World Academy of Science, Engineering and Technology
International Journal of Bioengineering and Life Sciences
Vol:1, No:5, 2007

An Advanced Approach Based on Artificial Neural Networks to Identify Environmental Bacteria

Mauro Giacomini, Stefania Bertone, Federico Caneva Soumetz, and Carmelina Ruggiero

***Abstract*—** Environmental micro-organisms include a large number of taxa and some species that are generally considered nonpathogenic, but can represent a risk in certain conditions, especially for elderly people and immunocompromised individuals. Chemotaxonomic identification techniques are powerful tools for environmental micro-organisms, and cellular fatty acid methyl esters (FAME) content is a powerful fingerprinting identification technique. A system based on an unsupervised artificial neural network (ANN) was set up using the fatty acid profiles of standard bacterial strains, obtained by gas-chromatography, used as learning data. We analysed 45 certified strains belonging to *Acinetobacter, Aeromonas, Alcaligenes, Aquaspirillum, Arthrobacter, Bacillus, Brevundimonas, Enterobacter, Flavobacterium, Micrococcus, Pseudomonas, Serratia, Shewanella* and *Vibrio* genera. A set of 79 bacteria isolated from a drinking water line (AMGA, the major water supply system in Genoa) were used as an example for identification compared to standard MIDI method. The resulting ANN output map was found to be a very powerful tool to identify these fresh isolates.

***Keywords*—** cellular fatty acid methyl esters, environmental bacteria, gas-chromatography, unsupervised ANN.

I. Introduction

IN many drinking water distribution systems microbiological quality control is limited to the identification of a few bacteria species. These are used as "indicators" of contamination, mostly coliform bacteria, whose identification is recommended [1], [2]. Heterotrophic bacterial count is mainly aimed to reduce interference with the detection of coliform bacteria. However, heterotrophic species that are normally considered nonpathogenic can increase risks in certain conditions. There is a growing belief that the heterotrophic bacteria group may contain opportunistic pathogens especially dangerous for young children, elderly people and immunocompromised individuals [3], [4]. For this reason it would be important to detect the presence of these opportunistic species.

Aquatic bacteria include a large number of taxa. Classification, that is the orderly arrangement of micro-organisms into taxonomic groups based on similarity, and identification, that is the determination as to whether an unknown bacterium belongs to one of the units defined in classification, are difficult problems for most of environmental micro-organisms. Convenient and accurate diagnostic schemes are unavailable for environmental bacteria, taxonomy is continuously rearranged with frequent emendation of description and nomenclature corrections, and many novel groups are continuously proposed, which need to be considered for approval and characterised in greater detail.

For identification purposes, traditional phenotypic and biochemical tests, e.g. API (Bio Merieux SA, Marcy-l'Etoile, France) and Biolog (Biolog, Inc., Hayward, Calif., USA) and systems are used by the majority of microbiological laboratories. However, the scarcity of phenotypic features in particular environmental bacterial groups often causes problems in identifying unknown strains. For these groups, alternative chemotaxonomic or genotypic methods can be useful [5].

Cellular fatty acid analysis by gas-chromatography is a rapid and reliable means for the identification of micro-organisms, provided that strict standardised culture conditions are used. Fatty acids are the major constituents of the lipid bilayer of bacterial membranes and lipopolysaccharides. The composition of cellular fatty acids is a very stable genetic trait and it is highly conserved within a taxonomic group. A large number of fatty acids can be found in bacteria, due to the variability which is present in the cellular fatty acid structure, chain length, double-bond positions and substitutions, the whole cell fatty acid (WCFA) analysis is successfully and extensively used for bacteria classification and identification purposes [6]. The method is rapid, cheap, simple, highly automated and its application is within fingerprinting technology. Fatty acid analysis carries information mostly from genus to species level [5].

Manuscript received July 28, 2003. This work was supported in part by the 'Parco Scientifico e Tecnologico della Liguria' under grant "Sperimentazione di un sistema per il riconoscimento rapido di batteri viventi in acqua".

M. Giacomini is with the University of Genoa, Department of Communication, Computer and System Sciences, Via Opera Pia, 13, 16145 Genoa, Italy (phone: +39-010-353-6546; fax: +39-010-353-2154; e-mail: mauro.giacomini@dist.unige.it).

S. Bertone is with R.I.L.A.B. s.r.l., Via Guerrazzi 24/12B, 16146 Genoa Italy (e-mail: administration@rilab.it).

F. S. Caneva is with University of Genoa, Department of Communication, Computer and System Sciences, Via Opera Pia, 13, 16145 Genoa, Italy (e-mail: fedecnv@dist.unige.it).

C. Ruggiero is with University of Genoa, Department of Communication, Computer and System Sciences, Via Opera Pia, 13, 16145 Genoa, Italy (e-mail: carmel@dist.unige.it).

Many laboratories that use fingerprinting technology, use computing facilities for storage and processing data. One of the most promising groups of classification methods is in the field of Artificial Intelligence, specifically artificial neural network (ANN), which are computational models of the brain that adapt their features during learning in order to reproduce the generalisation ability of the brain [7]-[8]. We have shown that ANN based programs are useful to classify and identify marine bacteria at genus level starting from the fatty acid profiles [9]-[11].

In the present work the performance of the unsupervised Kohonen neural network [12] for the automated identification of drinking water bacterial isolates was evaluated. The choice of this type of networks has been suggested by previous work of one author on biological data [7]. The identification was performed at genus level using the fatty acid profiles of standard strains, obtained by gas-chromatography, as learning data. The unsupervised ANN place the input patterns on a 2D finite plane divided into a finite number of areas (a square space, with an equal number of elements in rows and columns). Similar input examples are placed in the same output element, or in the very near elements. In this way, after learning, areas allocated to specific classes can be singled out. A pattern coming from an unclassified strain is classified according to the zone in which it is placed by the network.

The identification system was tested for the study of the genera composition of the cultivable heterotrophic bacterial community of a drinking water line from AMGA S.p.A., the major water supply system in Genova (Italy). We compared our results to the ones obtained with one of the most similar identification system present on the market: the MIDI Microbial Identification System (MIS; Microbial ID Inc., Newark, Del., USA).

II. Materials and Methods

2.1 Strains and culture conditions

For this study 122 aerobic heterotrophic bacteria were used. To learn the ANN 45 certified standard strains, belonging to 39 fresh water species and 14 genera among the most representative aerobic heterotrophic genera commonly found in fresh water, were used (Table 1).

To test the identification ANN, we used 79 fresh water bacteria coming from a drinking water line from AMGA S.p.A., the major water supply system in Genova (Italy). Fresh water samples were collected by using sterile 500 ml bottles at different sites and times from a drinking water line. From each sample, 1 ml (or a convenient dilution) was plated on agar (yeast extract 3 g, peptone 5 g, NaCl 5 g, agar 15 g, distilled water 1 l, pH 7.4) and incubated at 22 °C. After two or three days, depending on the speed of growth of the bacteria present in the sample, colonies with different morphologies were selected and considered different strains. A total of 79 isolates were chosen for this study.

For the WCFA analysis, growth temperature, cultivation medium and culture conditions were strictly standardised to assure reproducibility in the profile. Bacteria, both standard and environmental isolates, were grown as pure culture on nutrient agar at 22 °C: one single colony was inoculated in 200 ml of nutrient broth (yeast extract 3 g, peptone 5 g, NaCl 5 g, distilled water 1 l, pH 7.4) and incubated at 22 °C until the beginning of the stationary phase [13]. The bacteria were collected by centrifugation at 4,000 rpm for 20 min., washed twice with deionized water and freeze-dried.

2.2 Cellular fatty acid extraction and analysis

At least three analyses for each strain were performed for a total of 370 fatty acid analysed profiles. Fatty acid methyl esters (FAMEs) were extracted from freeze-dried bacteria by the standardised procedure described by Miller [14]. FAMEs were analysed by gas-liquid-chromatography (GLC) on a HP5890A gas chromatograph (Hewlett Packard) equipped with a flame ionization detector and an autosampler as previously described [9]. In short, the GLC settings were as follows: a fused silica capillary column (0.2 mm by 25 m; cross-linked 5 % methyl phenyl silicone; Hewlett Packard) and ultra high purity hydrogen (carrier gas) were used. The values of the other variables were: injector temperature, 250 °C; detector temperature, 300 °C; initial column temperature, 170 °C, increasing by 5 °C/min up to 270 °C in 20 min; carrier gas flow rate 50 ml/min; total analysis time 25 min; sample volume 1ml. The retention time data were used to calculate the equivalent chain length data. The fatty acids with a number of carbon atoms between 10 and 20 were identified by a HP 216 Personal Computer. A calibration mixture for capillary chromatography (Supelco Inc., Bellefonte, PA), containing a selection of methyl esters of the fatty acids commonly found in bacteria was used.

2.3 WCFA data elaboration

2.3.1 Artificial neural network based analysis

A competitive unsupervised Kohonen ANN [10], [12] was used for the classification of FAME profiles aimed to bacteria identification.

The learning set of FAME analyses has to contain non-contradictory profiles, and has to be complete and represent all the considered taxa. It was therefore decided to use profiles from certified strains coming from international collections for the learning set (Table 1).

The unsupervised ANN elaboration of FAME data shows the outputs on a bidimensional square divided into areas (output neurons). However, for better understanding, the output of an unsupervised ANN is not a limited square but should be considered as having the opposite sides connected to form a toroidal surface.

The number of output neurons should be of the same magnitude order of the input patterns in the learning phase. We chose to set the output neuron number to the first square number greater than the number of input patterns.

Once the learning phase is completed the areas do not move any more. The identification of unknown patterns is performed independently from the number of patterns to be

identified, as they are placed on the output map one at a time.

Specifically, the unsupervised ANN was implemented using Matlab 4.2 on a Pentium III with a learning phase of 5.000 steps. The computer time was approximately 5 minutes.

TABLE I
STANDARD STRAINS USED FOR ANN LEARNING

Genus	Species	Source [a]
Acinetobacter	*calcoaceticus*	LMG 1046^T
	junii	LMG 998^T
Aeromonas	*caviae*	LMG 3775^T
	hydrophila	LMG 2844^T
	media	LMG 9073^T
	salmonicida	LMG 14900^T
		LMG 3782^T
		LMG 3780^T
	sobria	LMG 3783^T
Alcaligenes	*latus*	LMG 3321^T
	xylosoxidans	LMG 1231^T
		LMG 1863
Aquaspirillum	*autotrophicum*	LMG 4326
	dispar	LMG 4329
	delicatum	LMG 4328
	serpens	LMG 3734
Arthrobacter	*crystallopoietes*	LMG 3819^T
	globiformis	LMG 3813^T
	histidinolovorans	LMG 3813^T
	ureafaciens	LMG 3812^T
Bacillus	*cereus*	LMG 6923^T
	licheniformis	LMG 6933^T
	mycoides	LMG 7128^T
Brevundimonas	*diminuta*	LMG 2089^T
	vesicularis	LMG 2350^T
Enterobacter	*aerogenes*	LMG 2094^T
Flavobacterium	*ferrugineum*	LMG 4021^T
	flevense	NCIMB 12056^T
	hydatis	NCIMB 2215^T
	johnsoniae	LMG 1341^T
	saccharophilum	NCIMB 2072^T
Micrococcus	*luteus*	LMG 4050^T
Pseudomonas	*alcaligenes*	LMG 1224^T
	mendocina	LMG 1223^T
	pseudoalcaligenes	LMG 1225^T
	putida	LMG 2257^T
		NCIMB 12708
		NCIMB 12182
Serratia	*fonticola*	LMG 7882^T
	marcescens	LMG 2792^T
	proteamaculans	LMG 7884^T
Shewanella	*baltica*	LMG 2250^T
	putrefaciens	LMG 2268^T
		LMG 2263
Vibrio	*ordalii*	NCIMB 2167^T

[a] LMG = Laboratorium Microbiologie Rijksuniversiteit Collection, Gent, Belgium; NCIMB = National Collection of Industrial and Marine Bacteria; T = Type strain.

2.3.2 *Microbial identification system MIDI*

The Microbial Identification System MIDI (MIS; Microbial ID Inc., Newark, Del., USA) is a fully automated, computerised, high resolution gas-chromatography system. The system gives the strain identification as a similarity index resulting from the best genus, species and subspecies match after the search in its microbial libraries. The system performs an automatic comparison of the fatty acid pattern of stored databases ("libraries") using pattern recognition software (MIS Software version 3.2).

The system we used in the present work had two libraries,

the TSBA (Rev. 3.2) and the CLIN (Rev. 3.2) libraries. We considered the TSBA library match as it is recommended for environmental bacteria, while the CLIN library is suited for clinical bacteria.

III. Results

3.1 Standard strain classification by ANN

The learning set consisted of all the standard strain profiles. All learning phase control parameters showed that the learning procedure was correct, that is all input patterns were correctly assigned to only one output neuron. The output was set as a square of 14 areas in 14 lines (196 areas) (Fig.1). More than one analysed standard strain was found to be located in the same area, the number is indicated inside the area with the first letter of the genus. Out of a total of 196 areas, 62 contain at least one standard analysis. The genera analysed are well separated and we can single out small or large zones corresponding to different genera. In "Fig." 1 the zones where the ANN put bacteria are colored in grey: numbers correspond to different analyses together with the first letters of the corresponding genus.

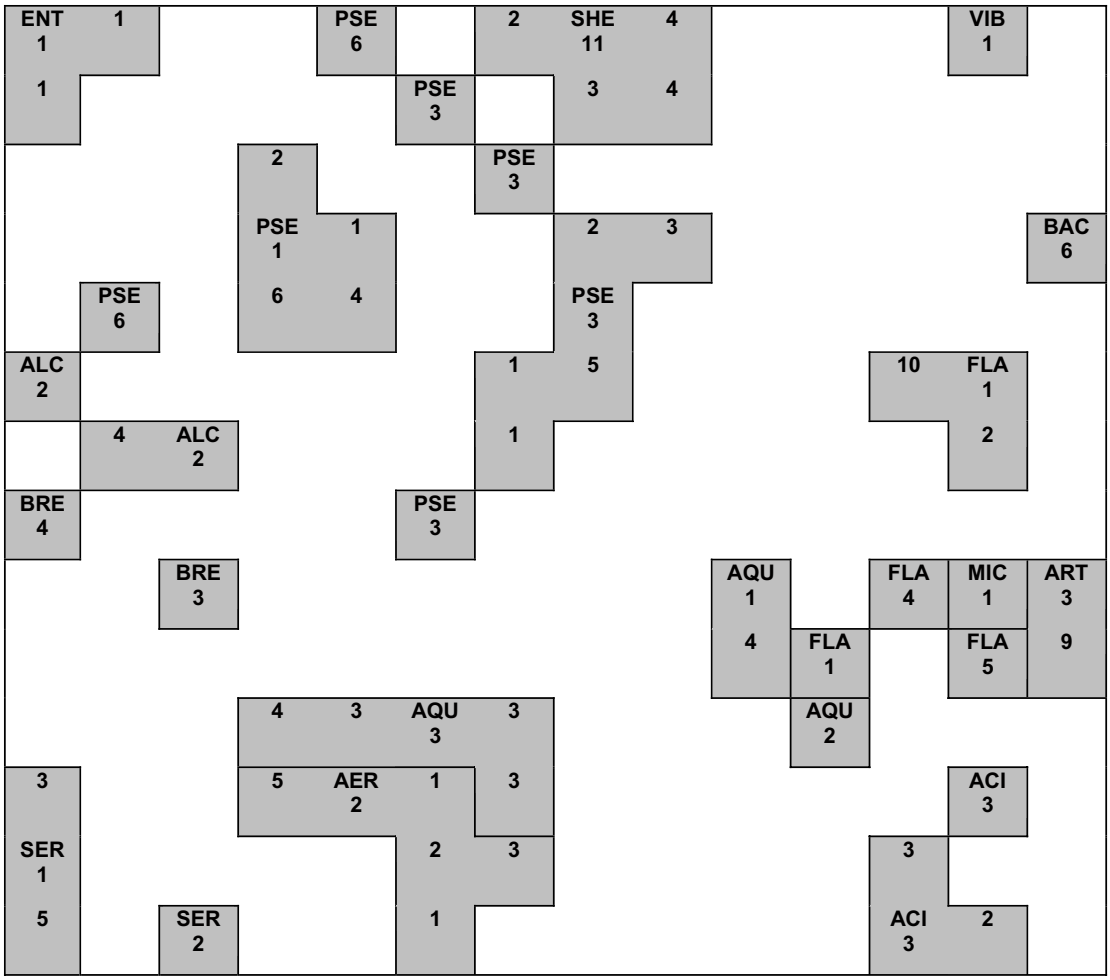

FIGURE 1: OUTPUT MAP OF THE ANN ELABORATION OF STANDARD STRAINS: ACI *Acinetobacter*; AER *Aeromonas*; ALC *Alcaligenes*; AQU *Aquaspirillum*; ART *Arthrobacter*; BAC *Bacillus*; BRE *Brevundimonas*, ENT *Enterobacter*; FLA *Flavobacterium*; MIC *Micrococcus*; PSE *Pseudomonas*; SER *Serratia*; SHE *Shewanella*; VIB *Vibrio*

3.2 Drinking water bacteria identification by ANN

After the elaboration of the standard strain profiles, the same output map was used to allocate the drinking water isolate profiles (Fig.2). On a total of 196 areas 42 contain at least one bacteria analysis. The numbers in "Fig." 2 indicate different strains allocated to the same area, with the first letters of the genus characterising the zone. Grey zones derive from the previous output map (Fig.1). The ANN used in the present work was found to be able to identify 70% of the strains analysed as sample, 55 out of 79 strains (Table 2).

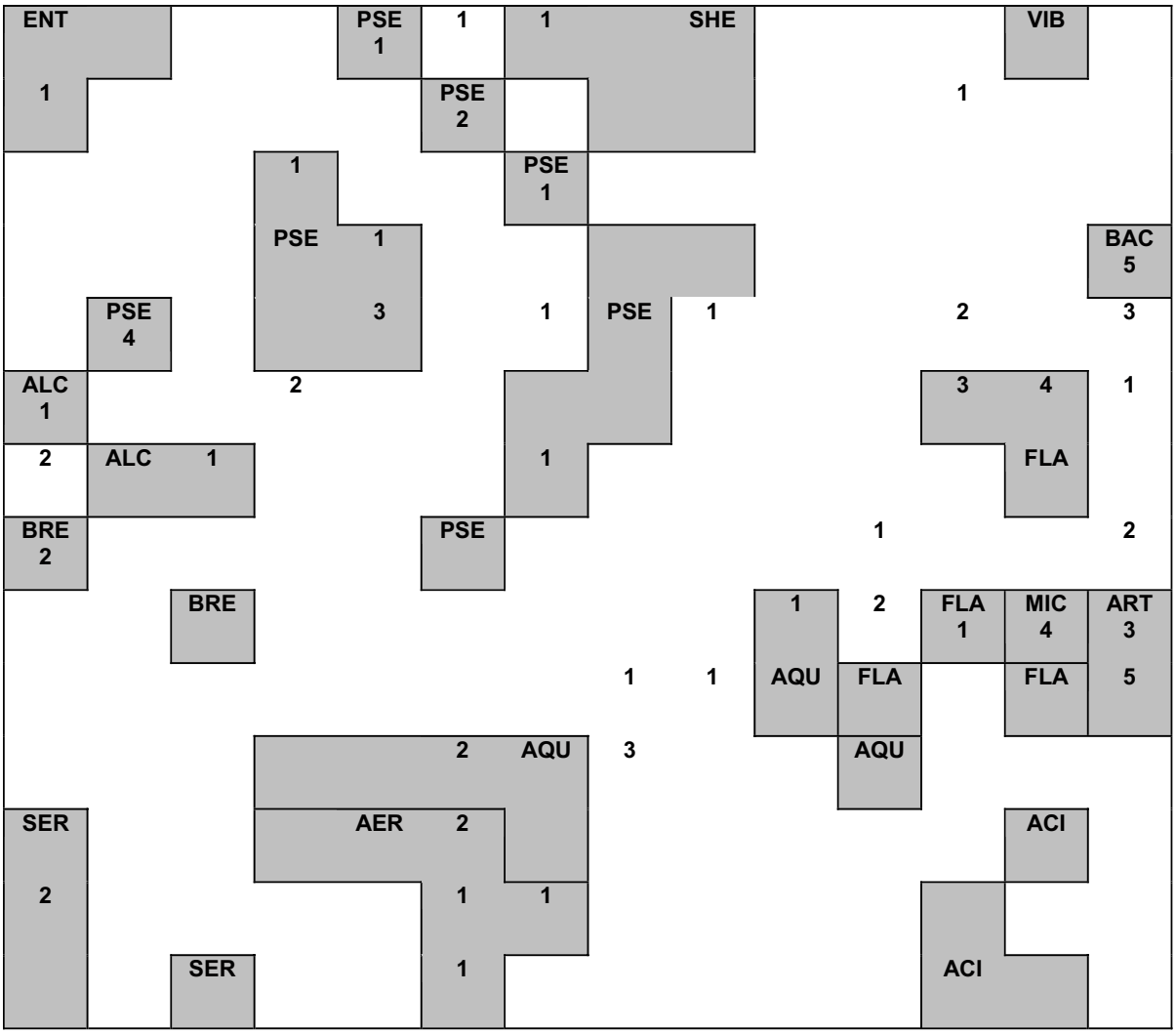

FIGURE 2. ALLOCATION OF THE FRESH ISOLATES INTO THE ANN OUTPUT MAP: numbers correspond to different strains together with the first letters of the genus: ACI *Acinetobacter*; AER *Aeromonas*; ALC *Alcaligenes*; AQU *Aquaspirillum*; ART *Arthrobacter*; BAC *Bacillus*; BRE *Brevundimonas*, ENT *Enterobacter*; FLA *Flavobacterium*; MIC *Micrococcus*; PSE *Pseudomonas*; SER *Serratia*; SHE *Shewanella*; VIB *Vibrio*.

The identification for the ANN corresponds to the allocation of the analysis to a zone which corresponds to a specific genus.

The majority of isolates belong to the *Pseudomonas* genus, 24 strains. A good quantity of *Flavobacterium* and *Arthrobacter* were found as well (8 strains each). Five strains were identified as *Aeromonas*, and other 5 as *Bacillus*. Four strains were identified as *Micrococcus* and 3 strains as *Aquaspirillum*. A few strains were identified as *Serratia*, *Alcaligenes* and *Brevundimonas* (2 strains each). Finally 1 strain was identified as *Shewanella*, 1 strain as *Enterobacter*, and no strains belonging to *Vibrio* and *Acinetobacter* were identified.

The remaining 24 strains analysed were allocated outside the zones corresponding to genera, therefore it was not possible to obtain a satisfactory identification with the ANN.

3.3 Drinking water bacteria identification by MIDI

MIS identification with a similarity index (SI) $\geq$ 0.500 were considered as a good match, while those with a SI ranging from 0.200 to 0.499 were considered as a low match. Four *Aeromonas* and 1 *Arthrobacter* strains were identified with a good SI (Tab. 2). One *Aeromonas*, 1 *Hydrogenophaga*, 6 *Arthrobacter*, 2 *Bacillus*, and 2 *Acidovorax* were identified with a low SI. Strain number 1331 gave a very good SI with three genera, *Enterobacter* (0.835), *Serratia* (0.20) and *Erwinia* (0.800). Moreover, 33 strains gave SI below 0.200, and so it was not possible to obtain a satisfactory identification with the MIDI system. Finally, for 26 strains there was no match with the TSBA and CLIN libraries.

TABLE II
IDENTIFICATION PERFORMED BY THE ANN COMPARED TO THE MIDI SYSTEM

Strain N°	ANN Identification	MIDI Identification
1297	*Aeromonas*	*Aeromonas* 0.893 [a]
1251	*Aeromonas*	*Aeromonas* 0.788
1309	*Aeromonas*	*Aeromonas* 0.662
1306	*Aeromonas*	*Aeromonas* 0.647
1280	*Aeromonas*	*Aeromonas* 0.398
1317	*Alcaligenes*	NM [b]
1335	*Alcaligenes*	*Enterococcus* 0.017
1301	*Aquaspirillum*	*Hydrogenophaga* 0.329
1300	*Aquaspirillum*	*Listonella* 0.177
1337	*Aquaspirillum*	NM
1272	*Arthrobacter*	*Arthrobacter* 0.565
1313	*Arthrobacter*	*Arthrobacter* 0.466
1256	*Arthrobacter*	*Arthrobacter* 0.371
1259	*Arthrobacter*	*Arthrobacter* 0.336
1296	*Arthrobacter*	*Arthrobacter* 0.291
1321	*Arthrobacter*	*Arthrobacter* 0.241
1253	*Arthrobacter*	*Arthrobacter* 0.208
1279	*Arthrobacter*	*Bacillus* 0.241
1274	*Bacillus*	*Bacillus* 0.257
1328	*Bacillus*	NM
1323	*Bacillus*	NM
1271	*Bacillus*	NM
1291	*Bacillus*	NM
1326	*Brevundimonas*	*Enterococcus* 0.031
1342	*Brevundimonas*	*Pseudomonas* 0.039
1290	*Enterococcus*	*Enterococcus* 0.254
1281	*Flavobacterium*	*Cytophaga* 0.139
1307	*Flavobacterium*	*Cytophaga* 0.135
1252	*Flavobacterium*	*Cytophaga* 0.033
1310	*Flavobacterium*	*Cytophaga* 0.085
1273	*Flavobacterium*	NM
1282	*Flavobacterium*	*Bacillus* 0.097
1322	*Flavobacterium*	NM
1333	*Flavobacterium*	NM
1303	*Micrococcus*	*Bacillus* 0.047
1327	*Micrococcus*	*Arthrobacter* 0.072
1254	*Micrococcus*	*Bacillus* 0.126
1270	*Micrococcus*	*Micrococcus* 0.155
1341	*Pseudomonas*	*Acidovorax* 0.374 (*Pseudomonas delafieldii*)
1316	*Pseudomonas*	*Acidovorax* 0.244 (*Pseudomonas delafieldii*)
1284	*Pseudomonas*	*Pseudomonas* 0.149; SEA: *Pseudomonas* 0.029
1314	*Pseudomonas*	*Acidovorax* 0.120 (*Pseudomonas delafieldii*)
1320	*Pseudomonas*	*Pseudomonas* 0.120;
1292	*Pseudomonas*	*Acidovorax* 0.075 (*Pseudomonas delafieldii*)
1308	*Pseudomonas*	*Pseudomonas* 0.050
1318	*Pseudomonas*	NM
1330	*Pseudomonas*	NM
1255	*Pseudomonas*	NM
1258	*Pseudomonas*	NM
1268	*Pseudomonas*	NM
1336	*Pseudomonas*	NM
1340	*Pseudomonas*	NM
1331	*Serratia*	*Enterobacter* 0.835, *Serratia* 0.820, *Erwinia* 0.800

1285	*Serratia*	*Variovorax* 0.073 (*Alcaligenes paradoxus*)
1332	*Shewanella*	*Sphingobacterium* 0.011
1294	UK [c]	*Gordona* 0.022
1289	UK	NM
1339	UK	NM
1343	UK	*Hydrogenophaga* 0.049
1269	UK	*Hydrogenophaga* 0.161
1275	UK	*Pseudomonas* 0.174
1266	UK	*Cytophaga* 0.033
1319	UK	NM
1329	UK	*Cytophaga* 0.173
1262	UK	NM
1325	UK	NM
1311	UK	*Arthrobacter* 0.011
1276	UK	*Arthrobacter* 0.021
1277	UK	*Bacillus* 0.125
1283	UK	NM
1295	UK	*Bacillus* 0.013
1298	UK	NM
1288	UK	*Streptoverticillium* 0.079
1312	UK	NM
1334	UK	*Enterococcus* 0.037
1338	UK	NM
1324	UK	NM
1286	UK	NM
1299	UK	*Pseudomonas* 0.198

[a] Identification similarity index (SI) obtained with the MIDI TSBA Database; [b] NM = no match with the MIDI TSBA or CLIN Database; [c] UK = unknown

IV. Discussion

Previous studies showed that FAME profiling can provide a useful tool for the identification of clinical and environmental bacteria [15], [16], and the MIDI system, present on the market for several years now, represents a powerful solution to perform this kind of analysis. To our knowledge, this is the first time that ANN has been applied to FAME classification for the identification of drinking water bacteria.

The unsupervised ANN processing of FAME data has given a good classification of the certified strains analysed at the genus level (Fig. 1). In fact, there is no overlapping of areas corresponding to different genera, and the output map shows separate zones corresponding to different taxa.

The identification power of the ANN was tested with a pool of 79 isolates coming from a water distribution line from AMGA, the major water supply system in Genova (Italy). It was possible to identify at genus level 70% of the isolates (Fig.2). The *Enterobacteriaceae* group of strains identified by the ANN, represent only 13 % of the fresh isolates analysed while the majority of strains belonged to other genera. In particular 16.6% *Pseudomonas* and 6% *Aeromonas* were found.

The remaining strains (about 30%) fell out of the zones corresponding to genera, however the position occupied by these strains can give useful information as well. For example, the strains allocated by the ANN in areas close to a genus, could belong to that genus. These strains could belong to a species not included in the set of bacteria used for the ANN learning. One of the prerequisites for the good performance of ANNs is that the learning set should be as complete as possible [17]. In this study, only 39 species belonging to 14 genera were analysed, but more genera and strains would be necessary to improve the identification power of the net; in fact, the more species and genera that are used for the learning phase, the more accurate and complete the identification of fresh isolates is.

Environmental microbiology is characterised by a continuous description of new species and genera and its subsequent rearrangement. Large amounts of data are likely to become available and analysis automation will increase its importance. FAME analysis carries information from genus to species level and nowadays has reached a high level of automation, so it is a fast method which allows the comparison and identification of large numbers of strains in a short period of time. Hence, suitable software development will be required to handle and process this large database.

The results herein presented show that ANNs represent a successful tool for bacteria classification by means of their FAME analyses. They can become a solid basis for a comprehensive artificial intelligence based system for drinking water bacteria identification, with special reference to the detection of opportunistic pathogens.

It would be worthwhile to build up a system of ANNs for identification of bacteria from genus to species. A comprehensive artificial intelligence based system for identification of drinking water bacteria coming from water treatment and supply systems consisting of multiple levels of ANNs that can be set up. In such a system, the first ANN could be like the one reported in the present study, i.e. for the identification at genus level, and the following connected networks, it could be specialised for each genus in order to identify at species and subspecies level.

REFERENCES

[1] Environmental Protection Agency (1989). National primary drinking water rules and regulations. U.S. EPA surface water rule. Filtration, disinfection, turbidity, Giardia lamblia, viruses, Legionella, heterotrophic bacteria, Federal Register, 54, pp.27486-27541.

[2] Environmental Protection Agency (1994). National primary drinking water regulation: Enhanced Surface Water Treatment Requirements: U.S. EPA Proposed Rule, Federal Register, 59, 159, pp.38832-38858.

[3] M. Handfield, P. Simard, M. Couillard, and R. Letarte, "Aeromonas hydrophila isolated from food and drinking water: hemagglutination, hemolysis and cytotoxicity for a human intestinal cell line (HT-29)", *Applied Environmental Microbiology*, vol. 62, pp. 3459-3461, 1996.

[4] P. Payment, E. Coffin, and G. Paquette, "Blood agar to detect virulence factors in tap water heterotrophic bacteria", *Applied Environmental Microbiology*, vol. 60, no. 4, pp. 1179-1183, 1994.

[5] P. Vandamme, B. Pot, M. Gillis, P. De Vos, K. Kersters, and J. Swings, "Polyphasic taxonomy, a consensus approach to bacterial systematics", *Microbiology Review*, vol. 60, pp. 407-438, 1996.

[6] K. Suzuki, M. Goodfellow, and A. G. O'Donnel, *Handbook of new bacterial systematics*. London: Academic Press Ltd, 1993.

[7] M. Giacomini, C. Ruggiero, M. Maillard, F. B. Lillo, and O. E. Varnier "Objective evaluation of two markers of HIV-1 infection (p24 antigen concentration and CD4+ cell counts) by a self organising neural network", *Medical Informatics*, vol. 21, pp. 215-228, 1996.

[8] D. T. Pham, and X. Liu, *Neural Networks for identification, prediction and control*. London: Springer-Verlag, 1997.

[9] S. Bertone, M. Giacomini, C. Ruggiero, C. Piccarolo, and L. Calegari, "Automated systems for identification of heterotrophic marine bacteria on the basis of their fatty acid composition", *Applied Environmental Microbiology*, vol. 62, pp.2122-2132, 1996.

10] M. Giacomini, C. Ruggiero, S. Bertone, and L. Calegari, "Artificial neural network identification of heterotrophic marine bacteria based on their fatty acid composition", *IEEE Transactions of Biomedical Engineering*, vol. 44, pp. 1185-1191, 1997.

11] M. Giacomini, C. Ruggiero, L. Calegari, and S. Bertone, "Artificial Neural Network based identification of environmental bacteria by gas-chromatographic and electrophoretic data", *Journal of Microbioly Methods*, vol. 1, pp. 45-54, 2000.

12] T. Kohonen, "The self -organizing map", *IEEE Proceedings*, vol. 78, no. 9, pp. 1464-1480, 1990.

13] M. P. Lechevalier, "Lipids in bacterial taxonomy: a taxonomist's view", *Clinical Microbiology Reviews*, vol. 5, pp. 109-210, 1977.

14] L. T. Miller, "Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids", *Journal of Clinical Microbiology*, vol. 16, pp. 584-586, 1982.

15] L. A. Briganti, and S. C. Wacker, "Fatty acid profiling and the identification of environmental bacteria for drinking water utilities", *American Water Works Association Research Foundation and American Water Works Association*, Denver, Colorado, 1995.

16] C. D. Norton, and M. W. LeChevallier, "A pilot study of bacteriological population changes through potable water treatment and distribution", *Applied Environmental Microbiology*, vol. 66, no.1, pp.268-276, 2000

17] D. L. Hudson and M. E. Cohen, *Neural networks and artificial intelligence for biomedical engineering*. New York: IEEE Press, 2000.