# Multidimensional Data Analysis of Essential Oils. Application to Ylang-Ylang (*Cananga odorata* Hook Fil. et Thomson, Forma Genuina) Grades Classification

Emile M. Gaydou,* Robert P. Randriamiharisoa, Jean-Pierre Bianchini,[1] and Jean-Richard Llinas

Chemical compositions of 44 essential oils of ylang-ylang (*Cananga odorata* Hook Fil. et Thomson, forma genuina) from Madagascar were analyzed by glass capillary gas chromatography (GC). Classification of these essential oils according to their commercial grades (first, second, third) using physical and chemical constants was compared to classification achieved by applying multidimensional data analysis to GC results. Thirty-two GC peaks were used for standardized principal-component analysis (PCA) and factorial discriminant analysis (FDA). The differentiation of the three groups was obtained by either PCA or FDA. By using stepwise FDA, we observed that only 10 compounds are needed for the correct classification of the learning set samples.

Oil of ylang-ylang is obtained by steam distillation of the flowers of the tree *Cananga odorata* Hook. Fil. et Thomson, forma genuina (family Anonaceae). Four grades are currently available: extra, first, second, and third. These grades are generally obtained in the Malagasy Republic by the separation of distillates collected in bottles containing from 0.3–0.4 L of products. Densities are noted as the distillation progresses for the determination of the grade. The qualities of the various grades are controlled before the exportation by a national laboratory of the Malagasy Republic, which used density at 20 °C and ester value in their determinations. From the producer point of view, distillation of ylang-ylang is not easy, and in order to produce constant qualities, the nature of the compounds and their analytical ranges are needed. With the development of analytical methods such as gas chromatography (GC) for the investigation of essential oils and by use of a glass or silica capillary column, finer separations are obtained. In a recent study of ylang-ylang, first grade, after fractionation, more than 52 compounds have been identified (Gaydou et al., 1986).

Since fingerprint chromatograms are commonly used by perfumers to control the essential oil qualities, we have determined in this study the range of variation of 32 GC peak area percentages for 44 ylang-ylang oil samples obtained during the years 1978–1982. Ylang-ylang production is classified in three grades in the Malagasy Republic: first, second, and third. For the determination of compounds playing important parts in the differentiation of the various ylang grades, we used in this study multivariate statistical methods to analyze the GC data. Multidimensional data analysis has been used and proven to be successful in a wide variety of chemical problems, and general reviews of the theory and application have been published (Powers and Keith, 1968; Kowalski, 1980; Schoenfeld and De Voe, 1976; Frank and Kowalski, 1982; Lebart et al., 1982; Delaney, 1984; Llinas and Ruiz, 1986; Ramos et al., 1986).

Laboratoire de Phytochimie and Ecole Supérieure de Chimie de Marseille, Université de Droit d'Economie et des Sciences, Centre de Saint Jérôme, avenue Escadrille Normandie-Niémen, 13397 Marseille Cédex 13, France.

[1] Present address: Département des Industries Agricoles et Alimentaires, Etablissement d'Enseignement Supérieur des Sciences Agronomiques, BP 175, Antananarivo 101, République Démocratique de Madagascar.

## EXPERIMENTAL SECTION

**Oils Samples.** A total of 44 samples of ylang-ylang oils (20 of the first grade, 13 of the second grade, 11 of the third grade) were obtained by hydrodistillation from flowers harvested during 1978–1982. The samples investigated were obtained from the principal producers of the Malagasy Republic (Nosy-Be, Ambanja, Mahajanga, Bevoay areas). These oils were supplied by the Service du Conditionnement et du Contrôle de la Qualité des Produits of Antananarivo (Malagasy Republic) who guarantee their authenticities. All samples were stored at −15 °C before analysis.

**Physical and Chemical Constants.** Densities ($d_{20}^{20}$) and ester values (saponification during 1 h) were determined according to NF T 75-111 and NF T 75-104 norms, respectively (Afnor, 1986).

**Gas Chromatography (GC).** A Delsi Model 30 gas chromatograph (Delsi, France), equipped with a fid and a glass injector, was used for the analyses. The column employed was an OV-101 WCOT glass capillary column (50-m length, 0.30-mm i.d., 0.15-$\mu$m phase thickness). The temperature was programmed from 90 to 220 °C at 2 °C min$^{-1}$. The inlet pressure of hydrogen used as carrier gas was 1.5 bar (linear velocity 40 mL s$^{-1}$). Peak areas were integrated by a Shimadzu ICR-1B integrator.

**Gas Chromatography–Mass Spectrometry (GC–MS).** Combined GC–MS were recorded on a Girdel 30 gas chromatograph linked to a Ribermag R-10-10B mass spectrometer and coupled with a Sidar data computer; the GC column was a 0.30-mm i.d. × 50-m length fused silica capillary column coated with OV 101, 0.15-$\mu$m phase thickness. The column temperature was programmed from 70 to 210 °C at 2 °C min$^{-1}$, carrier gas helium, ion source 220 °C, ionization voltage 70 eV.

**Identification of Compounds.** The various compounds were identified by comparison of their Kovats retention indices and mass spectra with our reported data (Gaydou et al., 1986).

**Statistical Analyses.** Two methods of multidimensional data analysis were used in the interpretation of data and classification of the commercial three grades of ylang-ylang oils. Principal-component analysis (PCA) was performed using a data set transformed into centered and reduced variables (standardized PCA). The data set was composed of 32 variables determined for the 44 oil samples. Factorial discriminant analysis (FDA) (Fischer, 1936; Mahalanobis, 1936) was performed to classify the oil sam-

0021-8561/88/1436-0574$01.50/0 

**Table I. Published Standards of Oil of Ylang-Ylang from Madagascar**

| property | grade | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | extra | | first | | second | | third | |
| | min | max | min | max | min | max | min | max |
| density ($d_{20}^{20}$) | | | | | | | | |
| Malagasy customs authorities | 0.950 | | 0.927 | | 0.917 | | 0.905 | |
| Afnor[a] | 0.950 | 0.965 | 0.933 | 0.945 | 0.923 | 0.929 | 0.906 | 0.921 |
| Grasse EOA[b] | 0.946 | 0.982 | 0.928 | 0.949 | 0.918 | 0.933 | 0.906 | 0.923 |
| refractive index at 20 °C | | | | | | | | |
| Afnor[a] | 1.5010 | 1.5090 | 1.5000 | 1.5100 | 1.5050 | 1.5110 | 1.5060 | 1.5130 |
| Grasse EOA[b] | 1.4980 | 1.5090 | 1.5000 | 1.5090 | 1.5057 | 1.5117 | 1.5070 | 1.5150 |
| opt rotation (deg) at 20 °C | | | | | | | | |
| Afnor[a] | −45 | −36 | −44 | −28 | −55 | −40 | −63 | −49 |
| Grasse EOA[b] | −40 | −25 | −60 | −44 | −68 | −40 | −67 | −35 |
| acid value | | | | | | | | |
| Afnor[a] | | <3 | | <3 | | <3 | | <3 |
| Grasse EOA[b] | | 2.8 | | 2.8 | | 2.8 | | 2.8 |
| ester value | | | | | | | | |
| Malagasy customs authorities | 125 | | 90 | | 58 | | 40 | |
| Afnor[a] | 125 | 160 | 90 | 120 | 65 | 80 | 38 | 58 |
| Grasse EOA[b] | 130 | 182 | 89 | 130 | 56 | 89 | 34 | 54 |

[a] Experimental norm NF T 75-246 (Afnor, 1986). [b] Grasse Essential Oil Association (1959).

**Table II. Density and Ester Values of Investigated Oils of Ylang-Ylang from Madagascar**

| property | grade | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | first[a] | | | | second[b] | | | | third[c] | | | |
| | min | max | mean | cv | min | max | mean | cv | min | max | mean | cv |
| density ($c_{20}^{20}$) | 0.930 | 0.955 | 0.942 | 0.008 | 0.918 | 0.938 | 0.926 | 0.005 | 0.913 | 0.923 | 0.918 | 0.003 |
| ester value | 90 | 130 | 106.3 | 11.5 | 58 | 89.7 | 72.1 | 11.2 | 42.4 | 56.0 | 51.4 | 3.8 |

[a] Determined from 20 samples. [b] Determined from 13 samples. [c] Determined from 11 samples.

ples into three categories corresponding to the three grades.

The search for compounds having the best discriminant power was realized with a stepwise FDA (Romeder, 1973; Rao, 1973). All processing was done by mean of the ANDON software on the computer (Hewlett-Packard HP 1000) of the Ecole Supérieure de Chimie de Marseille (Llinas, 1982).

## RESULTS AND DISCUSSION

**Physical and Chemical Constants of Ylang-Ylang Oils.** The oils are pale to deep yellow liquids having a harsh floral odor. The oils investigated were classified in three grades: (first, second, third) by the National Laboratory of the Malagasy Republic, on the basis of physical and chemical analyses. Samples of extra grade were not investigated since this grade was not produced during 1978–1982 in this country. The principal reason for fractionation of this essential oil is probably to ensure that the first oils emerging from the still may contain the finest parts in ester content. For fractionation to be good, the qualities should be within the strictest quality standards. As shown in Table I, the analytical ranges of published standards differ among various organizations: Malagasy customs autorities, Grasse Essential Oil Association, and Afnor. The French norm (Afnor, 1986) has been established to avoid the overlapping between some such as density. Table II gives the densities and ester values of the various grade samples of ylang-ylang investigated. The means of these samples are in agreement with the standards. Some samples of first and the second grades have a density or an ester value out of the limits of the standards of NF T 75-246 norm (Afnor, 1986) but are in agreement with Malagasy customs authorities (Table I).

**Composition of Ylang-Ylang Oils.** A typical chromatogram of the various grades of ylang-ylang oils is given in Figure 1. A total of 32 peaks were quantified for the 44 oil samples, which were all chromatographed on the same column. The same conditions were used so that all of the results were obtained under nearly identical conditions. Although many compounds have been identified (Gaydou et al., 1986), information about the content of each compound in the various grades is rarely encountered. In a recent survey, Buccellato (1982) has given the percentages of 25 compounds found in the four grades of these oils.

In our work, the compounds were identified as far as possible by GC–MS, Kovats retention indices, and comparison with the literature data (Gaydou et al., 1986). Table III lists the approximate concentrations (mean and range) of the major identified compounds (only two were not identified) in the oils. The percentages were determined by the method of peak area normalization and without the application of response factor corrections. Peaks were mostly identified by their accepted trivial names. The peak numbers in Table III show the elution order on the OV 101 column (Figure 1). The main compounds of the three grades of ylang-ylang oils are, for the sesquiterpenic compounds, $\beta$-caryophyllene and germacrene D. The content of the first sesquiterpene is always slightly higher than that of the second (7.4% vs 7.3% in the first; 14.3% vs 10.6% in the second; 17.5% vs 20.7% in the third). $\alpha$-Humulene + $\epsilon$-cadinene (2.5%, 4.6%, 4.7%), $\gamma$-cadinene + $\alpha$-farnesene (2.6%, 6.4%, 11.6%), and $\delta$-cadinene (1.5%, 3.7%, 4.0%) were found in a relatively higher content. As shown in Table III, the content of sesquiterpenes increased from the first to the third grade. Besides, the content of the light oxygenated compounds decreased from the first to the third grades: linalool (21%, 6.7%, 2.7%), geraniol (1.6%, 0.8%, 0.3%), benzyl acetate (5.5%, 1.4%, 0.7%), geranyl acetate (8.2%, 5.2%, 3.2%), methyl benzoate (4.1%, 1.1%, 0.4%), methyl salicylate (6.8%, 2.2%, and 0.7%), and *p*-methylanisole (6.3%, 1.8%, 0.7%). The percentages of some heavy oxygenated compounds such as (*E*,*E*)-farnesyl acetate, benzyl benzoate, and benzyl salicylate were quite similar from the first to the third grades. Therefore, multivariate statistical

**Table III. Relative Composition[a] of Ylang-Ylang Essential Oils from Madagascar**

| | | grade | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | first[b] | | | | second[c] | | | | third[d] | | | |
| no.[e] | peak | min | max | mean | cv | min | max | mean | cv | min | max | mean | cv |
| | | | | | | Monoterpenes | | | | | | | |
| 5 | α-pinene | 0.09 | 0.41 | 0.27 | 0.08 | 0.10 | 1.18 | 0.26 | 0.27 | 0.03 | 0.23 | 0.14 | 0.05 |
| 6 | β-pinene | 0.06 | 0.20 | 0.10 | 0.03 | 0.03 | 0.17 | 0.07 | 0.04 | 0.01 | 0.06 | 0.03 | 0.02 |
| 7 | myrcene | 0.05 | 0.21 | 0.11 | 0.04 | 0.02 | 0.09 | 0.05 | 0.02 | 0.01 | 0.11 | 0.04 | 0.03 |
| | | | | | | Sesquiterpenes | | | | | | | |
| 25 | β-caryophyllene | 1.10 | 11.21 | 7.38 | 2.57 | 1.74 | 19.62 | 14.32 | 4.75 | 14.81 | 21.45 | 17.46 | 1.93 |
| 27 | α-humulene + ε-cadinene | 1.39 | 3.65 | 2.52 | 0.64 | 3.43 | 5.84 | 4.62 | 0.81 | 3.86 | 5.82 | 4.70 | 0.56 |
| 29 | γ-muurolene | 0.51 | 1.90 | 1.12 | 0.44 | 1.52 | 3.83 | 2.83 | 0.60 | 2.02 | 3.21 | 2.52 | 0.36 |
| 30 | germacrene D | 0.10 | 13.48 | 7.26 | 4.26 | 1.51 | 19.26 | 10.58 | 5.13 | 15.13 | 25.13 | 20.67 | 3.26 |
| 32, 33 | γ-cadinene α-farnesene | 0.30 | 4.94 | 2.61 | 1.47 | 1.66 | 12.73 | 6.45 | 3.53 | 6.45 | 17.42 | 11.58 | 3.12 |
| 34 | α-copaene | 0.40 | 1.60 | 1.02 | 0.34 | 1.86 | 3.58 | 2.65 | 0.42 | 2.08 | 3.48 | 2.78 | 0.48 |
| 35 | α-muurolene | 0.21 | 1.06 | 0.54 | 0.23 | 0.70 | 1.88 | 1.27 | 0.33 | 0.78 | 1.63 | 1.14 | 0.27 |
| 37 | δ-cadinene | 0.23 | 2.78 | 1.51 | 0.64 | 2.12 | 5.17 | 3.70 | 0.88 | 3.09 | 4.76 | 3.98 | 0.45 |
| | | | | | | Oxygenated Compounds | | | | | | | |
| 3 | 3-methyl-2-buten-1-yl acetate | 0.06 | 0.40 | 0.21 | 0.08 | 0.01 | 0.20 | 0.05 | 0.06 | 0.00 | 0.02 | 0.00 | 0.01 |
| 4 | 2-methyl-3-buten-2-ol | 0.25 | 1.31 | 0.61 | 0.32 | 0.03 | 0.20 | 0.07 | 0.05 | 0.00 | 0.07 | 0.02 | 0.02 |
| 8 | 3-methyl-2-buten-1-ol | 0.05 | 0.20 | 0.12 | 0.04 | 0.02 | 0.11 | 0.04 | 0.03 | 0.00 | 0.02 | 0.00 | 0.01 |
| 8bis | bis(3-methyl-3-buten-1-ol) | 0.01 | 0.18 | 0.09 | 0.06 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.01 | 0.00 | 0.00 |
| 9 | p-methylanisole | 1.09 | 10.44 | 6.26 | 2.42 | 0.55 | 5.29 | 1.85 | 1.36 | 0.41 | 1.85 | 0.72 | 0.40 |
| 10 | 1,8-cineole | 0.35 | 0.83 | 0.54 | 0.15 | 0.10 | 0.33 | 0.19 | 0.07 | 0.00 | 0.10 | 0.05 | 0.03 |
| 11 | methyl benzoate | 1.73 | 5.57 | 4.13 | 0.99 | 0.59 | 2.32 | 1.06 | 0.50 | 0.18 | 0.62 | 0.38 | 0.13 |
| 12 | linalool | 11.69 | 29.96 | 20.84 | 5.20 | 3.86 | 12.15 | 6.71 | 2.48 | 1.29 | 4.75 | 2.74 | 1.03 |
| 13 | benzyl acetate | 3.29 | 8.02 | 5.52 | 1.35 | 0.61 | 3.08 | 1.37 | 0.67 | 0.36 | 1.23 | 0.67 | 0.26 |
| 14 | methyl salicylate | 1.74 | 10.44 | 6.38 | 2.08 | 0.55 | 5.29 | 2.24 | 1.64 | 0.35 | 1.85 | 0.71 | 0.41 |
| 16 | geraniol | 0.92 | 3.04 | 1.57 | 0.56 | 0.10 | 1.21 | 0.84 | 0.34 | 0.11 | 0.71 | 0.30 | 0.16 |
| 19 | geranyl acetate | 6.19 | 10.98 | 8.24 | 1.40 | 2.64 | 7.24 | 5.17 | 1.32 | 1.97 | 4.37 | 3.18 | 0.79 |
| 24 | (*E*)-cinnamyl acetate | 0.57 | 1.87 | 1.10 | 0.32 | 0.27 | 1.85 | 0.61 | 0.38 | 0.01 | 0.47 | 0.34 | 0.13 |
| 44 | δ-cadinol | 0.12 | 2.98 | 0.65 | 0.61 | 0.17 | 1.03 | 0.69 | 0.27 | 0.12 | 0.75 | 0.54 | 0.21 |
| 45 | α-cadinol | 0.19 | 2.00 | 0.76 | 0.43 | 0.45 | 1.90 | 1.31 | 0.50 | 0.75 | 1.54 | 1.17 | 0.22 |
| 46 | muurolol T + γ cadinol | 0.09 | 2.44 | 0.53 | 0.50 | 0.40 | 1.46 | 0.86 | 0.29 | 0.42 | 0.94 | 0.67 | 0.16 |
| 47 | (*E,E*)-farnesyl acetate | 0.49 | 7.82 | 2.26 | 1.79 | 0.72 | 6.23 | 2.97 | 1.31 | 1.60 | 3.07 | 2.44 | 0.36 |
| 49 | benzyl salicylate | 0.32 | 3.41 | 1.32 | 0.81 | 1.02 | 3.86 | 2.20 | 0.86 | 0.70 | 2.73 | 1.69 | 0.55 |
| 50 | benzyl benzoate | 4.34 | 14.85 | 8.60 | 3.07 | 5.31 | 12.29 | 8.41 | 2.09 | 5.90 | 12.80 | 8.69 | 1.74 |
| | | | | | | Unknown Compounds | | | | | | | |
| 22 | | 0.32 | 1.70 | 0.79 | 0.33 | 0.86 | 1.89 | 1.48 | 0.32 | 1.29 | 2.68 | 1.62 | 0.41 |
| 23 | | 0.27 | 0.68 | 0.45 | 0.12 | 0.52 | 1.79 | 0.86 | 0.30 | 0.62 | 0.97 | 0.79 | 0.09 |

[a] Determined by the method of peak area normalization using a nonpolar stationary phase (OV101). [b] Determined from 20 samples. [c] Determined from 13 samples. [d] Determined from 11 samples. [e] Peak number on the chromatograms (see Figure 1).

analyses were used to obtain further differentiation of the three grades from the GC composition of the various samples.

**Multidimensional Data Analysis.** The data set of the 44 ylang-ylang oil samples was composed of 20 samples of first grade, 13 samples of second grade, and 11 samples of third grade on the basis of their physicochemical properties. The correlation matrix reveals some intercorrelations between variables. Light-oxygenated compounds such as 3, 4, 8, and 14 (Table III) are positively correlated together ($r > 0.8$) and negatively correlated to most of sesquiterpenes (27, 29, 34, 35, 37). A second group of sesquiterpenes (germacrene D, γ-cadinene, α-farnesene) is highly positively correlated together and negatively correlated with geranyl acetate, and finally a third group of compounds composed of heavy oxygenated compounds is highly positively correlated together. Principal-component analysis (PCA) on reduced and centered variables was realized by diagonalization of the correlation matrix, and its main results are given in Table IV. The graphical representation of the projection of samples onto the two first principal components is shown in Figure 2. As indicated in Table IV, the first five components have a variance greater than unity account for 83% of the total information. The first component (58.3%) leads to a separation of the first-quality samples from the two other grades. This differentiation comes from the light-oxygenated compounds (3–24), which are highly positively correlated with this axis, and the sesquiterpenes (25–37 except germacrene D and γ-cadinene + α-farnesene), which are highly negatively loaded with this axis. This result is in agreement with the fact that light-oxygenated compounds are in higher amounts in the first quality of ylang-ylang oils and sesquiterpenes range in an opposite way. An appropriate differentiation between the second and the third grades is realized by the second component (11.1% of the total variance), which is highly positively correlated with germacrene D and γ-cadinene + α-farnesene and negatively correlated with heavy oxygenated compounds (Table IV). These heavy oxygenated compounds (44–50), the contents of which are higher in the second than in third qualities, are shown in Table III. A reverse phenomenom is observed for germacrene D and γ-cadinene + α-farnesene. The other axes are determined by particular trend in some compounds: axis 3 by benzyl benzoate and α-pinene, axis 4 by δ-cadinol and muurolol T + γ-cadinol, and axis 5 by benzyl salicylate (Table IV). For better discrimination of the three grades, on the basis of unambiguous assignment of samples grades, we used factorial discriminant analysis (FDA), first on the whole data set and second in a stepwise way in order to analyze the discrimination power of each compound concentration. Only two discriminant axes are built up for the three groups discrimination. The analysis was realized on centered and reduced variables with Mahalanobis' metric (1936). The affection of samples to the groups in done with

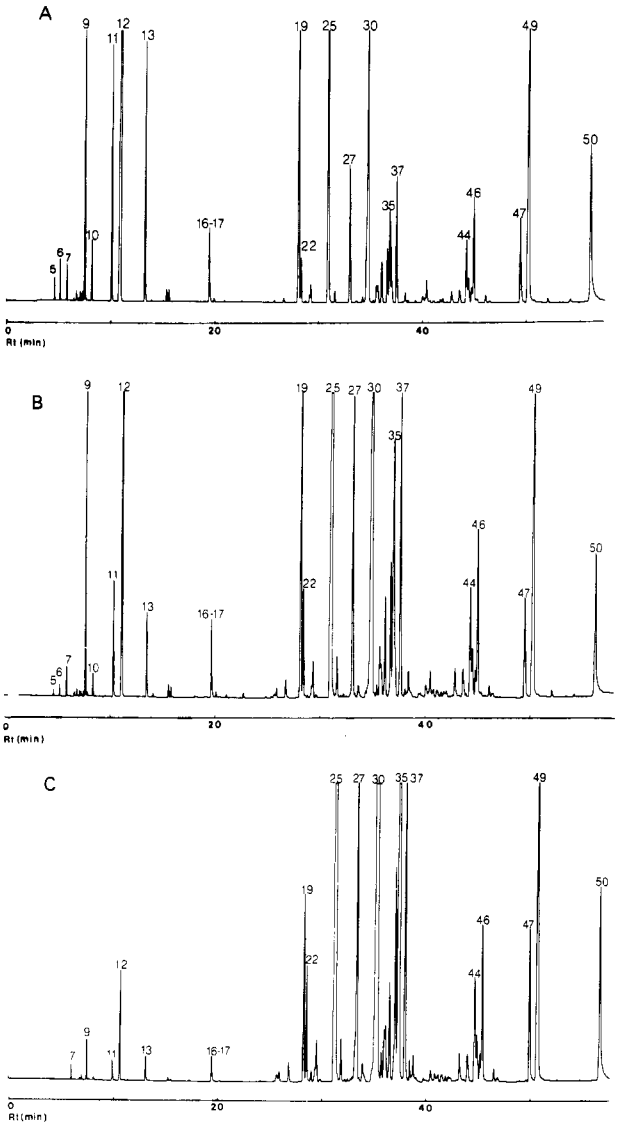


**Figure 1.** Typical gas chromatograms of ylang-ylang oils: (A) first grade; (B) second grade; (C) third grade. See Table III for peak identification.

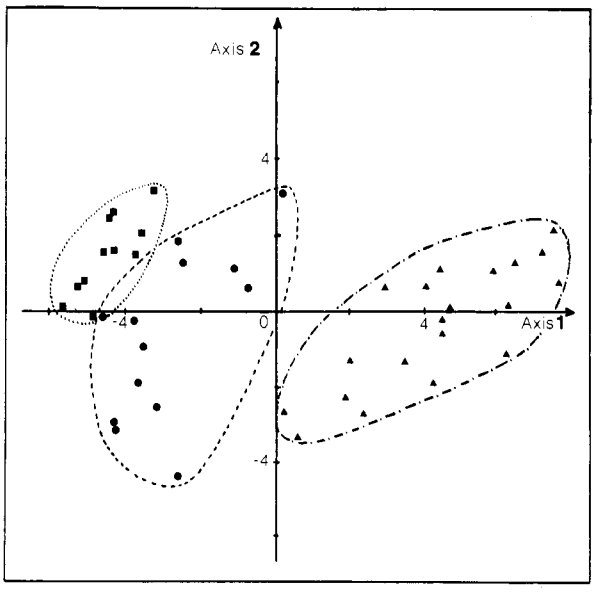


**Figure 2.** Graphic projection of the 44 ylang-ylang oil samples used in PCA from 32-dimensional space onto 2-dimensional plot preserving 69.4% of total variance: (▲) first grade; (●) second grade; (■) third grade.

**Table IV. Factor Loadings, Eigenvalues, and Percentage of Variance Using Standardized PCA for the Various Investigated Ylang-Ylang Oils[a]**

| variable[b] | component (axis) 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| | Monoterpenes | | | | |
| 5 | 0.26 | −0.18 | 0.59 | 0.08 | 0.42 |
| 6 | 0.72 | −0.01 | 0.34 | 0.07 | 0.22 |
| 7 | 0.76 | 0.15 | 0.45 | 0.04 | 0.06 |
| | Sesquiterpenes | | | | |
| 25 | −0.80 | 0.21 | 0.26 | −0.02 | 0.02 |
| 27 | −0.93 | −0.12 | 0.20 | −0.04 | 0.10 |
| 29 | −0.91 | −0.41 | 0.22 | 0.01 | 0.14 |
| 30 | −0.54 | 0.74 | −0.07 | 0.16 | −0.09 |
| 32,33 | −0.64 | 0.67 | −0.05 | 0.08 | −0.09 |
| 34 | −0.91 | 0.08 | 0.08 | 0.08 | 0.14 |
| 35 | −0.84 | −0.32 | 0.17 | −0.08 | 0.13 |
| 37 | −0.90 | 0.16 | 0.20 | 0.03 | −0.00 |
| | Oxygenated Compounds | | | | |
| 3 | 0.92 | 0.06 | −0.18 | 0.14 | 0.15 |
| 4 | 0.85 | 0.01 | −0.25 | −0.02 | 0.06 |
| 8 | 0.90 | −0.08 | 0.19 | 0.08 | −0.03 |
| 8bis | 0.71 | −0.08 | −0.11 | 0.13 | 0.13 |
| 9 | 0.92 | 0.14 | 0.15 | 0.05 | 0.02 |
| 10 | 0.87 | −0.24 | 0.12 | −0.08 | −0.14 |
| 11 | 0.98 | −0.02 | −0.02 | −0.04 | −0.04 |
| 12 | 0.95 | −0.13 | 0.11 | −0.10 | −0.12 |
| 13 | 0.97 | −0.02 | −0.16 | −0.03 | 0.00 |
| 14 | 0.95 | 0.15 | 0.06 | 0.05 | 0.02 |
| 16 | 0.75 | −0.22 | 0.36 | 0.04 | −0.08 |
| 19 | 0.78 | −0.45 | 0.15 | −0.20 | −0.04 |
| 24 | 0.73 | −0.09 | −0.24 | 0.22 | 0.31 |
| 44 | −0.15 | −0.58 | 0.05 | 0.61 | −0.45 |
| 45 | −0.71 | −0.60 | −0.01 | −0.10 | 0.02 |
| 46 | −0.45 | −0.53 | 0.03 | 0.64 | −0.22 |
| 47 | −0.37 | −0.71 | 0.03 | −0.35 | −0.03 |
| 49 | −0.37 | −0.23 | −0.29 | 0.43 | 0.64 |
| 50 | −0.76 | −0.57 | −0.56 | −0.28 | 0.08 |
| | Unknown Compounds | | | | |
| 22 | −0.84 | −0.13 | 0.11 | −0.14 | −0.04 |
| 23 | −0.71 | −0.12 | 0.13 | −0.09 | 0.08 |
| eigenvalues | 18.7 | 3.55 | 1.76 | 1.42 | 1.18 |
| %[c] | 58.3 | 11.1 | 5.5 | 4.4 | 3.7 |

[a] Determined for 44 samples. [b] The number of variables corresponds to the peak number on the chromatograms. For the identification see Table III. [c] Percentage of information (total variance).

a distance criterion of samples relative to the center of gravity. The factor loadings of variables with axes are given in Table V, and the graphical representation of samples is shown in Figure 3. The three groups corresponding to the three commercial grades of ylang-ylang oils are quite well differentiated with no overlapping, and so all the samples are correctly classified (100% of correct assignment). The first axis, associated to the eigenvalue 0.99 (discriminant power), induces a good separation of the three groups on the basis of the light-oxygenated compounds and sesquiterpene compounds. Further separation of the second- from the third-grade samples occurs along the second axis (discriminant power 0.94), which is highly correlated to the heavier oxygenated compounds. Supplementary samples of these three grades coming from the same Malagasy origin have been well classified and confirm the stability of the two discriminant factors in the operating field.

Stepwise FDA allows us to search for the more discriminant compounds. In this method a new variable (compound) is added in each step, and we can follow the evolution of the discriminant power of each axis and of their sum. The compounds, classified according to their rank of introduction in the factors, are given in Table V with

**Table V. Factor Loading of Variables Using FDA and Classification of Variables by Decreasing Order of Discrimination Power Using Stepwise FDA**

| | | FDA factor loading[b] | | stepwise FDA: evolution in discriminant power[c] | | |
|---|---|---|---|---|---|---|
| no.[a] | identification | axis 1 | axis 2 | axis 1 | axis 2 | no. misclassifications[d] |
| 11 | methyl benzoate | −0.95 | −0.30 | - | - | - |
| 30 | germacrene D | 0.95 | −0.30 | 0.87 | 0.29 | 8 |
| 14 | methyl salicylate | −0.97 | −0.24 | 0.88 | 0.46 | 5 |
| 47 | (*E*,*E*)-farnesyl acetate | 0.41 | 0.91 | 0.90 | 0.51 | 4 |
| 29 | γ-muurolene | 0.83 | 0.56 | 0.90 | 0.55 | 5 |
| 34 | α-copaene | 0.92 | 0.38 | 0.91 | 0.58 | 3 |
| 23 | unknown | 0.83 | 0.55 | 0.92 | 0.63 | 1 |
| 3 | 3-methyl-2-buten-1-yl acetate | −0.97 | −0.25 | 0.92 | 0.71 | 1 |
| 10 | 1,8-cineol | −0.98 | −0.20 | 0.93 | 0.73 | 1 |
| 45 | α-cadinol | 0.79 | 0.61 | 0.94 | 0.74 | 0 |
| 25 | β-caryophyllene | 0.98 | 0.18 | 0.95 | 0.75 | 0 |
| 24 | (*E*)-cinnamyl acetate | −0.99 | −0.13 | 0.96 | 0.75 | 0 |
| 37 | δ-cadinene | 0.93 | 0.35 | 0.96 | 0.78 | 0 |
| 50 | benzyl benzoate | 0.16 | 0.97 | 0.96 | 0.81 | 0 |
| 5 | α-pinene | −0.90 | 0.44 | 0.96 | 0.82 | 0 |
| 4 | 2-methyl-3-buten-2-ol | −0.93 | −0.37 | 0.96 | 0.85 | 0 |
| 13 | benzyl acetate | −0.94 | −0.33 | 0.96 | 0.87 | 0 |
| 46 | muurolol T + γ-cadinol | 0.57 | 0.82 | 0.97 | 0.87 | 0 |
| 6 | β-pinene | −0.99 | 0.09 | 0.97 | 0.89 | 0 |
| 49 | benzyl salicylate | 0.56 | 0.82 | 0.97 | 0.90 | 0 |
| 16 | geraniol | −0.99 | −.06 | 0.97 | 0.90 | 0 |
| 9 | *p*-methylanisole | −0.96 | −0.28 | 0.97 | 0.92 | 0 |
| 22 | unknown | 0.95 | 0.31 | 0.97 | 0.92 | 0 |
| 8 | 3-methyl-2-buten-1-ol | −0.99 | −0.14 | 0.97 | 0.93 | 0 |
| 35 | α-muurolene | 0.83 | 0.55 | 0.98 | 0.93 | 0 |
| 19 | geranyl acetate | −0.99 | −0.09 | 0.98 | 0.93 | 0 |
| 12 | linalool | −0.96 | −0.26 | 0.98 | 0.93 | 0 |
| 27 | α-humulene + ε-cadinene | 0.91 | 0.42 | 0.99 | 0.93 | 0 |
| 32, 33 | γ-cadinene + α-farnesene | 0.99 | −0.10 | 0.99 | 0.94 | 0 |
| 44 | δ-cadinol | −0.61 | 0.79 | 0.99 | 0.94 | 0 |
| 8bis | bis(3-methyl-3-buten-1-ol) | −1.00 | −0.01 | 0.99 | 0.94 | 0 |
| 7 | myrcene | −0.95 | −0.32 | 0.99 | 0.94 | 0 |

[a] Peak number on the chromatograms (see Figure 1). [b] Factor loadings using FDA with the 32 variables. [c] Discriminant power obtained with each subset of variables. [d] 44 samples used.

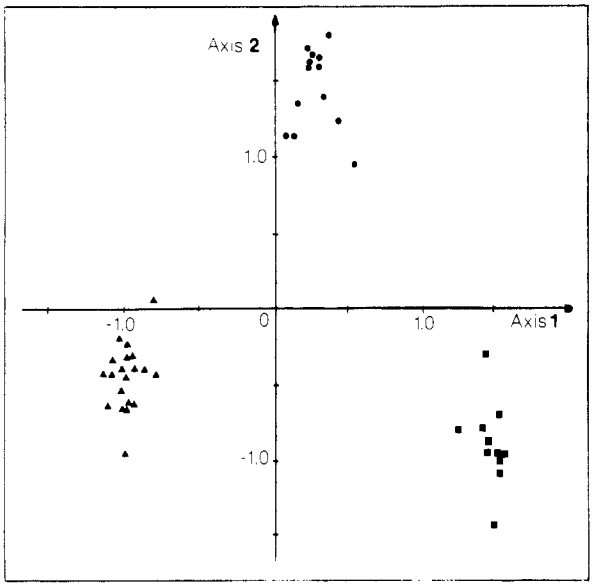


**Figure 3.** Two-dimensional plot of the three grades of ylang-ylang oil samples investigated in FDA: (▲) first grade; (●) second grade; (■) third grade.

the related sum of the discriminant power on the two axes together and the number of misclassified samples. With only the first six compounds (methyl benzoate, germacrene D, methyl salicylate, (*E*,*E*)-farnesyl acetate, γ-muurolene, α-copaene), 93% of the initial data set samples are well assigned. Finally all the samples of the learning set are well classified with only 10 variables. However, the classification of new samples will probably be better with the whole set of variables.

## CONCLUSION

Classification rules obtained either with PCA or with FDA analysis of the content of compounds using gas chromatography prove to be helpful in assigning ylang-ylang oil samples to the correct grade and to avoid, on another hand, adulteration. The original 32 chromatographic peaks could be reduced to as few as 10–15 peaks that still retained the necessary discriminating information. Among the more characteristic products, we have observed that some light-oxygenated compounds (3-methyl-2-buten-1-yl acetate, 1,8-cineole, methyl benzoate, methyl salicylate, (*E*)-cinnamyl acetate), some sesquiterpenes (germacrene D, γ-muurolene, α-copaene, β-caryophyllene), and some heavy oxygenated compounds ((*E*,*E*)-farnesyl acetate, γ-cadinol, benzyl benzoate) play an important role in the commercial-grade differentiation. Since geographic origin may affect the chemical composition of essential oils, a similar study of ylang-ylang of Comores should be realized.

## ACKNOWLEDGMENT

We express our gratitude to D. Andrianandrasana and B. Rahantamalala, Service de la Qualité et du Conditionnement des Produits, Antananarivo, Malagasy Republic, for providing authentic samples of ylang-ylang essential oils and to B. Descales, Ecole Supérieure de Chimie de Marseille, for helping in programming FDA.

**Registry No.** α-Pinene, 80-56-8; β-pinene, 127-91-3; myrcene, 123-35-3; β-caryophyllene, 87-44-5; α-humulene, 6753-98-6; ε-cadinene, 113474-59-2; γ-muurolene, 30021-74-0; germacrene D, 23986-74-5; γ-cadinene, 39029-41-9; α-farnesene, 502-61-4; α-copaene, 3856-25-5; α-muurolene, 10208-80-7; δ-cadinene, 483-76-1; 3-methyl-2-buten-1-yl acetate, 1191-16-8; 2-methyl-3-buten-2-ol, 115-18-4; 3-methyl-2-buten-1-ol, 556-82-1; *p*-methylanisole, 104-93-8; 1,8-cineole, 470-82-6; methyl benzoate, 93-58-3; linalool, 78-70-6; benzyl acetate, 140-11-4; methyl salicylate, 119-36-8; geraniol, 106-24-1; geranyl acetate, 105-87-3; (*E*)-cinnamyl acetate, 21040-45-9; δ-cadinol, 19435-97-3; α-cadinol, 481-34-5; T-muurolol, 19912-62-0; γ-cadinol, 50895-55-1; (*E*,*E*)-farnesyl acetate, 4128-17-0; benzyl salicylate, 118-58-1; benzyl benzoate, 120-51-4.

## LITERATURE CITED

Afnor *Recueil de Normes Françaises des Huiles Essentielles*, 2nd ed.; Association Française de Normalisation: Paris, 1986.

Buccellato, F. "Ylang Survey". *Perfumer Flavorist* **1982**, *7*, 9–12.

Chemical Society Analytical Methods Committee "Application of Gas-Liquid Chromatography to the Analysis of Essential Oils. Part VII. Fingerprinting of Essential Oils by Temperature-Programmed Gas-Liquid Chromatography Using Carbowax 20 M Stationary Phase". *Analyst (London)* **1980**, *105*, 262–273.

Delaney, M. F. "Chemometrics". *Anal. Chem.* **1984**, *56*, 261–271R.

Fischer, R. A. "The Use of Multiple Measurements in Taxonomics Problems". *Ann. Eug. Lond.* **1936**, *7*, 179–188.

Frank, I. E.; Kowalski, B. R. "Chemometrics". *Anal. Chem.* **1982**, *54*, 232–243R.

Gaydou, E. M.; Randriamiharisoa, R.; Bianchini, J. P. "Composition of the Essential Oil of Ylang-Ylang (*Cananga odorata* Hook Fil. et Thomson, Forma Genuina) from Madagascar". *J. Agric. Food Chem.* **1986**, *34*, 481–487.

Grasse Essential Oil Association *Methodes d'Analyse et Constantes Analytiques des Huiles Essentielles Adoptées par la Commission de Normalisation*, 2nd ed.; Syndicat National des Fabricants et Importateurs d'Huiles Essentielles et Produits Aromatiques: Grasse, 1959.

Kowalski, B. R. "Chemometrics". *Anal. Chem.* **1980**, *52*, 112–122R.

Lebart, L.; Fenelon, J. P.; Morineau, A. *Traitement des Données Statistiques*; Dunod: Paris, 1982.

Llinas, J. R. *Bibliothêque de Programmes*; Centre de Calcul de l'Ecole Supérieure de Chimie de Marseille: Marseille, 1982.

Llinas, J. R.; Ruiz, J. M. "Multivariate Analysis of Chemical Data Set". In *Computer Aids to Chemistry*; Vernin, G., Chanon, M., Eds.; Ellis Horwood: Chichester, 1986; pp 200–256.
Mahalanobis, P. C. "On the Generalized Distance in Statistics". *Proc. Natl. Inst. Sci. Ind.* **1936**, *12*, 49–55.
Powers, J. J.; Keith, E. S. "Stepwise Discriminant Analysis of Gas Chromatographic Data as an Aid in Classifying the Flavor Quality of Foods". *J. Food Sci.* **1968**, *33*, 207–213.
Ramos, L. S.; Beebe, K. R.; Carey, W. P.; Sanchez, E. M.; Erickson, B. C.; Wilson, B. E.; Wangen, L. E.; Kowalski, B. R. "*Chemometrics*". *Anal. Chem.* **1986**, *58*, 294–315R.
Rao, C. R. *Linear Statistical Inference and Its Applications*; Wiley: New York, 1973.
Romeder, J. M. *Methode de Discrimination*; Dunod: Paris, 1973.
Schoenfeld, P. S.; De Voe, J. R. "Statistical and Mathematical Methods in Analytical Chemistry". *Anal. Chem.* **1976**, *48*, 403–411R.

Received for review March 30, 1987. Accepted November 20, 1987. This work was supported in part by a grant from the Fonds d'Aide et de Coopération, France, for R.P.R.

# Changes in Chemical Composition of Burley Tobacco during Senescence and Curing. 2. Acylated Pyridine Alkaloids

Harold R. Burton,* Roger A. Andersen, Pierce D. Fleming, and Linus R. Walton

Acylated pyridine alkaloids, 2,3′-bipyridyl, and cotinine were identified in burley tobacco by capillary GC–MS, after comparison with authentic synthesized standards. A method was developed for the quantification of the minor alkaloids, and the method was used to determine influence of curing and plant maturity on accumulation of these alkaloids in burley tobacco. The compounds quantified were 2,3′-bipyridyl (**1**), cotinine (**2**), *N*′-formylanatabine (**3a**), and *N*′-formyl- (**4a**), *N*′-acetyl- (**4b**), *N*′-butanoyl- (**4c**), *N*′-hexanoyl- (**4d**), and *N*′-octanoylnornicotine (**4e**). Plant maturity had more influence on accumulation of acylated nornicotines than curing temperature. Data have also shown there is specificity for accumulation of acylnornicotines vs acylanatabines, even though the concentrations of nornicotine and anatabine were equivalent.

In recent years there has been increased interest in the chemical changes occurring during senescence and death of plants. This is particularly true for tobacco since the consumable product is obtained after vegetative growth, senescence, and curing. Recent reviews (Burton et al., 1983; Long and Weybrew, 1981; Enzell et al., 1977) summarize many of the chemical changes that occur during senescence and air-curing. During air-curing there is a decrease of nicotine (Burton et al., 1983, 1985; Enzell et al., 1977). Some of the decrease is due to oxidation of nicotine to cotinine and other oxidation products (Frankenburg et al., 1952). Nicotine decrease may be partially explained by formation of nornicotine via demethylation reaction (Bush, 1981; Enzell et al., 1977). There is no apparent increase of nornicotine between harvest and air-curing (Burton et al., 1985); however, the absence of net changes in nornicotine do not necessarily imply that nornicotine is stable but may undergo secondary reactions. Several nornicotine derivatives (Figure 1) [*N*′-formyl- (**4a**), *N*′-acetyl- (**4b**), *N*′-butanoyl- (**4c**), *N*′-hexanoyl- (**4d**), and *N*′-octanoylnornicotine (**4e**)] have been isolated from burley tobacco (Enzell et al., 1977; Matsushima et al., 1983; Strunz and Findlay, 1985). *N*′-Formyl- and *N*′-acetylanatabine and *N*′-formyl- and *N*′-acetylanabasine are known tobacco constituents (Miyano et al., 1979). Very little is known about the formation of these secondary tobacco alkaloids during senescence and curing. Therefore, this study was initiated to identify the minor alkaloids, develop a method for their quantification, and determine whether senescence and curing temperatures influenced their accumulation.

Department of Agronomy (H.R.B.), USDA—ARS, MSA (R.A.A., P.D.F.), and Department of Agricultural Engineering (L.R.W.), University of Kentucky, Lexington, Kentucky 40546.

## EXPERIMENTAL SECTION

**Plant Materials.** Commercially available burley tobacco (*Nicotiana tabacum* L. cv. KY 14) plants were grown at the Kentucky Agricultural Experimental Farm near Lexington in 1985. Recommended fertilization and cultural practices were followed during the growing season (Atkinson et al., 1982). After topping, tobacco was sprayed with a contact chemical (Offshoot T) to control sucker growth. Tobacco was harvested 1, 4, and 7 weeks (immature, mature, over-mature) after topping. At harvest, plants were stalk cut and stalks were speared on sticks. After wilting for 1 day in the field, half of the harvested tobacco (90 plants) was placed in a controlled environmental chamber and cured at 24 °C and 70% RH. The remaining harvested tobacco was placed in a controlled environmental chamber and cured at 32 °C and 83% RH. The use of 70 and 83% RH at 24 and 32 °C, respectively, was to maintain a constant drying rate during the curing process (Walton et al., 1982). Three replicate samples were taken from the top third of the plant at 0, 1, 2, 3, 5, 7, 9, 11, 14, 16, 19, and 21 days after each harvest to determine the changes in acylated pyridine alkaloids that occured during curing. Top stalk position was sampled because it cures more slowly. Therefore, chemical changes occurred over a longer time period. The midveins were removed, lamina were weighed, and leaf area was determined. The lamina were freeze-dried and reweighed to determine moisture content. Samples were ground to pass a 40-mesh screen and stored at −40 °C until analysis.

**Preparation of Standards.** *N*′-*Formylnornicotine* (***4a***). Nornicotine (Glenn and Edwards, 1978) (100 mg) was heated with an excess of formic acid (1 mL) for 8 h. Excess formic acid was removed under reduced pressure, the residue made alkaline with 10 mL of a 10% NaOH solution and extracted with 3 × 10 mL portions of ether. The ether extract was dried ($Na_2SO_4$) and the solvent removed to yield a straw-colored oil. After the residue was

0021-8561/88/1436-0579$01.50/0 