A Frequency Matrix for Probabilistic Identification of Some Bacilli

By FERGUS G. PRIEST* AND BRIAN ALEXANDER

Department of Brewing and Biological Sciences, Heriot-Watt University, Chambers Street, Edinburgh EH1 1HX, UK

(Received 13 May 1988; revised 29 July 1988)

A matrix comprising frequencies for positive results for 44 Bacillus taxa for 30 characters has been constructed. The 44 taxa include most of the common species and several clusters of environmental isolates including those described as B. firmus–B. lentus intermediates. The tests, which were chosen for their high diagnostic value, included some of the traditional tests used for identification of bacilli supplemented with a range of sugar fermentations and other characterization tests. The matrix was evaluated by identifying hypothetical median organisms, cluster representatives and a panel of 23 reference strains. All reference strains achieved Willcox probabilities above 0.995. Fifty-eight environmental isolates were also subjected to the 30 tests and identification was attempted. Forty-one strains (70%) achieved a Willcox probability >0.95, which was considered an acceptable identification, and were assigned to 12 taxa. If the SE of taxonomic distance was also considered in the identification score (an acceptable value being <7.0), the number of acceptable identifications was reduced to 34 (59%). It was encouraging that bacteria from garden soils identified to the common species such as B. subtilis, B. cereus and B. licheniformis whereas some of the bacteria from an estuarine habitat were identified as species such as B. firmus which are normally identified with that habitat.

INTRODUCTION

The aerobic endospore-forming bacteria of the genus Bacillus are important as pathogens of man and animals and have considerable value as sources of industrial enzymes, insect toxins for biological control and antibiotics (Debabov, 1982; Priest, 1988). They are distributed widely in the environment and are physiologically diverse with alkaliphilic, acidophilic, halophilic, thermophilic and psychrophilic species described in the literature (Claus & Berkeley, 1986). It is therefore surprising that little is known of their ecology (Slepecky & Leadbetter, 1983; Goodfellow & Dickinson, 1985); largely due to the difficulty of identifying environmental isolates.

The traditional identification schemes rely on the use of spore morphology to assign a Bacillus isolate to one of three groups based on spore shape and relative size in the sporangium. Physiological tests are then used in dichotomous keys or polythetic tables to identify the organism to the species level (Gordon et al., 1973; Hobbs & Cross, 1983). This approach has several drawbacks; in particular it is often difficult to decide spore morphology unequivocally and the monothetic nature of the keys leads to mis-identification. Moreover, the number of Bacillus species has increased dramatically over recent years and the traditional identification schemes cannot cope with the greater diversity of types.

In an attempt to improve identification of bacilli, Logan & Berkeley (1984) adopted the API 50CHB trays. A comprehensive database was established using these miniaturized kits supplemented with some traditional characters and a computerized identification scheme was developed (Bryant et al., 1985). However, the system is expensive for routine use.

Abbreviation: HMOs, hypothetical median organisms.
Frequency matrices contain results in the form of percentage positive reactions for the most diagnostic characters of the taxa in question. This information can be used for probabilistic identification of unknown strains (Willcox et al., 1980). This approach has been very successful for various Gram-negative bacteria (Holmes et al., 1986a, b; Dawson & Sneath, 1985) and has been increasingly used for Gram-positive organisms such as slow-growing mycobacteria (Wayne et al., 1980), Streptomyces (Williams et al., 1983), Streptovercillium (Williams et al., 1985) and Micrococcus species (Feltham & Sneath, 1982). In view of the need for a reliable, cost-effective and simple scheme for the identification of bacilli, we have prepared and tested a frequency matrix derived from a recent numerical classification of the genus (Priest et al., 1988).

METHODS

Organisms used. Reference strains of Bacillus (Table 2) were obtained from the German Collection of Micro-organisms, Braunschweig, FRG. Environmental isolates (Table 3) were obtained from samples of garden soil (EL strains) and estuarine muds (T strains) by plating heat-treated (80 °C for 10 min) suspensions in sterile physiological saline onto nutrient agar and incubating at 30 °C for 1 to 5 d. Colonies of various morphology were selected, taking care to avoid picking more than one type from any individual sample. Isolates were stored at −20 °C in 20% (v/v) glycerol.

Selection of taxa and tests. The clusters were based on those described by Priest et al. (1988). Forty-four of the 49 clusters were included. The test methods have been described in detail (Priest et al., 1988).

Construction and theoretical evaluation of the matrix. To assist the selection of the most diagnostic tests for the matrix, the CHARSEP program (Sneath, 1979b) was used initially. This calculates five separation indices for each test (see Williams et al., 1983). The DIACHAR program (Sneath, 1980a) was used to determine the most characteristic features of clusters. Having made an initial choice of tests, the OVERMAT program (Sneath, 1980c) was used to calculate overlap between taxa in the matrix and the matrix was further tested by calculation and subsequent identification of hypothetical median organisms using MOSTTYP (Sneath, 1980b).

Determination of identification scores. Strains were identified from the matrix using the MATIDEN program (Sneath, 1979a) which calculates five identification indices of which we routinely used Willcox probabilities (Willcox et al., 1973) and standard error (SE) of taxonomic distance (Sneath, 1978). Unless otherwise indicated, all test results were determined in this study.

RESULTS

Selection of characters

When the database, comprising result frequencies for 118 tests for 49 Bacillus clusters (see Priest et al., 1988) was examined using the CHARSEP program, 70 tests conformed to the criteria of CSP index >50% and VSP index >25% as suggested for acceptable diagnostic characters by Sneath (1979b). The DIACHAR program, which ranks the most diagnostic characters for taxa, was used to indicate the most characteristic features for all 49 taxa recovered in the original classification. Visual comparison of the CHARSEP and DIACHAR outputs was sufficient to enable a choice of 55 from the 70 ‘acceptable’ characters, to produce a 49 (taxa) × 55 (tests) matrix. Since carrying out 55 tests is expensive and time-consuming, a second smaller matrix comprising a subset of thirty characters (Table 1) was devised. This comprised 25 tests which had VSP scores of 34 to 87% (mean 65%) and which required some form of inoculation and five cellular morphology features that were included, not because of their diagnostic value (VSP scores 2 to 83%), but because this information will probably have been used to allocate the unknown organism to the genus Bacillus. The output from the DIACHAR program for this matrix indicated sums of diagnostic scores for taxa ranging from 12:34 to 18:46.

Theoretical evaluation of the matrices

Cluster overlap in the two matrices was estimated using OVERMAT. In the 49 × 55 matrix at the selected critical overlap level (Vp) of 5%, there was no significant overlap between any pair of taxa. In four cases, standard deviations (sds) of taxa were greater than 0:3, indicative of heterogeneity. These were unidentified clusters (taxa 10 and 17), B. psychrosaccharolyticus (taxon 12) and B. macquariensis (taxon 45) which contained three, two, two and two members respectively and had sds of 0:32, 0:39, 0:36 and 0:45.
When the 49 × 30 matrix was evaluated with OVERMAT, several pairs of taxa displayed significant overlap; these were almost exclusively restricted to the taxa with unacceptable SDs noted above. These four taxa again showed high SDs as did B. subtilis var. niger (taxon 16, three members, SD = 0·315) and the five taxa were deleted from the matrix until more strains and better diagnostic features become available. Taxon 27 also had high SD (six members, SD = 0·316) but was retained because it represented a major taxon in the original classification. In the resultant 44 × 30 matrix (Table 1), significant overlap was calculated between two pairs of taxa: B. licheniformis with taxon 21 (VE = 2·9%) and taxon 42 (B. fusiformis) with B. sphaericus (VE = 2·6%). It was considered that this would be acceptable for general usage.

The MOSTTYP program calculates hypothetical median organisms (HMOs) for each taxon in an identification matrix and then attempts to identify them. The results from this analysis were satisfactory with the larger matrix obtaining Willcox probabilities of 0·999 or 1·0 coupled with low SEs of taxonomic distance for all HMOs. The smaller matrix did not perform so well, but the Willcox probability was below 0·999 only for B. amyloliquefaciens (0·9987) and the HMO of taxon 27 (0·9989). The range for SE of taxonomic distance for this matrix was −1·0 to −6·9. The smaller matrix was therefore considered satisfactory for most purposes and we now use it routinely for identification of bacilli.

The final evaluation of the 44 × 30 matrix involved identifying representative organisms using the data that had been generated in the numerical classification study and on which the identification matrix was based. Most strains identified with a high Willcox probability (>0·999) and low SE of taxonomic distance (<3·5). Strains representing B. azotoformans, B. brevis, B. laterosporus, B. subtilis and taxon 27 achieved Willcox probabilities of 0·996 to 0·998. Equally importantly, with the exception of 'B. subtilis var. niger' strain S225 which was identified as B. licheniformis, single member clusters from the numerical classification which had not been included in the matrix failed to identify. Finally, some strains were examined with and without spore morphology data. The omission of these data had virtually no effect on the identification scores. For example, B. amyloliquefaciens, B. sphaericus and B. subtilis retained scores of 0·999; the largest reduction, from 0·999 to 0·997 was for B. licheniformis.

Practical evaluation of the selected matrix

We examined 22 reference strains including type and neotype cultures where possible, for the 30 tests. All these reference strains identified to their taxon with high Willcox probabilities and low SE of taxonomic distance (Table 2). Phenotypically, B. cereus, B. mycoides and B. thuringiensis are indistinguishable (Priest et al., 1988). Therefore the matrix does not contain separate entries for these bacteria and consequently B. mycoides and B. thuringiensis strains identified to B. cereus.

Bacillus strains were isolated from garden soils (EL strains) and estuarine muds (T strains) (see Methods). These strains were examined using the 30 tests listed in Table 1 and identified using MATIDEN. Of the 58 strains, 25 scored a Willcox probability >0·99 and 16 achieved a score of 0·95 to 0·99. Three strains scored 0·90 to 0·949 and 14 obtained low scores below 0·90. It therefore seemed reasonable to accept an identification as a Willcox probability >0·95 and using this criterion alone, 41 (70%) of the strains were identified.

If the SE of taxonomic distance is also considered in the identification, the results are not so good. A low SE of taxonomic distance indicates that the organism lies close to the centroid of the group and an acceptable score is below 3 (Sneath, 1979a). Only 11 strains achieved this. However, if the requirement is relaxed such that a Willcox probability of >0·95 coupled with a SE of taxonomic distance of <7·0 is accepted as an identification, then the results are more promising and 34 of the 58 strains (59%) were identified to 12 different taxa. Since many identification systems use the Willcox criterion alone, we do not consider that we are being over-lenient in setting these limits.

Discussion

Traditional schemes for the identification of aerobic, spore-forming bacteria have suffered the problems typically associated with monothetic systems, namely inability to proceed past a
Table 1. A frequency matrix for the identification of bacilli

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>. . .</th>
<th>1 2 3 4 5 6 7 8 9 11 13 14 15 18 19 20 21 22 23 24</th>
</tr>
</thead>
</table>

**Cell morphology:**
1. Length > 3 μm
2. Gram-positive
3. Spores oval
4. Spores central
5. Spores bulging
6. Aesculin
7. Casein
8. Hippurate
9. Starch
10. Urea
11. Chloramphenicol (4)
12. Nalidixic acid (32)
13. Polymyxin (16)
14. Streptomycin (8)
15. Cellobiose
16. Fructose
17. Galactose
18. Lactose
19. Mannose
20. Raffinose
21. Salicin
22. Xylose
23. Citrate
24. Succinate

**Resistance to (μg ml⁻¹):**
11. Chloramphenicol (4)
12. Nalidixic acid (32)
13. Polymyxin (16)
14. Streptomycin (8)

**Acid from:**
15. Cellobiose
16. Fructose
17. Galactose
18. Lactose
19. Mannose
20. Raffinose
21. Salicin
22. Xylose

**Utilization of:**
23. Citrate
24. Succinate

**Miscellaneous:**
25. Growth at 50 °C
26. Growth in 10% NaCl
27. Anaerobic growth
28. Nitrate reduction
29. Oxidase reaction
30. Voges–Proskauer test

particular point in the key because of a missing character and mis-identification due to an organism being aberrant in one or more features. Polythetic tables overcome these shortcomings but, because of the size of the genus *Bacillus* and the paucity of good diagnostic phenotypic data, they have their limitations when applied to this genus and are cumbersome. The database provided here offers *Bacillus* identifications using relatively few and inexpensive tests. The tests were chosen with two criteria in mind. From the diagnostic viewpoint, they should conform to the minimum VSP and CSP scores suggested by Sneath (1979b) and second, they should represent the full range of biological attributes. Thus biochemical, physiological, morphological and antibiotic resistance tests were included. Although the prevalence of plasmids in bacilli might argue against the use of antibiotic resistance tests, resistance to antibiotics such as
### Identification of bacilli

| Character | B. *formis* | Taxon 26 | Taxon 27 | Taxon 28 | Taxon 29 | B. *pantothenicus* | B. *cereus* | B. *simplex* | B. *pileolactis* | B. *cereus* | B. *aerolea* | B. *alvei* | B. *badius* | B. *pyocyaneus* | B. *bacillus* | B. *sphaericus* | B. *lentus* | B. *stearothermophilus* | B. *kaustophilus* |
|-----------|-------------|----------|----------|----------|----------|--------------------|-------------|-------------|----------------|-------------|-----------|---------|-----------|-------------|----------------|---------------|--------------|-------------|-------------------------|-----------------|
| 25        | 1           | 99       | 1        | 99       | 25       | 1                  | 80          | 1           | 75            | 1           | 99        | 27       | 99         | 1            | 50           | 99          | 89           | 1            | 99          | 99       | 1            | 1            |
| 26        | 25          | 99       | 99       | 99       | 50       | 78          | 99          | 50          | 67            | 71          | 20        | 1        | 99        | 99         | 99           | 25           | 33           | 99           | 60          | 67        | 50          | 60          |
| 28        | 88          | 99       | 83       | 99       | 99       | 1          | 99          | 99          | 99            | 57          | 80        | 82       | 1          | 99         | 99           | 1            | 50           | 99           | 80          | 67        | 1           | 1           |
| 29        | 5           | 1        | 17       | 1        | 1        | 99          | 1           | 1           | 99            | 1          | 80        | 99        | 99         | 1            | 99           | 99          | 89           | 1            | 60          | 33       | 99          | 99          |
| 30        | 6           | 1        | 1        | 1        | 99       | 99        | 99         | 99         | 99            | 99         | 50        | 1        | 1          | 99         | 99           | 99           | 33           | 1            | 1           | 1           | 1           | 1           |
| 32        | 8           | 88       | 50       | 1        | 50        | 1          | 99          | 99          | 99            | 33          | 86        | 1         | 91          | 1          | 50          | 99         | 1           | 11           | 99          | 99         | 99       | 99          | 99          |
| 33        | 9           | 99       | 99       | 63       | 50        | 99          | 11          | 40          | 99            | 1          | 1        | 1        | 1          | 1           | 1            | 1           | 1            | 1           | 1           | 1         | 99         | 99        | 99          | 99          |
| 34        | 10          | 1        | 1        | 1        | 1        | 60          | 1           | 1           | 1           | 1           | 1          | 1        | 1          | 1           | 1            | 1           | 1           | 1            | 1           | 1         | 1         | 1           | 1           |

Table 1 (continued)

1. Cluster designations are the same as those in Priest *et al.* (1988). Taxon 9 (previously *B. circulans*) has been named *B. amylolyticus* (Nakamura, 1984). The other unnamed taxa comprise the following strains: taxon 4, *B. circulans*; taxon 18, *B. subtilis*; taxa 21 and 24 to 29, isolates from marine muds and salt marshes; taxon 33, *B. megaterium*; taxon 41, *B. brevis*.

2. *B. cereus* includes *B. mycoides* and *B. thuringiensis*.

3. *B. badius* includes *B. freudenreichii*.

4. For details of *B. smithii* see Nakamura *et al.* (1988).

* Chloramphenicol is not commonly plasmid-borne in *Bacillus*, (Gryczan, 1982). For example, in bacilli chloramphenicol acetyltransferase and streptomycin resistance can be chromosomally encoded (Williams *et al.*, 1981; Yousten *et al.*, 1985) and offer useful identification criteria for *B. pumilus* and *B. sphaericus* respectively as well as other taxa. The morphological tests are
Table 2. Identification scores for reference strains using the 44 taxon \times 30 character matrix and characters determined in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>DSM no.*</th>
<th>Identity (cluster no.)</th>
<th>Willcox probability</th>
<th>SE of taxonomic distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. alvei</td>
<td>7</td>
<td>B. alvei (1)</td>
<td>0.999</td>
<td>2.784</td>
</tr>
<tr>
<td>B. amyloliquefaciens</td>
<td>29</td>
<td>B. amyloliquefaciens (14)</td>
<td>1.000</td>
<td>3.038</td>
</tr>
<tr>
<td>B. badius</td>
<td>23</td>
<td>B. badius (39)</td>
<td>0.999</td>
<td>2.324</td>
</tr>
<tr>
<td>B. brevis</td>
<td>30</td>
<td>B. brevis (37)</td>
<td>1.000</td>
<td>-0.219</td>
</tr>
<tr>
<td>B. cereus</td>
<td>31</td>
<td>B. cereus (11)</td>
<td>0.999</td>
<td>2.351</td>
</tr>
<tr>
<td>B. circulans</td>
<td>11</td>
<td>B. circulans (6)</td>
<td>0.995</td>
<td>6.763</td>
</tr>
<tr>
<td>B. coagulans</td>
<td>1</td>
<td>B. coagulans (47)</td>
<td>0.999</td>
<td>-0.228</td>
</tr>
<tr>
<td>B. firmus</td>
<td>12</td>
<td>B. firmus (25)</td>
<td>0.999</td>
<td>0.704</td>
</tr>
<tr>
<td>B. laterosporus</td>
<td>25</td>
<td>B. laterosporus (13)</td>
<td>0.997</td>
<td>1.800</td>
</tr>
<tr>
<td>B. lentus</td>
<td>9</td>
<td>B. lentus (44)</td>
<td>0.999</td>
<td>3.577</td>
</tr>
<tr>
<td>B. licheniformis</td>
<td>13</td>
<td>B. licheniformis (20)</td>
<td>0.999</td>
<td>0.683</td>
</tr>
<tr>
<td>B. macerans</td>
<td>24</td>
<td>B. macerans (5)</td>
<td>0.999</td>
<td>2.293</td>
</tr>
<tr>
<td>B. megaterium</td>
<td>32</td>
<td>B. megaterium (22)</td>
<td>1.000</td>
<td>4.989</td>
</tr>
<tr>
<td>B. mycoides</td>
<td>2048</td>
<td>B. cereus (11)</td>
<td>0.999</td>
<td>4.631</td>
</tr>
<tr>
<td>B. pantothenticus</td>
<td>26</td>
<td>B. pantothenticus (30)</td>
<td>1.000</td>
<td>6.963</td>
</tr>
<tr>
<td>B. polymyxa</td>
<td>36</td>
<td>B. polymyxa (8)</td>
<td>0.999</td>
<td>3.563</td>
</tr>
<tr>
<td>B. pumilus</td>
<td>27</td>
<td>B. pumilus (19)</td>
<td>1.000</td>
<td>2.235</td>
</tr>
<tr>
<td>B. psychrophilus</td>
<td>3</td>
<td>B. psychrophilus (40)</td>
<td>0.998</td>
<td>5.126</td>
</tr>
<tr>
<td>B. sphaericus</td>
<td>28</td>
<td>B. sphaericus (43)</td>
<td>0.999</td>
<td>2.993</td>
</tr>
<tr>
<td>B. steaothermophilus</td>
<td>22</td>
<td>B. steaothermophilus (48)</td>
<td>0.999</td>
<td>0.606</td>
</tr>
<tr>
<td>B. subtilis</td>
<td>10</td>
<td>B. subtilis (15)</td>
<td>0.999</td>
<td>2.625</td>
</tr>
<tr>
<td>B. thuringiensis</td>
<td>2046</td>
<td>B. cereus (11)</td>
<td>0.999</td>
<td>1.997</td>
</tr>
</tbody>
</table>

* German Collection of Micro-organisms.

Table 3. Summary of identification scores for the environmental isolates

<table>
<thead>
<tr>
<th>Identification</th>
<th>No. of isolates identified*</th>
<th>Willcox probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&gt;0.995</td>
</tr>
<tr>
<td>B. badius</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>'B. carotarum'</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>B. cereus</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>B. licheniformis</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>B. megaterium</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>B. polymyxa</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>B. pumilus</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>B. sphaericus</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B. fusiformis</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B. subtilis</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Taxon 27</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Taxon 29</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>34</strong></td>
<td><strong>23</strong></td>
</tr>
</tbody>
</table>

* An identification comprises a Willcox probability >0.95 and SE of taxonomic distance <7.0. Twenty-four of the 58 strains failed to reach these criteria.

supplementary since several did not achieve acceptable VSP and CSP levels and were not essential for successful identification. They were included simply because when handling an unknown Bacillus, these data are usually available from the initial examination of the strain. The number of tests was chosen as the minimum consistent with acceptable overlap figures. Little significant overlap was calculated using 5% as the critical value which is similar to the successful matrices for Streptomyces (Williams et al., 1983) and streptoverticillia (Williams et al., 1985). Moreover, the MOSTTYP program indicated efficient functioning of the matrix despite the relatively low number of tests.
The identification rate for reference strains was acceptable indicating that reliable reproducible tests were incorporated in the matrix. Identification of environmental isolates was 58%. This compares with 50% for field strains of 'coryneform' bacteria (Hill et al., 1978), 47% for slow-growing mycobacteria (Wayne et al., 1980) and 81.3% for Streptomyces (Williams et al., 1983). Although the figure for successful identifications of bacilli was lower than that for the streptomycetes, we adopted a more stringent Willcox probability than Williams et al. (0.95 against 0.85). If Williams et al. (1983) had used >0.995, the identification rate would have been reduced to 42%. This point regarding an 'acceptable identification score' is important in probabilistic identification since various levels of Willcox probability have been adopted in different studies. Generally, higher levels (>0.999) can be adopted for well-characterized taxa in which species are homogeneous such as the Enterobacteriaceae, but where species are more diverse and show less uniformity of test results a lower level is appropriate. Thus >0.85 was considered an acceptable identification for streptomycetes (Williams et al., 1983). We have approached this problem by setting an intermediate Willcox probability level (>0.95) and also a definite, but somewhat relaxed score for SE of taxonomic distance (<7.0). By ignoring the SE of taxonomic distance a higher success rate could be obtained (70%) but this was probably achieved at the expense of some mis-identifications. It is well-known that Willcox probabilities can sometimes be misleadingly high for an unrelated taxon when the correct taxon for the unknown is not represented in the matrix (Willcox et al., 1980). The SE of taxonomic distance does not have this problem and so by combining the two criteria a more reliable identification should be obtained.

It is difficult to assess the accuracy of the identifications since most of the environmental isolates could not be identified using traditional schemes. However, it is encouraging that 'common' bacteria such as B. licheniformis, B. pumilus and B. cereus were isolated from garden soils and halotolerant organisms related to B. firmus and B. badius, and including the various B. firmus–B. lentus 'intermediates' originally isolated from salt marsh soils (Gordon et al., 1977), were recovered from the estuarine mud samples. This indicates that the system is providing similar identifications to the traditional schemes and should be useful in industrial, medical and ecological applications.

We are grateful to Dr Colin Moore for assistance with computing. Part of this work was supported by a grant from Gesellschaft für Biotechnologische Forschung for research of relevance to the German Collection of Microorganisms.

REFERENCES


BASIC program for character separation indices from an identification matrix of percent positive characters. Computers and Geosciences 5, 349–357.

BASIC program for the most diagnostic properties of groups from an identification matrix of percent positive characters. Computers and Geosciences 6, 21–26.

BASIC program for determining the best identification scores possible for the most typical example when compared with an identification matrix of percent positive characters. Computers and Geosciences 6, 27–34.

BASIC program for determining overlap between groups in an identification matrix of percent positive characters. Computers and Geosciences 6, 267–278.


