Reports

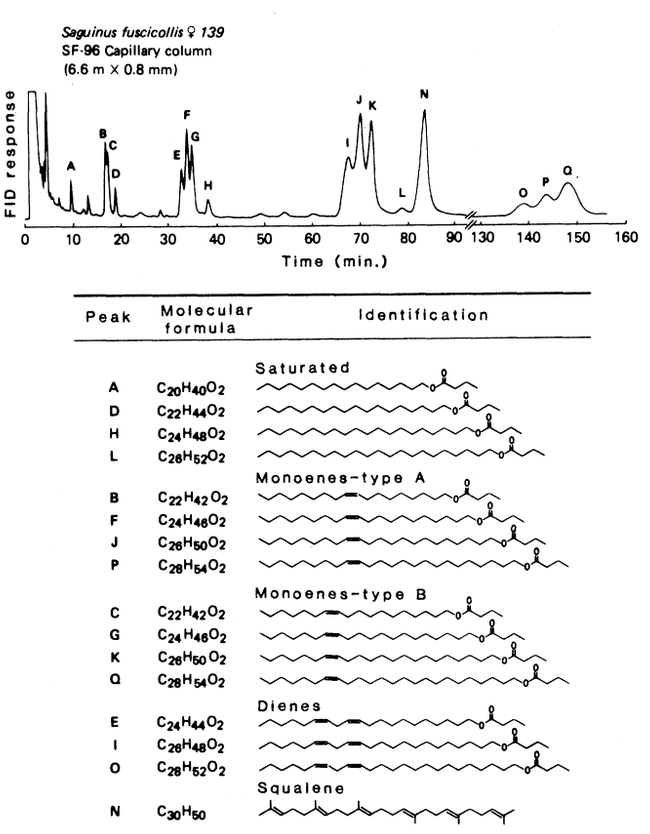
Computerized Pattern Recognition: A New Technique for the Analysis of Chemical Communication

Abstract. *Computerized pattern recognition techniques can be applied to the study of complex chemical communication systems. Analysis of high resolution gas chromatographic concentration patterns of the major volatile components of the scent marks of a South American primate,* Saguinus fuscicollis, *demonstrates that the concentration patterns can be used to predict the gender and subspecies of unknown donors.*

Chemical signals are important in transmitting information among conspecifics in many species of mammals (*1*, *2*). For example, work with several rodent species has established the presence of signals that alter behavior and endocrinological states of recipients (*2*). Chemical communication in primates, although perhaps of even greater complexity, appears to serve similar functions. Responses to conspecific messages are variable and depend on a large number of contextual factors such as sociosexual status, experience, and other sensory cues perceived by the recipient (*3*). In addition, chemical analyses indicate that these signals are part of highly complex mixtures. The isolation and identification of the biologically relevant portions of these mixtures have been difficult (*4*).

We have applied pattern recognition techniques to the analysis of the chemical signals utilized by primates, and now report that the identity of gender and subspecies in the tamarin monkey, *Saguinus fuscicollis*, can be directly correlated with specific concentration patterns of the major volatile constituents of the scent mark. This South American primate possesses specialized circumgenital-suprapubic scent glands that produce both sebaceous and apocrine secretions. This material, together with urine and possibly genital discharge, is deposited in the environment as scent marks. Behavioral studies have demonstrated that test animals can discriminate between material from males and females, from two different subspecies, and from individuals of differing reproductive condition or social status (*5*, *6*).

Chemical analysis of the scent mark has established the presence of 16 major volatile components, including 15 esters of *n*-butyric acid and squalene (*7*) (Fig. 1). The esters are of four types: saturated butyrates, monounsaturated *cis*-butyrates of two types (A and B) dependent on the location of the olefinic linkage, and *cis*, *cis*-diunsaturated butyrates (*8*). In addition to these esters and squalene, the scent material contains other highly volatile compounds that include short-



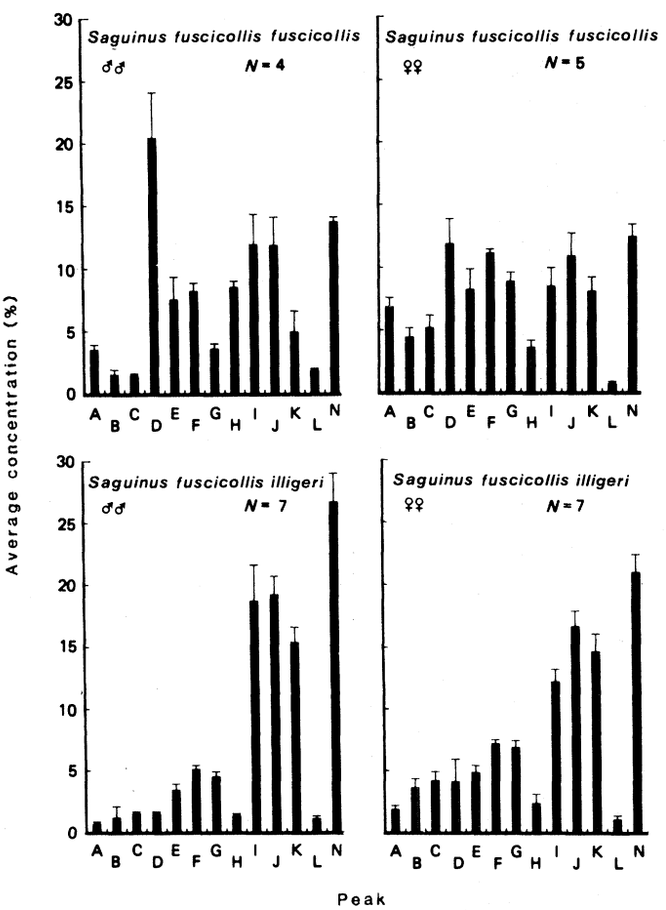


Fig. 1 (left). Identification of the major volatile components of the scent marks of *S. fuscicollis*. Fig. 2 (right). Concentrations of individual butyrates and squalene for males and females of two *S. fuscicollis* subspecies relative to total concentration.

chain organic acids, alcohols, aldehydes, alkanes, and esters (*9*).

No qualitative differences were apparent in the chemical composition of scent-mark volatiles between males and females, or between the three subspecies (*S. fuscicollis fuscicollis, S. fuscicollis illigeri*, and *S. fuscicollis nigrifrons*) maintained in our colony. It therefore appeared likely that quantitative differences in the components result in concentration profiles that identify gender and subspecies. In order to investigate this hypothesis, we monitored the concentrations of the esters and of squalene in the scent marks of males and females of two subspecies (*S. f. fuscicollis* and *S. f. illigeri*) by quantitative gas chromatography over a 6-year period. A data base of information concerning the relative concentrations of the butyrates and squalene for individuals of four donor types was compiled (Fig. 2) and used in pattern recognition analyses.

The technique of pattern recognition, which has been used to analyze gas chromatographic data (*10, 11*), starts with a set of measurements (descriptors) that classifies an object into its proper category. The object or pattern is represented as a point in *n*-dimensional space, where the axes of this space are defined by the set of descriptors used to characterize the object. For a binary classification problem (in the ideal case), one group will cluster in one region of this *n*-dimensional space, and the second group will cluster in another region. For this study, scent-mark material was collected from 23 male and 25 female *S. fuscicollis* donors (*12*). Of these 48 individuals, nine were *S. f. fuscicollis*, 14 were *S. f. illigeri*, one was *S. f. nigrifrons*, and 24 were hybrids (mainly between *S. f. fuscicollis* and *S. f. illigeri*). An average of 18 ± 4 (standard deviation) scent collections, each including two to three scent marks, were analyzed for each individual (*13*).

High-resolution capillary gas chromatography with an SF96 coated, open-tubular glass column was used to analyze scent-mark material. Integration of peak areas and calculation of relative retention times was done with a Perkin-Elmer PEP-II computer interfaced to the gas chromatograph. Concentrations of the butyrate esters and squalene were normalized and expressed as percent of total to indicate relative concentrations (*14*). Individual means and standard errors were then calculated for each animal.

Of the 48 donors, 40 comprised a training set from which measurements were taken in order to generate classification schemes (discriminants). Measurements

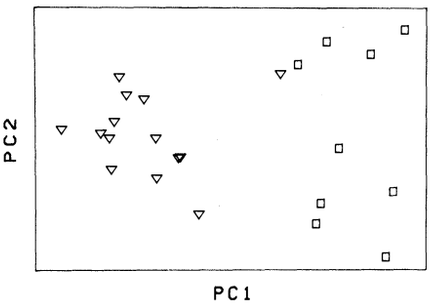


Fig. 3. A principal component representation of the pattern space defined by peaks A, D, F, G, H, K, and N (in Fig. 1). The cumulative variance accounted for by the first two principal components is 88 percent. The squares represent *S. f. fuscicollis* and the triangles are *S. f. illigeri* individuals.

obtained from the remaining eight animals (prediction set) were used to test the validity of the classification schemes. Members of the prediction set were chosen by random lot. The data were autoscaled (standardized) and examined by linear discriminant analysis (*11*) and the SIMCA pattern recognition method (*15*). A discriminant that correctly classified 38 members of the training set was developed from compounds B and C. This discriminant was tested on the prediction set and gave 100 percent correct classification. In the SIMCA study, a principal component model for each class was developed from compounds B, C, and H. Only four animals in the training set were misclassified, and all of the members of the prediction set were correctly classified.

Pattern recognition analysis of the scent prints for subspecies information was also successful. Because the number of animals available for this study was smaller than for the gender analysis (nine *S. f. fuscicollis* and 14 *S. f. illigeri* donors), a prediction set was not used. A principal component analysis (*16*) was performed on all 13 descriptors and a decision function was developed from the first principal component that misclassified only one animal. In the SIMCA study a one principal-component model for each class was developed from compounds A, D, F, G, H, K, and N that misclassified only one animal. The best two-dimensional display of the *n*-dimensional space where $n = 13$ (the number of peaks) and each point represents one animal is shown (Fig. 3). The two subspecies were well separated in this two-dimensional principal-component space.

These results demonstrate that information derived solely from the butyrate and squalene composition of the scent marks could categorize the donor as to gender and subspecies. This implies a direct relationship between the concentration of these compounds and the identity of gender and subspecies.

We have tested the discriminatory responses of subject animals to fractions of natural scent marks from males and females that contained only the major volatiles (butyrates and squalene). This material was obtained by preparative gas chromatographic fractionation. When the eluate was collected in two fractions, one containing the butyrates and squalene, the other the more highly volatile components, test animals did not differentiate male from female with either fraction. They did, however, exhibit a good, although not statistically significant, trend to investigate preferentially male scent marks in both fractions. However, when the entire eluate was collected in one fraction, a statistically significant preference for male over female material was obtained ($\alpha = 0.016$, Walsh test).

These results indicate that the butyrates and squalene alone do not contain sufficient information for gender discrimination. A number of additional components (presumably of higher volatility) in conjunction with the butyrates and squalene appear to be necessary to elicit the complete discriminative response. Our application of this novel technique to the problem of deciphering the chemosensory code of a primate suggests that computerized pattern recognition will be important in the evaluation of the informational content of highly complex chemical communication systems.

AMOS B. SMITH, III
ANNE M. BELCHER
Monell Chemical Senses Center and *Department of Chemistry, University of Pennsylvania, Philadelphia 19104*
GISELA EPPLE
Monell Chemical Senses Center and *German Primate Center, Göttingen, Federal Republic of Germany*
PETER C. JURS
BARRY LAVINE
Department of Chemistry, Pennsylvania State University, University Park 16802

References and Notes

1. M. Birch, Ed., *Pheromones* (North-Holland, Amsterdam, 1974).
2. D. Müller-Schwarze and R. M. Silverstein, Eds., *Chemical Senses* (Plenum, New York, 1983); D. M. Stoddart, Ed., *Olfaction in Mammals* (Academic Press, New York, 1980).
3. G. Epple, *Communication by Chemical Signals* (in press).
4. R. P. Michael, R. W. Bonsall, D. Zumpe, *Vitam. Horm. (N.Y.)* **34**, 137 (1976); D. A. Goldfoot, M. A. Kravetz, R. W. Goy, S. K. Freeman, *Horm. Behav.* **7**, 1 (1976); R. P. Michael, D. Zumpe, M. Richter, R. W. Bonsall, *ibid.* **9**, 296 (1977).
5. G. Epple, N. F. Golob, M.-S. Cebul, A. B. Smith, III, *Chem. Senses* **6**, 377 (1981).
6. Our bioassay procedures most frequently in-

volve "choice tests" during which the animals are allowed to investigate two stimulus plates, each carrying a different type of scent. Hansen scores for marking, sniffing, and contacts are computed and analyzed by nonparametric statistics. A statistically significant preference for one class of mark over another (such as male over female) indicates that information for discriminating these two classes is found in the mark [S. Siegel, *Nonparametric Statistics* (McGraw-Hill, New York, 1956].
7. A. B. Smith, III, R. G. Yarger, G. Epple, *Tetrahedron Lett.* **13**, 983 (1976); R. G. Yarger *et al.*, *J. Chem. Ecol.* **3**, 45 (1977).
8. As occurs in other mammalian fatty acid systems, each member of a class differs from the next by a two-carbon unit. The stereospecific total synthesis of each butyrate has been achieved [N. F. Golob, R. G. Yarger A. B. Smith III, *J. Chem. Ecol.* **5**, 543 (1979)]
9. A. M. Belcher, G. Epple, J. G. Kostelc, A. B. Smith, III, paper presented at the Association for Chemoreception Sciences, Sarasota, Fla., April 1982.
10. L. Kryger, *Talanta* **28**, 871 (1981).
11. P. C. Jurs and T. L. Isenhour, *Chemical Applications of Pattern Recognition* (Wiley-Interscience, New York, 1975).
12. The Monell Center maintains a successful breeding colony of *Saguinus fuscicollis*. Our resident population of 150 animals lives in family groups or pairs in wire mesh cages equipped with natural branches and sleeping boxes. Visual contact is limited by solid metal walls on two sides of each cage to avoid aggressive arousal in these easily stressed animals.
13. Material was collected by allowing the monkey to scent mark a 60-cm-long frosted glass plate. The animals are accustomed to this procedure and usually mark within seconds of introducing the plate into their home cage. After collection, each scent-marked plate was rinsed with 30 ml of a solution containing methylene chloride and methanol in a ratio of 3 to 1. For gas chromatographic analysis, the solvent mixture was removed by rotary evaporation at reduced pressure; the residue was then dissolved in a small amount of hexane and evaporated under nitrogen to a suitable volume (~15 μl).
14. The three slowest eluting compounds (O, P, and Q) were not included in the analysis, as computer integration of their relatively broad peaks gave inconsistent results.
15. S. Wold, *Pattern Recognition* **8**, 127 (1976).
16. B. B. Jackson, *Multivariate Data Analysis—An Introduction* (Irwin, Homewood, Ill., 1983).
17. Supported by NSF grants BNS 78-19875 and BNS 83-00937 (A.B.S.), NIH training grant 5 T32 NSO7176-03 (A.M.B.), and NSF grant CHE-8202620 (P.C.J.).

30 January 1984; accepted 13 December 1984

Flight of Winter Moths Near 0°C

Abstract. *Some noctuid winter moths fly at near 0°C by maintaining an elevated (30° to 35°C) thoracic muscle temperature. Geometrid winter moths sustain themselves in free flight at subzero muscle temperatures. However, the temperature characteristics of citrate synthase and pyruvate kinase from both of these different kinds of moths and from a sphinx moth that flies with a muscle temperature of 40°C are nearly identical. Furthermore, mass-specific rates of energy expenditure of both kinds of winter moths are also similar at given thoracic temperature (near 0°C). The geometrids that are able to fly with a thoracic temperature near 0°C do so largely because of unusually low wing-loading, which permits a low energetic cost of flight.*

Most moths that have been investigated are highly endothermic during flight (*1, 2*), and some species maintain a thoracic temperature (T_{Th}) near 45°C (*3*), up to 35°C above ambient temperature (T_A). The endothermic moths shiver (*4*) before flight until the power output of the muscles is sufficient for takeoff (*5*). The lower the T_A, the more energy the moths must expend for any given preflight warm-up or the more heat they must retain in their flight muscles to continue flight, or both (*6*). We investigated the question of why certain noctuid and geometrid moths are able to fly in the winter at zero or even subzero temperatures.

In one group of winter moths, the Cuculiinae (Noctuidae), the adults emerge in the fall (*7*) and fly throughout the winter when T_A approaches 0°C. Some of these moths (*Eupsilia morrisoni, E. tristigmata, E. sidus*, and *E. vinulenta*) were captured with sugar bait (*8*) in Maine and Vermont in February, and they warmed up by shivering in the laboratory at T_A as low as −3°C. They continued to shiver until T_{Th} exceeded 30°C (*9*), their minimum T_{Th} for sustained flight. Thoracic temperatures were maintained between 30° and 39°C during continuous free flight at T_A from 3° to 21°C (Fig. 1). Prolonged shivering at low T_A (*10*), a 1.0-mm-thick layer of insulating pile on the thorax, and greatly reduced heat flow from the thoracic muscles to the abdomen (*11*) allow these moths to resist the cold.

A second group of winter moths (Geometridae) also emerge in the fall. We examined two species, *Operophtera bruceata* Hulst. and *Alsophila pometaria* Harris that fly in late November (in northeastern United States) until the first heavy snowfalls stop their activity. In these moths the females are wingless, and the males (which do not feed) fly both at night and in the daytime. We found *O. bruceata* flying in the field in daytime at T_A as low as −3°C.

The winter geometrids are small moths (usually <10 mg) and they are uninsulated. They were at no time observed to either shiver or bask. Thoracic temperatures of flying moths were within one Celsius degree of T_A (Fig. 1). What allowed these insects to fly with a muscle temperature below the freezing point of water? Did they have specialized enzyme systems that operate at subzero temperatures? Many insects survive such temperatures by becoming inactive and adapting biochemically (*12*). But to our knowledge no insect has been shown to operate its wings in flight while its thoracic muscles are cooled to 0°C or lower.

We analyzed (*13*) the catalytic efficiency of citrate synthase (E.C. 4.1.3.7) in noctuid and geometrid winter moths over temperature (Fig. 2). Overall activities of the enzyme per gram of thorax fell into a common range (50 to 100 units, measured at 15°C), even though the noctuids flew with a T_{Th} near 35°C, while the geometrids flew with a T_{Th} as low as −3°C. Similarly, the thermal response of the enzyme in a sphinx moth, *Manduca sexta*, which normally flies with a T_{Th} near 40°C (*6*), also fell into the same range. These results contrast with the available data in other ectotherms, which show a pronounced increase in enzyme activity associated with decreasing temperature (*14*).

Enzyme activities in all moth species that we examined revealed identical behaviors over a range of temperatures, surprisingly independent of the T_{Th} necessary for flight. All data points fall on lines with slopes that are not significantly different from each other, giving the four moths enzymes with equal energies (E_a) and enthalpies ($\Delta H^{\ddagger}$) (*15*) of activation over the entire temperature range (E_a = 10,140 cal mol^{-1}; standard devi-

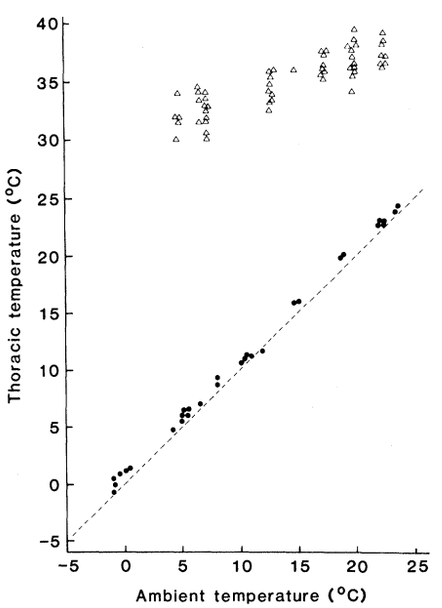


Fig. 1. Thoracic muscle temperatures of *Eupsilia* spp. (△) and *Operophtera bruceata* (●) in free flight, as a function of ambient temperature. Dashed line indicates the isothermal. Note that *O. bruceata* flies with thoracic temperature close to ambient temperature down to −3°C. The thoracic temperatures indicated also reflect the range of ambient temperature over which the moths were able to remain in continuous free flight.