This paper not to be cited without prior reference to the autor

TAXONOMY OF BACILLUS ISOLATES FROM MARINE SEDIMENTS

By A.Boeyé, M.Aerts and G.Clerx

Our numerical taxonomic study of the Bacillus isolates from North Sea soils
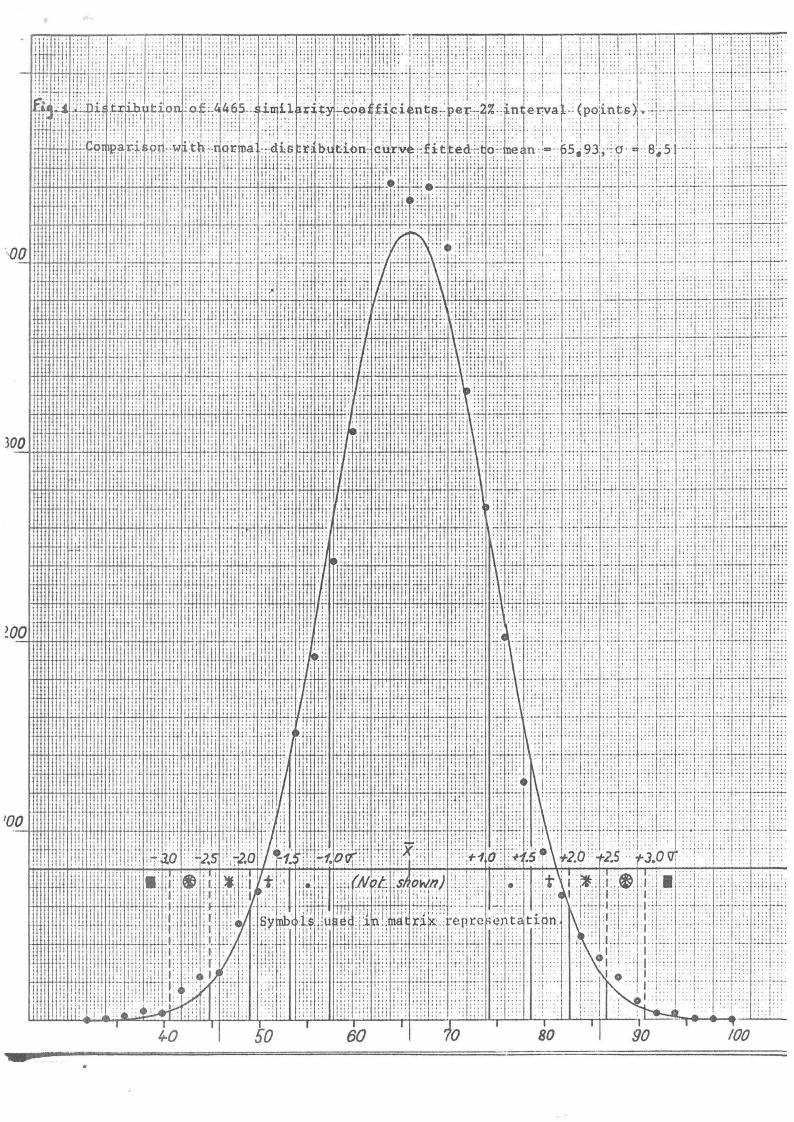was extended in several ways.

1) The analysis of 28 new isolates was completed and the total of completely
   analyzed now stands at 95.
2) Additional tests were introduced and applied to all isolates :
   - growth at pH 5.7
   - formation of a pellicle (N.B., pH 7, 30°C)
   - hydrolysis of pectate
   - hydrolysis of chitin

The latter properties were examined because they are especially relevant to
the ability of the bacteria to decompose organic matter. For the same reason,
several attempts were made at developing a good test for the decomposition of
cellulose and alginate, but the results proved to be of insufficient reprodu-
cibility.

The present number of encoded data now stands at 62.

3) A special effort was made at developing computer programs allowing the analysis
   of the data.
a) A better clustering program was developed.
b) A program was developed to analyse the relations between preformed clusters.

c) Another was developed to draw similarity matrices showing the similarity coefficients (S.C.) which differ from the average by any chosen multiple of the standard diviation. Since these double-tailed matrices show similarities as well as dissimilarities they are especially useful to find the optimal order corresponding to the best over-all picture.

To illustrate this method, the distribution of the 4465 different S.C. which exist between 95 elements is shown in figure 1. The observed distribution closely approached the theoretical (normal) distribution curve. The S.C. forming the left and right tails of this distribution are represented by different symbols in the combined matrices shown in Fig.2. It can be seen that the ordering is already satisfactory since the highest values lie close to the diagonal and the lowest close to a corner, but a close inspection will easily reveal some inversions which would still improve the picture. We feel that the usual clustering procedures are especially valuable for ordering closely related elements, while on the other hand a procedure based on the visual representation and correction by trial and error may ultimately prove superior for the purpose of establishing the over-all picture; it may be especially useful to deal with overlapping clusters.

d) A program was also developed to determine a posteriori the distribution of each character to the formation of the clusters. In a novel approach, we equated this contribution to the average increase of the "internal" S.C (between any element of a cluster and the other elements of the same cluster) minus the average increase of the "external" (=all other)S.C.

This analysis is expected to help us select a set of specially useful tests, so that future analyses may be restricted to a relatively small number of tests. The evaluation of the results will be done as soon as the analysis of a new series of isolates, now in progress, is completed.
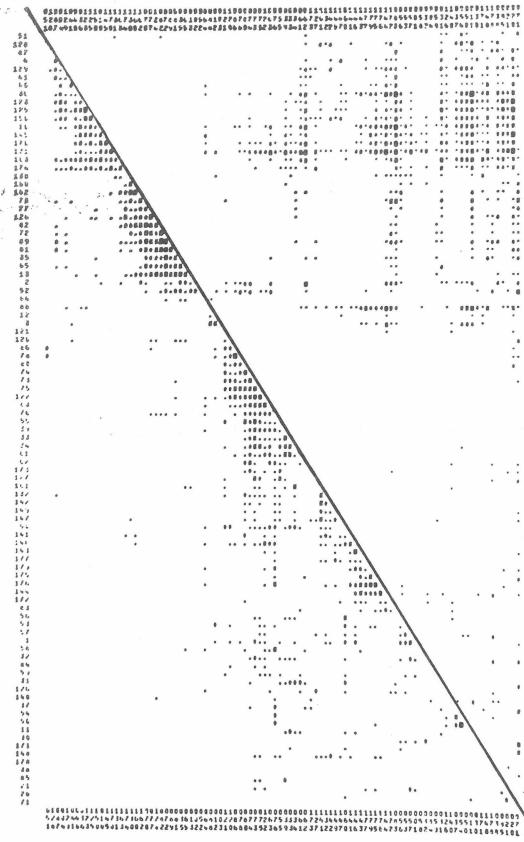
Fig. 1. Distribution of 4465 similarity coefficients per 2% interval (points). Comparison with normal distribution curve fitted to mean = 65.93, σ = 8.51

Fig. 2