CURRENT MICROBIOLOGY Vol. 32 (1996), pp. 77–84

Current Microbiology
An International Journal

Identification and Discrimination of Oral Asaccharolytic *Eubacterium* spp. by Pyrolysis Mass Spectrometry and Artificial Neural Networks

Royston Goodacre,[1] Sarah J. Hiom,[2] Sarah L. Cheeseman,[2] David Murdoch,[3] Andrew J. Weightman,[4] William G. Wade[2]

[1]Institute of Biological Sciences, University of Wales, Aberystwyth, Dyfed, SY23 3DA, UK
[2]Department of Oral and Dental Science, University of Bristol, Lower Maudlin Street, Bristol, BS1 2LY, UK
[3]Department of Medical Microbiology, Bristol Royal Infirmary, Bristol, BS2 8HW, UK
[4]School of Pure and Applied Biology, University of Wales College of Cardiff, PO Box 915, Cardiff, CF1 3TL, UK

Abstract. Curie-point pyrolysis mass spectra were obtained from 29 oral asaccharolytic *Eubacterium* strains and 6 abscess isolates previously identified as *Peptostreptococcus heliotrinreducens*. Pyrolysis mass spectrometry (PyMS) with cluster analysis was able to clarify the taxonomic position of this group of organisms. Artificial neural networks (ANNs) were then trained by supervised learning (with the back-propagation algorithm) to recognize the strains from their pyrolysis mass spectra; all *Eubacterium* strains were correctly identified, and the abscess isolates were identified as un-named *Eubacterium* taxon C_2 and were distinct from the type strain of *P. heliotrinreducens*. These results demonstrate that the combination of PyMS and ANNs provides a rapid and accurate identification technique.

The oral asaccharolytic *Eubacterium* species are a diverse group of organisms that have been implicated in periodontitis [23, 29, 31], advanced caries [21], and dentoalveolar abscesses [32]. Their slow growth on complex artificial media and requirement for strict anaerobiosis make them difficult to isolate, and in consequence they have been poorly studied. In addition to *Eubacterium brachy*, *Eubacterium nodatum*, *Eubacterium timidum*, and *Eubacterium saphenum*, a number of other groups have been described on the basis of conventional biochemical tests and protein profiles [20, 23, 28–30]. One of these groups, *Eubacterium* C_2, has been strongly associated with oral abscesses [32]. Other workers have identified a group of Gram-positive anaerobic cocco-bacilli isolated from abscesses at various body sites as *Peptostreptococcus heliotrinreducens*, on the basis of enzyme profiles [24]. The profiles described for these strains resembled those of *Eubacterium* C_2 [30].

The ideal method for the examination of the relationships between bacterial strains would have minimum sample preparation, would analyze samples directly (that is, would not require reagents), would be rapid, automated, quantitative, and relatively cheap. PyMS is an instrument-based technique that satisfies these requirements. Pyrolysis is the thermal degradation of complex molecules in a vacuum causing their cleavage to smaller, volatile fragments separable on the basis of their mass-to-charge ratio (m/z) so as to produce a pyrolysis mass spectrum, which can then be used as a ''chemical profile'' or fingerprint of the complex material analyzed. PyMS has been applied to the characterization and identification of a variety of microorganisms and their products [7, 17, 22] and, because of its high discriminatory ability [8], presents a powerful fingerprinting technique applicable to any (micro-)biological material. One of the major advantages that PyMS has over other diagnostic methods, such as ELISA [3] and nucleic acid probing [27], is that it is rapid, both for a single sample (typical sample time is less than 2 min) and in the (automated) throughput of samples (300 or more in a working day).

Recently there have been a number of important

Correspondence to: W.G. Wade

advances in the field of PyMS, notably the application of the new numerical techniques of artificial neural networks (ANNs) [19, 26, 33] to pyrolysis mass spectra to gain quantitative, as well as qualitative, information about the chemical constituents of microbial (and other) samples analyzed. The first demonstration of the ability of ANNs to discriminate between biological samples from their pyrolysis mass spectra was by Goodacre et al. [10, 11], who successfully used PyMS and ANNs for the assessment of the presence of lower-grade seed oils as adulterants in extra virgin olive oils. It has now also been shown that this supervised learning method is effective for the rapid identification of strains of *Mycobacterium* [6], *Propionibacterium* [14], and *Streptomyces* [5]. It is also possible to use the combination of ANNs and PyMS for the quantification of indole production in bacteria [9], of recombinant protein production in whole cells of *Escherichia coli* [12], of tertiary mixtures of microbial cultures [13], and in the rapid and quantitative screening of cultures and fermentor broths for metabolite overproduction [15].

The aim of this study was to use PyMS to examine a collection of oral asaccharolytic *Eubacterium* strains, together with abscess isolates previously identified as *P. heliotrinreducens*, in order to clarify the taxonomic position of this group. Furthermore, we investigated the ability of artificial neural networks to identify *Eubacterium* spp. by their pyrolysis mass spectra.

Materials and Methods

Organisms and cultivation. Details and origins of the organisms are given in Table 1. Strains were cultured on Fastidious Anaerobe agar (Lab M, Bury, UK) plus 5% sheep blood and incubated anaerobically in an anaerobic cabinet in an atmosphere of N_2 80%, CO_2 10%, H_2 10% for 72 h. The bacteria were harvested with a nichrome wire loop and suspended in phosphate-buffered saline to 20 mg/ml. The samples were then frozen at −20°C until they were analyzed by PyMS at the University of Wales, Aberystwyth.

Strains were tested in the API Rapid ID32A kit and for metabolic end products by gas chromatography as previously described [30].

Pyrolysis mass spectrometry. 5-μl aliquots of the bacterial suspensions were evenly applied onto iron-nickel foils. Prior to pyrolysis the samples were oven-dried at 50°C for 30 min. Samples were run in triplicate. The pyrolysis mass spectrometer used was the Horizon Instruments PYMS-200X (Horizon Instruments Ltd., Ghyll Industrial Estate, Heathfield, E. Sussex) as described by Aries and associates [1]. Full operational details have been described previously [13, 14]. The sample tube carrying the foil was heated, prior to pyrolysis, at 100°C for 5 s. Curie-point pyrolysis was at 530°C for 3 s, with a temperature rise time of 0.5 s. Data were collected over the m/z range 51 to 200 and normalized to the total ion count to remove any influence of sample size. Three replicate spectra were obtained for each strain.

Multivariate data analysis. The normalized data were processed with the GENSTAT package [25] as previously described [8]. In essence, the first stage was the reduction of the data by principal components analysis, keeping only those principal components whose eigen values accounted for more than 0.1% of the total variance. Canonical variates analysis then separated the samples into groups on the basis of the retained principal components and some a priori knowledge of the appropriate number of groupings (that is, the replicate pyrolysis mass spectra). Next, a percentage similarity matrix was constructed by transforming the Mahalanobis' distance between a priori groups in canonical variates analysis with the Gower similarity coefficient S_G [16]. Finally, hierarchical cluster analysis was employed to produce a dendrogram, using average linkage clustering [18].

Artificial neural networks. All ANN analyses were carried out with a user-friendly neural network simulation program, NeuralDesk version 1.2 (Neural Computer Sciences, Lulworth Business Centre, Nutwood Way, Totton, Southampton, Hants), which runs under Microsoft Windows NT (and Windows 3.1) on an IBM-compatible PC. In-depth descriptions of the modus operandi are given elsewhere [9, 13, 14].

The algorithm used was standard back-propagation [26]. This algorithm employed processing nodes (neurons or units) linked by abstract interconnections (connections or synapses). Connections each have an associated real value, termed the weight, that scaled signals passing through them. Nodes summed the signals feeding to them and output this sum to each driven connection scaled by a ''squashing'' function.

For training the ANNs, each of the inputs was the normalized triplicate pyrolysis mass spectrum derived from the six type strains and the three cluster groups (details are given in Table 2) and was paired with each of the desired outputs. These were binary encoded such that *E. nodatum* was coded as 100000000, *E. timidum* as 010000000, *E. brachy* as 001000000, *Eubacterium* C_1 as 000100000, *Eubacterium* C_2 as 000010000, *Eubacterium* New 1 as 000001000, *E. lentum* as 000000100, *E. saphenum* as 000000010, and *P. heliotrinreducens* as 000000001. These nine training pairs collectively made up the training set. The input was applied to the network, which was allowed to run until an output was produced at each output node. The differences between the actual and the desired output, taken over the entire training set, were fed back through the network in the reverse direction to signal flow (hence back-propagation) modifying the weights as they went. This process was repeated until an acceptable level of error was achieved.

The structure of the ANN used in this study to analyze pyrolysis mass spectra consisted of three layers: 150 input nodes, 9 output nodes (one for each strain), and one ''hidden'' layer containing 8 nodes (i.e., a 150-8-9 architecture). Before training commenced, the values applied to the input and output nodes were normalized between 0 and 1, and the connection weights were set to small random values [33]. Each epoch represented 1289 connectionweight updatings and a recalculation of the root mean squared (RMS) error between the true and desired outputs over the entire training set. A plot of the RMS error vs. the number of epochs represents the ''learning curve'' and was used to estimate the extent of training. Finally, after training to an RMS error of 0.001, all 35 pyrolysis mass spectra were used as the ''unknown'' inputs (test data); the network then calculated its estimate, and for each sample the winning node in the output layer was taken as its identity.

Table 1. Metabolic end-products and API rapid ID32 A profiles of bacterial strains included in the study

Identifier[a]	Species/group	Strain number	Metabolic end-products	API rapid ID32A profile	Source (reference)[b]
A	*E. nodatum:*	ATCC 33099^T	a b l s	0000 0020 00	ATCC
B		W2192	a b l s	0000 0000 00	BDH
C	*E. timidum:*	ATCC 33093^T	l s pa	0000 0200 00	ATCC
D		W557	l s pa	0000 0200 00	BDH
E		W690	l s pa	0000 0240 00	BDH
F		W693	l s pa	0000 0240 00	BDH
G		W2847	l s pa	0000 0200 00	BDH
H	*E. brachy:*	ATCC 33089^T	a ib iv ic	0000 0000 00	ATCC
I		W858	a ib iv ic	0000 0000 00	BDH
J	*Eubacterium* C_1:	W1471	b	0000 0400 00	BDH [32]
K		W687	a b l s	0000 0400 00	BDH
L		W1475	b l s	0000 0400 00	BDH
M		W1470	a b l s	0000 0400 00	BDH
N	*Eubacterium* C_2:	SC142	l s	2000 0737 05	BDH [32]
O		SC108	l s	2000 0737 05	BDH
P		W1365	l s	2000 0737 05	BDH
Q		W733	l s	2000 0737 05	BDH
R		W2848	l s	2000 0237 05	
S	*Eubacterium* New 1:	SC68	a b l s	0000 0120 01	BDH [4]
T		SC88P	a b l s	0000 0120 01	BDH
U		SC41B	a b l s	0000 0120 01	BDH
V		SC37	a b l s	0000 0120 01	BDH
W		SC87K	a b l s	0000 0120 01	BDH
X	Identity unknown:	SC3D	l s	0000 0100 00	BDH
Y	*E. lentum:*	NCTC 11813^T	—	2000 0000 00	NCTC
Z	*Eubacterium* S-group:	15P7	l s	2000 0337 05	Dr T. Sato [28]
a	*E. saphenum:*	ATCC 49989^T	b l s	0000 0000 00	Dr T. Sato
b	*P. heliotrinreducens:*	NCTC 11029^T	l s	2000 0237 05	NCTC
c		SBH463	l s	2000 0237 05	Strains from abscesses isolated at St Bartholomews Hospital, London
d	*P. heliotrinreducens:*	SBH481	l s	2000 0237 05	
e		SBH462	l s	2000 0237 05	
f		SBH403	l s	2000 0237 05	
g		SBH477	l s	2000 0237 05	
h	Identity unknown:	W1219	l s	2000 0237 05	BDH
i	Identity unknown:	GF10	l s	2000 0237 05	Dr. B Olsvik

[a] This was the identifier for the GENSTAT analysis and is referred to in Fig. 1.
[b] ATCC = American Type Culture Collection; BDH = Bristol Dental Hospital.

Results and Discussion

Each of the 35 strains, each represented by three replicate spectra, was coded to give 35 individual groups (see Table 1); the resulting canonical variates (CVA) plot and dendrogram are shown in Figs. 1 and 2, respectively. In the dendrogram (Fig. 2) it can be seen that at 95% relative similarity the bacteria were grouped into five clusters: Cluster 1 was the only heterogeneous group and comprised the type strains of *E. lentum*, *E. saphenum*, *P. heliotrinreducens*, the two *E. nodatum* strains, the five *Eubacterium* C_2 strains, the *Eubacterium* S-group strain, the five clinical abscess isolates, and the strains SC3D, W1219, and GF10; Cluster 2 consisted of the five *E. timidum* strains; Cluster 3 comprised the four strains of *Eubacterium* C_1 [32]; Cluster 4, the five strains of *Eubacterium* New 1 [4]; finally, the two strains of *E. brachy* were quite different from the other strains (only 40% relative similarity) and were recovered together in Cluster 5. These clusters can also be seen in the CVA plot (Fig. 1), where 87.8% of the total variation is displayed. Although Cluster 1 and Cluster 2 appear to be mixed in this CVA plot, they can easily be separated if a different 3D angle is viewed (results not shown).

Because of the heterogeneity of the 19 strains in Cluster 1 at this level of differentiation, the relationship between these strains was investigated further (Fig. 3). It can be seen that, with the exception of *Eubacterium* C_2 strains SC142 and W1365, which

Table 2. Identity of the bacteria used in the test set and training set as judged by artificial neural networks

Strain	Estimates from artificial neural networks (standard deviation)[a]									Percentage error[b]
	E. nodatum	*E. timidum*	*E. brachy*	*E.* C_1	*E.* C_2	*E.* New 1	*E. lentum*	*E. saphenum*	*P. heliotrinr.*	
E. nodatum ATCC 33099^T [c]	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06
E. nodatum W2192	0.98	0.00	0.05 (0.09)	0.00	0.00	0.00	0.00	0.00	0.00	0.76
E. timidum ATCC 33093^T [c]	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07
E. timidum W557	0.00	1.00	0.00	0.00	0.02	0.01	0.00	0.00	0.01	0.46
E. timidum W690	0.07 (0.11)	0.88 (0.14)	0.08 (0.10)	0.39 (0.42)	0.00	0.00	0.00	0.00	0.00	7.33
E. timidum W693	0.08 (0.12)	0.88 (0.13)	0.14 (0.16)	0.37 (0.41)	0.00	0.00	0.00	0.00	0.00	7.93
E. timidum W2847	0.00	1.00	0.00	0.00	0.02	0.00	0.00	0.00	0.01	0.35
E. brachy ATCC 33089^T [c]	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
E. brachy W858	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08
Eubacterium C_1 W1471[c]	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.06
Eubacterium C_1 W687	0.00	0.00	0.00	1.00	0.00	0.01	0.00	0.00	0.00	0.12
Eubacterium C_1 W1475	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.01	0.00	0.15
Eubacterium C_1 W1470	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.13
Eubacterium C_2 SC142[c]	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.06
Eubacterium C_2 SC108	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.10
Eubacterium C_2 W1365	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.07
Eubacterium C_2 W733	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.09
Eubacterium C_2 W2848	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.18
Eubacterium New 1 SC68[c]	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.06
Eubacterium New 1 SC88P	0.00	0.00	0.00	0.00	0.00	0.92	0.00	0.00	0.00	0.92
Eubacterium New 1 SC41B	0.00	0.00	0.00	0.02	0.00	0.99	0.00	0.00	0.00	0.39
Eubacterium New 1 SC37	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.11
Eubacterium New 1 SC87K	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.11
E. lentum NCTC 11813^T [c]	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.08
E. saphenum ATCC 49989^T [c]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.07
P. heliotrinr. NCTC 11029^T [c]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.06
SBH463	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.01	0.32[d]
SBH481	0.00	0.00	0.00	0.00	0.85 (0.09)	0.00	0.07 (0.11)	0.00	0.09 (0.11)	3.47[d]
SBH462	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.17[d]
SBH403	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.17[d]
SBH477	0.00	0.00	0.00	0.00	1.00	0.01	0.01	0.00	0.00	0.27[d]
Eubacterium S-group 15P7	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.07[d]
W1219	0.07 (0.07)	0.08 (0.17)	0.00	0.57 (0.38)	0.00	0.00	0.00	0.00	0.00	—
GF10	0.13 (0.09)	0.00	0.00	0.20 (0.41)	0.00	0.00	0.15 (0.31)	0.04	0.00	—
SC3D	0.11 (0.13)	0.00	0.00	0.00	0.00	0.00	0.03	0.04	0.00	—

[a] The values given are the averages from training five ANNs with different random starting weights. The values in brackets are the standard deviations for the averages and are displayed only if the standard deviation was greater than 0.05.
[b] This percentage error was calculated between the estimates from ANNs and the output that was expected. No values are given for SC3D, W1219, and GF10 since no identy was known.
[c] The pyrolysis mass spectra from these bacteria were used to train five ANNs; the values above are the averages of five different ANNs.
[d] For the abscess isolates and *Eubacterium* S-group 15P7 the expected output was thought to be to *Eubacterium* C_2; the percentage error was, therefore, calculated using this.

were recovered together at >95% relative similarity, all the *Eubacterium* C_2 strains, the abscess isolates, and the *Eubacterium* S-group strain were recovered in one sub-cluster at a level of >95% relative similarity. The type strains of *E. saphenum*, *E. lentum*, and *P. heliotrinreducens* were recovered separately. The unidentified strains W1219 and GF10 clustered together and were 92% similar to SC3D; this group was 87%

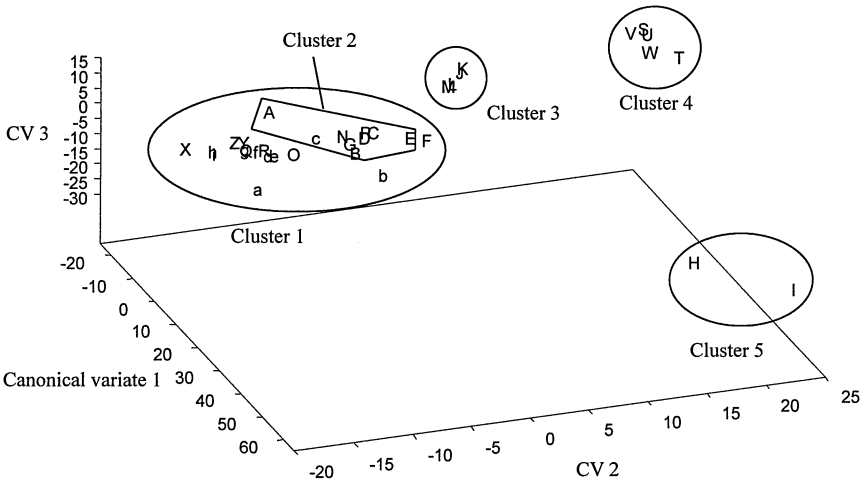

Fig. 1. Pseud-3D canonical variates analysis plot based on PyMS data analyzed by GENSTAT showing the relationship between the 34 *Eubacterium* strains and *Peptostreptococcus heliotrinreducens*. The first, second, and third canonical variates accounted for 73.7%, 8.4%, and 5.7% (total 87.8%) of the total variance respectively. The identity of the strains encoded as letters are listed in Table 1.

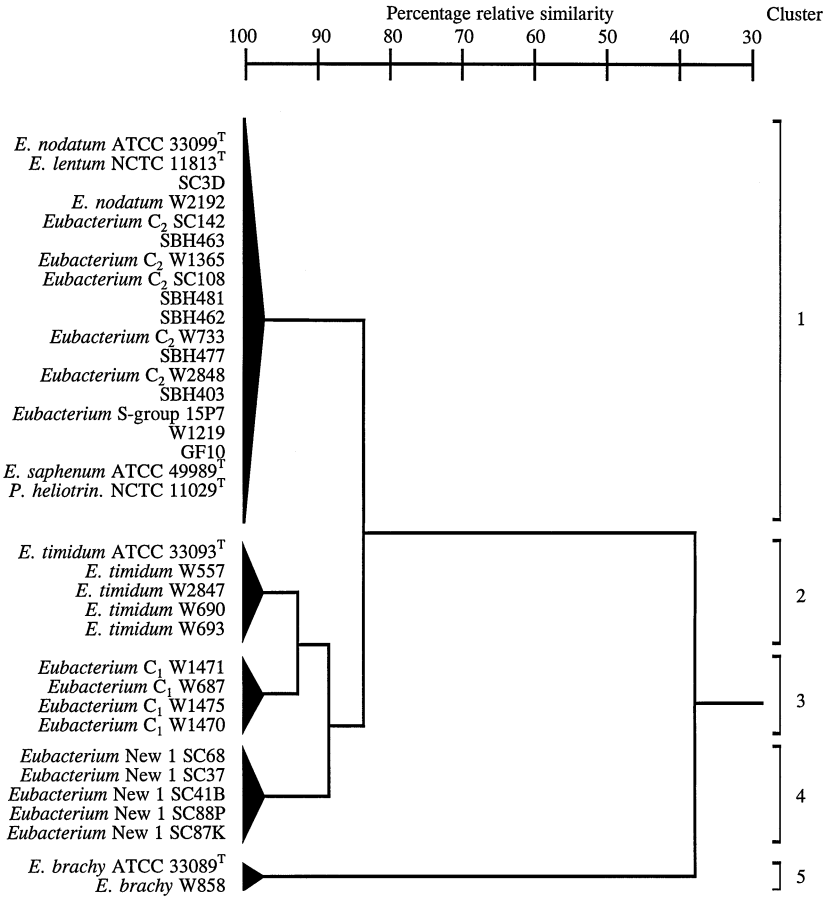

Fig. 2. Dendrogram representing the relationships between the 34 *Eubacterium* strains and *Peptostreptococcus heliotrinreducens*, based on PyMS data analyzed by GENSTAT.

similar to *E. nodatum* ATCC 33099^T. The two *E. nodatum* strains did not cluster together, and strain W2192 was outlying. PyMS is a technique that highlights small phenotypic differences between bacteria; in order to ascertain whether the results observed were because W2192 was contaminated or whether these two strains are phylogenetically different, DNA homology studies or 16S rRNA sequencing would have to be performed [2].

The clustering obtained from the PyMS spectra was entirely consistent with the metabolic end product and enzyme profile data (Table 1), further confirming the value of the technique for taxonomic use.

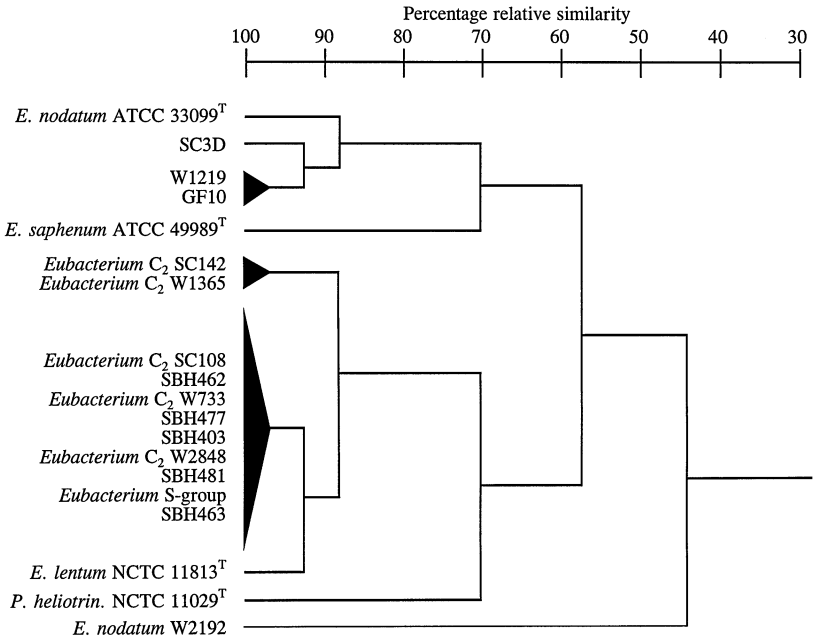

Fig. 3. Dendrogram representing the relationships between the 19 strains in Cluster 1 (from Figs. 1 and 2), based on PyMS data analyzed by GENSTAT.

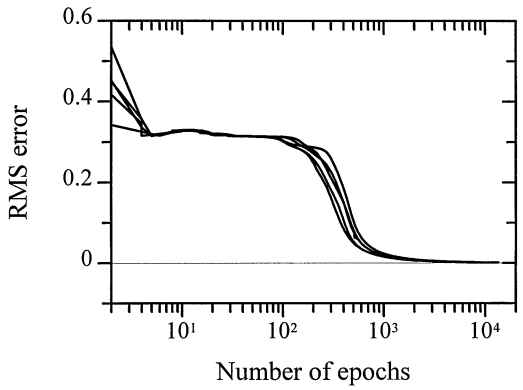

Fig. 4. The learning curves for 150-8-9 artificial neural networks employing the standard-back propagation algorithm, trained to assess the identity of the *Eubacterium* strains. Training was carried out five times with different randomized starting connection weights.

To identify the abscess isolates, ANNs were trained with the normalized ion intensities from the averaged pyrolysis mass spectra from the training set (the six type strains and the three cluster groups; Table 2). The nine strains were binary encoded at the 9 output nodes as described above. The 150-8-9 ANNs were trained with the standard back-propagation algorithm, and the effectiveness of training was expressed in terms of the average RMS error between the actual and the desired outputs; examples of these ''learning curves'' are shown in Fig. 4. Training was stopped after the average error had reached 0.001, and the network was interrogated by use of the pyrolysis mass spectral data from all bacteria from both the training and test sets. Training was carried out five times, using randomized, small initial values for the starting weights. The learning curves from these runs were found essentially to superimpose, despite the randomized starting connection weights, showing that training was executed in a reproducible and rapid manner. To reach an RMS error of 0.001, training typically took between $10 \cdot 10^3$ and $13 \cdot 10^3$ epochs, and the actual time taken to train was only 6–8 min.

When training had ceased, the network was interrogated and the estimate output for each sample. As expected, the network's estimate of the bacterial identity of the training set was the same as the known identities (Table 2). The results of the network's final analysis of the unknown test set (given as the average of the outputs for the five training runs) are shown in Table 2; the percentage error between the ANNs estimates and the output that was expected is also given. The identity of the known *Eubacterium* strains was correctly judged for *E. brachy*, *Eubacterium* C_1, *Eubacterium* C_2, *Eubacterium* New 1, and *E. nodatum*, with less than 1% error in these estimates. It is interesting that *E. nodatum* W2192 was correctly identified despite being recovered separately from the type strain in the dendrogram (Fig. 3). Although all of the *E. timidum* strains were correctly identified, for strains W690 and W693 the error was 7.33% and 7.93% respectively, and the node for *E. timidum*

scored 0.88 while the node for *Eubacterium* C_1 scored approximately 0.4. These strains were identified as *E. timidum* because the node's activation for *E. timidum* was significantly greater than for any of the others. It is noteworthy that in the dendrogram of these strains (Fig. 2) the *E. timidum* and *Eubacterium* C_1 strains clustered closely together and were 92% similar. The five abscess isolates were identified non-subjectively as *Eubacterium* C_2. SBH481 scored 0.85 at the *Eubacterium* C_2 node; this was again taken to be its identity because this was the node with the highest score. *Eubacterium* S-group 15P7 was also unequivocally identified as *Eubacterium* C_2. Strains SC3D and GF10 were not identified since none of the nodes were activated. W1219 was also unidentified even though the activation at the *Eubacterium* C_1 node was 0.57 because the standard deviation in this value was high (0.38).

The abscess isolates studied here had been previously identified as *P. heliotrinreducens*, which is a species normally isolated from the sheep rumen [24]. The present study demonstrated that these strains were dissimilar from the type strain of *P. heliotrinreducens* but, together with the *Eubacterium* S-group strain, clustered with the un-named taxon *Eubacterium* C_2, which has been found frequently in oral abscesses [32].

This study clearly showed that PyMS can discriminate between the oral asaccharolytic *Eubacterium* spp. and that there was congruence between the PyMS results and the proposed new groupings for *Eubacterium* C_1 and *Eubacterium* C_2 [30] and *Eubacterium* New 1 [4]. Back propagation neural networks were also trained successfully with PyMS data to identify these strains. We conclude that the combination of PyMS and ANNs provides an objective, rapid, and accurate discriminatory technique.

ACKNOWLEDGMENTS

R. Goodacre was supported by the Chemicals & Pharmaceuticals Directorate of the UK BBSRC, under the terms of the LINK scheme in Biochemical Engineering, in collaboration with Horizon Instruments, Neural Computer Sciences, and Zeneca Bioproducts plc. S.L. Cheeseman was in receipt of an MRC Research Studentship.

Literature Cited

1. Aries RE, Gutteridge CS, Ottley TW (1986) Evaluation of a low-cost, automated pyrolysis mass spectrometer. J Anal Appl Pyrol 9:81–98
2. Austin B, Priest F (1986) Modern bacterial taxonomy. Wokingham: Van Nostrand Reinhold, pp 74–81
3. Chantler SM, McIllmurray MB (1987) Labelled-antibody methods for the detection and identification of microorganisms. Methods Microbiol 19:273–332
4. Cheeseman SL, Weightman AJ, Wade WG (1995) 16S ribosomal ribonucleic acid sequence comparisons of oral asaccharolytic *Eubacterium* spp. In: Duerden BI, Brazier J, Eley A, Hudson M, Wade WG (eds) Medical and dental aspects of anaerobes. Cambridge: Science Reviews, in press
5. Chun J, Atalan E, Ward AC, Goodfellow M (1993) Artificial neural network analysis of pyrolysis mass-spectrometric data in the identification of *Streptomyces* strains. FEMS Microbiol Lett 107:321–325
6. Freeman R, Goodacre R, Sisson PR, Magee JG, Ward AC, Lightfoot NF (1994) Rapid identification of species within the *Mycobacterium tuberculosis* complex by artificial neural net work analysis of pyrolysis mass spectra. J Med Microbiol 40:170–173
7. Goodacre R (1994) Characterisation and quantification of microbial systems using pyrolysis mass spectrometry: introducing neural networks to analytical pyrolysis. Microbiol Eur 2:16–22
8. Goodacre R, Berkeley RCW (1990) Detection of small genotypic changes in *Escherichia coli* by pyrolysis mass spectrometry. FEMS Microbiol Lett 71:133–138
9. Goodacre R, Kell DB (1993) Rapid and quantitative analysis of bioprocesses using pyrolysis mass spectrometry and neural networks: application to indole production. Anal Chim Acta 279:17–26
10. Goodacre R, Kell DB, Bianchi G (1992) Neural networks and olive oil. Nature 359:594
11. Goodacre R, Kell DB, Bianchi G (1993) Rapid assessment of olive oil adulteration using pyrolysis mass spectrometry and artificial neural networks. J Sci Food Agric 63:297–307
12. Goodacre R, Karim A, Kaderbhai MA, Kell DB (1994a) Rapid and quantitative analysis of recombinant protein expression using pyrolysis mass spectrometry and artificial neural networks: application to mammalian cytochrome b_5 in *Escherichia coli*. J Biotechnol 34:185–193
13. Goodacre R, Neal MJ, Kell DB (1994b) Rapid and quantitative analysis of the pyrolysis mass spectra of complex binary and tertiary mixtures using multivariate calibration and artificial neural networks. Anal Chem 66:1070–1085
14. Goodacre R, Neal MJ, Kell DB, Greenham LW, Noble WC, Harvey RG (1994c) Rapid identification using pyrolysis mass spectrometry and artificial neural networks of *Propionibacterium acnes* isolated from dogs. J Appl Bacteriol 76:124–134
15. Goodacre R, Trew S, Wrigley-Jones C, Neal MJ, Maddock J, Ottley TW, Porter N, Kell DB (1994d) Rapid screening for metabolite overproduction in fermentor broths using pyrolysis mass spectrometry with multivariate calibration and artificial neural networks. Biotechnol Bioeng 44, in press
16. Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857–874
17. Gutteridge CS (1987) Characterization of microorganisms by pyrolysis mass spectrometry. Methods Microbiol 19:227–272
18. Gutteridge CS, Vallis L, MacFie HJH (1985) Numerical methods in the classification of microorganisms by pyrolysis mass spectrometry. In: Goodfellow M, Jones D, Priest F (eds) Computer-assisted bacterial systematics. London: Academic Press, pp 369–401
19. Hecht-Nielsen R (1990) Neurocomputing. Reading, Massachusetts: Addison-Wesley
20. Hill GB, Ayers OM, Kohan AP (1987) Characteristics and sites of infection of *Eubacterium nodatum*, *Eubacterium timidum*,

Eubacterium brachy, and other asaccharolytic *Eubacteria*. J Clin Microbiol 25:1540–1545
21. Hoshino E (1985) Predominant obligate anaerobes in human carious dentine. J Dent Res 64:1195–1198
22. Meuzelaar HLC, Haverkamp J, Hileman FD (1982) Pyrolysis mass spectrometry of recent and fossil biomaterials. Amsterdam: Elsevier
23. Moore WEC, Holdeman LV, Cato EP, Smibert RM, Burmeister JA, Ranney RR (1983) Bacteriology of moderate (chronic) periodontitis in mature adult humans. Infect Immun 52:510–515
24. Murdoch DA, Mitchelmore IJ (1989) Isolation of *Peptostreptococcus heliotrinreducens* from human polymicrobial abscesses. Lett Appl Microbiol 9:223–225
25. Nelder JA (1979) Genstat Reference Manual. Scientific and Social Service Program Library: University of Edinburgh
26. Rumelhart DE, McClelland JL, The PDP Research Group (1986) Parallel distributed processing. Experiments in the microstructure of cognition. Cambridge, Mass. MIT Press
27. Saano A, Lindström K (1990) Detection of rhizobia by DNA-DNA-hybridization from soil samples: problems and perspectives. Symbiosis 8:61–73
28. Sato T, Hoshino E, Uematsu H, Noda T (1993) Predominant obligate anaerobes in necrotic pulps of human deciduous teeth. Microb Ecol Health Dis 6:269–275
29. Uematsu H, Hoshino E (1992) Predominant obligate anaerobes in human periodontal pockets. J Periodont Res 27:15–19
30. Wade WG, Slayne MA, Aldred MJ (1990) Comparison of identification methods for asaccharolytic *Eubacterium* species. J Med Microbiol 33:239–242
31. Wade WG, Moran J, Morgan JR, Addy M (1992) The effects of antimicrobial acrylic strips on the subgingival microflora in chronic periodontitis. J Clin Periodontol 19:127–134
32. Wade WG, Lewis MAO, Cheeseman SL, Absi EG, Bishop PA (1993) An unclassified *Eubacterium* taxon in acute dentoalveolar abscess. J Med Microbiol 40:115–117
33. Wasserman PD (1989) Neural computing: theory and practice. New York: Van Nostrand Reinhold.