178. Sommerville, B.A., McCormick, J.P. and Broom, D.M. 1994. Analysis of human sweat volatiles: an example of pattern recognition in the analysis and interpretation of gas chromatograms. *Pestic. Sci.*, 41, 365-368.

Analysis of Human Sweat Volatiles: an Example of Pattern Recognition in the Analysis and Interpretation of Gas Chromatograms*

Barbara A. Sommerville†, James P. McCormick & Donald M. Broom

Department of Veterinary Medicine, University of Cambridge, Madingley Rd, Cambridge CB3 0ES, UK

(Revised manuscript received 12 January 1994; accepted 8 March 1994)

Abstract: The analysis of the volatile components of human body odour is described with a view to establishing how individual identity can be reflected in an odour profile. In order to get good chromatographic replication, a recirculating system was built to load the samples of axillary sweat head-space into Tenax concentrating traps. The loading temperature and humidity could be controlled and trace volatiles concentrated. The head-space samples were desorbed from the trap and analysed by GLC. The GC/FID was interfaced with an ATARI ST 1040 computer which stored the output on disc. Pairs of chromatograms were then compared in a pattern-matching programme which allowed irrelevant differences, such as variations in running time and sample size, to be corrected and then various match parameters to be measured. These were alignment coefficient, profile correlation, Euclidian distance and box car distance. The whole and selected parts of the chromatograms could be matched so that each range of volatiles could be assessed for its contribution to an identity signal.

The pattern of sweat volatiles from two pairs of identical twins showed significantly higher match correlations than that of unrelated people. The range of compounds eluting between 80 and 150°C showed the highest match correlations.

These techniques could be used to match chromatograms from any source, for example, in drugs and pesticide research, environmental monitoring, food, flavour and perfume analysis or clinical and forensic investigations.

1 INTRODUCTION

Head-space sampling and the trapping and concentration of trace volatiles is commonly employed in environmental control, semiochemical and medical research and in the food, drink, flavour and perfume industries. Most plants[1] and animals[2] emit odorous volatiles of a species-specific nature and in many animals the pattern is specific to a group, family or an individual.[3] In most cases, chromatographic analysis reveals a complex mixture of many volatiles which it may not be necessary to identify individually, provided that a typical pattern can be recognized and matched. Unfortunately, so many factors affect the profile of a gas chromatogram that it is notoriously difficult to get acceptable replication from a succession of analyses of the same sample[4] and, without this, pattern recognition is difficult. To address this initial problem, we have designed a recirculating system which allows the temperature, humidity and selection of volatility range to be controlled while trapping and concentrating a head-space sample.[4] Using this apparatus we have repeated the experiments carried out by Sommerville *et al.*,[5] which showed that the pattern of volatiles in human axillary sweat reflects a genetic influence and can be used to distinguish identical twins from unrelated people.

The ultimate aim of this project was to establish the composition of the body odour signature which enables dogs to recognize and discriminate individual people.

* Paper presented at the symposium Non-labelled Techniques—New Directions in Drug and Pesticide Research organized by D. M. Livingstone, R. M. Hyde & M. G. Ford on behalf of the Physicochemical and Biophysical Panel of the SCI Pesticides Group and held on 14 October 1993 at 14/15 Belgrave Square, London SW1.

† To whom correspondence should be addressed.

Pestic. Sci. 0031-613X/94/$09.00 © 1994 SCI. Printed in Great Britain

This knowledge could be employed forensically to match a suspect to an odour collected at the scene of a crime[6] or, possibly, to detect emotional disturbance in a lie or fear detector or as a basis for the diagnosis of certain metabolic diseases.[7] The analytical techniques employed could be used in many of the fields mentioned above.

2 EXPERIMENTAL METHODS

The sweat donors were two pairs of twins who had been shown to be identical (monozygous) by DNA profiling. One pair were teenage boys and the other adult women. The axillary sweat was collected by squares of woven cotton fabric pinned to the armpit area of a T-shirt which was worn for about eight hours. The squares were then removed and packed in aluminium foil within a labelled polyethylene bag. The cotton squares were washed, dried and heated to 150°C for 10 min before sending to the donors. The donors used odour-free soap to wash their armpits during the 48 hours prior to sampling and their hands before handling the cotton squares. The apparatus shown in Fig. 1[4] was used to load the sweat volatiles into the Tenax traps. The oven was set at 70°C and 120 ml of dry air recirculated for 2 min to load the first trap and then, immediately, a second trap was loaded for 3 min and lastly a third trap for 5 min. Each trap was desorbed at 320°C and the sample injected onto a 12-m capillary column (BP 10/OV 1701; int. diameter 0·22 mm; ext. diameter 0·33 mm; film thickness 0·25 μm) using an S.G.E. updated head space injector system.[8]

The programme used was 40°C to 190°C ramping at 6°C min^{-1}. The gas chromatograph/flame ionization detector (FID) was interfaced with an ATARI ST 1040 computer which stored the FID output on disc. The first load of sample was dominated by the high-volatile compounds and the third by the low-volatiles as shown in Fig. 2. The analyses detailed below were carried out on the second load of sample in which the medium-volatile range was well represented, since earlier work had shown this part to be important in individual body odour discrimination, whether carried out by a dog bioassay or by chromatogram pattern matching.[5]

The pattern matching programme[9] allowed the two chromatograms in Fig. 2 (A & B) to be 'tailored' to match as closely as possible. The smaller sample could

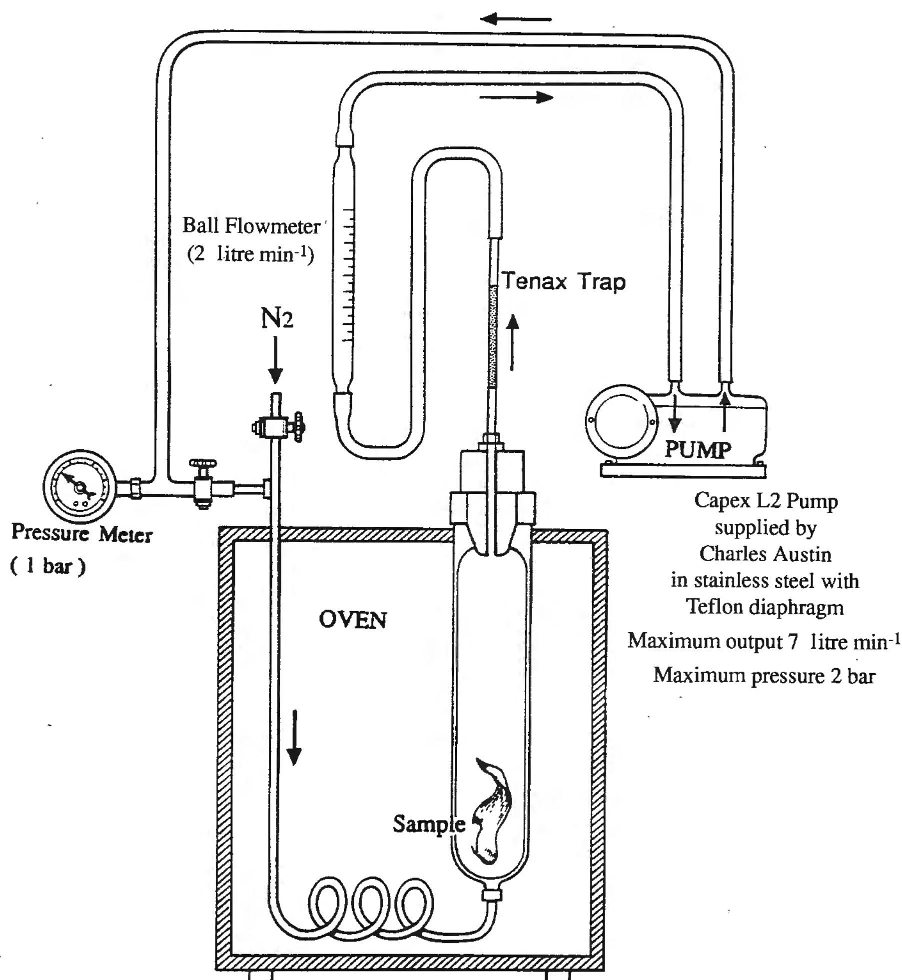

Fig. 1. A recirculating system for loading Tenax traps (see Ref. 4).

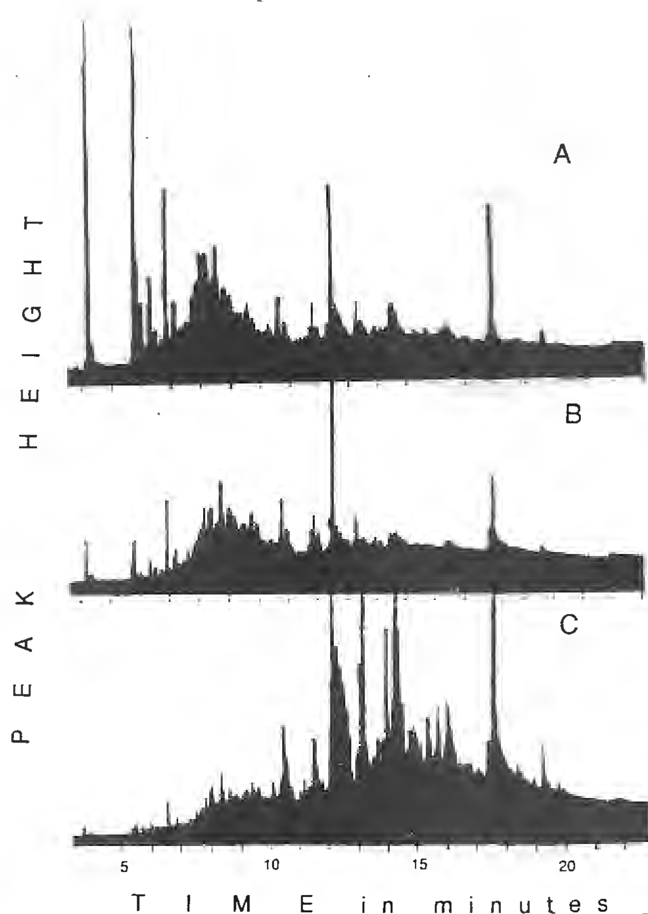

Fig. 2. Three chromatograms showing the different range of volatiles trapped during: A. the first load of 2 min from time 0 to 2 min; B. the second load of 3 min from time 2 to 5 min; C. the third load of 5 min from time 5 to 10 min.

be scaled up, common peaks could be aligned to correct for slight differences in running time, spurious peaks (interference or marker spikes) could be removed, the area under the base-line could be reduced and base-line trends corrected. Having made the necessary adjustments, measurements of how well the two chromatograms matched could be made. Four match parameters could be applied over a specified length of time (t) on the pair of chromatograms.

2.1 Alignment coefficient

The alignment coefficient measures the ratio of the number of peaks aligned in A and B to the total number and may be set to consider only peaks over a specified minimum height. A bad match is indicated by a low number and a perfect match scores 1.

2.2 Profile correlation

The profile correlation gives a similarity measurement of the shapes of A and B between a specified time interval t of the whole chromatogram run T, where $0 < t_0 < t_1 < T$, with a perfect match again rated as 1 and poorer matches lower than this. It is calculated as:

$$\frac{\int_{t_0}^{t_1} u_t v_t \,\mathrm{d}t}{\left[\int_{t_0}^{t_1} u_t^2 \,\mathrm{d}t \int_{t_0}^{t_1} v_t^2 \,\mathrm{d}t\right]^{0.5}}$$

where $u_t = B_t - B$ and $v_t = A_t - A$ ($B = B_t\,\mathrm{d}t/T$ and $A = A_t\,\mathrm{d}t/T$).

2.3 Euclidian distance

The Euclidian distance is a measure of dissimilarity which summates the difference in height of each corresponding peak and is useful where there are a few major differences between A and B. A and B may be normalized to unit area before matching. A perfect match gives a measurement of 0 and poor matches a high figure. It is calculated as:

$$\left[\int_{t_0}^{t_1} w_t(B_t - A_t)\,\mathrm{d}t\right]^{0.5}$$

where w_t is an optional Gaussian weighting function.

2.4 Box car distance

The box car distance is also a dissimilarity measure based on peak height and is useful where there are a lot of small differences. A and B may be normalized to unit area before matching. A perfect match gives a measurement of 0 and a poor match a higher figure. It is calculated as:

$$\int_{t_0}^{t_1} w_t|B_t - A_t|\,\mathrm{d}t$$

Two sweat samples collected two to four days apart were run for each of the four donors. Pairs of chromatograms were matched using different combinations of the second-load runs and the results analysed using the unpaired, one-tailed Students' t-test[10] assuming that the twins would be better matched than the unrelated pairs.

3 RESULTS AND DISCUSSION

Table 1 summarizes the results of matches between the chromatograms. There was a significant ($P < 0.007$–0.0001) difference between the twin matches and the unrelated matches on all four match parameters. The largest difference was obtained with the profile correlation, indicating that identity signals may be conveyed by variations in the concentration of certain ranges of volatiles. The difference in the alignment coefficient indicates that some part of the identity signal depends upon an absence or presence of certain compounds. The approximately two-fold differences seen in the Euclidian and box car distances confirm that variations in the

TABLE 1
Results of Pattern Analysis to Measure the Degree of Match of Pairs of Six Different Chromatograms

	Twins[a] *mean*[b] (±*S.D.*)	*Non-related*[a] *mean*[b] (±*S.D.*)	*One-tailed t-test t-value* (P^b)
Alignment coefficient	0·74 (±0·07)	0·41 (±0·14)	5·18 (0·0002)
Profile correlation	0·94 (±0·08)	0·08 (±0·28)	7·33 (0·0001)
Euclidian distance	0·13 (±0·04)	0·26 (±0·06)	4·26 (0·0009)
Box car distance	0·22 (±0·08)	0·42 (±0·14)	2·96 (0·007)

[a] $n = 6$ in each case (±S.D.).
[b] $df = 10$.

amount of certain compounds is an important factor in identity signalling.

As a test of the dependability of this technique and the effect of genetic influence, colleagues in the Department of Immunology at Kiel University asked us to analyse six unlabelled sweat samples from people of known antigenic makeup. We easily identified the single pair of identical twins and demonstrated a trend for similarity ranking on the basis of some common HLA Class 1 antigens. An account of this study is in preparation.

4 CONCLUSION

It seems that human identity is signalled by qualitative and quantitative differences in the sweat volatiles. We are in the process of examining how these differences are influenced by the major histocompatibility complex of genes.

The clear demonstration of a marked similarity in the sweat volatiles of identical twins shows that the methods used in this study have potential. They could be employed for comparing and contrasting chromatograms in many applications and also used to give some indication of the chemical basis of the matching.

ACKNOWLEDGEMENTS

We should like to express our deep gratitude to the sweat donors who have gone to a great deal of trouble on behalf of this project over the past six years. We should like to thank Dr Alan Bleasby who wrote the pattern-matching programme, Dr Josephine Pemberton who carried out the DNA profiling and Dr Nicholas Zavazava and his colleagues of Christian Albrechts Universität zu Kiel for collaboration. Last but not least, we should like to thank the Wellcome Trust who have funded the project for the past three years.

REFERENCES

1. Macku, C. & Jennings, W. G., Production of head-space volatiles by ripening bananas, *J. Agric. Food Chem.*, **35** (1987) 845–8.
2. Albone, E. S., *Mammalian Semiochemistry.* Wiley, Chichester, 1984.
3. Belcher, A., Epple, G., Kuderling, I. & Smith, A. B., Volatile components of scent material from cotton top tamarin (*Sanguinas o. oedipus*): a chemical and behavioural study, *J. Chem. Ecol.*, **14** (1986) 1367–84.
4. Sommerville, B. A., Darling, F. M. C., McCormick, J. P., Waterhouse, J. & Broom, D. M., A recirculating system for concentrating volatile samples, *Chromatographia*, **1** (1994) in press.
5. Sommerville, B. A., Green, M. A. & Gee, D. J., Using chromatography and a dog to identify some of the compounds in human sweat which are under genetic influence. *Chemical Signals in Vertebrates*, Vol. 2, ed. D. W. Macdonald, D. Muller-Schwartze & S. E. Natynczug. Oxford University Press, Oxford, 1990, pp. 634–9.
6. Szinak, J., Identification of odours. *Internat. Crim. Police Rev.*, March 1985, 58–63.
7. Liddell, K., Smell as a diagnostic marker, *Postgrad. Med. Jour.*, **52** (1976) 136–8.
8. Scientific Glass Engineering UK Ltd, Potters Lane, Kiln Farm, Milton Keynes MK11 3LA, UK.
9. Marshall, R. J., Bleasby, A. J., Turner, R. & Cooper, E. D., A computer system for analysis of chromatographic data, *Chemom. Intell. Lab. Syst.*, **1** (1987) 285–95.
10. Statview SE + Publ. Abicus Concepts, Inc., Berkeley, California, 1991.