IDENTIFICATION OF STAPHYLOCOCCUS AUREUS INFECTIONS BY VOLATILE CHEMICAL HEADSPACE ANALYSIS

J.W. Gardner[1], L. Beeby[2], M.J. Chappell[1], F. Udrea[3], J.W. Yates[1] and C. G. Dowson[2]

[1]School of Engineering and [2]Department of Biological Sciences
University of Warwick, Coventry, CV4 7AL
United Kingdom
Email: j.w.gardner@warwick.ac.uk

[3]Department of Engineering
University of Cambridge, Cambridge CB2 PZ
United Kingdom

ABSTRACT
An investigation has been made into the identification of different strains of bacteria through the analysis of their volatile chemical headspace. Cultured samples were prepared and compared with clinical blood samples with known infection. Two different analytical instruments are combined to sample the volatile headspaces of samples, namely a mass spectrometer and a conducting polymer based electronic nose. Data were analysed using both linear parametric (principal components analysis) and non-linear non-parametric (radial basic function) techniques. Our results show that it is possible to discriminate between different strains of bacteria and in some cases between methicillin-resistant *S. aureus* (MRSA) and methicillin-sensitive *S. aureus* (MSSA). However, tests on inoculates prepared from closely related clinical isolates of *S. aureus* characterised by multilocus sequence typing (MLST), suggest that the acquisition of methicillin resistance *per se* cannot be detected by its volatile headspace alone. Thus the ability to discriminate between these bacterial strains appears to be more related to their overall genetic evolutionary distance or underlying differences in metabolism. This observation shows that the discriminating power of an electronic nose is limited to a subset of the subgroups of strains of MRSA and MSSA.

KEY WORDS
Intelligent instrumentation, Electronic Nose, MRSA.

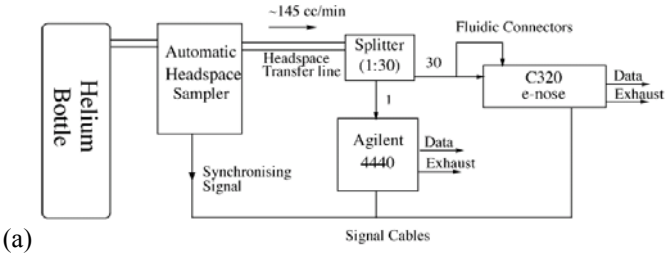
1. Introduction

Bacterial infections are responsible for a wide variety of medical conditions – some of which are minor and others are life threatening. Two important human pathogens are *Staphylococcus aureus* (*S. aureus*) which is a Gram-positive spherical bacterium and the Gram negative rod shaped bacterium *Escherichia col (E. coli)*. It has been estimated that about 1/3 of all healthy individuals carry *S. aureus* in their nostrils whilst *E. coli* is commonly present in large numbers in the human gut. *S. aureus* has recently been reported as the most commonly acquired hospital infection [1] and can be deadly to patients whose health is compromised by other conditions. Of widespread concern has been the emergence of new strains of bacteria that are resistant to common antibiotics. It is now estimated that over 95% of *S. aureus* infections are resistant to first line antibiotics (e.g. penicillin) and 30% are also resistant to methicillin. This so-called methicillin-resistant strain of *S. aureus* (MRSA) was first reported in 1961 and is now the predominant bacterium in critically ill patients with some 31% of longer stay patients becoming colonised or admitted with MRSA [2]. This temporal expression of genes has also resulted in the emergence of new strains of *E. coli*; the most renowned being the strain of *E. coli* 0157 that produces verocytotoxins and can lead to food poisoning or even death for young and old people.

The identification of different bacterial types by their volatile chemical headspace (or "smell") was first reported by Gardner and Hines in 1998 [3]. The original experiments employed an electronic nose comprising an array of metal oxide resistive gas sensors and a back-propagation neural network. The original results based on metal oxide type gas sensors were extended to conducting polymer type resistive gas sensors. In this later study, the University of Warwick together with Heartlands Hospital (Birmingham) analysed a set of samples cultured from ENT and eye infections using a commercial Cyranose C320 system [2]. This work has now been extended further and the commercial 32-element conducting polymer electronic nose system has been used to analyse swabs taken directly from patients in a clinical environment. In this case the results showed that the electronic nose had potential to screen rapidly patients for bacterial type infections with a success rate of *ca* 96% [4]. A significant number of technical papers have been published on the different biomedical applications of electronic nose technology and interested readers are referred to a recent review [5].

An alternative technological approach to a sensor-based electronic nose system is that of conventional analytical

instrumentation, i.e. mass spectrometry. At the University of Warwick, a commercial quadrupole mass spectrometer (Chemical Sensor 4440, Agilent Technology) has been combined in-line with a 32-element polymer based electronic nose (Cyrano C320) for the rapid automated screening of biological samples; see Figure 1 [6]. The automated test system has been employed to study infected samples of both human blood and urine.



(a)

(b)

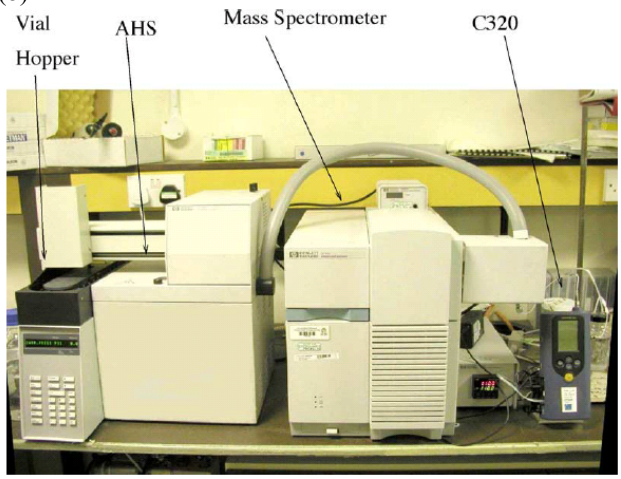


Fig. 1: (a) Schematic arrangement and (b) photograph of system used to analyse the volatile chemical headspace of infected specimens in vials [6].

Our study here is concerned with the identification of different strains of the bacterium *S. aureus* by its characteristic volatile chemical headspace or, to be more specific, differentiation between the clinically important subgroups of MRSA and methicillin-resistant *S. aureus* (MRSA). Methicillin resistance is determined genetically and has been located to the genetic element of the *Staphylococcus* cassette chromosome mec (SCCmec) [7]. MecA is the methicillin resistant gene that is included within this cassette and encodes a penicillin binding protein resistant to β-lactam antibiotics that is not present in susceptible strains, but has been acquired by multiple, genetically different strains of MSSA [8]. So the most important question to ask is:

Can we discriminate between these genetically most similar strains of MRSA and MSSA and identify the acquisition of methicillin resistance simply by the analysis of volatile chemical headspace?

2. Experimental

The bacterial strains employed in the experimental work were obtained from strains located in the research group of Prof. Dowson at the University of Warwick and isolates obtained from Prof. Enright at Bath University (UK). Unlike in previous reports, isolates of *S. aureus* are studied that are closely related and characterised by multilocus sequence typing (MLST). This method is used to trace the evolutionary lineage of the strains and so determine genetic similarity. The proximity of the strains is determined by the sequences of *ca* 450 base pair (bp) internal fragments of seven housekeeping genes. The isolates obtained for this study were 15 pairs of *S. aureus*, each pair consisting of one MSSA strain and one closely related MRSA strain. These fifteen pairs of isolates belong to five different clonal complexes (CC) and so it is likely that isolates in each of these clusters have descended from a single ancestral genotype [8,9]. Table 1 shows the strain collection numbers of the isolates used.

Table 1: Strain collection references of isolates.

Organism	*S. epidermidis*	*S. aureus*	MRSA
Strain no.	2445	1967	1049

Master plates were obtained by streaking each strain onto an agar plate containing brain and heart infusion (BHI) and incubated for a period of 16 h at a temperature of 37 ºC and atmosphere of 5% CO_2. These master plates were then frozen and used to create single pure colonies to re-streak on further BHI plates

It is important to create a culture with a known value of colony forming units (cfu) per ml and so frozen aliquots of both MRSA and *S. aureus* were created and isolated (numbered 1-4) of the 15 pairs of susceptible (MSSA) and resistant (MRSA) isolates (Table 2). These 15 pairs belonged to five different sequence types (ST5, 8, 22, 30 and 45) [1]. For each strain three universals of 10 ml of HBI broth were prepared and inoculated with one colony. The inoculated bottles were incubated at 37ºC for 16 h whilst shaken at 75 rpm. Following incubation, one culture bottle of each strain was removed, vortexed and pipetted in 1 ml aliquots into 1.5 eppendorfs. The aliquots were then formed into a pellet by centrifuging at 13,000 rpm for a period of 2 min and the supernatant removed. A similar procedure was employed to process the other two universal samples. Finally, 1 ml of BHI with 15% glycerol was added to the pellet, vortexed and frozen at −80ºC. The sterile sample bottles (BacT/ALERT® SA, Biomerieux UK Ltd) contained 40 ml of a media formulation of pancreatic digest of casein (1.7% w/v), papaic digest of soybean meal (0.3% w/v), sodium polyaneatholesulfonate (0.035% w/v), pyidoxine HCl (0.001% w/v) with other amino acids and carbohydrate substrates in purified water. The cultured bottles were injected with 10 ml of the same sterile blood sample to represent the clinical procedure of adding a patient's

blood sample. Finally the bottles were inoculated using the frozen bacterial pellets and incubated at 37ºC while shaking at 75 rpm.[1] Samples were then transferred as 1.5 ml aliquots in 10 ml vials. Blanks of BacT/Alert SA and blood (defibrinated horse blood, Oxoid Ltd and human blood) were analysed alongside inoculated samples to act as controls. In the case of frozen isolates (pairs 1-4), 1 ml of known volume of bacteria was injected into the culture bottle; however, for isolate pairs 5-30 no frozen stock was available so 1 ml of BacT/ALERT media was removed and inoculated with a fresh 16 h incubated culture and returned to the full volume of media. Inoculates were created by serial dilution in phosphate buffered saline solution and ranged from 10^3 to 10^8 cfu per ml.

Table 2: Sequence types of the 30 different MRSA and MSSA isolates

ST*	Isolate	MSSA	Isolate	MRSA
5	1	D181	2	SLOV30
	3	D10	4	POL MR1
	5	D102	6	EMRSA3
8	7	CAN 6909 0655	8	EMRSA 2
	9	CAN 5619 1074	10	EMRSA 7
	11	CAN 6519 1387	12	99ST6509
22	13	CAN 7111 1743	14	C720
	15	CAN 7819 1058	16	99ST18126
	17	CAN 7819 1096	18	99ST18128
30	19	D107	20	SWED A017934/97
	21	H73	22	H312
	23	D517	24	EMRSA 16
45	25	CAN 6115 0443	26	SWED CCUG 41787
	27	CAN 6825 0622	28	GERM 825/96
	29	CAN 7114 0177	30	GERM 792/96

*ST is Sequence Type. Isolates with the same sequence type have identical sequences at seven genes around their chromosomes are therefore are very closely related and regarded as being part of the same clone [1]. Differences between MRSA and MSSA isolates with the same ST will essentially be due to the acquisition of methicillin resistance.

In addition 24 blood samples were obtained from the Walsgrave Hospital (Coventry, UK) and represented the three most common metabolic blood disorders of high bilirubin levels, high glucose levels, and high urea levels. These samples had been recorded positive and thus examined for micro-organism presence. Table 3 lists the micro-organisms found present in the 24 specimens.

Table 3: Micro-organisms found present in blood of 24 patients from the Walsgrave hospital, UK.

Ref. No.	Micro-organisms present	Ref. No.	Micro-organisms present
1	Coagulase – *Staph.*	13	*E. coli & Strep*
2	Coagulase – *Staph.*	14	Micrococcus
3	Coagulase – *Staph.*	15	MRSA
4	Coagulase – *Staph.*	16	Coagulase – *Staph.*
5	Micrococcus	17	Coagulase – *Staph.*
6	MRSA	18	Coagulase – *Staph.*
7	Coagulase – *Staph.*	19	Gram + *Cocci*
8	Pneumococcus	20	MRSA
9	Coagulase – *Staph.*	21	Gram + *Cocci*
10	Gram + *Bacilli*	22	Pneumococcus
11	Coagulase – *Staph.*	23	*Enterococcus*
12	MRSA	24	*Streptococci*

The volatile headspace samples were formed by the injection of helium gas into 10 ml sterile flat-bottom vial (Agilent Technologies, Inc.) containing 1.5 ml aliquots of the culture. All samples were run in triplicates with blanks of BacT/ALERT SA and blood were run alongside inoculated samples. The vials (see Figure 2) are held at 80ºC by the Agilent 7694 headspace sampler and then injected into the quadrapole mass spectrometer and Cyranose C320 unit. The mass spectrometer (Agilent 5973N) analysed the mass content of the sample headspace and was set to a range of 46 to 550 Daltons. The time required to sample each vial was approximately 15 min. The mass abundances were recorded on a PC and analysed subsequently using different pattern recognition techniques. Each mass can be considered as a type of pseudo-sensor and so the instrument may be regarded as a chemical sensor array with 504 sensors.

The Cyrano 320 logs the resistance of each of the 32 polymer sensors housed within the unit. The time series data were pre-processed by determining the resistance before and after samples were introduced. This reduced the dimensionality of the basic problem to 32 features, each one being the fractional change in resistance.

[1] Similar setting to those at a local hospital.

Fig. 2: Photograph of 10 ml vials with aluminium crimped cap and PTFE/silicone septum and 1.5 ml aliquots of blood or blank media.

3. Analysis of results

The experimental data were logged in the form of tab delimited text files and read into MATLAB v5.3 software for multivariate analysis. The data-sets were analysed using different multivariate techniques. Firstly, principal components analysis (PCA) is used to reduce the high dimensionality of the problem (504 in the case of mass spectrometer) to 3 dimensions by linear vectorial decomposition. Distinct separation of samples in a PCA plot shows that the problem is linearly separable and that the sensors are strongly correlated. Secondly, a k-nearest neighbours (KNN) cluster analysis is employed and is a non-parametric, non-linear approach that links together the samples closest in distance (Mahalanobis) from each other in multidimensional space. Finally, Sammon mapping is a non-linear projection method from the area of multidimensional scaling and again reveals any structure present in the data.

Work that has been reported [10] showed some encouraging results for urine samples analysed by both the mass spectrometer and the polymer C320 unit. However in this study the Cyranose C320 proved to be unstable and so the analysis of the results has been restricted to those data obtained from the Agilent 4440 mass spectrometer.

Figure 3 shows a 3-D PCA plot of the headspaces for isolates 3 and 4 that were cultured in BacT/ALERT bottles and removed from incubation after 8 h and 24 h, respectively. The data-points for these isolates almost entirely overlap whilst those for the blank specimens (BacT and horse blood) are almost entirely distinct in linear principal component space. This demonstrates that the volatile headspace is similar at incubation periods of 8h and 24 h. It should be noted that the abundance data are autoscaled and so any increase in mass abundance that generally accompanies an increase in biomass is removed.

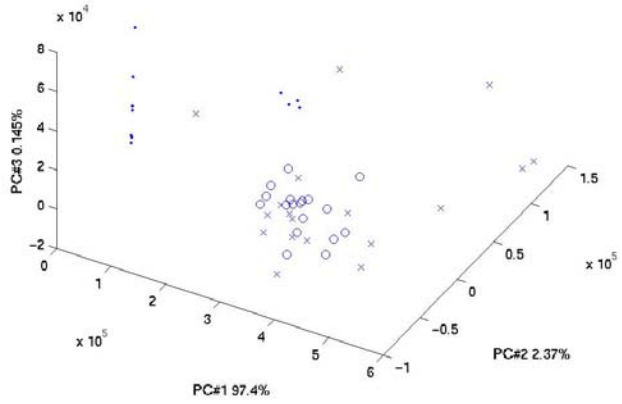


Fig. 3: PCA showing the sensor response is the same for 8 h and 24 h growth times. Key: isolates 3 (○) & 4 (×), and blank BacT/horse blood (∗).

Next all 30 (15 pairs) of the MRSA and MSSA isolates were cultured for a period of 24 h and so should be in the stationary phase of their growth cycle. Again the data were analysed using principal components analysis and the results are shown below in Figure 4.

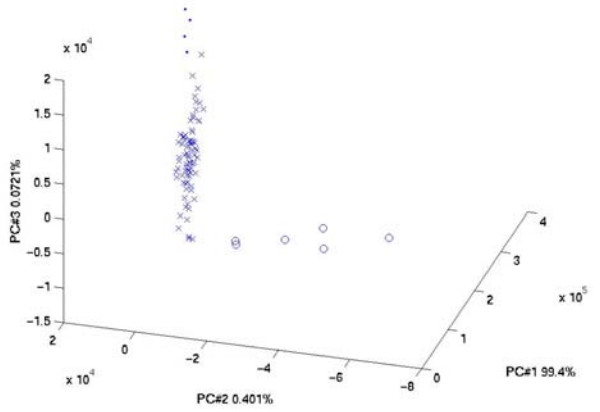


Fig. 4: PCA showing the sensor response to all isolates 1-30 (×), blank BacT/horse blood (∗) and standard BHI (○).

In this case, it is evident that, once again, the controls of BacT/Horse blood and brain heart infusion are distinct from the isolates. It can also be seen that the 30 related isolates cluster together in one principal component plane (#3) and represent only 0.072% of the variance in the data-set.

The genetic closeness of the MRSA and MSSA isolates is known through the previously carried out multilocus sequence typing of the strains. A KNN cluster analysis was carried out of the sensor responses for all 30 isolates but it was not possible to discern the two subgroups of MRSA and MSSA from the resultant dendrogram.

The experiment was then repeated but this time two unrelated strains (Warwick collection) of MRSA and MSSA were cultured and tested. The PCA plot shows distinct clusters of the two strains (see Figure 5) that could be separated out using non-linear methods. This result is consistent with previous work that has shown that unrelated strains of MRSA and MSSA may be separated out using the volatile chemical headspace.

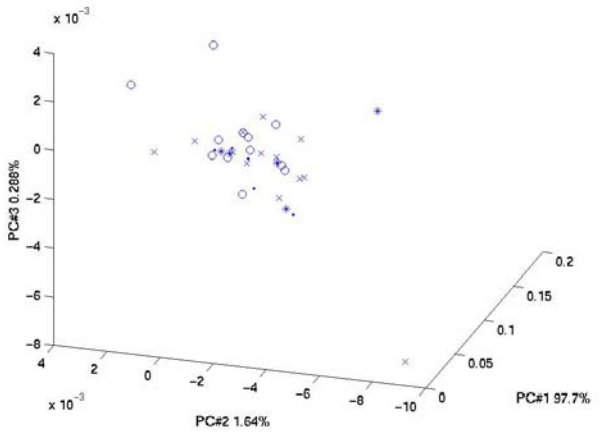


Fig. 5: PCA showing the sensor response to unrelated MRSA (×) and MSSA (○) cultured samples.

Figure 6 shows a PCA plot of the 24 clinical blood samples taken from patients at Walsgrave Hospital (UK). The samples received were all in the stationary growth phase of the micro-organism and of differing ages. Although the PCA plot shows little discrimination between most of the different pathogens, *enterococcus* does appear to be distinct.

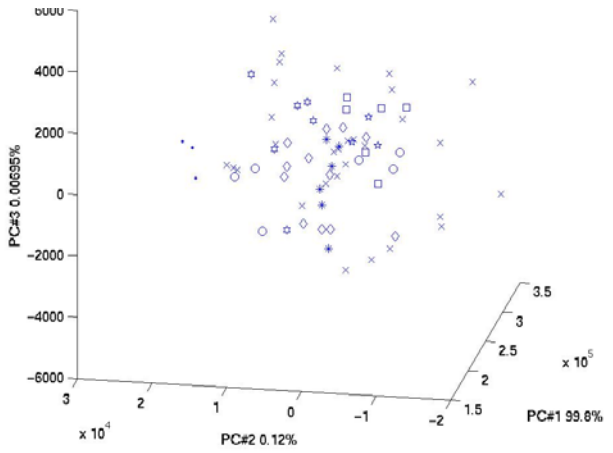


Fig. 6: PCA showing the sensor response to clinical blood samples with infections: enterococcus (•), gram +ve cocci (○), CoN Staphylococci (×), micrococcus (∗), pneumococcus (□), MRSA (◊), & gram +ve bacilli (◊).

Analysis of blood samples with different bacterial strains is shown in Figure 7 and further non-linear data analysis has been carried out [10] that combines a Sammon mapping algorithm with a radial basis function (RDF) neural network [11].

The Sammon mapping algorithm seeks a projection of the data to a lower dimension that preserves the distance between data points. This algorithm is suitable for pre-processing of data for an RBF neural network because the arguments of the RBFs are the distances between the chosen exemplars and the input.

The RDF functions used here are Gaussian with the cluster centres determined by a non-linear optimisation algorithm and these form the hidden layer in a multilayer perceptron network. The weights were found by a support vector machine [11] approach. This approach seeks to minimize the prediction error of the neural network while also minimizing the complexity of the resulting network. The complexity is minimized via the number of cluster centres used. The use of this optimization criterion with nonlinear functions in the hidden layer allows the training of a neural network where the number of exemplars is smaller than the input dimension of the network.

Using this powerful combination of an RBF network and Sammon mapping, it was possible to discriminate between <u>all</u> of the different micro-organisms and the growth medium.

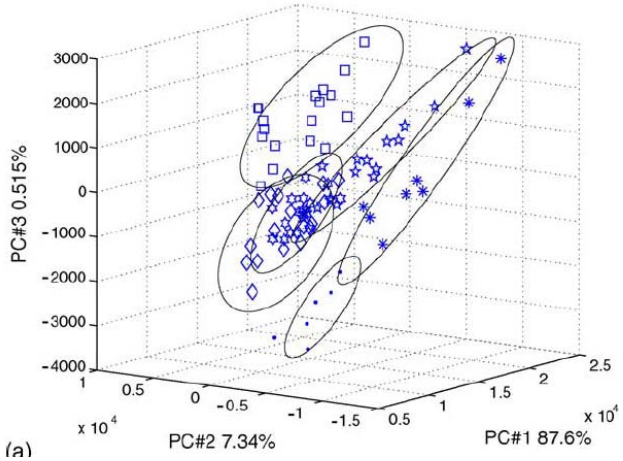


Fig. 7: PCA showing the sensor response to different infected blood samples: MRSA (○), MSSA/NCTC (+), *Staphylococcus epidermis* (*), *S.warneri* (◊), *S.simularis* (□), *S.haemalyticus* (∗), *S.lugdenenesis* (¤) & growth medium (•).

4. Conclusion

A detailed study has been undertaken to determine the extent by which a pathogen can be identified from the volatile chemical headspace of cultured blood samples.

Our results show that it is possible to discriminate between different bacterial pathogens with a success level of 100% using Sammon mapping coupled with a radial basis function neural network. Our results were obtained using a mass spectrometer and are similar to other studies that have been reported based upon solid-state volatile-sensitive sensor arrays [2-4].

However, and more importantly, we have studied very close genetic subgroups of the important pathogen *S. aureu*s, namely MRSA and MSSA. Our results suggest that the property of methicillin-resistance, as inferred by the presence of the mecA gene, does not influence significantly the volatile products released. Consequently it is only possible to discriminate between strains of MRSA and MSSA that are genetically far apart. Clearly, this limits the applicability of an electronic nose to screen specifically for MRSA infections in patients.

5. Acknowledgements

The authors thank the Engineering and Physical Research Council for the financial support of James Yates and Agilent Technologies (Delaware, USA) for the donation of the mass spectrometer and headspace autosampler.

References:

[1] M.C. Enright, N.P. Day, C.E. Davies, S.J. Peacock & B.G. Spratt, Multilocus sequence typing for the characterisation of methicillin-resistant and methicillin-sensitive clones of S. aureus, *Journal of Clinical Microbiology, 38*, 2000, 1008-1015.

[2] P. Boilot, E.L. Hines, J.W. Gardner, R. Pitt, S. John, J. Mitchell & D. Morgan, Classification of bacteria responsible for ENT and eye infections using the Cyranose system, *IEEE Sensors Journal, 2*, 2002, 218-229.

[3] J.W. Gardner, M. Craven, E.L. Hines & C. Dow, Prediction of bacteria type and growth phase by an electronic nose with a multi-layer perceptron network, *Meas. Sci. Technol., 9*, 1998, 120-127.

[4] R. Dutta, D. Morgan, N. Baker, J.W. Gardner & E.L. Hines, Identification of staphylococcus aureus infections in hospital environment: electronic nose based approach, *Sensors and Actuators B, 109*, 2005, 355-352.

[5] T.C. Pearce, S. Schiffman, T. Nagle & J.W. Gardner (eds.), Ch. 18 of *Handbook of Machine Olfaction* (Wiley-VCH), 2003). ISBN 3-527-30358-8.

[6] J.W. Yates, Black box and mechanistic modelling of electronic nose systems, *PhD thesis*, University of Warwick, Coventry, UK, 2004.

[7] B. Berger-Bachi & S. Rohrer, Factors influencing methicillin resistance in staphylocci, *Archives in Microbiology, 178*, 2002, 165-271.

[8] M.C. Enright, D.A. Robinson, G. Randle, E.J. Feil, H. Grundman & B.G. Spratt, The evolutionary history of methicillin-resistant S. aureus, *Proc. National Acad. Sci., 11,* 2002, 7687-7692.

[9] N.P. Day, C.E. Moore, M.C. Enright et al., A link between virulence and ecological abundance in natural populations of S. aureus, *Science, 292*, 2001, 114-116.

[10] J.W Yates, M.J. Chappell, J.W. Gardner, C.S. Dow, C. Dowson, A. Hamood, F. Bolt & L. Beeby, Data reduction in headspace analysis of blood and urine samples for robust bacterial identification, *Computer Methods and Programs in Biomedicine, 79*, 2005, 259-271.

[11] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, 1999.