Classification and Identification of *Mycobacterium africanum* by Pyrolysis Mass Spectrometry

By GERHAN WIETEN,^{1*} JOHAN HAVERKAMP,¹ DICK G. GROOTHUIS,² LUCAS G. BERWALD² AND HUGO L. DAVID³

¹ FOM-Institute for Atomic and Molecular Physics, Kruislaan 407, 1098 SJ Amsterdam, The Netherlands

² National Institute of Public Health, Antonie van Leeuwenhoeklaan 1, 3721 MA Bilthoven, The Netherlands

³ Institut Pasteur, 28 Rue du Dr. Roux, 75724 Paris, Cedex 15, France

(Received 8 June 1983)

Pyrolysis mass spectrometry was used to classify and identify strains of Mycobacteriumafricanum and of M. tuberculosis, M. bovis and M. bovis BCG. The multicharacter mass pyrograms were evaluated by computerized data handling procedures that were suited for classification and identification. The results revealed considerable heterogeneity among the African strains, which was shown to be linked to the geographic distribution of the strains. On the basis of a routine mass spectrometric identification key the African strains were identified without exception as belonging to, what is referred to as the 'Tuberculosis complex' (i.e. the clinically relevant group formed by strains of M. tuberculosis, M. bovis and M. bovis BCG). Classification of the strains by means of discriminant analysis indicated an intermediate clustering for the majority of the African strains and overlap for some African strains with in particular M. bovis. It was concluded that from the mass spectrometric data a species status for the group of African strains was not justifiable.

INTRODUCTION

Pyrolysis mass spectrometry (Py-MS) is an analytical technique that provides 'biochemical fingerprints' (mass pyrograms) of complex organic samples, such as whole bacterial cells or bacterial cell fractions (Irwin, 1979; Gutteridge & Norris, 1979; Meuzelaar *et al.*, 1982; Wieten *et al.*, 1983). In mycobacteriology (Wieten *et al.*, 1981*a*, *b*) Py-MS has been successfully applied for the identification of *Mycobacterium tuberculosis* and the related (sub-)species *M. bovis* and *M. bovis* BCG [as in previous papers the term 'Tuberculosis complex' will be used as a convenient, though nomenclaturally irrelevant reference to these three (sub-)species]. It was also shown that strains of the 'Tuberculosis complex' could be well separated from other species of mycobacterium. In other studies the heterogeneity and classification of several mycobacterial species, including *M. leprae*, has been investigated (Meuzelaar *et al.*, 1977; Wieten *et al.*, 1979, 1982). The influence of culturing and sample handling on the pyrolytic fingerprints has been extensively studied (Meuzelaar *et al.*, 1976; Wieten *et al.*, 1983).

Pyrolysis mass spectrometric procedures are well suited for automation, offering the possibility of processing large numbers of samples rapidly for taxonomical purposes. Sample preparation is extremely straightforward involving direct sampling from the culture tubes, and the sample quantities required are low $(5-20 \mu g \text{ organic material})$. For the evaluation of mass spectrometric data, various computerized procedures for multivariate statistical analysis are available (Meuzelaar *et al.*, 1982; Wieten *et al.*, 1983).

Mycobacterium africanum (Castets et al., 1969) is a mycobacterial species which is described as

closely related to the 'Tuberculosis complex'. From taxonomical as well as clinical points of view its status as a separate species has been disputed (David *et al.*, 1978; Wayne & Diaz, 1979; Van den Groen & Pattyn, 1975). The present paper describes the analysis by Py-MS of representatives of *M. africanum*, originating from different geographical areas and of reference strains and clinical isolates of *M. tuberculosis* and the related (sub-)species *M. bovis* and *M. bovis* BCG.

METHODS

Bacterial strains. From the culture collection of the Institute Pasteur, Paris, 10 strains of Mycobacterium africanum were selected which had all been extensively characterized in a previous taxonomic study (David et al., 1978). The geographical origin of the strains, the previously used strain codes (David et al., 1978) and the codes used in this paper (in brackets) were as follows. Originating from Dakar: ATCC 25420 (DI), HB 1702 (DII), HB 2827 (DIII), HB 5656 (DIV); from Yaoundé: RT 611 (YI), RT 6002 (YII); from Rwanda: S 3237 (RI), RT 6856 (RII), RT 6833 (RIII) and from Burundi: RT 8815 (BI). The strains were mailed under code to the National Institute of Public Health (NIPH), Bilthoven, The Netherlands for subculturing and sample preparation. From M. tuberculosis, M. bovis and M. bovis BCG three strains each were analysed, all of which were previously described and were used during several years as references in various Py-MS studies (Wieten et al., 1979, 1981 a, b, 1982). In addition three strains of M. tuberculosis, three of M. bovis and two of M. bovis BCG were selected from the collection of recent clinical isolates of the NIPH.

Sample preparation for Py-MS analysis. Streak cultures were made in duplicate on standard Löwenstein-Jensen medium and incubated at 37 °C for 3 weeks. A small amount of culture was scraped off the slants and emulsified in a droplet of sterile distilled water. The tip of a ferromagnetic wire was turned slowly in the suspension. In this way sufficiently reproducible amounts of bacterial mass remained attached to the wires as could be deduced from the total spectrum intensity. From each suspension duplicate samples were taken and thus a total of four samples from each strain were prepared for Py-MS analysis. Individual wires were mounted in separate glass housings, coded and autoclaved for 30 min at 120 °C. Samples were stored overnight at 4 °C and transferred the next day at ambient temperature from the NIPH to the FOM-Institute for Py-MS analysis.

Pyrolysis mass spectrometry. A detailed description of the fully automated Py-MS instrument that was used for this study is given elsewhere (Meuzelaar *et al.*, 1976; Wieten *et al.*, 1982). Samples were analysed in one batch with a sample turnover rate of 30 samples h^{-1} . Curie point pyrolysis *in vacuo*, was carried out by means of high frequency inductive heating of the Fe/Ni sample wires (Curie-point temperature 510 °C) with a temperature rise time of 0·1 s and a total heating time of 1·0 s. The mixture of volatile products that resulted from pyrolysis diffused via a heated (150 °C) expansion chamber into the quadrupole mass spectrometer. Low energy (14 eV) electron impact ionization was used. The mass-range was 15 to 160 a.m.u. and the scan-time for this range 0·1 s. Successive scans were accumulated in a signal averaging mode for a 10 s period, whereafter spectra were transferred to a large time-shared computer facility (C.D.C. 7600) for data processing.

Data processing. The computerized multivariate procedures, used to evaluate the mass spectrometric data, were discriminant analysis (DA), combined with non-linear mapping to visualize multidimensional differences in mass pyrograms and a key-mass identification procedure. Prior to data evaluation, all mass pyrograms were subjected to a normalization procedure in order to correct for variations in sample size, so that peak intensities were expressed as percentage of total ion current (Meuzelaar *et al.*, 1982; Wieten *et al.*, 1983).

Discriminant analysis is a procedure which displays dissimilarities among mass pyrograms in a quantitative way (Windig et al., 1981, 1983). Dissimilarities among mass pyrograms obtained from samples that are highly related, such as mycobacterial cells, usually reflect differences in the relative amount of one or more chemical components. In general a single chemical component gives rise to a suite of pyrolysis fragments and therefore to a sequence of mass peaks (subpattern) in the mass pyrogram of the total sample. Thus in a set of mass pyrograms a difference in the relative amount of a chemical component will be reflected by correlated changes in a number of peak intensities. When applied on standardized data sets, DA can be used to define new variables, i.e. discriminant functions, that describe the correlations among mass peaks. Each function represents a linear combination of the original mass peaks. The calculation of discriminant functions is determined by the ratio of between group to within group variances, a group being defined as the set of replicate mass pyrograms of one bacterial strain. Thus the first discriminant function, D1, represents the best possible differentiation between individual bacterial strains, whilst D2 represents the linear combination of mass peaks that accounts maximally for the remaining variance, etc. The pattern of mass peaks described by a discriminant function can be visually represented by a discriminant spectrum. Due to normalization of mass pyrograms a discriminant spectrum is made up from positive and negative parts. The pattern displayed by a discriminant spectrum can be used for the chemical interpretation of dissimilarities among mass pyrograms (Windig et al., 1983; Wieten et al., 1983). Discriminant scores, i.e. the 'intensities' of discriminant functions in individual mass pyrograms, provide a direct means for quantitative determination of dissimilarities among mass pyrograms. These dissimilarities are visualized by plotting the discriminant scores in either one, two or three dimensional plots. If more than three discriminant functions are to be considered simultaneously, non-linear mapping of the discriminant scores can be used as a visualization technique.

Non-linear mapping as developed by Kruskal (1964) and adapted for Py-MS data by Eshuis *et al.* (1977) is an iterative data fitting procedure that can be used to visualize multidimensional data by means of two dimensional plots. The goodness of fit (stress factor) is determined by a set of rules that emphasize relative dissimilarities amoung mass pyrograms (nearest neighbour criterion), rather than absolute dissimilarities. Consequently a non-linear map is a qualitative aid for data evaluation and should not be used for quantitative interpretation of dissimilarities among mass pyrograms.

The Py-MS screening procedure (key mass identification) that was used, identifies unknown strains as either representatives of the 'Tuberculosis complex' or as so-called 'atypical' strains and has been described in detail previously (Wieten *et al.*, 1981*b*). The identification was based on the intensities of an empirically determined, batch-independent set of key masses, i.e. m/z 50, 59, 71, 81, 98 and 116. One-way analysis of variance was used as a measure of the dissimilarity between the averaged mass pyrograms from the replicate analyses of each unknown strain and a hypothetical mean mass pyrogram of the 'Tuberculosis complex'. The latter was deduced from the mass pyrograms of the standard set of nine reference strains. The critical dissimilarity value for identification was based on the dissimilarity between the references and the hypothetical mean of the 'Tuberculosis complex', combined with an empirical factor to correct for incomplete covering of the 'Tuberculosis complex' heterogeneity by the reference strains. The whole procedure was fully computerized.

RESULTS

Evaluation of the mass pyrogram of M. africanum and the members of the 'Tuberculosis complex' was carried out in a stepwise manner by DA, using discriminant scores to display quantitative dissimilarities among mass pyrograms.

First the mass pyrograms of M. tuberculosis, M. bovis and M. bovis BCG were evaluated by DA and the plot of discriminant scores for the first and second discriminant function is given in Fig. 1. From this plot it will be clear that the strains from M. tuberculosis and M. bovis BCG form two homogeneous and closely related clusters. Among the M. bovis strains heterogeneity is apparent with some of the M. bovis strains appearing to be closely related to M. tuberculosis and others being highly dissimilar. Notably there is no overlap between the three clusters of the 'Tuberculosis complex.'

The mass pyrograms of the *M. africanum* strains were added to those of the 'Tuberculosis complex' and the complete data set was re-evaluated by DA. Figure 2 represents the plot of discriminant scores for the first and second discriminant function, together accounting for 67% of the total variance in the mass pyrograms. The relationships amoung the strains of the 'Tuberculosis complex' remain essentially unchanged as revealed by comparison with Fig. 1. The *M. africanum* strains segregate in a cluster which is more heterogeneous than the *M. tuberculosis* or *M. Bovis* BCG cluster. The majority of the strains are very closely related to the 'Tuberculosis complex', some overlapping with either *M. tuberculosis* or *M. bovis*. The relationship of the *M. africanum* strains to *M. tuberculosis* and *M. bovis* was investigated in greater detail by a separate re-evaluation of the combined mass pyrograms of *M. tuberculosis, M. bovis* BCG and *M. africanum* and of *M. bovis* BCG could be well differentiated from the *M. africanum* strains, whereas *M. tuberculosis* and *M. africanum* remained as two very close, but non-overlapping clusters. *Mycobacterium bovis* and *M. africanum* were so closely related that overlap between these two clusters persisted.

To investigate the heterogeneity among the African strains the sub-set of M. africanum mass pyrograms was reanalysed by DA. A plot of the scores of the first discriminant function D1, accounting for 63% of the total variance in this sub-set is given in Fig. 3. The plot shows a clear subclustering tendency among the M. africanum strains. When the coding of the strains was broken a correlation between the two subclusters and the geographical sites of primary isolation was obvious. All strains with low discriminant scores represented isolates from Western Africa (Dakar and Yaoundé), whereas those with high discriminant scores represented Central African



Fig. 1. Ordination of samples from *M. tuberculosis* (\bigcirc), *M. bovis* (\square) and *M. bovis* BCG (\bigcirc) relative to the first and second discriminant function. These functions account for 80% of the total variance in this data set. Each point represents a sample; interconnected samples represent replicate analyses of single strains. Note the non-overlapping distribution of the (sub-)species and the differences in heterogeneity.

isolates (Rwanda and Burundi). From this plot it was also clear that this discriminant function did not differentiate among the strains within each of the two subclusters. For a complete view on the M. africanum heterogeneity additional discriminant functions had to be considered. Ninety-eight percent of the total variance in the subset of M. africanum mass pyrograms was accounted for by the first four discriminant functions. A non-linear map based on the discriminant scores of D1 to D4 is shown in Fig. 4. Again the two subclusters corresponding to the sites of primary isolation are apparent. In addition all strains in each of the subclusters are well differentiated, but further subclustering tendencies were not observed.

The discriminant function D1, that described the geographically based heterogeneity among the *M. africanum* strains, represents the linear combination of characters (i.e. mass peaks) shown in the discriminant spectrum of Fig. 5. The positive part (D1+) represents the pattern of mass peaks that were relatively dominant in the Central African strains, whereas the negative part (D1-) represents the pattern of characters that had relatively higher abundancies in the West African strains. In both parts characteristic character sequences could be observed that are indicative of the biochemical origins of the dissimilarities between Central and Western African strains. In D1+ two sequences were obvious, i.e. m/z 42, 68, 70, 81, 82, 97, 98, 116 and m/z43, 56, 58, 68, 82, 84, 96, 98, 114. These sequences can be tentatively assigned to fragments of deoxyribose and ribose, respectively (Meuzelaar *et al.*, 1982) and may point to nucleic acid constituents. The peak at m/z 59 was commonly observed in cytoplasm fractions of mycobacteria (Haverkamp *et al.*, 1983; Wieten, 1983) and points to either electron impact fragments of neutral carbohydrates or nitrogen containing pyrolysis fragments of *N*acetylamino sugars. The peak sequences observed in D1- point to predominantly protein



Fig. 2. Ordination of samples from *M. tuberculosis* (\bigcirc), *M. bovis* (\square) and *M. bovis* BCG (\bigcirc) and *M. africanum* (\blacktriangle), relative to the first and second discriminant function of this data set. The functions account for 67% of the total variance in the data set. Note the elongated clustering of the *M. africanum* strains and the unchanged relationships between *M. tuberculosis*, *M. bovis* and *M. bovis* BCG.



Fig. 3. Ordination of the samples from the *M. africanum* strains relative to the first discriminant function of this data set (63%) of total variance). Individual points represent analysis of individual samples. The plot shows the existence of two sub-clusters of African strains and their correlation to the geographical origin of the strains.



Fig. 4. Non-linear map based on the first four discriminant functions that were calculated for the M. africanum data set (98% of total variance described). The map shows the two geographical subclusters and in addition the existence of dissimilarities among individual strains. Stress: 8%.



Fig. 5. Discriminant spectrum representing the linear combination of characters that form the first discriminant function of the *M. africanum* data set. The upper part of the spectrum shows the set of characters that are relatively dominant in the Rwanda/Burundi subcluster, the lower part, those characters that are relatively dominant in the Dakar/Yaounde subcluster.

Table 1. Identification of the M. africanum strains and the clinical isolates of the 'Tuberculosis complex' by analysis of variance on the basis of six key-masses

Strains were ranked according to the dissimilarity values between each strain and a hypothetical mean spectrum calculated from the reference strains in the 'Tuberculosis complex'. The critical identification value is the product of the reference dissimilarity value and the empirical factor.

| Rank | Strain code | Dissimilarity value | Microbiological identification | Rank | Strain code | Dissimilarity value | Microbiological identification |
|------|----------------|------------------------|-----------------------------------|------|----------------|------------------------|-----------------------------------|
| 1 | 11 | 1 2·74 | M. tuberculosis | 10 | 14 | 70 ·96 | M. bovis |
| 2 | 12 | 21.37 | M. tuberculosis | 11 | 15 | 73.53 | M. bovis |
| 3 | 10 | 24·99 | M. tuberculosis | 12 | 17 | 77.96 | M. bovis BCG |
| 4 | 18 | 27.25 | M. africanum (DI) | 13 | 25 | 80.94 | M. africanum (YII) |
| 5 | 23 | 46.38 | M. africanum (RII) | 14 | 24 | 81.85 | M. africanum (YI) |
| 6 | 16 | 50.94 | M. bovis BCG | 15 | 27 | 101.82 | M. africanum (RIII) |
| 7 | 19 | 51.91 | M. africanum (RI) | 16 | 13 | 121.19 | M. bovis |
| 8 | 21 | 55.65 | M. africanum (DIII) | 17 | 22 | 125.66 | M. africanum (DIV) |
| 9 | 26 | 60.18 | M. africanum (BI) | 18 | 20 | 1 56 ∙08 | M. africanum (DII) |

Reference dissimilarity value, 62.74; empirical correction value, 2.55; critical identification value, 159.99.

constituents (Meuzelaar et al., 1982), e.g. m/z 67, 69, 83 (pyrrole compounds), 92, 106, 120 (alkyl benzenes), 94, 108, 122 (alkyl phenols), 110, 124, 138 (thiophenol compounds) and 34, 48 (H₂S and CH₃-SH).

From a clinical point of view procedures for primary screening of the 'Tuberculosis complex' will also have to be successful for the identification of *M. africanum*. A screening procedure for the 'Tuberculosis complex' that was based on a pyrolysis mass spectrometric identification key has been reported previously (Wieten *et al.*, 1981 *b*). Peak intensities measured on a small set of key masses in the mass pyrograms of unknown strains and a standard set of reference strains were evaluated by one-way analysis of variance. Recently the original set of key masses, i.e. m/z 31, 50, 58, 59, 71, 86 and 98, was modified so as to obtain improved identification results, in particular for *M. bovis*. When this revised set of key-masses, i.e. m/z 50, 59, 71, 81, 98 and 116, was tested on approximately 200 coded strains, 93% were correctly identified, errors of first order not being observed, i.e. not a single strain was incorrectly typed as 'atypical' (Wieten, 1983).

All *M. africanum* strains, as well as the eight recent clinical isolates from the 'Tuberculosis complex' were subjected to the above mentioned key-mass screening procedure. The nine 'Tuberculosis complex' strains that had been used in previous studies as references were again used as such. A copy of the original output (except for the column 'microbiological identification', which was added afterwards), is presented in Table 1. The critical value for identification, 159.99 was calculated from the average reference dissimilarity value and the correction value for the 'Tuberculosis complex' heterogeneity (Wieten *et al.*, 1981*b*). The data reveal that all strains of *M. tuberculosis*, *M. bovis* and *M. bovis* BCG, as well as *M. africanum* strains have dissimilarities below the critical value for identification and were thus identified as belonging to the 'Tuberculosis complex'. In accordance to the results provided by DA (see Fig. 2), the strains DII and DIV that were located furthest away from the centre of the 'Tuberculosis complex', appear in Table 1 as closest to the critical value for identification. The observed heterogeneity among *M. africanum*, which was found to be partly dependent on a similar set of features as the set of key-masses (see Fig. 5, D1+), did not interfere with the identification of these strains as members of the 'Tuberculosis complex'.

DISCUSSION

Pyrolysis has been defined as a transformation of molecules through the agency of heat alone (Levy, 1966). Since molecules tend to break at specific points the pyrolysis fragments can be expected to give characteristic information about the original sample. Each mass pyrogram used

in this study contains about 150 individual characters, which may be partly correlated due to common biochemical origins. The overall composition of closely related bacterial strains can be expected to be more or less identical and therefore dissimilarities between mass pyrograms will be quantitative rather than qualitative. Considering these aspects, a comparison of mass pyrograms by computerized multivariate data analysis is indispensible in order to ensure optimal use of the information enclosed in the patterns. DA can be described as a general applicable procedure that makes simultaneous use of the full information content of the data set. DA applied to Py-MS data approximates to procedures that are commonly used in numerical taxonomy. Notably, DA uses weighting with respect to the ratios of between-group to withingroup variance to obtain the best view of the relationships among groups. Since a group was defined as the set of replicate mass pyrograms of a single bacterial strain, the ultimate effect of weighting will be that the best view is obtained on dissimilarities among individual strains. In other words, a priori knowledge with respect to the (assumed) species status of individual strains has not been used (this in contrast to the application of DA for identification purposes). Applied in this way DA will lead to what may be considered as a 'natural classification'.

By conventional microbiological methods the close relationship between M. africanum strains and the 'Tuberculosis complex' has since long been established and various characters have been observed linking M. africanum to either M. tuberculosis or M. bovis (e.g. Meissner & Schröder, 1969; Van den Groen & Pattyn, 1975; David et al., 1978; Wayne & Diaz, 1979). The present results obtained by means of the key mass identification procedure confirm this close relationship and at the same time lend support to the previously reported effective use of this screening procedure for the identification of the 'Tuberculosis complex'. Thus, although M. africanum strains had not been involved in establishing the empirical, diagnostic mass-key, all strains tested in this study were identified as members of the 'Tuberculosis complex'. It should be noticed however (see Table 1) that the M. africanum strains added considerably to the heterogeneity of the 'Tuberculosis complex'.

Evaluation by DA of the full information content of the mass pyrograms provides a more detailed view on the heterogeneity among the African strains and their taxonomical relationship to the other (sub-)species in this study. The observed correlation between differences among the M. africanum strains and the geographical origin of primary isolation (Figs 3 and 4) is of epidemiological interest. This geographically based heterogeneity represents a major proportion of the total variation among the *M. africanum* strains. This is deduced from the high percentage of the total M. africanum strain variance described by the discriminant function plotted in Fig. 3, i.e. 63%. The interpretation of the dissimilarities between the two geographical clusters in terms of biochemical components, i.e. ribose, deoxyribose and protein components, was based on comparison to model-compound mass pyrograms (Meuzelaar et al., 1982). To allocate these components to specific molecular entities of the intact bacterial cells is still tentative and requires a detailed study of *M. africanum* cellular fractions to be conclusive. Such a study is not presently available, but mass pyrograms from cellular fractions of other mycobacteria (Wieten, 1983) indicate that ribose and deoxyribose components are derived mainly from the cytoplasma and are likely to represent fragments from RNA and DNA, respectively. The protein fragments may originate mainly from the cell wall, but contributions from the cytoplasma can not as yet be excluded.

The taxonomical relationship between the African strains and the 'Tuberculosis complex' can be deduced from Figs 1 and 2. It is important to note that the relationships among M. tuberculosis, M. bovis and M. bovis BCG, as they can be observed from Fig. 1, correspond to previous results from various Py-MS studies (Meuzelaar *et al.*, 1976; Wieten *et al.*, 1979, 1981 a, b, 1982). The reliability of the present classification is also demonstrated by comparison of the results depicted in Figs 1 and 2, from which it can be concluded that addition of the M. africanum mass pyrograms to the data set of Fig. 1, followed by renewed calculation of discriminant functions did not lead to any important changes in the relationships among the 'Tuberculosis complex' strains, thereby illustrating the stability of the classification. From Fig. 2 the majority of the African strains were shown to be closely related to M. bovis, M. tuberculosis and M. bovis BCG, and some of the M. african strains could not be classified separately from the M. bovis and *M. tuberculosis* clusters. On detailed inspection of the data presented in Fig. 2 and of those of the separate DA evaluation of the combined *M. tuberculosis/M. africanum* and *M. bovis/M. africanum* data sets, the African strains emerge as more closely related to *M. bovis* than to *M. tuberculosis*.

Thus, from a taxonomical point of view, it can be stated that the group of strains designated as M. africanum, although heterogeneous, can be clearly regarded as a member of the 'Tuberculosis complex'. But given the observed overlap with in particular M bovis, it can also be concluded that the Py-MS data do not indicate the species status for the African strains. Provided the integrity of the M. africanum group is confirmed by other studies a designation as a subspecies (of M. bovis) may be appropriate.

The results obtained on Py-MS analysis are in general agreement with the previous numerical taxonomy study of David et al. (1978), but some differences should be mentioned. Py-MS showed an intermediate clustering of the larger part of the African strains, a general closer relationship to M. bovis, and finally an obvious difference in overlap of the M. africanum subcluster with M. bovis and M. tuberculosis. With respect to the latter, David et al. (1978) observed that the Rwanda and Burundi strains clustered with M. tuberculosis and the Dakar and Yaoundé strains with M. bovis. In contrast to this, based on Py-MS data the Rwanda and Burundi strains appear more closely related to M. bovis, whereas the Dakar and Yaoundé strains were observed as intermediate. These discrepancies may be explained by different methods that were used to investigate the taxonomic structure of the data, i.e. clustering methods versus a weighted ordination method. Moreover there is no a priori reason to expect two classifications that are based on two different character sets to be congruent, unless the numbers of characters are sufficiently large and ensure considerable overlap. Thus, the generally high level of congruence between the outcome of DA on the multicharacter set of mass pyrograms and the results of the previous numerical taxonomy study was not altogether unexpected, but the observed dissimilarities suggest that pooling of the data sets might lead to an even more stable classification.

Finally a remark with respect to the easy applicability of Py-MS seems justified. To obtain a raw data set, of the size used in this study, no more than one day is needed, including preparation of the Py-MS samples. This high efficiency of Py-MS, combined with the broad applicability in various areas of microbiology (Meuzelaar *et al.*, 1982; Wieten *et al.*, 1983) recommends rewarding use of Py-MS as an independent analytical tool.

The authors are greatly indebted to J. G. Baas for preparing the samples and to Mrs A. Tom and Mrs B. Brandt for performing the pyrolysis mass spectrometric analysis. This investigation was supported by the Foundation for Fundamental Research on Matter (F.O.M.) and the Ministry of Health and Enivronmental Hygiene in the Netherlands.

REFERENCES

- CASTETS, M., RIST, N. & BOISVERT, H. (1969). La variété africaine du bacille tuberculeux humain. Médecine d'Afrique Noire 16, 321–322.
- DAVID, H. L., JAHAN, M. T., JUMIN, A., GRANDRY, J. & LEHMAN, E. H. (1978). Numerical taxonomy analysis of Mycobacterium africanum. International Journal of Systematic Bacteriology 28, 467–472.
- ESHUIS, W., KISTEMAKER, P. G. & MEUZELAAR, H. L. C. (1977). Some numerical aspects of reproducibility and specificity. In *Analytical Pyrolysis*, pp. 151–166. Edited by C. E. R. Jones & C. G. A. Cramers. Amsterdam: Elsevier.
- GUTTERIDGE, C. S. & NORRIS, J. R. (1979). The application of pyrolysis techniques to the identification of microorganisms. *Journal of Applied Bacteriology* 47, 5-43.
- HAVERKAMP, J., WIETEN, G., BOERBOOM, A. J. H.,

DALLINGA, J. W. & NIBBERING, N. M. (1983). Pyrolysis collisional activated dissociation mass spectrometry of organic model compounds and bacterial samples. In *Proceedings of the Fifth International Symposium on Analytical Pyrolysis*. Edited by K. J. Voorhees & H. L. C. Meuzelaar. Seven Oaks: Butterworth (in the Press).

- IRWIN, W. J. (1979). Analytical pyrolysis an overview (Part II). Journal of Analytical and Applied Pyrolysis 1, 89–122.
- KRUSKAL, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrica* 29, 1–29.
- LEVY, R. L. (1966). Pyrolysis gas chromatography, review of the technique. *Chromatographic Reviews* 8, 48-89.
- MEISSNER, G. & SCHRÖDER, K. (1969). Ueber sog.

afrikanische Mykobacterienstämme aus dem tropischen West-Afrika. Zentralblatt für Bakteriologie, Parisitenkunde, Infektionskrankheiten und Hygiene (Abteilung I, Originale A) 211, 69-81.

- MEUZELAAR H. L. C., KISTEMAKER, P. G., ESHUIS, W. & ENGEL, H. W. B. (1977). Progress in automated and computerized characterization of microorganisms by pyrolysis mass spectrometry. In *Rapid Methods and Automation in Microbiology*, pp. 225– 230. Edited by S. W. B. Newson & H. H. Johnston. Oxford: Learned Information.
- MEUZELAAR, H. L. C., HAVERKAMP, J. & HILEMAN, F. D. (1982). Pyrolysis Mass Spectrometry of Recent and Fossil Biomaterials. Amsterdam: Elsevier.
- VAN DEN GROEN, G. & PATTYN, S. R. (1975). Comparison of Mycobacterium africanum, M. tuberculosis and M. bovis by their utilization of carbon and nitrogen sources. Annales Société Belge de Médicin Tropique 55, 647-651.
- WAYNE, L. & DIAZ, G. A. (1979). Reciprocal immunologic distances of catalase derived from Mycobacterium avium, M. tuberculosis and closely related species. International Journal of Systematic Bacteriology 29, 19-24.
- WIETEN, G. (1983). Studies on classification and identification of mycobacteria by pyrolysis mass spectrometry. Ph.D. thesis, FOM Institute, Amsterdam.
- WIETEN, G., HAVERKAMP, J., ENGEL, H. W. B. & TARNOK, I. (1979). Pyrolysis mass spectrometry in mycobacterial taxonomy and identification. In *Twenty-five Years of Mycobacterial Taxonomy*, pp. 171–189. Edited by G. P. Kubica, L. G. Wayne & L. S. Good. Atlanta: CDC Press.

- WIETEN, G., HAVERKAMP, J., MEUZELAAR, H. L. C., ENGEL, H. W. B. & BERWALD, L. G. (1981a). Pyrolysis mass spectrometry: a new method to differentiate between the mycobacteria of the 'Tuberculosis Complex' and other mycobacteria. *Journal of General Microbiology* 122, 109-118.
- WIETEN, G., HAVERKAMP, J., ENGEL, H. W. B. & BERWALD, L. G. (1981b). Application of pyrolysis mass spectrometry to the classification and identification of mycobacteria. *Reviews of Infectious Diseases* 3, 871–877.
- WIETEN, G., HAVERKAMP, J., BERWALD, L. G., GROOTHUIS, D. G. & DRAPER, P. (1982). Pyrolysis mass spectrometry: its applicability to mycobacteriology, including *Mycobacterium leprae*. Annales de Microbiology 133B, 15-27.
- WIETEN, G., MEUZELAAR, H. L. C. & HAVERKAMP, J. (1983). Analytical pyrolysis in clinical and pharmaceutical microbiology. In Gas Chromatography/ Mass Spectrometry Applications to Microbiology. Edited by G. Odham, L. Larsson & P. A. Märdh. London: Plenum Press (in the Press).
- WINDIG, W., KISTEMAKER, P. G. & HAVERKAMP, J. (1981). Chemical interpretation of differences in pyrolysis mass spectra of simulated mixtures of biopolymers by factor analysis with graphical rotation. Journal of Analytical and Applied Pyrolysis 3, 199-212.
- WINDIG. W., HAVERKAMP, J. & KISTEMAKER, P. G. (1983). Chemical interpretation of sets of pyrolysis mass spectra by discriminant analysis and graphical rotation. *Analytical Chemistry* 55, 81–87.