Liquid Chromatographic Analysis of Non-Volatile Strawberry Compounds

J. J. Mangas[1] / J. Moreno[1] / B. Suárez[1] / A. Picinelli[1] / D. Blanco[2]*

[1] Centro de Investigación Aplicada y Tecnología Agroalimentaria, 33300 Villaviciosa, Asturias, Spain
[2] Departamento de Química Física y Analítica, Universidad de Oviedo, 33006 Oviedo, Spain

Key Words

Column liquid chromatography
Non-volatile strawberry compounds
Multivariate analysis

Summary

A procedure is presented for simultaneous extraction of organic acids, vitamin C and sugars from strawberries and analysis by reversed-phase (RP) HPLC and cation-exchange (calcium form) chromatography. Recoveries from strawberries spiked at different levels ranged from 96 to 103 % for carbohydrates, 92 to 112 % for organic acids, and 84 % for vitamin C; the repeatability of the method evaluated as the relative standard deviation in the optimum range was < 5 % for sugars and 7.5 % for organic acids and vitamin C. Multivariate analysis techniques, such as Bayes and SIMCA analyses, have enabled correct classification of the strawberries tested but we were unable to establish adequate models for typifying them according to their geographic origin.

Introduction

The strawberry belongs to the Rosaceae family, genus Fragaria. Generally, garden strawberries have larger fruits and are sourer but much less aromatic than wild strawberries. In this sense, numerous breeding programs have been developed to obtain cultivars with a higher vitamin C level and total solids concentration, and a more aromatic profile. [1].

Among the different compounds present in the fruits, organic acids, vitamin C, carbohydrates and anthocyanins are the most frequently assayed owing to their contribution to the fruits' sensory and nutritional properties. Organic acids, moreover, determine pH, inhibit the action of enzymes and are chelating agents of iron and copper, and therefore impede chemical precipitation and oxidation [2]. Vitamin C is one of the principle nutritional components of strawberries, and many people consume strawberries because of this [3]. Carbohydrates occur in fruit as storage reserves. Their nutritional effect is related to the daily intake – an excess of carbohydrates might promote a higher concentration of blood lipids, but a low intake can, on the other hand, give rise to metabolic disorders [4].

The red colour of strawberries originates from anthocyanins. Pelargonidin-3-glucoside and cyanidin-3-glucoside are considered the major strawberry anthocyanins [5]. Knowledge of the level of red pigments is necessary to control the sensory quality of strawberry product derivatives; it has been reported that the initial pigment content of strawberry juice can affect its colour stability [6]. Bakker et al. [7] established differences between 39 strawberry genotypes with regard to colour measurements such as L*, Hue angle and Chroma, the last being well correlated with anthocyanin concentration.

The determination of organic acids and vitamin C in different fruits (whortleberry, blackberry, red currant, etc.) has been accomplished by liquid chromatography, usually reversed-phase (RP), with UV detection [8,9], although ion-exchange chromatography with an ION-300 column and refractive index detection has also been used for determining carboxylic acids, sugars, glycerol and ethanol in wine and grape must [10]. Other detectors can be used for HPLC analysis of ascorbic acid. For instance, HPLC with electrochemical detection has been used to monitor vitamin C in orange juice; this method was 100 times more sensitive than HPLC with UV-Vis detection at 245 nm [11].

Chemometric methods are increasingly employed for differentiation and characterisation of fruit species [12–14], alcoholic beverages [15–17], etc. For instance, cluster analysis is a common technique for searching for 'natural' groups, principal components analysis (PCA) is used to reduce the dimensional data structure, *K*-nearest neighbour (KNN) and linear discriminant analysis (LDA) are employed as classification methods, and Bayes, soft independent modelling of class analogy (SIMCA) and partial least-squares (PLS) analyses are modelling techniques.

0009-5893/98/02 197-06 $ 03.00/0

The aim of this work was to quantify the main organic acids, carbohydrates, vitamin C and total anthocyanins in strawberries to enable differentiation of strawberries according to their geographic origin by means of multivariate chemometric techniques.

Experimental

Reagents

Analytical standard-grade quinic, malic, shikimic, ascorbic, fumaric and citric acids, and sucrose, glucose and fructose were purchased from Sigma-Aldrich (Madrid, Spain) and cyanidin-3-galactoside chloride from Extrasynthese (Genay, France). Standard solutions were prepared in Milli-Q water for sugars, and in 10^{-2} M potassium dihydrogen phosphate buffer solution of adjusted to pH 2.5 with phosphoric acid for organic acids and vitamin C; the highest concentrations of standards used were: 0.52 mM for quinic acid, 1.50 mM for malic acid, 0.11 mM for shikimic acid, 0.28 mM for ascorbic acid, 2.00 mM for citric acid, 0.04 mM for fumaric acid, 5.84 mM for sucrose and 11.11 mM for glucose and fructose.

Preparation of Strawberry Samples

Thirty four samples of strawberries from the Principality of Asturias (Spain) and seventeen samples from other regions of Spain were picked in the springtime of 1996. Samples were freeze-dried and ground in a mill, and approximately 0.25 g was extracted with methanol – 0.014 N HCl (85:15, *v/v*; 10 mL) for 2 h at 4 °C. The liquid phase was then removed and the residue washed with the solvent mixture used for extraction (3 × 10 mL). The total hydroalcoholic extract was evaporated (t = 30 min; T = 30 °C) to eliminate the methanol. A helical nitrogen flow (P_N = 1.8 atm) generated by vortexing (Zymark Turbo Vap^{TM} Evaporator) was employed for solvent removal. The aqueous extract obtained was adjusted to 50 mL with water. Each sample was extracted three times.

An aliquot of the aqueous extract was adjusted to pH 1.0 with 2 N HCl before determination of total anthocyanins.

For analysis of sugars, C_{18}-RP Sep-Pak classic cartridges (360 mg; Waters, Milford, MA, USA) were preconditioned by sequential passage of methanol (5 mL) and water (5 mL) and the aqueous extract (approximately 6 mL) was loaded on to the cartridge. The first 2 mL were discarded, and the remainder was filtered through a 0.45 µm cellulose acetate (CA) membrane (Lida; Teknokroma, Spain) before HPLC analysis.

For organic acid and vitamin C analysis, the aqueous extract was micro-filtered directly through a 0.45 µm (CA) membrane before chromatography.

Apparatus and Conditions

Chromatography was performed with an HPLC system (Waters Chromatography Division, Milford, USA) comprising a 712 automatic injector, an M510 pump, a Millennium v.2.1 software data module, a differential refractometer (for sugar analysis) and an LC spectrophotometer (for organic acid and vitamin C analysis). Separation of sugars was performed by the method described elsewhere [18]. Analysis of organic acids and vitamin C was accomplished isocratically at a flow rate of 0.5 mL min^{-1} by means of a Spherisorb C_{18} column (250 mm × 4.6 mm i.d.; 3 µm particle size) at 18 °C; the mobile phase was 10^{-2} M potassium dihydrogen phosphate buffer solution adjusted to pH 2.5 with phosphoric acid. Column effluents were monitored at 206 nm. All solvents were HPLC grade and de-gassed with helium before use. Quantification was performed by the external standard method.

Total anthocyanins were determined by spectrophotometry with a Perkin-Elmer (Norwalk, CT, USA) Lambda 5 UV-Vis instrument. The aqueous extract was adjusted to pH 1.0 and absorbance was measured at 500 nm using a solution of 0.03 mM cyanidin 3-galactoside chloride in 0.1 N HCl as the calibration standard.

Statistical Methods

The data were processed by means of the PARVUS statistical package [19]. We constructed a data matrix where rows (51) represented strawberries and columns (9) corresponded to organic acids (malic, shikimic, citric and fumaric), vitamin C, sugars (sucrose, glucose and fructose) and total anthocyanins. The data were autoscaled and categorised as category 1 (strawberries from Candamo-Principality of Asturias), and category 2 (strawberries from other regions of Spain).

Results and Discussion

Table I lists the mean values obtained for the amounts of sugars, vitamin C, organic acids and total anthocyanins, with other descriptive statistics.

Sugars

A typical chromatogram of major carbohydrates (fructose, glucose and sucrose) in strawberries is shown in Figure 1; the presence of methanol in this chromatogram is a consequence of insufficient removal of the alcohol during evaporation. Fructose was the most abundant sugar. The average recoveries for carbohydrates were between 96 and 103 %; the relative standard deviation (RSD) was < 5 %.

Organic Acids and Vitamin C

A method was optimized for simultaneous analysis of organic acids and vitamin C in strawberries by RPHPLC. The different chromatographic conditions

Table I. Descriptive statistics of variables ($n = 51$).

Variable	Mean (% of DM[a])	RSD[b]	Skewness	Kurtosis	Min	Max	Median	Range
Sucrose	13.41	32.66	0.34	0.07	4.53	24.10	12.18	19.57
Glucose	21.49	9.17	-0.29	0.62	16.41	26.40	21.73	9.99
Fructose	27.26	8.25	-0.53	0.58	20.71	31.62	26.92	10.91
Anthocyanins	0.33	24.24	0.64	0.10	0.19	0.53	0.33	0.34
Malic acid	1.50	27.33	0.63	-0.16	0.67	2.53	1.41	1.86
Shikimic acid[c]	5.69	14.11	2.91	9.34	0.00	66.09	0.00	66.09
Ascorbic acid	0.39	28.21	-0.92	1.70	0.07	0.66	0.42	0.59
Citric acid	9.40	15.64	-0.35	-1.05	6.63	11.87	9.75	5.24
Fumaric acid[d]	0.15	125.22	6.29	44.56	0.05	1.01	0.13	0.96

[a]Dry Matter (DM). [b]Relative standard deviation. [c]mg kg^{-1} of DM. [d]g Kg^{-1} of DM.

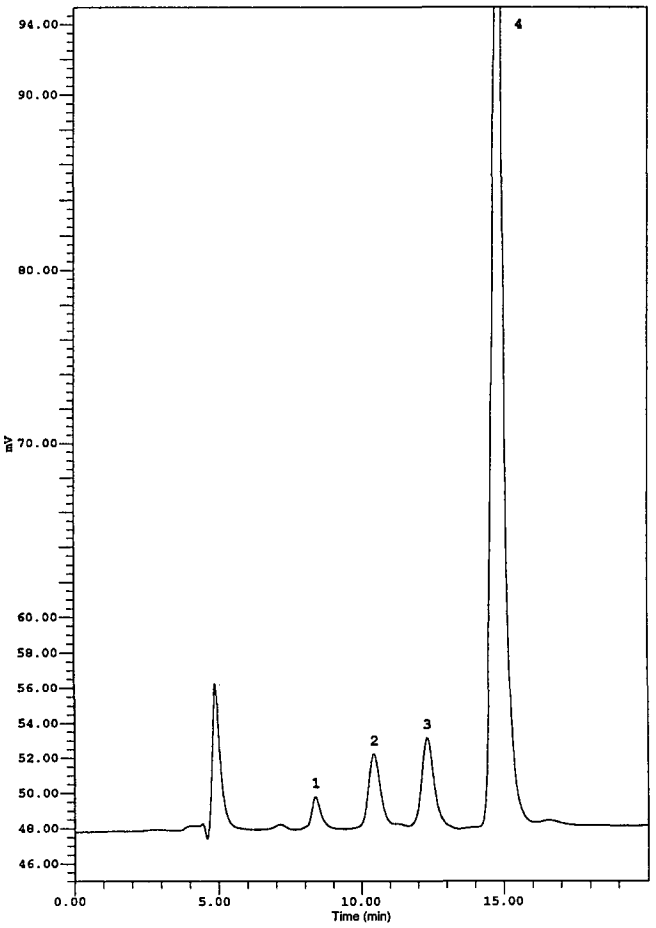

Figure 1
Chromatogram of carbohydrates in strawberries. 1,sucrose; 2, glucose; 3, fructose; 4, methanol. Chromatographic conditions: column, Sugar Pak I (300 mm × 6.5 mm i.d.; 10 µm particle size); flow rate, 0.5 mL min^{-1}; temperature, 90 °C; mobile phase, aqueous solution of calcium salt of EDTA (50 mg L^{-1}); detection, refractive index.

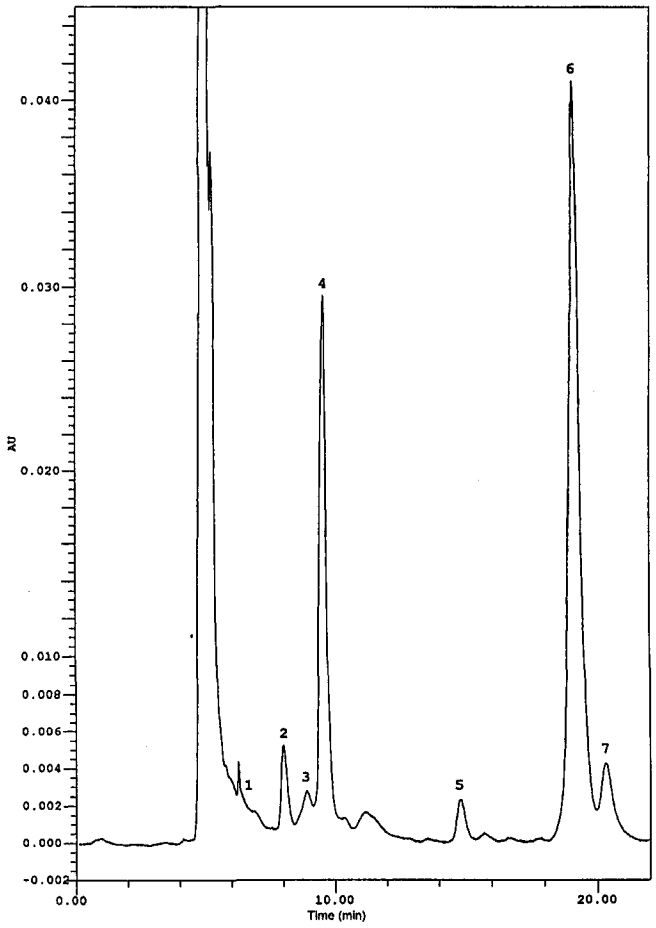

Figure 2
Chromatogram of organic acids and vitamin C in strawberries. **1** = quinic acid; **2** = malic acid; **3** = shikimic acid; **4** = ascorbic acid; **5** = unknown; **6** = citric acid; **7** = fumaric acid. Chromatographic conditions: column, Spherisorb C_{18} (250 mm × 4.6 mm i.d.; 3 µm particle size); flow rate, 0.5 mL min^{-1}; temperature, 18 °C; mobile phase, 10^{-2} M potassium dihydrogen phosphate buffer adjusted to pH 2.5 with phosphoric acid; detection, spectrophotometric ($\lambda = 206$ nm).

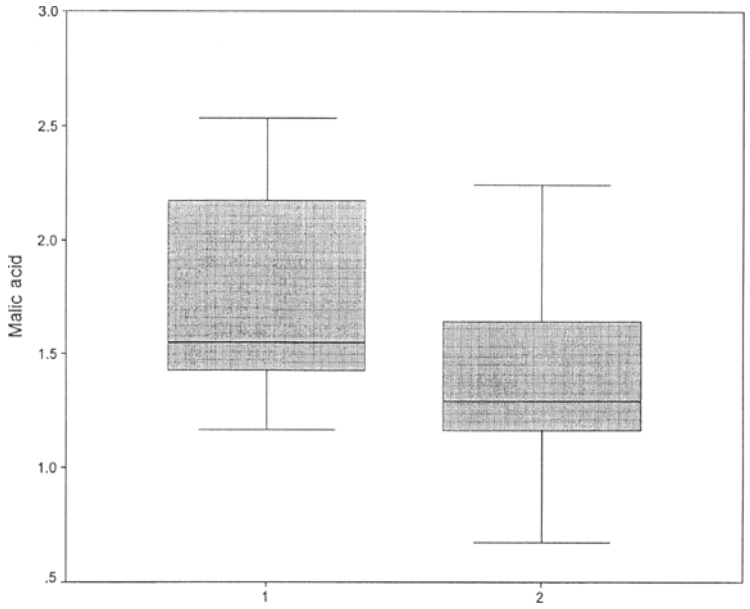

Figure 3
Multiple Box-Whisker plots for malic acid (% dry matter). Class 1: Spain. Class 2: Principality of Asturias.

studied were temperature, in the range 10–50 °C; ionic strength, 10^{-1}–10^{-3} M; and pH, 2.25–2.70. Ionic strength did not influence the capacity factor (k) within the interval studied but temperature did. The capacity factor of all analytes decreased when temperature increased; this was particularly critical for ascorbic and shikimic acids, because severe overlapping was observed ($\alpha = k_{ascorbic}/k_{shikimic} = 1$) between 20 and 25 °C. The separation factor or selectivity (α) decreased above 25 °C and increased below 20 °C. Only citric acid was influenced by pH within the range studied. Consequently, to achieve suitable peak separation the chromatographic conditions established were temperature, 18 °C; ionic strength, 10^{-2} M and pH 2.5. A typical chromatogram of the main organic acids and vitamin C in strawberries is shown in Figure 2. It is apparent that resolution is adequate ($R_S \geq 1$) for all the peaks. The major organic acids found were citric and malic; quantification of quinic acid was not possible because of its low concentration in the samples and elution next to an unknown peak of the elution front. Recoveries for quinic, malic, shikimic, citric and fumaric acids ranged between 92 and 112 %; for vitamin C it was 84 %. The relative standard deviation between extractions was < 7.5 %.

Statistical Analysis

The Fisher weight (the ratio between inter-centroid variance and intra-category variances) was evaluated to ascertain the most discriminant variables. Malic acid and fructose were found to be the variables with the highest Fisher values (0.3158 and 0.212) that of glucose was lowest (0.00987). Figure 3 shows Box-Whisker plots constructed for malic acid (the most discriminant variable) and the two categories; it is apparent that it is not possible to differentiate between the two classes with this variable; this is in accordance with results obtained when stepwise linear discriminant analysis was performed, because the Wilks' lambda obtained taking into account an F-to-enter of 3.84 and an F-to-remove of 2.71 was 0.8728, thus only 13 % of total variance is explained by within-groups differences. Once more, the most relevant variable was malic acid.

The use of classification techniques such as linear discriminant analysis (LDA establishes a boundary between two classes from a hyperplane that includes all the points at the same distance from the centroid of each class) and the K nearest neighbour ($K = 3$) method (this method selects the K nearest objects to the unknown sample and classifies these in the group where the majority of the K samples belong) did not enable suitable differentiation between the two categories, correct classifications found being 78 % and 76 %, respectively. For cross-validation of the LDA method, the original training set was split into three groups for cancellation, each group for cancellation was removed from the training set, and the remaining objects were used to build the classification rule. The results obtained showed that correct predictions ranged between 60 and 80 %.

Bayes analysis and SIMCA are modelling techniques. These methods build boundaries between one class and the rest of universe. The Bayesian methods classify the objects into the class with sufficiently high density, once probability density functions have been calculated for each category. SIMCA also considers each class separately, establishing a principal components model for each category. Table II shows the classification matrices obtained using LDA, KNN, Bayes and SIMCA analyses. It is apparent that the classification capacities of the modelling techniques are higher (98 % for Bayes and 88 % for SIMCA) than the LDA and KNN methods. At

Table II. Classification matrices for LDA, KNN, Bayes and SIMCA methods.

True category	LDA Assigned category			KNN ($K = 3$) Assigned category			Bayes Assigned category			SIMCA Assigned category		
	1	2	Hits (%)	1	2	Hits (%)	1	2	Hits (%)	1	2	Hits (%)
Category 1	13	4	76.5	13	4	76.5	17	0	100	14	3	82.4
Category 2	7	27	79.4	8	26	76.5	1	33	97.1	3	31	91.2
Overall			78.4			76.5			98			88.2

Table III. Sensitivities and specificities for the two categories using Bayes and SIMCA methods.

Classes	Bayes		SIMCA	
	Sensitivities	Specificities	Sensitivities	Specificities
Category 1	100	91.2	88.4	85.3
Category 2	97.1	17.6	91.2	35.3

the same time, when Bayes analysis was validated using five cancellation groups correct predictions ranged between 60 and 70 %. The differences detected in the classification capacity between the LDA method and Bayesian analysis could be because of equality variance-covariance matrices of the classes.

To evaluate the modelling capacity of the variables set, sensitivities and specificities were estimated using Bayes and SIMCA techniques (Table III). Sensitivity ($1 - \alpha$) and specificity ($1 - \beta$) are related to the first-class error (α) and the second-class error (β), where α = pr{rejecting H_0/H_0 being true} and β = pr{accepting H_0/H_0 being false} (the null hypothesis (H_0) $\Rightarrow$ the object belongs to the class considered, and the alternative hypothesis (H_a) $\Rightarrow$ the object does not belong to the class considered). As is apparent, sensitivities were high for the two categories using both modelling methods, so that each of the classes accept the majority of the samples categorized initially in each class. However, the specificity of category 2 is very low, irrespective of the model considered. In consequence, it is not possible to model this category correctly on the basis of the variables evaluated.

To obtain a more precise differentiation of the strawberries on the basis of their geographical origin, therefore, other variables, such as trace metal content [20], should be considered for establishing more suitable pattern recognition.

Conclusions

RPHPLC and cation-exchange chromatography are a rapid convenient techniques for simultaneous analysis of sugars, organic acids and vitamin C in strawberries. These analytical methods, and several chemometric techniques, have enabled us to classify strawberries correctly on the basis of their geographic origin; however, it was not possible to establish suitable pattern recognition on the basis of the set of variables evaluated.

Acknowledgements

This work was made possible by financial support from the Autonomic Administration of the Principality of Asturias. We are also grateful for the help of Miguel A. Fueyo (Head of Department of Horticulture and Fruticulture, CIATA) and his collaborators in the sampling of strawberries.

References

[1] *G. Konja, T. Lovric*, Berry fruit juices, in S. Nagy, C. Chen, P. Shaw (Editors) Fruit Juice Processing Technology, Agscience Inc., Auburndale, Florida, 1993.

[2] *N. A. M. Eskin*, Biochemistry of food spoilage: enzymatic browning, in Biochemistry of Foods, Academic Press, San Diego, 1990.

[3] *R. S. Kirk, R. Sawyer*, Appendix 13, in Pearson's Composition and Analysis of Foods, Longman Scientific and Technical, UK, 1991.

[4] *L. W. Aurand, A. E. Woods, M. R. Wells*, Carbohydrates, in Food Composition and Analysis, Van Nostrand Reinhold, New York, 1987.

[5] *C. F. Timberlake, P. Bridle*, Distribution of anthocyanins in food plants, in P. Markakis (Editor) Anthocyanins as Food Colors, Academic Press, New York, 1982.

[6] *G. Skrede, R. E. Wrolstad, P. Lea, G. Enersen*, J. Food Sci. **57**, 172 (1992).

[7] *J. Bakker, P. Bridle, S. Bellworthy*, J. Sci. Food Agric. **64**, 31 (1994).

[8] *M. A. Romero Rodríguez, M. L. Vázquez Oderiz, J. López Hernández, J. Simal Lozano*, J. Chromatogr. Sci. **30**, 433 (1992).

[9] *H. S. Lee, G. A. Coates*, J. Micronutrient Anal. **3**, 199 (1987).

[10] *M. Calull, R. M. Marcé, F. Borrull*, J. Chromatogr. **590**, 215 (1992).
[11] *C. W. Wilson, P. E. Shaw*, J. Agric. Food Chem. **35**, 329 (1987).
[12] *M. Defernez, E. Katherine Kemsley, R. Wilson*, J. Agric. Food Chem. **43**, 109 (1995).
[13] *K. McRae, P. Lidster, A. DeMarco, A. Dick*, J. Sci. Food Agric. **50**, 329 (1990).
[14] *M. Forcen, A. Berna, A. Mulet*, Lebensm. Wiss. Technol. **26**, 54 (1993).
[15] *M. Cruz, J. A. Saez, J. López*, Analyst **118**, 801 (1993).
[16] *J. J. Mangas, R. Rodríguez, J. Moreno, D. Blanco*, J. Agric. Food Chem. **44**, 268 (1996).
[17] *J. J. Mangas, R. Rodríguez, J. Moreno, D. Blanco*, Lebensm. Wiss. Technol. **29**, 357 (1996).
[18] *D. Blanco, M. D. Gutiérrez, J. J. Mangas, A. Noval*, Chromatographia **25**, 701 (1988).
[19] *M. Forina, R. Leardi, C. Armanino, S. Lanteri*, PARVUS. An Extendible Package of Programs for Data Exploration, Classification and Correlation, Elsevier, Amsterdam, 1988.
[20] *S. Nikdel, S. Nagy, J. Attaway*, in S. Nagy, J. Attaway, M. Rhodes (Editors) Adulteration of Fruit Juice Beverages, Marcel Dekker, New York, 1988.

Received: Apr 15, 1997
Revised manuscript received: Sep 29, 1997
Accepted: Oct 17, 1997