Jour. Microbiol. June 1995 p. 95-102 Vol. 33, No. 2

Construction of Probability Identification Matrix and Selective Medium for Acidophilic Actinomycetes Using Numerical Classification Data

Chi Nam Seong*, Seok Kyu Park[1], Michael Goodfellow[2], Seung Bum Kim[3], and Yung Chil Hah[3]

Department of Biology, College of Natural Sciences, Sunchon National University, Chonnam 540-742, Korea, [1]Department of Food and Nutrition, College of Natural Sciences, Sunchon National University, Chonnam 540-742, Korea, [2]The Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, U.K., and [3]Department of Microbiology and Research Center for Molecular Microbiology, Seoul National University, Seoul 151-742, Korea

(Received April 20, 1995/Accepted June 5, 1995)

A probability identification matrix of acidophilic *Streptomyces* was constructed. The phenetic data of the strains were derived from numerical classification described by Seong *et al.* The minimum number of diagnostic characters was determined using computer programs for calculation of different separation indices. The resulting matrix consisted of 25 clusters versus 53 characters. Theoretical evaluation of this matrix was achieved by estimating the cluster overlap and the identification scores for the Hypothetical Median Organisms (HMO) and for the representatives of each cluster. Cluster overlap was found to be relatively small. Identification scores for the HMO and the randomly selected representatives of each cluster were satisfactory. The matrix was assessed practically by applying the matrix to the identification of unknown isolates. Of the unknown isolates, 71.9% were clearly identified to one of eight clusters. The numerical classification data was also used to design a selective isolation medium for antibiotic-producing organisms. Four chemical substances including 2 antibiotics were determined by the DIACHAR program as diagnostic for the isolation of target organisms which have antimicrobial activity against *Micrococcus luteus*. It was possible to detect the increased rate of selective isolation on the synthesized medium. The results show that the numerical phenetic data can be applied to a variety of purposes, such as construction of identification matrix and selective isolation medium for acidophilic actinomycetes.

Key words: probability identification matrix, acidophilic *Streptomyces*, HMO, diagnostic, cluster overlap, identification score, selective medium

Conventional numerical taxonomy has been one of the most effective modern methods used to establish relationships between actinomycetes at both sub-generic and generic levels (2). Classifications derived using numerical taxonomic methods have high information contents and are polythetic, that is they are based on a complete set of recorded characters rather than on the presence or absence of single characters.

The introduction of numerical taxonomy has led to improvements in the classification of many actinomycete taxa (2). Numerical classification also provide an ideal basis for the construction of identification matrices (22). Such matrices ideally contain the minimum number of characte rs, selected from the data base, necessary to discriminate between the relevant taxa.

To date, a few probabilistic identification matrices have been designed for the identification of actinomycetes. One of the examples of workable, polythetic identification system is that devised by Williams *et al.* (27) for the identification of *Streptomyces* isolates. Using this identification matrix, over 81% of the unknown isolates were identified to cluster groups. Although it was theoretically and practically sound, minor clusters and single-membered clusters were excluded in the matrix. Therefore, separate matrices were constructed for major and minor

* To whom correspondence should be addressed.

clusters, and a diagnostic table for single-membered clusters were also produced by Langham *et al.* (9). The data in those matrices was derived from previous numerical classification of *Streptomyces* species (26). A more comprehensive numerical survey of *Streptomyces* species was carried out by Kämpfer *et al.* (6), and the results of this numerical study were used to construct a single probability matrix for members of all major and minor clusters (7).

Moreover, diagnostic characters derived from numerical classification data are useful for designing a selective medium for a specific taxon (4, 23). One of the approaches to the search for new metabolites is concerned with the rare actinomycetes (5, 12).

Acidophilic *Streptomyces* are widely distributed in a variety of soils and play an important role in turnover of organic matter. A number of acidophilic strains are known to produce important metabolites such as antifungal substances and exoenzymes (3, 5). Comprehensive numerical classification of those organisms was carried out by Lonsdale (11). In this work, representative strains of each cluster were confirmed with the analysis of chemical compounds of the cell such as fatty acids, quinones and amino acids.

For the 166 acidophilic strains, including 64 representatives from previous works (11), a wider range of phenetic tests were made by the authors (13, 14). All of the isolates produced an extensively branched substrate and aerial mycelia with L,L-diaminopimelic acid, and were supposed to belong to genus *Streptomyces*. The strains were recovered as 14 major and 21 minor clusters and 35 single member clusters at 86.5% SSM similarity level. Each cluster was filled with strains originated from single soil sources.

A workable identification system of acidophilic *Streptomyces* is, therefore, required in a number of fields such as taxonomy, ecology, and biotechnology (1, 12). The primary aim of our study is to construct a probability identification matrix of acidophilic *Streptomyces* using numerical classification data (14). The numerical phenetic data was assessed for its application potential in designing a selective medium to the isolated antibiotic-producing acidophilic strains.

Materials and Methods

Construction of probability matrix

Details of the origins, cultivation, and preservation of strains, together with the definition and composition of clusters, were described earlier (13, 14).

Selection of clusters: The clusters used in the construction of the identification matrix were selected from the original classification. Of the 35 clusters, 10 minor ones were deleted because of much overlap with each other (13). Twenty clusters included in this matrix contained three or more strains defined at the 86.5% SSM similarity level. Single member strains CN627 and CN 713 were grouped with cluster 16 and 17, respectively, because of their relatedness to parent clusters. Three minor clusters 20, 34 and 35 containing only two strains were also included in the matrix, since they show distinctiveness with low inter-cluster similarities (13, 14). Finally 25 clusters encompassing 110 strains were used in the construction of matrix.

Selection of the most diagnostic characters: To determine the minimum number of characters necessary for discrimination between clusters, selection was made from 229 unit characters used in the construction of numerical classification. As a first step, selection of reduced number of characters was achieved with the CHARSEP program (17), which calculates the usefulness of different characters for separating groups. This program provides several indices for each character. The index which proved to be most useful was the variance separation potential (VSP) derived from Sneath and Johnson (21). The 55 best scoring characters out of the 229 original ones were initially selected. These 55 characters were next assessed by applying MOSTTYP program (19), which determines the theoretically best identification scores for the hypothetical median organism (HMO) of each cluster. The DIACHAR program (18), which ranks the diagnostic scores of each character for each cluster, was used to determine the most diagnostic characters for each particular cluster.

Theoretical evaluation of the identification matrix

The matrix was further evaluated using the OVERMAT program (20) designed to determine overlap between groups in a matrix. For each pair of groups, a disjunction index (W) and a corresponding nominal overlap (V_G) is calculated ranging from 1 for complete overlap to 0 for complete disjunction. The significance of the determined overlap is assessed, using a non-central t-statistic, against a selected critical overlap value (V_0). In this case the chosen critical value was 5%.

The identification scores were calculated using the MATIDEN program (16), which provides several identification scores for known and unknown strains of q taxa and m characters. Percentages in the matrix, with 0 changed to 1 and 100 changed to 99, are converted to proportions P_{ij} for the ith character of taxon J. The character states of an unknown strain (u) are input and compared with each taxon in turn. Three identification coefficients were calculated.

1. Willcox probability (p) (25). This is likelihood of u against taxon J divided by the sum of the likelihoods of u against all q taxa ($L_{uj}/\sum qL_{Uj}$). The nearer the score approaches to 1.0, the better is the fit of an unknown with a taxon in the matrix.

2. Taxon distance (d), which is given by $\sqrt{[\sum(ui-P_{ij})^2/m]}$. This expresses the distance of an unknown from the centroid of the group with which it is compared.

3. The standard error of the taxonomic distance, which is given by the constant (c) in the equation $d=\bar{d}_J+cs_{ij}$, where $\bar{d}_J$ is the mean distance of OUTs of taxon J from the centroid and siJ is the standard deviation of the distance (d). An acceptable score is less than 2.0 to 3.0, and about the half of the members of a taxon will have negative scores, indicating a stronger relationship to the centroid as the average. This program provides identification scores to the best group and the two next best alternatives.

The MOSTTYP program (19) was used to calculate the identification coefficients for the HMO of each taxon in the matrix.

The next stage in the evaluation of the matrix was achieved by the input of the results of randomly selected members of each cluster in the matrix.

Practical evaluation of the identification matrix

This matrix was further evaluated by applying it to the identification of 32 unknown isolates. The characters in the matrix were determined for these isolates, using the methods of Seong *et al.* (14). The test data was assessed against this matrix using the MATIDEN program to determine identification scores.

Formulation of selective isolation medium

An attempt was made to design a selective isolation medium for the strains showing antimicrobial activity against *M. luteus*. With the aid of DIACHAR program, initial selection of the most diagnostic characters was achieved for cluster 35 against all of the clusters defined in the numerical classification was achieved. Workability of the medium was determined by applying this medium to the isolation of antibiotic-producing strains from soil samples. The strains showing antibiosis were identified with the probability matrix constructed in this study. Sample treatment and isolation procedures were described in Seong *et al.* (14).

Results

Construction of identification matrix

The matrix (Table 1) consisted of 25 phena versus 53 characters derived from numerical classification data. The characters covered a wide range of attributes, including morphology, pigmentation, antibiosis, antibiotic resistance, growth tolerances, and utilization of carbon and nitrogen sources. Most characters showed VSP indices < 30%, and thus exceeded the value of 25% recommended by Sneath (17). Some tests with high VSP scores were excluded due to practical difficulties (e.g. utilization of dimethylamine and aesculine, white spore mass color). While one of characters (utilization of succinic acid) had low scores in the DIACHAR program, it was included as it was useful for differentiating particular clusters. Finally, the VSP scores for the tests that were included ranged from 75.73% (utilization of serine) to 11.67% (utilization of succinic acid).

Theoretical evaluation of the identification matrix

A summary of scores obtained from the calculations of the programs CHARSEP, MOSTTYP, DIACHAR, and OVERMAT is given in Table 2. Sums of diagnostic scores determined by DIACHAR program ranged from 16.07 (Cluster 29) to 23.47 (Cluster 18). The OVERMAT program showed only 12 overlaps (from 300 possibles) taking a 5% level and a confidence interval of P=90. The theoretically best identification score for the HMO of each cluster was calculated using MOSTTYP program. Willcox probabilities for all phena exceeded 0.987. Scores for taxonomic distances ranged from 0.092 (Cluster 3) to 0.220 (Cluster 22). Standard errors of the taxonomic distances were negligible in all cases. The identification scores for randomly selected cluster representatives, provided by the MATIDEN program, were inferior to those for the HMO. Nevertheless, all the representatives were identified to the correct phenon with Willcox probabilities <0.955, low taxonomic distance (0.145 to 0.318), and standard errors (−2.047 to 2.104).

Practical evaluation of the identification matrix

For practical evaluation, a minimum Willcox probability of 0.85 was selected as the prime criterion for the identification of unknown strains (9, 27). Two additional criteria adopted for a successful identification were as follows: 1) first group scores substantially better than those against the next best two alternatives; 2) 'characters against' should be zero or few. Examples to illustrate the range of scores obtained for identified and non-identified isolates are given in Table 3 and 4. The scores for the identification coefficients generally followed the expected pattern, with those for taxonomic distance, and its standard errors increasing as the Willcox probabilities decreased. Of the 32 strains isolated from a variety of soils, 23 (71.9%) were identified to one of 8 clusters (Table 4).

Table 1. A percentage positive probability matrix for acidophilic *Streptomyces* defined in the clusters of Seong *et al.* (14)

Cluster	1	3	4	10	11	12	13	14	15	16	17	18	19	20	21	22	24	25	27	28	29	30	33	34	35
No. of strains	4	3	6	3	3	3	5	3	3	3	4	4	3	2	9	3	8	15	7	3	3	4	5	2	2
Character																									
Degradation of:																									
Elastin	1	1	1	1	1	1	1	33	1	33	25	75	67	50	1	99	1	7	1	1	1	1	1	1	50
Guanine	1	1	99	33	67	67	99	1	1	1	99	99	99	99	90	99	88	80	71	99	67	99	60	1	1
Tyrosine	1	99	1	1	33	1	1	99	67	67	50	99	67	1	90	67	25	1	1	1	1	50	1	1	99
Xanthine	1	1	1	1	1	1	1	1	1	67	25	99	1	1	90	33	1	1	1	1	1	1	1	1	99
DNA	99	99	83	99	33	99	99	99	99	99	99	75	99	99	99	99	99	87	99	67	67	1	60	99	99
Growth in the presence of:																									
Adenine (4%, w/v)	1	1	1	1	33	1	20	99	1	99	50	75	67	1	78	1	13	1	1	1	1	1	1	1	99
Cobalt chloride (50 μg/ml)	99	99	83	33	99	99	99	1	1	99	99	99	1	99	99	99	99	99	71	99	99	99	99	99	99
Copper sulfate (50 μg/ml)	99	99	99	67	99	1	1	99	99	99	99	99	99	99	99	99	99	93	99	99	99	99	80	99	99
Thallus acetate (10 μg/ml)	1	1	83	99	1	1	60	99	99	99	99	50	1	99	33	33	1	87	57	67	33	99	1	50	99
2 : 3 : 5-Triphenyl tetrazolium chloride (10 μg/ml)	50	99	99	67	99	1	80	99	99	99	99	99	99	99	99	99	99	87	99	99	99	99	99	1	50
Growth at: pH 3.5	99	99	99	99	33	1	1	99	99	99	99	99	33	1	67	67	88	99	99	99	99	99	99	99	99
Antibiotic resistance to:																									
Ampicillin (Na) (4μg/ml)	99	99	99	99	1	99	99	99	67	99	99	99	99	50	99	99	99	99	99	99	99	99	99	99	50
Cephapirin (Na) (32 μg/ml)	1	99	1	1	1	33	1	1	1	1	99	99	99	1	11	67	1	1	1	1	1	1	1	1	99
Chloramphenicol (4 μg/ml)	1	1	1	1	33	67	99	1	1	33	50	99	67	1	33	1	1	27	1	1	1	1	1	1	99
Gentamycin (SO_4) (32 μg/ml)	75	99	1	1	1	1	1	1	1	67	99	25	1	1	1	1	50	1	1	1	1	25	1	1	50
Nalidicic acid (32 μg/ml)	1	99	1	1	1	1	60	1	1	1	99	99	99	1	1	1	1	1	1	1	1	1	1	1	50
Neomycin (SO_4) (4 μg/ml)	99	99	99	99	99	33	1	99	99	99	99	99	99	99	99	33	99	99	99	99	99	99	99	99	99
Oleandomycin (PO_4) (32 μg/ml)	1	1	1	1	1	1	1	99	1	99	1	99	99	1	99	1	13	20	1	1	1	1	20	1	99
Polymyxin B (SO_4) (32 μg/ml)	99	99	99	99	99	1	1	99	99	99	99	99	99	50	99	1	99	99	99	99	99	99	99	99	99
Rifampicin (4 μg/ml)	1	1	1	1	1	33	1	1	67	33	1	25	67	1	1	99	13	13	1	1	33	1	1	1	99
Torbramycin (SO_4) (32 μg/ml)	99	1	1	1	1	1	60	1	1	1	75	1	1	1	1	1	1	7	1	1	1	1	1	1	50
Utilization as nitrogen sources:																									
Glycine	99	99	99	99	99	99	60	33	99	99	50	25	99	99	1	99	88	93	99	67	99	99	99	99	99
L-Histidine	99	99	83	99	99	33	1	99	99	99	99	99	99	99	99	99	88	20	57	33	1	99	1	99	99
Potassium Nitrate	1	99	67	99	99	1	40	67	1	99	99	99	99	50	67	99	99	99	99	99	99	99	99	99	99
L-Serine	1	1	1	1	33	1	40	1	1	67	1	1	33	1	1	1	99	99	99	67	67	1	40	50	99
D-Tryptophan	1	1	1	1	1	1	1	1	99	67	1	1	99	1	56	1	1	27	29	1	1	75	20	1	99
Utilization as carbon sources:																									
D(+) Fucose	1	1	1	1	1	1	1	1	1	1	1	1	1	1	78	1	1	73	29	1	1	99	20	1	1
D(+) Mannose	1	1	17	1	1	1	1	1	1	67	1	1	33	1	1	1	99	99	86	99	1	99	60	1	99
D(−) Ribose	1	1	1	33	1	1	1	1	1	1	1	1	1	1	78	1	38	99	86	67	1	99	20	50	99
D(+) Xylose	1	1	1	1	67	1	1	1	1	33	1	1	1	50	1	1	99	99	99	99	67	99	60	1	99
D(+) Melibiose	25	1	17	33	1	1	1	1	33	33	1	1	1	1	1	1	99	99	99	99	67	99	40	1	99
D(+) Trehalose	99	99	67	99	99	1	40	99	99	33	99	1	99	1	1	99	99	99	99	99	99	99	99	1	99
D(+) Turanose	75	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	99	99	99	99	33	99	1	99	99
D(+) Raffinose	1	1	67	1	1	1	1	1	1	1	1	1	33	1	1	1	50	99	99	1	67	99	20	1	99
Arbutin	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	99	99	99	99	33	99	80	1	1
2-Keto-D-gluconic acid	1	33	17	99	99	1	99	99	99	99	75	99	99	99	99	99	99	99	99	99	99	99	99	99	99
Dulcitol	1	1	1	1	1	1	1	1	1	1	1	1	1	99	99	1	88	99	57	99	1	25	20	50	1
L-Alanine	25	67	50	99	67	1	60	99	99	99	99	99	99	99	99	67	99	99	99	99	99	99	99	99	99
L-Isoleucine	99	99	17	99	99	1	99	67	99	99	99	75	99	99	99	33	99	99	99	99	99	99	99	99	99
L-Ornithine	1	1	1	1	1	1	1	1	1	1	1	1	1	99	99	1	38	99	29	33	1	75	20	1	99
L-Serine	1	1	1	1	1	1	1	1	1	1	1	1	1	1	90	1	99	99	29	33	33	1	60	1	1
D-Tryptophan	1	1	1	1	1	1	1	1	1	1	1	1	33	1	99	1	99	99	29	33	67	1	20	1	1
Fumaric acid (Na)	25	99	99	99	99	33	1	99	99	99	99	99	99	99	99	99	99	53	71	67	67	99	40	99	99
Succinic acid (Na)	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	7	71	99	67	99	1	99	99
Citric acid	99	99	83	99	99	33	60	99	99	99	99	99	99	99	99	99	99	33	99	67	67	99	1	99	99
Butyric acid (Na)	1	1	1	1	1	1	1	67	1	1	50	99	99	1	90	1	13	1	1	1	1	1	1	1	99
Propionic acid (Na)	1	66	1	1	1	1	1	67	1	33	99	99	99	99	99	1	1	1	1	1	1	1	1	1	99
DL-Lactic acid	1	99	1	33	99	1	60	99	99	99	75	99	99	99	99	1	99	99	99	99	99	99	99	99	99
Diffusible pigmentation:																									
Green	1	99	1	1	1	1	1	33	1	1	1	50	1	1	1	99	1	1	1	1	1	1	1	1	1
Orange/Yellow	1	1	1	1	1	1	1	1	1	1	25	1	1	1	1	1	13	7	1	67	1	1	1	1	1
Spore chain morphology:																									
Retiaculum-apertum	1	1	1	1	1	1	40	33	1	1	99	99	99	99	33	1	1	1	1	1	1	1	20	50	99
Spria	1	1	1	1	1	1	1	67	99	99	1	1	1	1	67	1	1	1	43	1	1	1	80	1	1
Antimicrobial activity against:																									
M. luteus	1	67	1	1	1	1	1	1	1	99	25	1	1	1	67	1	1	1	1	1	1	1	1	1	99

Formulation of selective isolation medium

The most diagnostic or selective characters for cluster 35 were provided by the DIACHAR program (Table 5). The characters were further assessed as to whether or not the character states are common in other clusters. The two characters (resistance to adenine, spiral spore chain) were excluded, because the positive percentage value of 5 clusters are same as 100. The following characters are also deleted because of practical difficulties in the isolation stage: degradation of xanthine, growth in the presence of 5% NaCl, and antimicrobial activity against *Aspergillus niger*. Utilization of melibiose was not included as a character in this experiment. The negative selective character guanine degradation was also discarded. Therefore, 4 characters - resistance to antibiotics (Rifampicin, chloramphenicol), potassium tellurite, and thallous acetate were selected to design a selective medium for the strains of cluster 35 (Fig. 1). Effectiveness of each substance for the selective isolation of target strains was confirmed by detecting the resistance of representatives to each of the chemicals (Fig. 1). Starch casein agar (8) supplemented with Rifampicin (4 μg/ml), Chloramphenicol (4 μg/ml), potassium tellurite (10 μg/ml), and thallous acetate (10 μg/ml) were used to isolate antibiotic-producing strains of cluster 35.

Table 2. A summary of scores obtained in the construction and assesment of identification matrix.

Program	Information	Range of sores
CHARSEP	VSP (%)	11.67 to 75.73
MOSTTYP	Hypothetical median organism:	
	Willcox probability	0.987 to 1.000
	Taxonomic distance (*d*)	0.092 to 0.220
	Standard error of *d*	−4.423 to −2.940
DIACHAR	Sum of diagnostic sores	16.07 to 23.47
OVERMAT	Number of clusters with 5% overlap	12
MATIDEN	Identification coefficients for cluster representatives:	
	Willcox probability	0.955 to 1.000
	Taxonomic distance (*d*)	0.145 to 0.318
	Standard error of *d*	−2.047 to 2.104

Table 3. Examples of identification scores for unknown isolates

Origin of isolates	Cluster identification	Willcox probability	Taxonomic distance (*d*)	Standard error of *d*
Pinus forest soil	Cluster 1	>0.999	0.205	−1.683
Bamboo forest soil	Cluster 35	>0.999	0.281	−1.385
Pinus forest soil	Cluster 27	>0.999	0.264	−1.284
Pinus forest soil	Cluster 24	>0.999	0.292	−0.856
Ginseng cultivate soil	Cluster 21	0.997	0.424	1.968
Ginseng cultivate soil	Cluster 27	0.869	0.478	2.834
Coal mine waste	Cluster 24	0.873	0.511	2.877
Coal mine waste	Not identified	0.844	0.533	3.827
Pinus forest soil	Not identified	0.790	0.534	4.632

Table 5. An example of best selective and diagnostic tests for cluster 35 against all of the clusters

Character	% positive	ID score	Difference
Sodium chloride (5%, w/v)[a]	100	0.947	0.935
Phenol (1000 μg/ml)[a]	100	0.939	0.919
Xanthine[b]	100	0.900	0.844
M. luteus[c]	100	0.884	0.814
Adenine[a]	100	0.858	0.767
Spira[d]	100	0.825	0.709
Melibiose[e]	100	0.811	0.686
* Rifampicin (4 μg/ml)[a]	100	0.801	0.669
* Potassium tellurite (10 μg/ml)[a]	100	0.781	0.634
* Chloramphenicol (4 μg/ml)[a]	100	0.754	0.592
Oleandomycin (PO_4) (32 μg/ml)[a]	100	0.754	0.591
Guanine[b]	0	0.732	0.558
* Thallous acetate (10 μg/ml)[a]	100	0.728	0.552

The substances marked are selected to design a medium.
[a], degradation; [b], resistance; [c], antibiosis; [d], spore chain morphology; [e], utilization of carbon sources.

Table 4. Identification of unknown isolates by probability matrix

Cluster identification	No. of isolates in cluster	Willcox probability (p)			
		$p \geq 0.995$	$0.995 > p \geq 0.990$	$0.989 > p \geq 0.850$	$p < 0.85$
Cluster 1	2	1		1	
Cluster 18	2	1	1		
Cluster 21	3	3			
Cluster 24	4	3	1		
Cluster 25	4	2	1	1	
Cluster 27	6	2		4	
Cluster 33	1			1	
Cluster 35	1	1			
Total	23	13	2	8	9
	(71.9%)	(40.6%)	(6.2%)	(25.0%)	(28.1%)

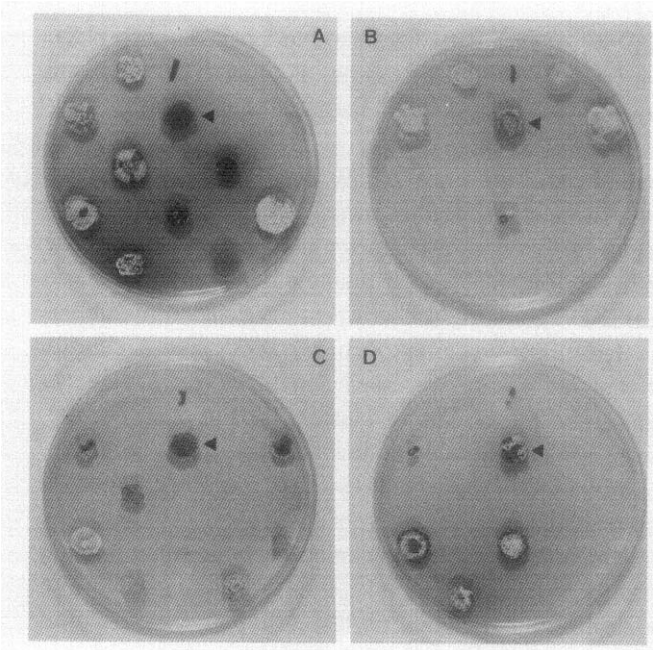

Fig. 1. Photographs showing the resistance of the target organism, CN633 (closed triangle) to the addition of (A) Rifampicin (4 μg/ml), (B) Chloramphenicol (4 μg/ml), (C) thallous acetate (10 μg/ml), and (D) potassium tellurite (10 μg/ml).

Table 6. Effect of selective medium on the isolation of antibiotic-producing strains

	Non-selective medium (a)	Selective medium (b)	Ratio (b/a)
No. of isolates (ml^{-1})	4.5×10^5	1.2×10^3	0.03
No. of antibiotic-producing strains	4(32)	19(32)	4.75
No. of strains identified to cluster 35	1(4)	10(19)	2.11

* Numerals in the parentheses are the number of strains tested.

Application of the selective isolation medium

When selective isolation of organisms from various soils was carried out using synthetic medium, the number of isolates was decreased and the growth of microorganisms was retarded. However, increased isolation rate of antibiotic-producing strains was found. A large portion (10 out of 19) of the anti-Micrococcal isolates were identified to cluster 35 (Table 6).

Discussion

Numerical classifications result in objective groupings based on a large number of phenetic characters. The large data content of numerical classification is important and can be applied to a variety of purposes. However, a large number of characters sometimes leads to restrictions in practical usefulness. To be practicable, the identification matrices are to contain the minimum number of characters needed to discriminate between the clusters (15, 22). Such probabilistic matrices have been designed for *Streptomyces* and related genera (7, 9, 27). The matrices were shown to be practically sound, and are widely used to identify unknown isolates (28).

However, only a few taxonomic studies were concerned with the acidophilic actinomycetes (5, 11, 13, 14). Thus, construction of a sound and workable classification and identification system is urgently required. The primary aim of this study is to construct a single probability matrix for the identification of acidophilic *Streptomyces*. The character states of clusters defined by numerical phenetic methods were used in this approach. Selection of clusters and reduction of diagnostic characters were achieved by using specifically designed computer programs (16, 19, 20, 25). Ten minor clusters containing only 2 strains were excluded in the matrix because of significant overlap using 5% as the critical value. Finally, fifty-three characters against 25 clusters encompassing 110 strains were included in the probability matrix.

The matrix was evaluated theoretically and found to be relatively sound. The identification scores of HMO and representatives of each cluster in this matrix were superior to those of previous matrices designed by Langham *et al.* (9) and Kämper and Kropepenstedt (7). It may be thought that the clusters of this matrix are defined at the high similarity level of 86.5% with S_{SM} in numerical classification (14), while the clusters of previous analyses are recognized at 77.5% (26) and 81.5% (6) SSM similarity level, respectively. The overlaps (4.0%) between clusters were similar or inferior to those of Kämper and Kropepenstedt (7) and Langham *et al.* (9). For the major cluster (cluster 25), significant overlaps with minor clusters (cluster 27, 28, 29, 30) were observed. The large overlaps may be due to high inter-cluster similarity more than 84.3% between those clusters (13). The practical evaluation of the matrix against soil isolates was a relatively limited exercise. The overall identification rate for 32 strains, (71.9%) at a Willcox probability level of 0.85, was inferior to those of previous works (7, 9, 27). Despite this identification rate was comparable with other probabilistic systems, such as 70.8% for Gram-negative non-fermentative rods (10) and 47% for slowly-growing mycobacteria (24). One of the reasons for the low identification rate depicted is that the matrix contained insufficient data, since 56 strains from the numerical analysis are excluded in the probability matrix. Otherwise, the clusters included in the matrix could be illustrated to be defined more tightly.

Some problems remained in the identification of aci-

dophilic *Streptomyces*. The determination of spore chain morphology is not easy, and is inadequate for identification. Basically, large numbers of strains from a wider range of sources should be included in the original classification. Moreover, nomenclature of the defined strains are ought to be carried out urgently. With the aid of polyphasic classification methods such as chemotaxonomy and molecular phylogenetic analysis (1).

A broad range of numerical phenetic data can be accessed for a variety of purposes in addition to construction of an identification matrix (23). An attempt was carried out in this study to design a selective medium for the isolation of antibiotic-producing strains. Detection of increased rates of selective isolation on a designed medium will help to improve practicable methods for screening of useful strains. But, problems still remain in formulating a selective medium. Combination of 4 chemicals inhibited the growth and sporulation not only of unrecommended but also of target strains. It brought about difficulties in recognizing the colony as a target strain in the initial selection stage. To resolve these problems, it is advisable to design a selective medium with a variety of carbon sources.

It must be emphasized that, for the acidophilic actinomycetes, more comprehensive and comparative taxonomic studies derived by various methods are needed. Despite this discrepancy, numerical classification data was found to be useful in constructing the probability identification matrix and in designing a selective medium.

Acknowledgement

This paper was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation, 1993.

References

1. **Goodfellow, M. and A.G. O'Donell,** 1993. Roots of bacterial systematics, p. 3-41. *In* Goodfellow, M. and A.G. O'Donell (ed.), Handbook of new bacterial systematics. Academic Press, London.
2. **Goodfellow, M. and T. Cross,** 1984. Classification, p. 7-64. *In* Goodfellow, M., M. Mordarski, and S.T. Williams(ed.), Biology of the Actinomycetes. Academic Press, London.
3. **Hagedorn, C.,** 1976. Influences of soil acidity of *Streptomyces* populations inhabiting forest soils. *Appl. Environ. Microbiol.* **32**, 368-375.
4. **Kang, H.I.,** 1989. Taxonomic study on the acidophilic and neutrotolerant actinomycetes. Ph. D. Thesis. Seoul National University.
5. **Khan, M.R. and S.T. Williams,** 1975. Studies on the ecology of actinomycetes in soil. VIII. Distribution and characteristics of acidophilic actinomycetes. *Soil Biol. Biochem.* **7**, 345-348.
6. **Kämpfer, P. and R.M. Kroppenstedt,** 1991. Probabilistic identification matrix of *Streptomyces* using miniaturized physiological tests. *J. Gen. Microbiol.* **137**, 1893-1902.
7. **Kämpfer, P., R.M. Kroppenstedt, and W. Dott,** 1991. A numerical classification of the genera *Streptomyces* and *Streptoverticillium* using miniaturized physiological tests. *J. Gen. Microbiol.* **137**, 1831-1891.
8. **Kuster, E. and S.T. Williams,** 1964. Selection of media for the isolation of actinomycetes. *Nature* **202**, 928-929.
9. **Langham, C.D., S.T. Williams, P.H.A. Sneath, and A. M. Mortimer,** 1989. New probability matrices for identification of *Streptomyces*. *J. Gen. Microbiol.* **135**, 121-133.
10. **Lapage, S.P., S. Bascomb, W.R. Willcox, and M.A. Curtis,** 1983. Identification of bacteria by computer: general aspects and perspectives. *J. Microbiol.* **77**, 273-290.
11. **Lonsdale, J.T.,** 1985. Aspects of the biology of acidophilic actinomycetes. Ph. D. Thesis, University of Newcastle upon Tyne.
12. **Nolan, R.D. and T. Cross,** 1988. Isolation and screening of actinomycetes, p. 1-26. *In* Goodfellow, M., S.T. Williams, and M. Mordarski (ed.), Actinomycetes in biotechnology. Academic Press, London.
13. **Seong, C.N.,** 1992. Numerical taxonomy of acidophilic and neutrotolerant actinomycetes isolated from acid soil in Korea. Ph. D. Thesis. Seoul National. University.
14. **Seong, C.N., M. Goodfellow, A.C. Ward, and Y.C. Hah,** 1993. Numerical classification of acidophilic actinomycetes isolated from acid soil in Korea. *Kor. J. Microbiol.* **31**, 355-363.
15. **Sneath, P.H.A.,** 1978. Classification of microorganisms, p. 1-31. *In* Norris J.R. and M.H. Richmond (ed.), Essays in microbiology. Wiley, Chichester.
16. **Sneath, P.H.A.,** 1979. Basic program for identification of an unknown with presence-absence data against an identification matrix of present positive characters. *Computers Geosci.* **5**, 195-213.
17. **Sneath, P.H.A.,** 1979. Basic program for character separation indices from an identification matrix of present positive characters. *Computers Geosci.* **5**, 349-357.
18. **Sneath, P.H.A.,** 1980. Basic program for the most diagnostic properties of groups from an identification matrix of present positive character. *Computers Geosci.* **6**, 21-26.
19. **Sneath, P.H.A.,** 1980. Basic program for determining the best identification scores possible from the most typical examples when compared with an identification matrix of present positive characters. *Computers Geosci.* **6**, 27-34.
20. **Sneath, P.H.A.,** 1980. Basic program for determining the overlap between groups in an identification matrix of present positive characters. *Computers Geosci.* **6**, 267-278.
21. **Sneath, P.H.A. and R. Johnson,** 1972. The influence on numerical taxonomic similarities of errors in microbiological tests. *J. Gen. Microbiol.* **72**, 377-392.
22. **Sneath, P.H.A. and R.R. Sokal,** 1973. Numerical taxo-

nomy: The principles and practice of numerical classification. Freeman, San Fransisco.

23. **Vickers, J.C., S.T. Williams, and G.W. Ross,** 1984. A taxonomic approach to selective isolation of streptomycetes from soil, p. 553-561. *In* Ortiz-Ortiz, L., L.F. Bojalil, and V. Yakoleff (ed.), Biological, biochemical and biomedical aspects of actinomycetes. Academic Press, Orlando.
24. **Wayne, L.G., E.J. Krichevsky, L.L. Love, R. Johnson, and M.I. Krichevsky,** 1980. Taxonomic probability matrix for use with slow-growing mycobacteria. *Int. J. Syst. Bacteriol.* **30**, 528-538.
25. **Willcox, W.R., S.P. Lapage, S. Bascomb, and M.A. Curtis,** 1973. Identification of bacteria by computer: theory and programming. *J. Gen. Microbiol.* **77**, 313-330.
26. **Williams, S.T., M. Goodfellow, G. Alderson, E.M.H. Wellington, P.H.A. Sneath, and M.J. Sackin,** 1983. Numerical classification of *Streptomyces* and related genera. *J. Gen. Microbiol.* **129**, 1743-1813.
27. **Williams, S.T., M. Goodfellow, E.M.H. Wellington, G. Alderson, J.C. Vickers, P.H.A. Sneath, M.J. Sackin, and A.M. Mortimer,** 1983. A probability matrix for identification of some Streptomycetes. *J. Gen. Microbiol.* **129**, 1815-1830.
28. **Yun, N.R. and Y.N. Lee,** 1994. Identification of soil *Streptomyces* spp. producing antifungal substance. *Kor. J. Microbiol.* **32**, 449-454.