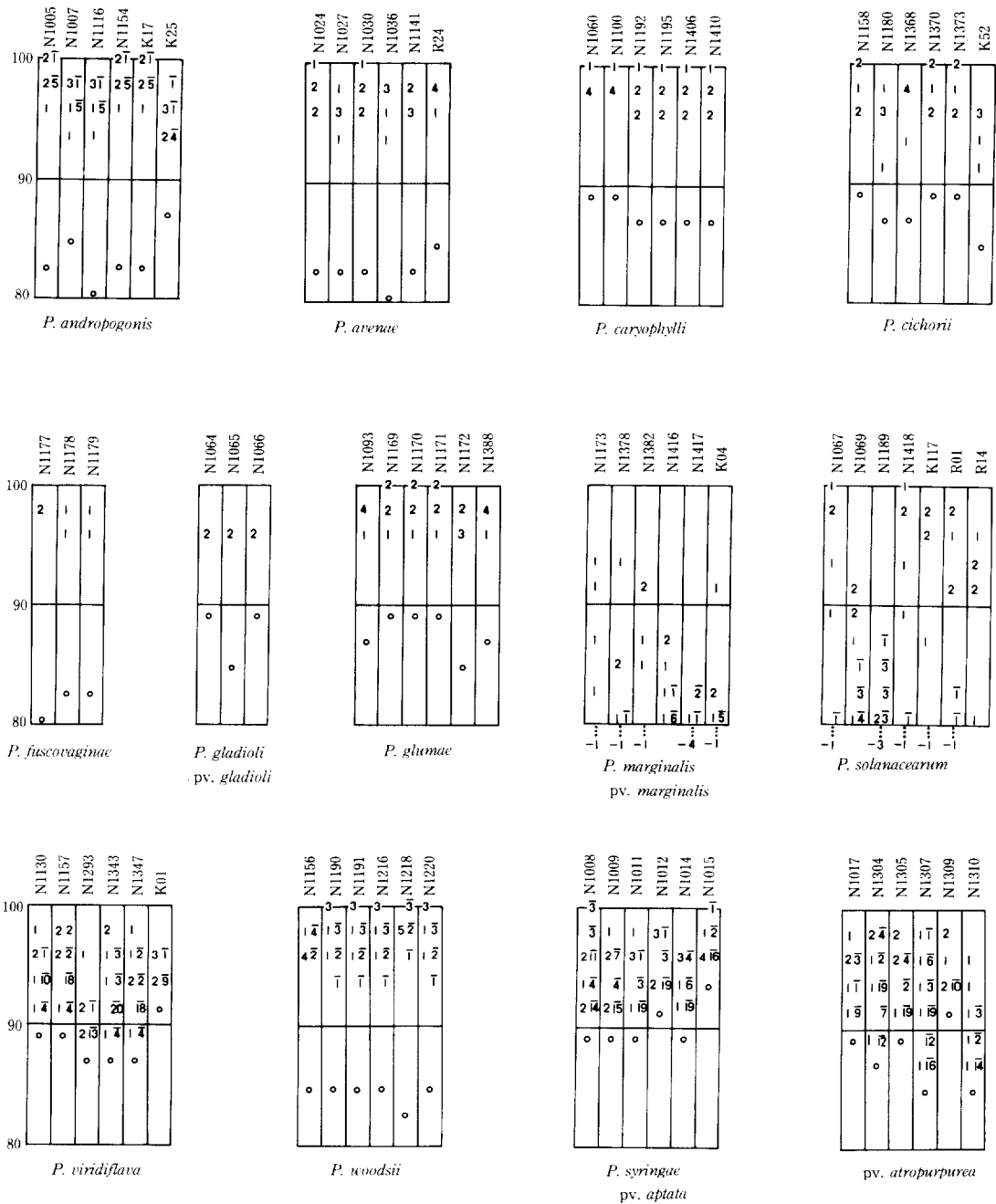
TEST 2	ケンサ コウモク ハ 30イジョウ ヒツヨウデス			

TEST 3	グラム ハンノウ (-), OF(O) デハ アリマセンノデ ケンサクシマセン			

TEST 4	SM ガイトウ スル キン ハ ファイル サレテ イマセン			
TEST 4	PM ガイトウ スル キン ハ ファイル サレテ イマセン			

TEST 5	SM 97.8% (45/46) PM 93.3% (14/15)	N 1235	P. ANDROPOGONIS	
ファイルキン ノベ 10,000 キンカブ イジョウ ハ インサツ サレマセン、パーセンテージ ヲ アゲルカ、アラタメテ データ ヲ オイレクダサイ、				

図-3 出力結果(その3)。各ケンサ菌についてファイル菌との比較結果が印刷される。上の例は単純一致率SM値と陽性一致率PM値の両方でリストアップしたときのものである。なお、入力や出力にいくつかの制限を設定したので、図中にかたかなで示した内容が印刷されることもある。



図一 4 単純一致率でみた各種(pathovar)の細菌学的性質の変異幅 (その1)

図中の数字は縦軸の%で一致する菌株数。バーがついていない数字は、同種(pathovar)と同定されており、リストアップされることが期待される菌株数。バーがついている数字は、異種(pathovar)と同定されており、リストアップが期待されない、いわば“ノイズ”の菌株数。○はノイズの位置のみを示す。80%ラインより下の数字は同種(pathovar)と同定されているにもかかわらず、一致率が80%未満だった菌株数。

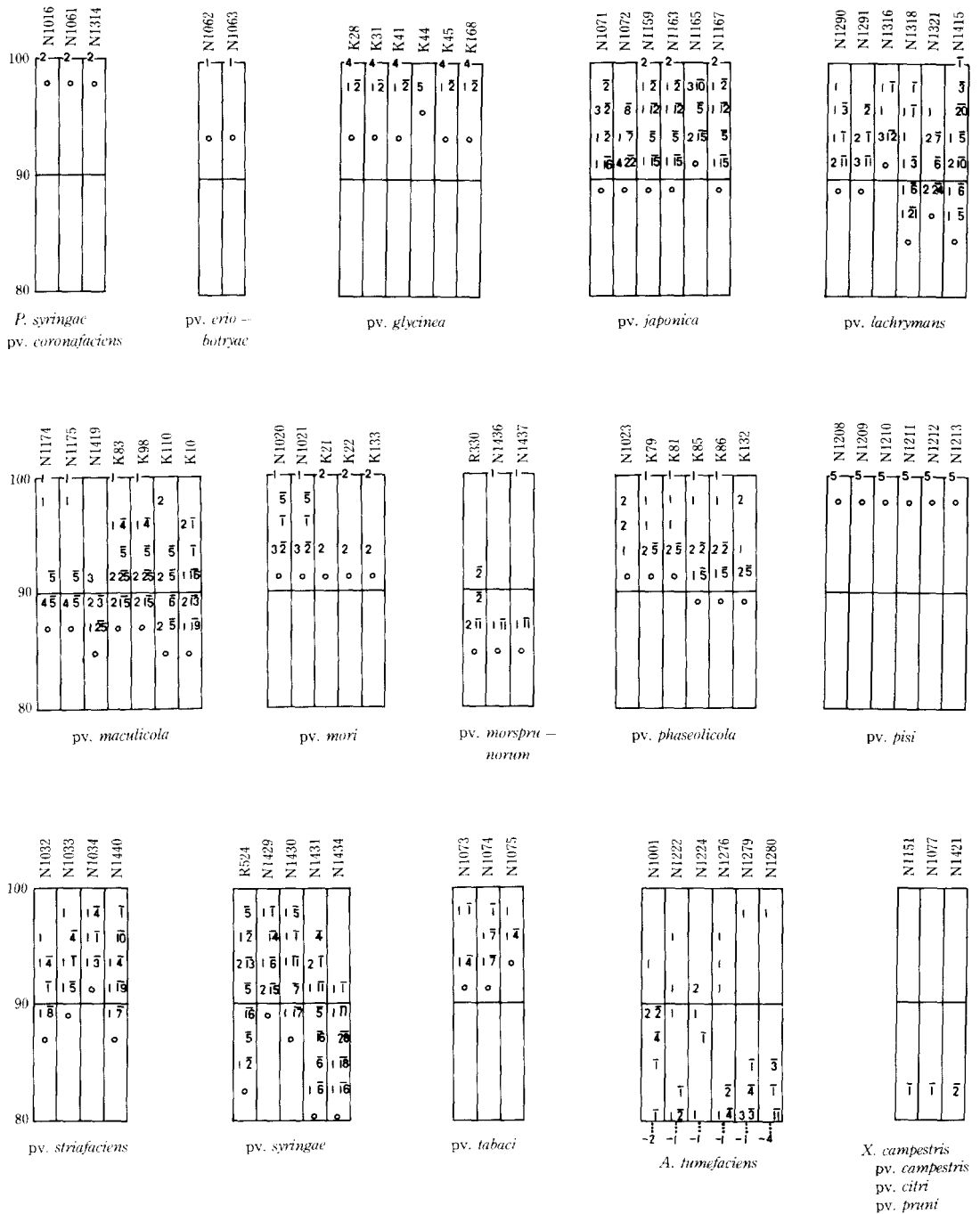
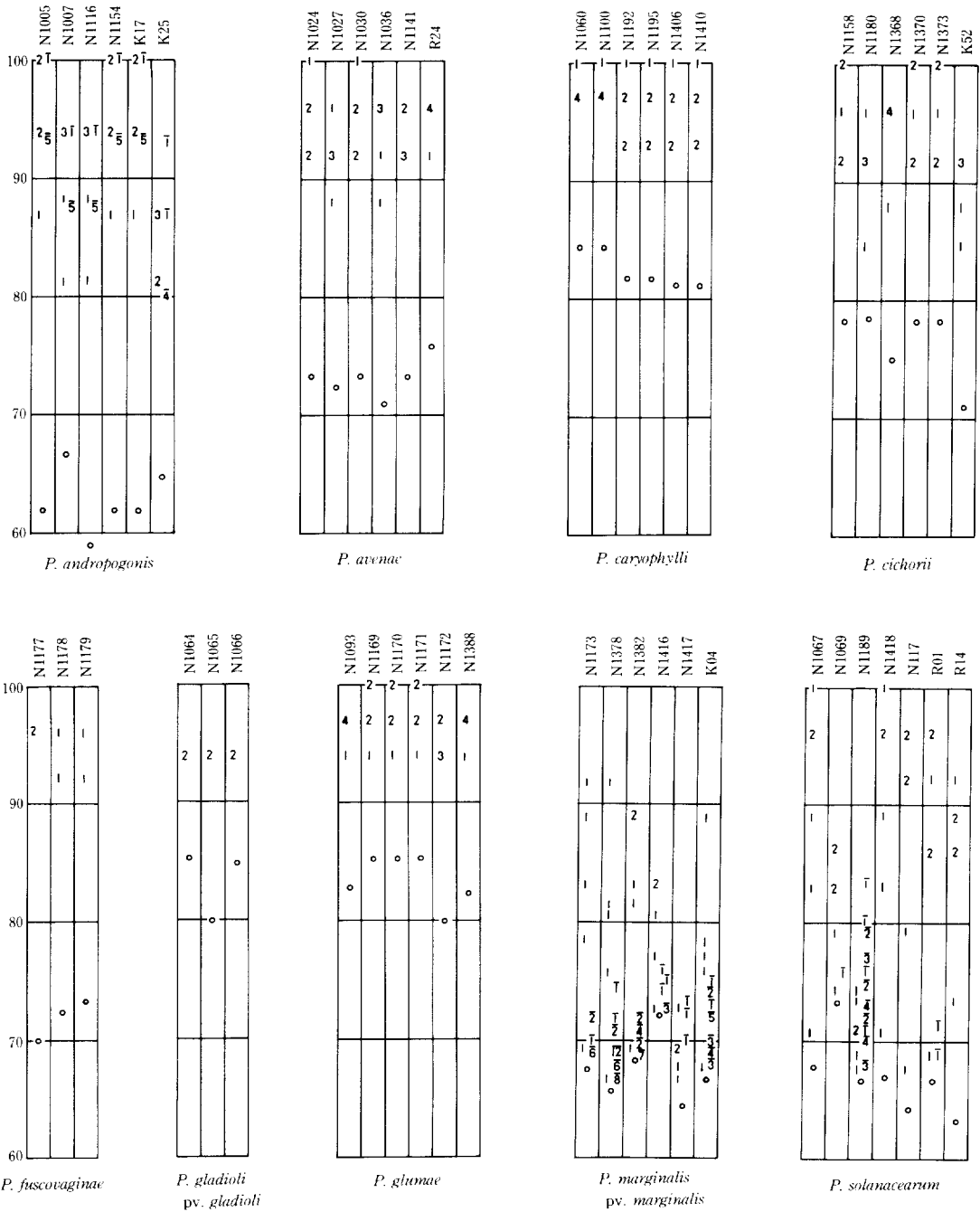


図-5 単純一致率でみた各種(pathovar)の細菌学的性質の変異幅 (その2)



図一六 陽性一致率でみた各種(pathovar)の細菌学的性質の変異幅 (その1)

陰性で一致した項目は除外され、算出式の分母がまちまちになり、一致率は小さきみになる場合があるため、その差が小さくて菌株数が書けない場合にはまとめてある。

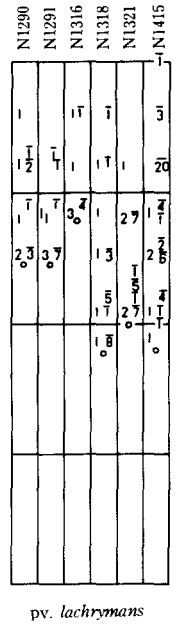
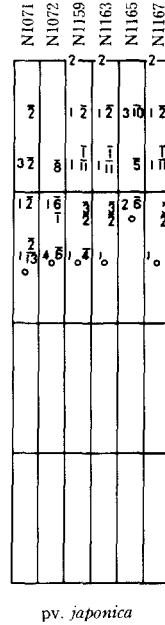
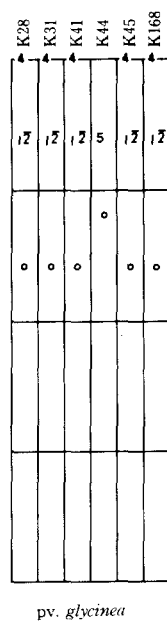
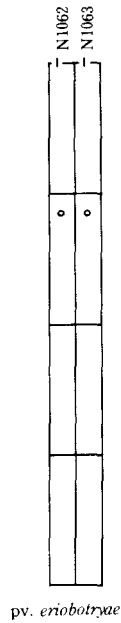
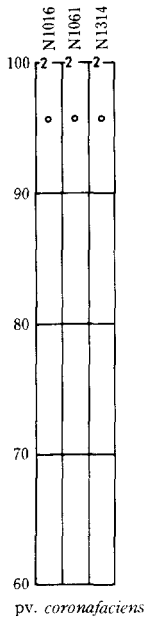
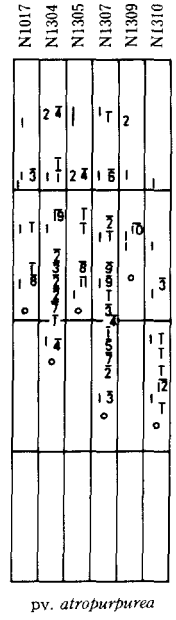
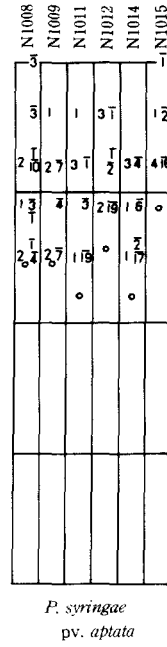
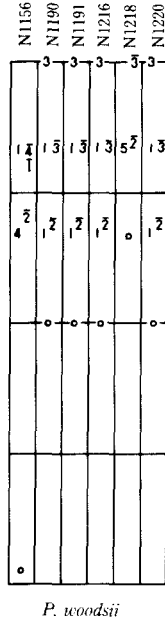
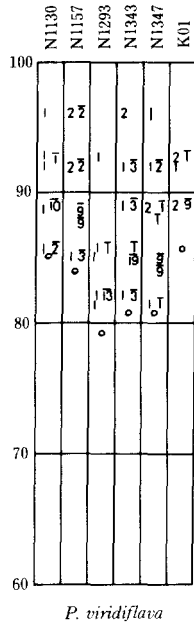
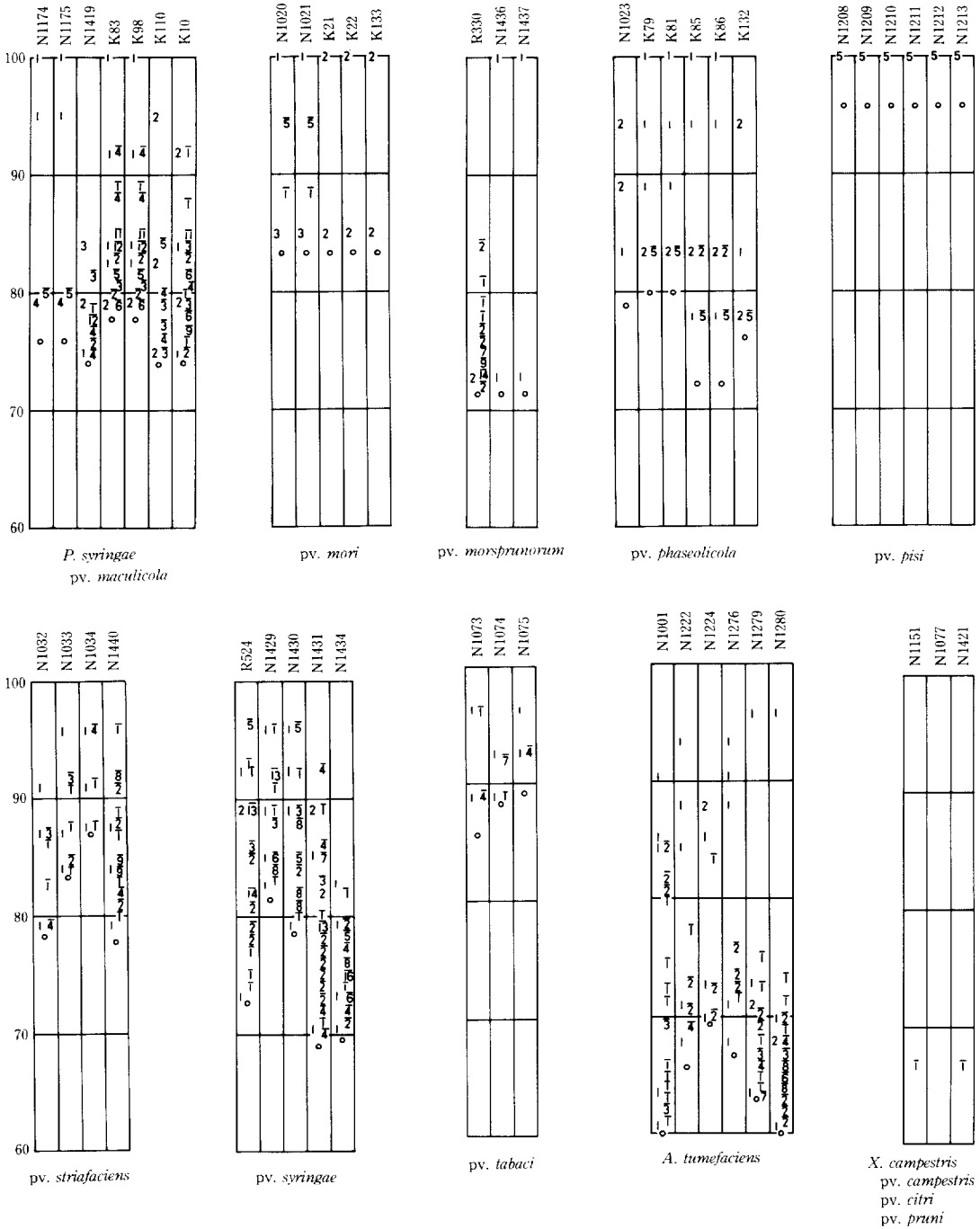


図-7 陽性一致率でみた各種 (pathovar) の細菌学的性質の変異幅 (その2)



図一8 陽性一致率でみた各種(pathovar)の細菌学的性質の変異幅 (その3)

あった。*P. andropogonis* と *P. woodsii* はそれぞれ互いにノイズとしてリストアップされた。*P. marginalis* pv. *marginalis* と *P. solanacearum* それぞれの中で一致率が低い菌株も見られた。*P. viridiflava* では *P. syringae* 群菌のノイズが入った。

図-4～5に見られるように、*P. syringae* 群菌のうち、pv. *coronafaciens*, pv. *eriotrotyae*, pv. *glycinea*, pv. *japonica*, pv. *maculicola*, pv. *mori*, pv. *morsprunorum*, pv. *phaseolicola*, pv. *pisi* では pathovar まで検索可能であったが、それ以外ではノイズが多く、pathovar までの検索は不可能な場合が多かった。

Agrobacterium 属菌の場合、一致率が低い菌株もあった。

(2) 陽性一致率による検索：陽性一致率はその算出式から、単純一致率と等しいかそれよりも小さな値しかとらないが、単純一致率の場合と同じ傾向であった。すなわち、単純一致率では変異幅が小さくよくまとまっている種では90～100%、まとまっていない種でも80～100%に分布していた菌が、陽性一致率では図-6～8に見られるように、それぞれ80～100%、60～100%と広い範囲にわたって分布した。このような違いは見られたがリストアップされる菌株の順序は、単純一致率の場合とほとんど変わることなく、単純一致率の場合に言えたことがそのまま陽性一致率の場合にもあてはまった。

V 考 察

図-4～8にみられるように、各細菌学的性質に重みを付けなくても、単純一致率あるいは陽性一致率を求め、ケンサ菌を検索することは一部の細菌については十分可能であった。しかしノイズが多くリストアップされる菌については、すでにその菌群の識別に役立つことが知られている検査項目を選んでもう一度検索しなおしてみる必要があるであろう。

P. andropogonis と *P. woodsii* はそれぞれ互いにノイズとしてリストアップされているが、西山(1979)によれば後者は前者の異名と考えられるので、もし同一種とするならば、それらは互いにノイズではなくなる。

P. viridiflava では *P. syringae* 群菌のノイズが入り、リストアップされた結果のみから正しく判定されうるか疑問であった。しかしそれは BILLING(1970)の *P. viridiflava* は *P. syringae* に近縁であるという結果と一致している。

P. syringae 群菌の場合、pathovar まで検索しようとしてもノイズが多く、検索は不可能な場合が多かった。し

かしノイズの多くは *P. syringae* の他の pathovar であり、pathovar の定義(細菌学的性質はほとんど同じで寄生性が異なるもの)からそれは当然のことであり、分類体系の問題点として議論されるべきではないことは述べるまでもない。pathovar まで識別するためには細菌学的性質だけでは不十分なので、まず細菌学的性質で検索した後にさらに寄生性に重み付けをして検索できるようにすれば、この問題は解決されるであろう。ただ、寄生性がわかればわざわざそのようにするまでもなく、細菌学的性質で検索した段階ですぐに pathovar も判定される。

Agrobacterium 属菌の場合、比較的ノイズが少なかったが、それは *A. tumefaciens* 1種類だけであり、*A. rhizogenes* が含まれていなかったことにもよると考えられる。

Xanthomonas 属菌の場合、各 pathovar 1菌株しか用いなかったので pathovar 内変異幅は明らかにすることはできなかったが、pathovar 間の一致率に関しては普段の同定で感じているよりも低いので、pathovar まで検索できると考えられた。

ただし、上述の検索可能性については、ケンサ菌が植物病原細菌であるという前提があって初めてあてはまることであると考えられる。植物に病原性を示さなくとも、細菌学的性質からはそれに似ている細菌が自然界に多数存在している可能性はきわめて高いからである。

46項目の細菌学的性質をもとに、単純一致率と陽性一致率で細菌を検索することは一部の細菌では可能であったが、分類まで議論するためには、今回の菌種(pathovar)数、菌株数、検査項目数では不十分である。分類体系は、今後さらにそれらを増やして、より安定した結果が得られるようにしたうえで議論されねばならない。

以上のような一つの表として表された細菌学的性質の情報は、細菌の分類や同定にあたって、あるいはある特定の形質をもった菌株を探すのにきわめて便利ではあるが、次のような問題が残されている。

(1) データ集積にあたり、まず検査方法、検査期間、判定基準を統一することが必要である。細菌学的性質の調査は従来の方法に従う限りかなりの時間と労力が必要とされ、このような大きな表は研究者一人で埋められるものではない。また、多くの研究者のデータが集積されたときに初めて、大きな価値を持つのである。そのために他の研究者のデータを入力するのは当然であるが、その際に研究者あるいは研究室間で検査方法が異なるという問題がある。これは最も重要な問題であると考えられ

る。植物病原細菌の同定に用いられる検査項目の場合、検査方法がわずかに異なっても同じ結果が得られることが多く、検索には大きな影響を及ぼさないと考えられる。しかし、個々の検査項目についてみるならば、たとえば糖からの酸の産生や有機酸の利用能をみるときに加える糖や有機酸の濃度によって異なった結果が得られる場合がある。あるいは糖からの酸の産生をみるのか利用(生育)をみるのかは研究者によって異なり、それらのデータを比較できない場合がしばしばある。またラクトースからの酸の産生やアルギニンジヒドロラーゼ活性のように、1週間後の調査では陰性であっても2週間後には陽性を示すものがあり、調査期間によって判定結果が異なるものもある。さらにオキシダーゼ活性の検査では、菌株によって判定が非常に困難なものがあり、わずかな実験条件の違いによっても大きく左右され、またその判定は研究者間で異なることになる。このように異なった検査方法、検査期間、判定基準のデータを一つの表にしたところで、個々の性質の比較や菌株の検索は正確にはなされない。したがって、まずこれらを統一するのか、あるいは検査方法などが違う場合はどの程度の違いまでは許容されて同じ検査と見なされるのか、あるいは当面は方法が違う検査結果は別な検査とみなしてしまうのか、等について十分検討されなければならない。

検査項目や検査方法については、Bergey's manual of systematic bacteriology (KRIEG編, 1984), Laboratory guide for plant pathogenic bacteria (SCHAAD編, 1980), 医学細菌同定の手引き (COWAN, 1974), Abstracts of microbiological methods (SKERMAN, 1969) あるいは西山(1978)や後藤ら(1984a-d)を参考にして検討するのが妥当であろう。

なお細菌の分類と同定には様々な性質が調べられているべきであるが、検査項目や検査方法を統一することは同時にそれ以外の性質がおざりにされる危険性をはらんでいる。それらが統一されたとしても、それは学問の進歩によって新たな性質が分類学上重要であることがわかったとき、その性質を採用できるような柔軟性をもったものでなければならない。また、あまりに厳格なものであると入力できるデータが非常に狭められてしまう、ということも考慮されていなければならない。

(2) 調査結果のコンピュータへの入力、ここでは未調査0、陰性1、弱陽性2、陽性3として行ったが、それだけでは同定のための野帳の内容が十分に表記されない。何度行っても陽性か陰性が判定できない性質がある。そのような性質であっても、ある特定の細菌群ではは

きりした結果を与え、同定上有効である場合が多く、除くことはできない。そこで陽性とも陰性とも判定できないが未調査ではないことを表せるようにしておく必要がある。また弱陽性2は最終調査で弱陽性となったことを表すものであるが、それは中間調査では陰性1の場合もありうることを示唆している。しかし陽性3の場合にはそのような情報を与えない。中間調査結果を用いて早期に検索したい場合には、遅れて(たとえば2週間後)陽性となるような性質では、本来一致するはずのものが一致しない。

これらの場合を表すために、次のように区分するのが妥当であると考えられる。0:未調査, 1:陰性, 2:陰性とも陽性とも判定できない(この場合0と同じく2も除外して検索), 3:弱陽性, 4:中間調査で陰性, 判定不可または弱陽性, 最終調査で陽性, 5:すみやかに陽性となった。

(3) アメリカの National Institute of Health では細菌や糸状菌などについて現在約9,000項目のデータが集積できるプログラムが作られている(ROGOSAら, 1971)。しかし細菌のみに関しても、検査項目や検査方法は各分野ごとにあるいは属で異なっており、そこにデータを集積するならばほとんどが埋められないまま空白として残されるであろう。他分野の細菌とも比較できるようにそれらが共通化できれば好ましいが、たとえば生育速度、生育温度、栄養要求性、好・嫌気性など根本的に異なっていることがあり、それは非常に困難な状況にある。はたしてそれが可能であるのか、あるいは本当に好ましいことなのかどうかは難しい問題である。当面は植物病原細菌の分野で属ごとにファイルを作り、あまりに異なる細菌は一つのファイルに含めない方がよいのではないかと考えられる。

おわりに、今回作成されたプログラムによって出力されるものは同定結果ではないことを指摘しておきたい。細菌の分類は単に性質の一致率にもとづいているのではないので、コンピュータの出力結果は単なる参考値あるいは類似細菌のら列にすぎない。それから種あるいは pathovar を判断するのは人間である。また普段から植物病原細菌を同定している研究者は検査菌がどの種に“近縁であるか”は比較的容易に見当がつくので、このプログラムはその後さらに似ている菌株がないかを検索したり、特定の性質を持った菌株を検索したり、あるいはデータ集積をしたりするためのものである。

なおこのデータとプログラムは、農林水産研究計算センターにアクセスできる端末のあるところでは利用可能

である。

VI 引用文献

- 1) AYERS, S. H., P. RUPP and W. T. JOHNSON (1919) : A study of the alkali-forming bacteria in milk. *Bull. U. S. Dep. Agr.*, **782**, 1-38
- 2) 畔上耕児ら(1984a) : コンピューター利用による植物病原細菌のデータ集積と分類同定 I 理論と方法. *日植病報*, **50**, 84 (講要).
- 3) ———(1984b) : 同上 II 実施例. 同上, **50**, 84 (講要).
- 4) BILLING, E. (1970) : *Pseudomonas viridiflava* (BURKHOLDER, 1930; CLARA, 1934). *J. appl. Bact.*, **33**, 492-500
- 5) COLWELL, R. R. et al. (1968) : Computer analysis of relationships among phytopathogenic bacteria. *Phytopathology*, **58**, 1207-1215
- 6) COWAN, S. T. (1965) : Principles and practice of bacterial taxonomy—a forward look. *J. gen. Microbiol.*, **39**, 143-153
- 7) ———(1968) (駒形和男, 杉山純多訳) : 微生物分類用語事典, p.91, 118-119, 学会出版センター, 東京
- 8) ———(1970) : Heretical taxonomy for bacteriologists. *J. gen. Microbiol.*, **61**, 145-154
- 9) ———(1971) : Sense and nonsense in bacterial taxonomy. *Ibid.*, **67**, 1-8
- 10) ———(1974) (坂崎利一訳) : 医学細菌同定の手引き, 2版, p.1-335, 近代出版, 東京
- 11) DYE, D. W. (1981) : A Numerical taxonomic study of the genus *Erwinia*. *N. Z. J. Agric. Res.*, **24**, 223-229
- 12) 後藤正夫(1980) : 新植物細菌病学, p.27-37, ソフトサイエンス社, 東京
- 13) ———・龍川雄一(1984a) : 植物病原細菌同定のための細菌学的性質の調べかた(1). *植物防疫*, **38**, 339-344
- 14) ———・———(1984b) : 同上(2), 同上, **38**, 385-389
- 15) ———・———(1984c) : 同上(3), 同上, **38**, 432-437
- 16) ———・———(1984d) : 同上(4), 同上, **38**, 479-484
- 17) 金子太吉(1982) : 微生物の化学分類実験法 (駒形和男編), p.309-346, 学会出版センター, 東京
- 18) KRIEG, N. R. ed.(1984) : *Bergey's manual of systematic bacteriology*. Vol. 1, p. 1-964, Williams & Wilkins, Baltimore
- 19) LAPAGE, S. P. et al. eds. (1975) : *International code of nomenclature of bacteria*. 1975 rev., American society for microbiology, Washington, D. C.
- 20) 西山幸司(1978) : 植物病原細菌簡易同定法の試案. *植物防疫*, **32**, 283-288
- 21) ———ら(1979) : *Pseudomonas andropogonis* によるチューリップ黒腐病. *日植病報*, **45**, 668-674
- 22) ROGOSA, M., M. I. KRICHEVSKY and R. R. COLWELL (1971) : Method for coding data on microbial strains for computers (ed. AB). *Int. J. Syst. Bacteriol.*, **21**, 1A-175A
- 23) 坂野勲(1974) : 微生物の新しい分類学 (長谷川武治編), p.51-105, 講談社サイエンティフィック, 東京
- 24) SANDS D. C., M. N. SCHROTH and D. C. HILDEBRAND (1970) : Taxonomy of phytopathogenic pseudomonads. *J. Bacteriol.*, **101**, 9-23
- 25) SCHAAD, N. W. ed. (1980) : *Laboratory guide for identification of plant pathogenic bacteria*. p.1-72, Bacteriology Committee of American Phytopathological Society, Minnesota
- 26) SKERMAN, V. B. D. ed. (1969) : *Abstracts of Microbiological Methods*. p.1-883. John Wiley & Sons, New York
- 27) SNEATH, P. H. A. (1957a) : Some thoughts on bacterial classification. *J. gen. Microbiol.* **17**, 184-200
- 28) ———(1957b) : The application of computers to taxonomy. *Ibid.*, **17**, 201-226
- 29) ———and R. R. SOKAL (1973) : *Numerical Taxonomy - The principles and practice of numerical classification*. p.114-187, Freeman, San Francisco
- 30) 鈴木茂(1982) : 応用統計ハンドブック (応用統計ハンドブック編集委員会編), p.404-416, 養賢堂, 東京
- 31) 矢島敬二(1976) : 統多変量解析法 (奥野忠一ら著), p.207-237, 日科技連, 東京
- 32) 柳田友道(1980) : 微生物科学 1 分類・代謝・細胞生理, p.56, 学会出版センター, 東京
- 33) 吉沢正(1981) : 多変量解析法 (奥野忠一ら著), 改定版, p.391-411, 日科技連, 東京

Spectra of Plant Pathogenic Bacterial Species (Pathovars) and Storage and Retrieval of the Data by Computer

Koji AZEGAMI*, Koushi NISHIYAMA*, Yasumasa WATANABE* and Tokuji FUKUDA**

Summary

One hundred and forty four reference isolates of plant pathogenic bacteria, i.e. 26 species (pathovars) of *Pseudomonas* (135 isolates), *Agrobacterium tumefaciens* (6) and 3 pathovars of *Xanthomonas campestris* (3), were examined for 46 bacteriological characteristics and the data were stored in a computer. The program for retrieval of the data was written in FORTRAN, and the similarity of every pair of isolates in all possible combinations was computed. The similarities between the isolates were measured by simple and positive-only matching coefficients. Then, for every isolate, all the reference isolates having high similarities with it were listed in order by the retrieval function. The spectrum of each species (pathovar) including various isolates was represented in a graph based on the matching coefficients.

As for the pseudomonads which do not belong to the *P. syringae* group, with the exception of *P. marginalis* pv. *marginalis*, *P. solanacearum* and *P. viridiflava*, each species was so well categorized that the variation in the bacteriological characteristics of each species and the "noise (the isolate which is listed by the computer-aided retrieval because of the high matching coefficient although it has been conventionally identified as another species and is not expected to be listed)" were small enough to enable the retrieval based on the characteristics. All the noises accompanying *P. andropogonis* consisted of *P. woodsii*, and vice versa. Most of the noises accompanying *P. viridiflava* were pathovars of *P. syringae*. These results agree with those of NISHIYAMA et al. (1979) and BILLING (1970) respectively.

As for most of the pathovars of *P. syringae*, the noises were too large to enable the retrieval of pathovars. Most of the noises accompanying a pathovar corresponded to another pathovar of *P. syringae*, indicating that the pathogenicity must be weighted for successful retrieval.

Simple and positive-only matching coefficients between the isolates of the relatively homogeneous species ranged between 90–100% and 80–100% respectively, while those of the heterogeneous species ranged between 80–100% and 60–100% respectively.

* National Institute of Agro-Environmental Sciences, Yatabe, Tsukuba, Ibaraki, 305 Japan.

** Tropical Agriculture Research Center. *Ibid.*