Nereis. Revista Iberoamericana Interdisciplinar de Métodos, Modelización y Simulación	12	19-42	Universidad Católica de Valencia San Vicente Mártir	Valencia (España)	ISSN 1888-8550

Molecular Classification of *Eucalyptus camaldulensis* and *Mentha pulegium* Oils Components

Clasificación molecular de los componentes de los aceites de *Eucalyptus camaldulensis* y de *Mentha pulegium*

Fecha de recepción y aceptación: 7 de febrero de 2020, 24 de abril de 2020

DOI: 10.46583/nereis_2020.12.572

Francisco Torrens Zaragozá[1*]

[1] Institut Universitari de Ciència Molecular. Universitat de València.
[*] Correspondencia: Universitat de València. Institut Universitari de Ciéncia Molecular. Edifici d'Instituts de Paterna P. O. Box. 22085. 46071 Valencia. España. *E-mail*: torrens@uv.es

ABSTRACT

Eucalyptus and *Mentha* remain flavours in agro-food manufacturing. Oils and components present antifungal potency *vs*. decay of fruits; *E. camaldulensis* is led by 1,8-cineole and α-pinene, unlike *M. pulegium*, which is led by pulegone. The antifungal activity of *M. pulegium* is three times more frequent than that of *E. camaldulensis*. The phytochemicals present synergy. Categorization is recommended on the basis of *information entropy*. The quantity of C-C double bonds, O-atoms and cycles cluster structures. The procedure undergoes a combinatorial upsurge. Nevertheless, following *equipartition conjecture*, one gets a criterion for selection. Entropy allows clustering phytochemicals according to cluster and principal component analyses. In the periodic classification, components in the same column show similar features. Phytochemicals in the same row present utmost similarity.

KEYWORDS: *periodic table, periodic property, periodic law, periodic classification, information entropy.*

RESUMEN

El *Eucalyptus* y la *Mentha* se siguen utilizando como sabores en la fabricación agroalimentaria. Sus aceites y componentes presentan actividad antifúngica frente a la descomposición de frutas; *E. camaldulensis* está encabezado por 1,8-cineol y α-pineno, pero *M. pulegium* lo está por pulegona. La actividad antifúngica de *M. pulegium* resulta tres veces la de *E. camaldulensis*. Los fitoquímicos presentan sinergia. Se recomienda la categorización sobre la base de la *entropía informacional*. La cantidad de dobles enlaces C-C, átomos de O y ciclos agrupa las estructuras. El procedimiento sufre una explosión combinatoria. Sin embargo, según la *conjetura de equipartición*, se puede obtener un criterio para la selección. La entropía permite agrupar fitoquímicos de acuerdo con el análisis de grupo y los componentes principales. En la clasificación periódica, los componentes en la misma columna muestran características parecidas. Los fitoquímicos presentan también en la misma fila la máxima semejanza.

PALABRAS CLAVE: *tabla periódica, propiedad periódica, ley periódica, clasificación periódica, entropía informacional.*

INTRODUCTION

The therapeutic benefits of essential oils (EOs) are known since antiquity; however, the scientific study of the power of aromatic and medicinal plants has arisen in the last few years, with the aim of looking for alternatives to the chemical substances that jeopardize human health and the environment [1]. Many works proved different bioactivities of aromatic and medicinal plants (*e. g*., antifungal [2-4], antibacterial [5,6], antioxidant [7], insecticide [8-10] powers) by which EOs could serve as preserving agent of food products; EOs from certain aromatic plants present growth and toxigenesis inhibition of bacteria and fungi which are responsible for food infections [11-14]. *Eucalyptus camaldulensis* and *Mentha pulegium* are aromatic and medicinal plants that are used in traditional medicine with antimicrobial purposes [15-18]. River red gum spread throughout the world because of its utilization in reforestations. Its wood is mainly destined for paper-pulp fabrication. Oil from *E. camaldulensis* is well known for relieving the symptoms of cold and influenza. Pennyroyal is used to cure cold, throat illness, cough, bronchitis, lung infection and chill; it is an excellent digestive. Chemical composition and antimicrobial activity of *E. camaldulensis* and *M. pulegium* EOs were described [19-22]. Antifungal activity of EOs was studied *vs*. fungi responsible for apple rot [23]. 1,8-Cineole (eucalyptol, *cf*. fig. 1c) is a colourless liquid phytochemical. It is a cyclic ether and a monoterpenoid. Cloez (1870) called the leading part of *E. globulus* EO *eucalyptol*. *Eucalyptus oil* is not 1,8-cineole, its major component, which is present in camphor laurel, bay leaves, tea tree, mugwort, sweet basil, wormwood, rosemary, sage, etc. It is produced by fractional distillation of *Eucalyptus* EO. It is used as flavouring and a medicine component, but it is poisonous at high doses. It has a fresh camphor-like odour and a spicy, cooling flavour. It is insoluble in water although miscible with ether, ethanol and chloroform. It develops crystal adducts with halogen acids, *o*-cresol, resorcinol and H_3PO_4, which creation is helpful for purification. (*R*)-(+)-Pu1egone (fig. 1e) is a phytochemical obtained from EOs of various plants (*e. g*., catnip *Nepeta cataria*, *Mentha piperita*, pennyroyal). It is a monoterpene. It is a clear colourless oily liquid and shows pleasant odour similar to pennyroyal, peppermint and camphor. It is used in flavouring agents, perfumery and aromatherapy. Science, technology and applications of oils [24], as well as antimicrobials and antifungals [25], were reviewed. Post-emergent herbicidal activity of *Eucalyptus* spp. EOs was informed [26-28]. A procedure makes associations between molecular structures and biosignificance possible [29,30]. The initial position is meant to utilize information entropy theory in pattern recognition. Entropy is expressed on the basis of a *similarity matrix* among phytochemicals. As entropy is feebly discerning for categorization, a stronger idea is included: its *equipartition conjecture* [31]. In earlier publications, the method was used in the classification of polyphenols [32], flavonoids [33], legumes [34], stilbenoids [35], triterpenoids, steroids [36], isoflavonoids [37], lactones [38,39] and artemisinins [40]. The aim of this report is to find features that help distinguish the chemical structures of EOs phytochemicals. This work uses chemical descriptors to EOs based on chemical structures. The objective of this study is to corroborate the value of the descriptors through their ability to distinguish between components, as well as to compare with other phytochemicals, oils and anti-inflammatory drugs.

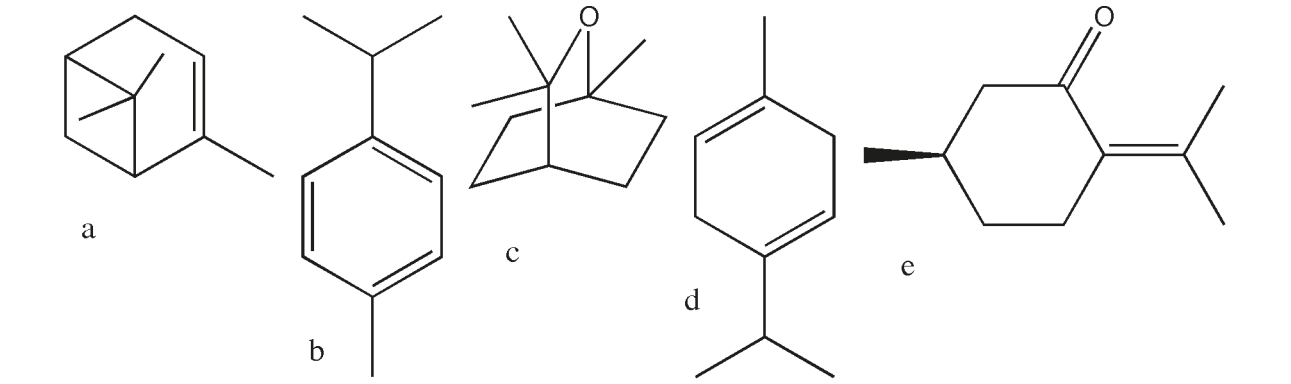

Fig 1. Molecular structures: (a) α-pinene, (b) *p*-cymene, (c) 1,8-cineole (eucalyptol), (d) γ-terpinene and (e) pulegone.

COMPUTATIONAL METHOD

The important question in clustering is to describe *similarity indices* if more than a few decisive factors of contrast are concerned. The first step in counting the similarity idea, for EOs components, is to record the most significant chemical features of the compounds. The *vector of properties* $\bar{i} = <i_1,i_2,\ldots i_k,\ldots>$ must be connected with each feature *i*, whose elements match dissimilar characteristics in the molecule, in a hierarchy consistent with the significance of their abundance. If the *m-th* feature is more important for abundance than feature *k-th*, then $m < k$. Elements i_k are 1 or 0, consistent if an alike feature of rank *k* is present in compound *i* as contrasted to the reference. The examination comprises four characteristics in the five major constituents: presence of lesser than three C=C ($C=C_{012}$), occurrence of lesser than two C=C ($C=C_{01}$), existence of one O (O_1) and occupancy of one cycle (cyc_1, fig. 1). Supposedly, *chemical features* of a constituent can be *ranked*, which consists in their input to composition in the subsequent sequence of dropping importance: $C=C_{012} > C=C_{01} > O_1 > cyc_1$. Descriptor $i_1 = 1$ indicates $C=C_{012}$ ($i_1 = 0$ for $C=C_3$), $i_2 = 1$ indicates $C=C_{01}$ ($i_2 = 0$ for either $C=C_2$ or $C=C_3$), $i_3 = 1$ signifies O_1 ($i_3 = 0$ for O_0) and $i_4 = 1$ indicates cyc_1 ($i_4 = 0$ for cyc_2). In (*R*)-(+)-pulegone, quantity of C=C, O atoms and cycles is one. Its linked vector is <1111>. Pulegone was chosen as a *reference* due to its abundance. Table 1 lists the vector properties connected with five major components of *E. camaldulensis*/*M. pulegium* EOs. Vector <1100> is associated with α-pinene since it shows $C=C_1$, O_0 and cyc_2.

Table 1. Vector property ($C=C_{012}$,$C=C_{01}$,O_1,cyc_1) and major chemical composition of oils

Component	*E. camaldulensis* [%]	*M. pulegium* [%]
1. α-pinene $C_{10}H_{16}$ <1100>	28.30	0.37
2. *p*-cymene $C_{10}H_{14}$ <0001>	6.50	–
3. 1,8-cineole $C_{10}H_{18}O$ <1110>	42.30	–
4. γ-terpinene $C_{10}H_{16}$ <1001>	7.30	–
5. (*R*)-(+)-pulegone $C_{10}H_{16}O$ <1111>	–	80.28
Total	84.40	80.65

By r_{ij} ($0 \leq r_{ij} \leq 1$) we indicate the similarity index of two components connected with vectors $\bar{i}$ and $\bar{j}$, in that order. The similarity relationship is described through *similarity matrix* $\mathbf{R} = [r_{ij}]$. The similarity index among two compounds $\bar{i} = \langle i_1, i_2, \ldots i_k, \ldots \rangle$ and $\bar{j} = \langle j_1, j_2, \ldots j_k, \ldots \rangle$ is described:

$$r_{ij} = \sum_k t_k (a_k)^k \quad (k = 1,2,\ldots) \tag{1}$$

where $0 \leq a_k \leq 1$ and $t_k = 1$ if $i_k = j_k$; however, $t_k = 0$ if $i_k \neq j_k$. The description allocates a weight $(a_k)^k$ to all properties entailed in the explanation of compound *i* or *j*. The *grouping algorithm* uses the *stabilized* similarity matrix gotten by the *max-min composition rule* o:

$$(\mathbf{R} \circ \mathbf{S})_{ij} = \max_k \left[\min_k (r_{ik}, s_{kj}) \right] \tag{2}$$

where $\mathbf{R} = [r_{ij}]$ and $\mathbf{S} = [s_{ij}]$ are matrices of the same kind, as well as $(\mathbf{RoS})_{ij}$ indicates the (i,j)-*th* element of the **RoS** matrix [41-44]. On using the max-min composition rule iteratively so that $\mathbf{R}(n+1) = \mathbf{R}(n)$ o $\mathbf{R}$, there is an integer *n* such that: $\mathbf{R}(n) = \mathbf{R}(n+1) = \ldots$ The $\mathbf{R}(n)$ matrix is the *stabilized similarity matrix*. The significance of stabilization lies on the information that, in the categorization, it produces a separation into disjoint groupings. The stabilized matrix is indicated by $\mathbf{R}(n) = [r_{ij}(n)]$. The *grouping rule* ensues: *i* and *j* are allocated to the same cluster if $r_{ij}(n) \geq b$. The grouping of *i* indicated $\hat{i}$ is collection of species *j* that assures rule: $r_{ij}(n) \geq b$. Matrix of clusters is:

$$\hat{\mathbf{R}}(n) = \left[\hat{r}_{\hat{i}\hat{j}} \right] = \max_{s,t} (r_{st}) \quad (s \in \hat{i},\ t \in \hat{j}) \tag{3}$$

where *s* indicates every descriptor of molecules fitting in grouping $\hat{i}$ (likewise for *t* and $\hat{j}$). Law (3) indicates the greatest similarity index among compounds of two dissimilar clusters. In information theory, the *information entropy h* indicates the astonishment that a source emitting sequences provides [45,46]. Let's consider the utilization of a qualitative mark assay to decide the presence of Fe in a water sample. Lacking an account of assaying, the analyst should start by supposing that both results 0/1 (Fe absent/present) turn out to be equiprobable with likelihoods 1/2. If two elements may be present in the sample (*e. g.*, Fe, Ni), there are four potential results, varying from neither (0,0) to both being present (1,1) with equal likelihoods $1/2^2$. What results from the four possibilities can be established using two assays, each with two observable states. Similarly with three elements, eight possibilities exist, each one having a probability of $1/2^3$; three assays are required. The next model clearly connects the information required to solve it. The quantity of options is state to the power of 2. The power to which 2 should be elevated to provide the quantity of options *N* is defined as the base 2 logarithm of that quantity. Information is described quantitatively in the conditions of the base 2 logarithm of the quantity of likely analytical results: $I = H = \log_2 N = \log_2 1/p = -\log_2 p$, where *I* is the information included in a solution, provided that there were *N* options, *H*, the early information resulting from the necessity to assume *N* options and *p*, the probability of every result if every one

of the N options is evenly probable to happen. The equation is widespread to the case in which the option of all results is not equal. If one is acquainted with history that some metals are more probable to be present than others, the expression is regulated so that the logarithms of the suitably weighted particularized options are added: $H = -\Sigma\, p_i \log_2 p_i$, where $\Sigma\, p_i = 1$. It assumed initial instance but, at present, history proved that 90 % of the samples did not include Fe. Degree of information is computing utilizing: $H = -(0.9 \log_2 0.9 + 0.1 \log_2 0.1) = 0.469$ bit. For an only occasion happening with option p, the degree of surprise is proportional to $-\ln p$. By extending outcome to a chance variable X (that takes N likely values $x_1, \ldots, x_N$ with options $p_1, \ldots, p_N$), mean astonish obtained on finding out X value results: $-\Sigma\, p_i \ln p_i$. Information entropy connected with similarity matrix **R** is:

$$h(\mathbf{R}) = -\sum_{i,j} r_{ij} \ln r_{ij} - \sum_{i,j} \left(1 - r_{ij}\right) \ln\left(1 - r_{ij}\right) \tag{4}$$

We indicate by C_b the collection of classes and by $\widehat{\mathbf{R}}_b$ the similarity matrix at classification level b. The information entropy assures the next features. (1) $h(\mathbf{R}) = 0$ if either $r_{ij} = 0$ or $r_{ij} = 1$. (2) $h(\mathbf{R})$ is maximum if $r_{ij} = 0.5$, that is to say, when imprecision is maximum. (3) $h\left(\widehat{\mathbf{R}}_b\right) \leq h(\mathbf{R})$ for any b, in other words, classification leads to entropy loss. (4) $h\left(\widehat{\mathbf{R}}_{b_1}\right) \leq h\left(\widehat{\mathbf{R}}_{b_2}\right)$ if $b_1 < b_2$, that is to say, entropy is a monotone function of grouping level b. In the categorization procedure, all *hierarchical trees* match entropy reliance on the classification level and the h–b plot is gotten. The *equipartition conjecture of entropy production* of Tondeur and Kvaalen is suggested as a criterion for selection, between dissimilar alternatives resulting from categorization between hierarchical dendrograms. Consistent with the guess, for a certain duty, the flow sheet with the most excellent arrangement turns out to be the one in that the entropy production results most consistently dispersed, that is to say, neighbouring a type of equipartition. One goes on similarity utilizing *information entropy* in place of the thermodynamic one. Equipartition entails a linear reliance next to scale b, so that *equipartition line* is:

$$h_{\text{eqp}} = h_{\max} b \tag{5}$$

As the categorization is separate, the method of stating equipartition is a regular staircase function. The best alternative is decided to result the one minimizing the sum of squares of the differences:

$$\text{SS} = \sum_{b_i} \left(h - h_{\text{eqp}}\right)^2 \tag{6}$$

Learning procedures alike to the ones found in *stochastic methods* are put into practice [47]. Let's assume a certain separation into groupings as *good* from convenient remarks, which matches a *reference* similarity matrix $\mathbf{S} = [s_{ij}]$ provided for the same weights $a_1 = a_2 = \ldots = a$ and for a random quantity of fictional features. After that, we assume the same collection of molecules as in the good categorization and real characteristics. Degree of similarity r_{ij} is calculated with Eq. (3) producing

R matrix. Quantity of features for **R** and **S** matrices differs. Learning procedure resides in finding clustering outcomes for **R** matrix, as near as achievable to the *good* categorization. The first weight a_1 is found as constant and just next weights a_2, a_3,… are given to chance changes. A newer similarity matrix is obtained by utilizing Eq. (3) as well as newer weights. The distance among separations into groupings typified by **R** and **S** matrices is:

$$D = -\sum_{ij}\left(1-r_{ij}\right)\ln\frac{1-r_{ij}}{1-s_{ij}} - \sum_{ij} r_{ij}\ln\frac{r_{ij}}{s_{ij}} \quad \forall 0 \le r_{ij}, s_{ij} \le 1 \tag{7}$$

The description was proposed by the one presented in information theory by Kullback to gauge the distance among two likelihood distributions [48]. It is the distance among **R** and **S** matrices. As for any matrix there is a matching categorization, two clusterings are evaluated by the distance, which results into a positive number that moves towards $d = 0$ as the similarity among **R** and **S** matrices rises. The outcome of the procedure is a collection of weights that make categorization possible. The algorithm was used in the production of complicated trees utilizing information entropy [49,50]. Our code MolClas is an easy, dependable, effective and quick program for molecular clustering, founded on the equipartition conjecture of entropy production, consistent with Eqs. (1)-(7). It reads the quantity of features and molecular characteristics. It permits optimization of coefficients. It electively reads initial coefficients and quantity of iterations. Correlation matrix is computed by algorithm or read from input archive. Code MolClas computes feature similarity matrix in symmetric storage form. It uses an illustrative correlation representation for partial correlation chart. It calculates categorization, computes distances among classes, calculates clusters similarity matrices, analyzes categorizations information entropy, optimizes coefficients, carries out single and complete-linkage hierarchical cluster analyses, and charts grouping charts. Code MolClas was created not simply to examine equipartition conjecture of entropy production, but to look at the world of molecular categorization too.

CALCULATION RESULTS

The Oils composition table (table 1) was taken from Hmiri et al. Chromatography identified 31 (red river gum) *vs*. 17 compounds (pennyroyal). The major components of *E. camaldulensis* EO (fig. 1) are 1,8-cineole, α-pinene, γ-terpinene and *p*-cymene, which compares to *E. camaldulensis* EO from Sidi Amira (Morocco), richer in 1,8-cineole/*p*-cymene and poorer in α-pinene/γ-terpinene. Oil of *E. camaldulensis* from Cameroon is richer in 1,8-cineole, *p*-cymene and γ-terpinene, and poorer in α-pinene. Oils of *E. camaldulensis* from Taiwan present chemotypes: dominated by (1) 1,8-cineole and (2) α-pinene, *p*-cymene and α-phellandrene; however, in Spain, EO is dominated by spathulenol, *p*-cymene and cryptone. In Cyprus, *E. camaldulensis* EO showed ethanone/1,8-cineole dominance. Oil of *M. pulegium* was characterized by pulegone, which is similar to most works in Morocco. Works in Tunis showed that pulegone is a major component of *M. pulegium* EO. Works in Iran/Morocco evidenced EO major substances: piperitone/piperitenone; while in Algeria, *M. pulegium* EO was characterized by monoterpenes (α/β-pinenes, camphene, sabinene, α-terpinene, myrcene). Antifungicity of *E. camaldulensis*/*M. pulegium* EOs *vs*. two apples-rot fungi was taken from Hmiri

et al.; *M. pulegium* EO results more potent than *E. camaldulensis*: it inhibited *Alternaria alternata*/*Penicillium expansum* growth from 10μL while in eucalyptus EO both resisted 30μL; the difference is composition: pennyroyal is mostly formed by ketone pulegone, while eucalyptus is made of terpenic oxide 1,8-cineole. Ketones are more active than terpenic oxides. Pennyroyal EO inhibits *Penicillium* growth (20μL); *M. pulegium* EO inhibited dermatophytes growth from $2\mu g \cdot mL^{-1}$. *Botrytis cinerea*, another apples-rot fungus, and others (*Aspergillus flavus*, *A. niger*, *Fusarium culmorum*, *F. oxysporum*, *Trichoderma*) are sensitive to pennyroyal EO. Oils of *E. camaldulensis* from Taiwan were active *vs*. 10 fungi from China, inactive *vs*. *A. alternata*. Antifungicity of *E. camaldulensis* EOs (1,8-cineole) is because of monoterpene, which mechanisms on inhibitory growth of fungal/vegetal cells remain obscure; however, among effects on biomembranes, deleterious results on mitochondrial membranes inhibit mitochondrial energetic metabolism that affects physiological processes in cell. Pulegone/1,8-cineole inhibit mycelium growth at greater concentrations than EOs because of minor-constituents synergism. The antifungal activity, categorized as 0.3333 for *E. camaldulensis* and 1.0 for *M. pulegium* EOs, is fitted *vs*. major components pulegone/1,8-cineole. Best lineal fit turns out to be:

$$I = -0.00688 + 0.0126\%\text{pulegone} + 0.00801\%1,8 - \text{cineole}, \text{MAPE} = 1.22\% \quad \text{AEV} = 0.0002 \qquad (8)$$

Here, mean absolute percentage error (MAPE) is 1.22 % and approximation error variance (AEV), 0.0002. However, fit improves if the second major component in *E. camaldulensis* (α-pinene) is added:

$$I = -0.00377 + 0.0125\%\text{pulegone} + 0.0119\%\alpha - \text{pinene}, \text{MAPE} = 0.70\% \quad \text{AEV} = 0.0001 \qquad (9)$$

and AEV decays by 50 %. The matrix **R** of Pearson correlation coefficients was computed among all couples of properties $<i_1,i_2,i_3,i_4>$ for all five proximates in Table 1. The correlations are displayed in the partial correlation chart, which includes high ($r \geq 0.75$), medium ($0.50 \leq r < 0.75$), low ($0.25 \leq r < 0.50$) and *zero* ($r < 0.25$) partial autocorrelations. The couples of compounds with the highest partial associations present a similar property. Notwithstanding, the outcomes must be obtained with concern since Entry 5 in Table 1, with invariant property <1111>, presents a standard deviation $\sigma = 0$, producing the utmost partial association of one with all components, which results an artifact. After using the equipartition conjecture of entropy production, **R** matrix upper triangle turns out to be:

$$\mathbf{R} = \begin{pmatrix} 0.938 & 0.125 & 0.813 & 0.625 & 0.750 \\ & 0.938 & 0.000 & 0.438 & 0.063 \\ & & 0.938 & 0.500 & 0.875 \\ & & & 0.938 & 0.563 \\ & & & & 0.938 \end{pmatrix}$$

Some of the associations are rather elevated, for instance, R_{35} = 0.875. They are displayed in the partial correlation chart, which includes three high (*cf*. fig. 2, *red lines*), three medium (*orange*), one low (*yellow*) and three *zero* (*black*) partial associations. Two out of four high partial associations of Entry 5 (Table 1) are fixed: association 4–5 is a medium association and 2–5, a *zero* partial association.

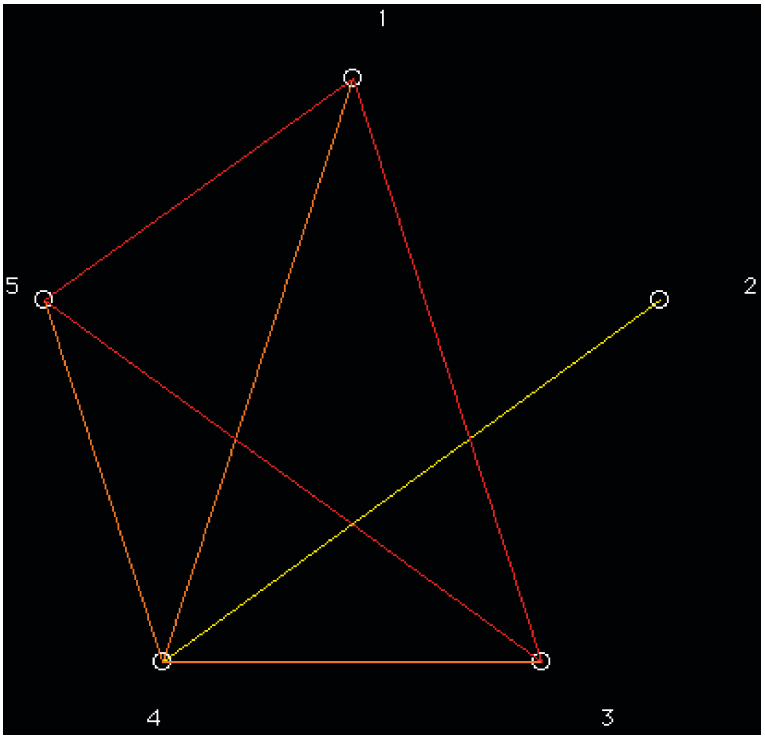

Fig. 2. Partial correlation chart: High (*red*), medium (*orange*) and low (*yellow*) correlations.

Classification law for equal weights a_k = 0.5 for grouping level $0.63 \leq b_1 \leq 0.75$ allows three classes:

C–b_1 = (1,3,5)(2)(4)

The groupings result with entropy h–**R**–b_1 = 4.15. The binary classification tree (*dendrogram*, *cf*. fig. 3) corresponding to $<i_1,i_2,i_3,i_4>$ and C–b_1 is calculated [51,52]. It gives a dual classification of Table 1 that divides clusters 1, 2 and 3 with 3, 1 and 1 compounds, respectively [53]. It presents dissimilar performance of compounds varying according to the quantity of C=C bonds. For instance, compounds (1,3,5) with lesser than two C=C bonds are classified into the same grouping. The proximates fitting in the same cluster result greatly connected in the partial correlation chart (fig. 2).

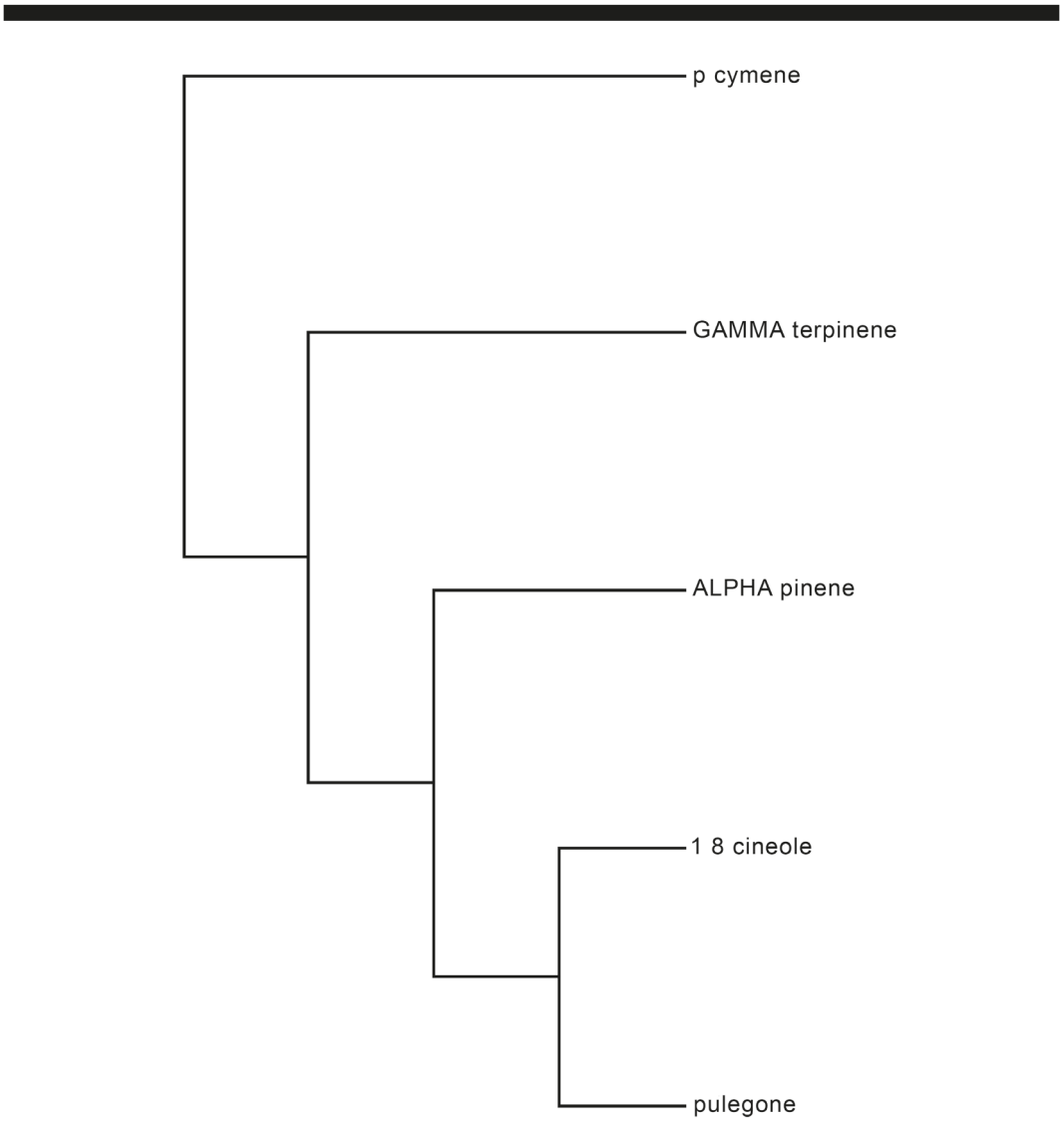

Fig. 3. Dendrogram of five components of essential oils of *E. camaldulensis* and *M. pulegium*.

Figure 4 shows picture of grouping as a radial tree. Dissimilar performance of compounds varies according to quantity of C=C. Clusters result in accord with partial correlation chart and binary tree (figs. 2 and 3). Again, substances (1,3,5) with lesser than two C=C are grouped into the same class.

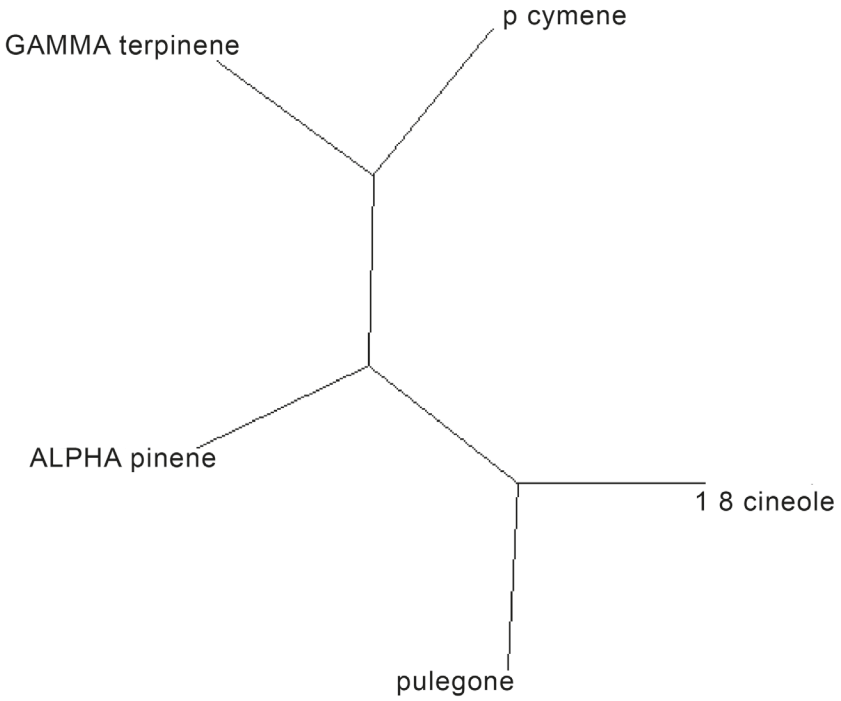

Fig. 4. Radial tree of five components of essential oils of *E. camaldulensis* and *M. pulegium*.

For the classification level b_2, with $0.13 \leq b_2 \leq 0.62$, the collection of clusters result:

C–b_2 = (1,3,4,5)(2)

Two groupings appear and entropy drops to h–**R**–b_2 = 1.84, matching $<i_1,i_2,i_3,i_4>$ and C–b_2 splitting into clusters 1 and 2 with 4 and 1 units, respectively. Again, substances (1,3,5) with lesser than two C=C bonds are classed together. Molecules in the same class result greatly connected in the partial correlation chart and dendrogram (figs. 2 and 3). Table 2 lists the relative examination of collection in 1–5 clusters, varying according to partial correlation chart, as well as binary/radial trees (figs. 2-4).

Table 2. Level of classification, quantity of classes and entropy for the property of oil components

Classification level *b*	Number of classes	Entropy *h*
1.00	5	10.68
0.87	4	7.19
0.75	3	4.15
0.62	2	1.84
0.12	1	0.23

The SplitsTree code makes analyzing information of cluster analysis (CA) possible [54]. On the base of *split decomposition*, it gets as a contribution a *distance matrix* and creates a chart that symbolizes associations among clusters. For perfect information, the chart is a dendrogram, while lesser model data will generate a dendrogram-like system that can be understood as a likely indication for contradictory information. Furthermore, as split decay does not try to compel information onto a dendrogram, it provides a high-quality sign of how *dendrogram*-like is certain information. The splits graph for all five proximates in Table 1 (*cf*. fig. 5) expose no contradictory association among classes. It presents the dissimilar performance of compounds varying according to the quantity of C=C bonds, in accord with the partial correlation chart, as well as binary and radial trees (figs. 2-4).

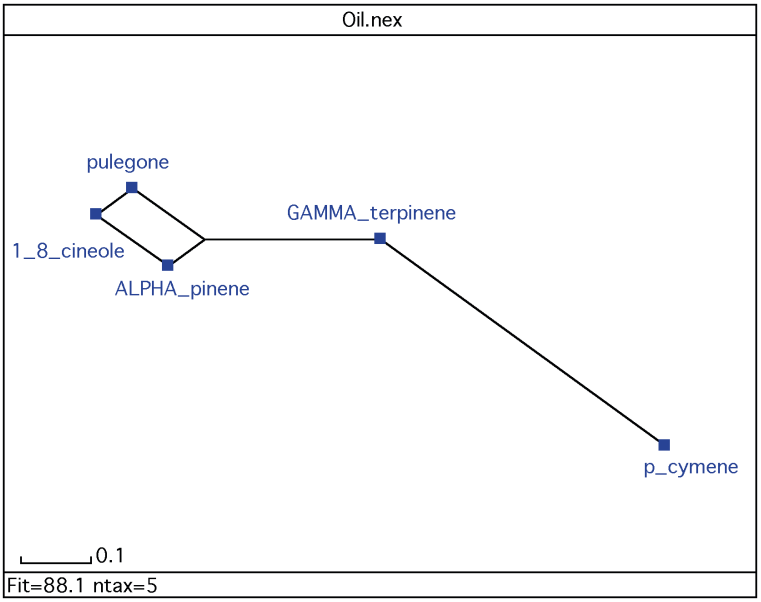

Fig. 5. Splits graph of five components of essential oils of *E. camaldulensis* and *M. pulegium*.

In data of QSPR, the archive includes lesser than a hundred items and more than 1000 *X*-variables. There are so many *X*-variables that nobody discovers by *inspection* models, tendencies, groups, *etc*. in items. *Principal component analysis* (PCA) is a method helpful to *sum up* all data included in

X-matrix and set it in an comprehensible form [55-60]. The PCA acts by factoring the **X**-matrix as a product of two lesser matrices **P** and **T**. The *loading matrix* (**P**) with data concerning variables includes a few vectors, the principal components (PCs) that are obtained as linear combinations (LCs) of the first *X*-variables. The *score matrix* (**T**), with data concerning the items, is such that all items are explained in terms of projections onto PCs in its place of the first variables: **X** = **TP'** + **E**, where **'** indicates the transpose matrix. The data not included in matrices stays as *unexplained* X-*variance* in the *residual matrix* (**E**). Every PC_i is a novel feature stated as an LC of the first characteristics x_j: $PC_i = \Sigma_j b_{ij} x_j$. The newer features PC_i are called *scores*/*factors* whereas the b_{ij} coefficients are named *loadings*. The scores are sorted consistent with their data content with regard to the whole variance between all items. The *score-score plots* present the places of the items in the newer feature scheme, whereas the *loading-loading plots* point to sites of characteristics, which stand for items in the newer features. The PCs show two attractive characteristics. (1) They are taken out in dropping sort of significance. The first PC F_1 every time includes more knowledge than the second F_2, F_2 more than the third F_3 and so on. (2) All PC result orthogonal to each other. No association exists among the knowledge included in dissimilar PCs. A PCA was carried out for the properties of EOs components. The significance of the F_1–F_4 factors of PCA for $\{i_1,i_2,i_3,i_4\}$ is listed in Table 3. First factor F_1 represents 63 % of explained variance (37 % of error), two first factors (F_1, F_2), 83 % of variance (17 % of error), three first factors (F_1, F_2, F_3), 97 % of variance (3 % of error) and so on.

Table 3. Significance of principal component analysis factors for composition of the five oils proximates

Factor	Eigenvalue	Percentage of variance	Cumulative percentage of variance
F_1	2.50000000	62.50	62.50
F_2	0.83333333	20.83	83.33
F_3	0.53745748	13.44	96.77
F_4	0.12920919	3.23	100.00

The factors loadings of PCA are listed in Table 4.

Table 4. Principal component analysis loadings for the composition of the five oil proximates

	PCA factor loadings [1]			
Property	F_1	F_2	F_3	F_4
i_1	0.48867778	0.00000000	**-0.86382434**	0.12248073
i_2	0.59850559	0.00000000	0.22976546	**-0.76746263**
i_3	0.44887921	**-0.70710678**	0.31702944	0.44497167
i_4	-0.44887921	**-0.70710678**	-0.31702944	-0.44497167

[1] Loadings larger than 0.7 are typed in boldface.

The profile from F_1 to F_4 of PCA for the property is reported in Table 5. For the F_1 factor, the i_2 variable presents the utmost weight in the profile; nevertheless, the F_1 factor cannot be cut to two variables $\{i_1,i_2\}$ lacking 40 % of the error. For the F_2 factor, the $\{i_3,i_4\}$ variables show 100 % weight and F_2 can be cut to both variables with 0 % of the error. For the F_3 factor, the i_1 variable reveals the utmost weight; notwithstanding, the F_3 factor cannot be cut to two variables $\{i_1,i_3\}$ lacking 15 % of the error. For the F_4 factor, the i_2 variable displays the utmost weight; nevertheless, the F_4 factor cannot be cut to two variables $\{i_2,i_3\}$ lacking 21 % of the error. In short, factors from F_1 to F_4 can be assumed LCs of the pairs of variables $\{i_1,i_2\}$, $\{i_3,i_4\}$, $\{i_1,i_3\}$ and $\{i_2,i_3\}$ with 40 %, 0 %, 15 % and 21 % of the errors.

Table 5. Profile of principal component analysis factors for composition of the five oil proximates

	Percentage of i_1 [1]	Percentage of i_2	Percentage of i_3	Percentage of i_4
F_1	23.88	35.82	20.15	20.15
F_2	0.00	0.00	**50.00**	**50.00**
F_3	**74.62**	5.28	10.05	10.05
F_4	1.50	**58.90**	19.80	19.80

[1] Percentages larger than 50 % are typed in boldface.

Scores plot of PCA F_2–F_1 for components (*cf.* fig. 6) presents dissimilar performance according to C=C bonds. It differentiates three groupings: cluster 1 with 3 compounds ($F_1 > F_2 \approx 0$, *right*), class 2 with 1 substance ($F_1 << F_2 = 0$, *left*) and grouping 3 with 1 molecule ($F_1 < F_2 = 0$, *middle*).

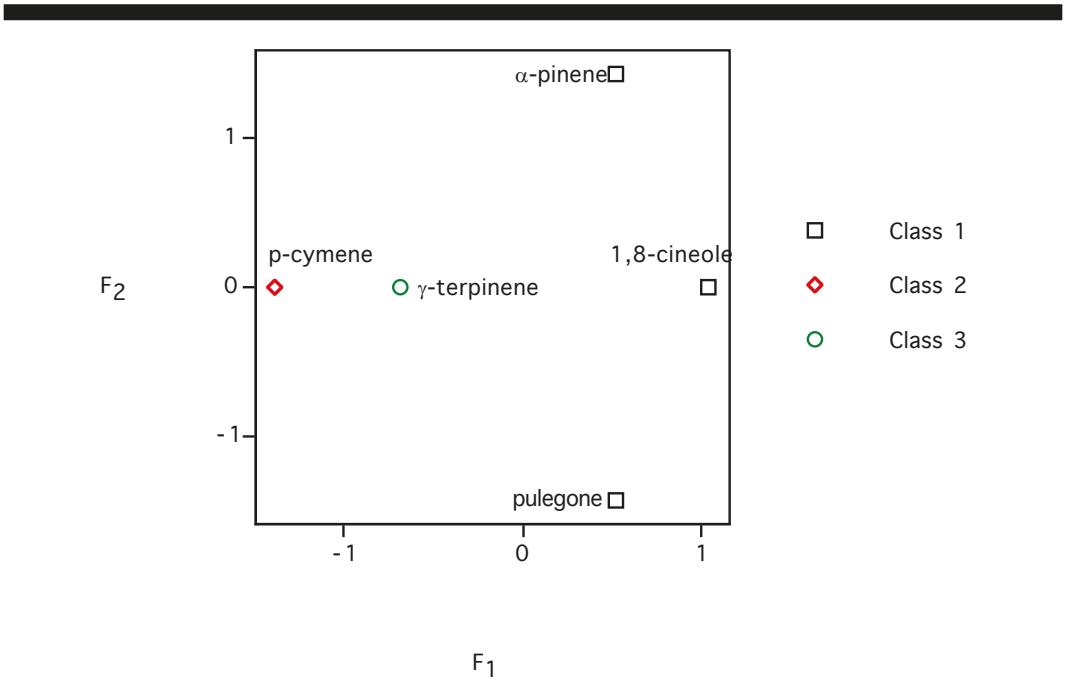

Fig. 6. Principal component analysis F_2 *versus* F_1 scores plot for the five components of oils.

From factors loadings of PCA, F_2–F_1 loadings plot (*cf.* fig. 7) illustrates properties. As an accompaniment to scores plot (fig. 6) for loadings (fig. 7), it results that compounds in grouping 1 situated in the right show an input of $C{=}C_{01}$ positioned on the same side. Substance in cluster 2 in the left presents a more marked input of cyc_1. Molecule in class 3 on the middle has more pronounced contribution of $C{=}C_{012}$. Two groupings of properties result evidently differentiated in loadings plot: grouping 1 $\{C{=}C_{012}, C{=}C_{01}\}$ ($F_1 > F_2 = 0$, *top*) and cluster 2 $\{O_1, cyc_1\}$ ($0 \approx F_1 > F_2$, *bottom*).

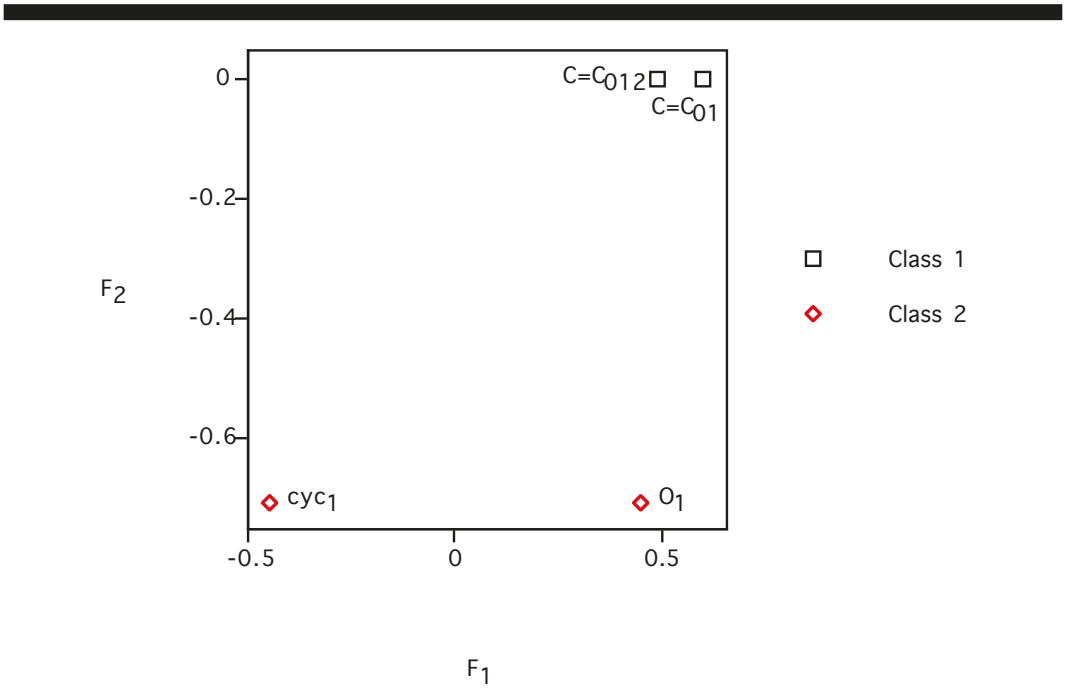

Fig. 7. Principal component analysis F_2 *versus* F_1 loadings plot for the five components of oils.

In the place of all five proximates in Table 1 in $\Re^4$ space of the four properties, assume the four properties in $\Re^5$ space of the five constituents. Matrix **R** upper triangle among couples of properties is:

$$\mathbf{R} = \begin{pmatrix} 0.969 & 0.906 & 0.406 & 0.094 \\ & 0.969 & 0.469 & 0.031 \\ & & 0.969 & 0.531 \\ & & & 0.969 \end{pmatrix}$$

Some of the associations are high, for instance, $R_{12} = 0.906$. Properties dendrogram (*cf.* fig. 8) divides $\{C{=}C_{012}, C{=}C_{01}\}$ (class 1) from $\{O_1, cyc_1\}$ (grouping 2), in accord with loadings plot of PCA (fig. 7).

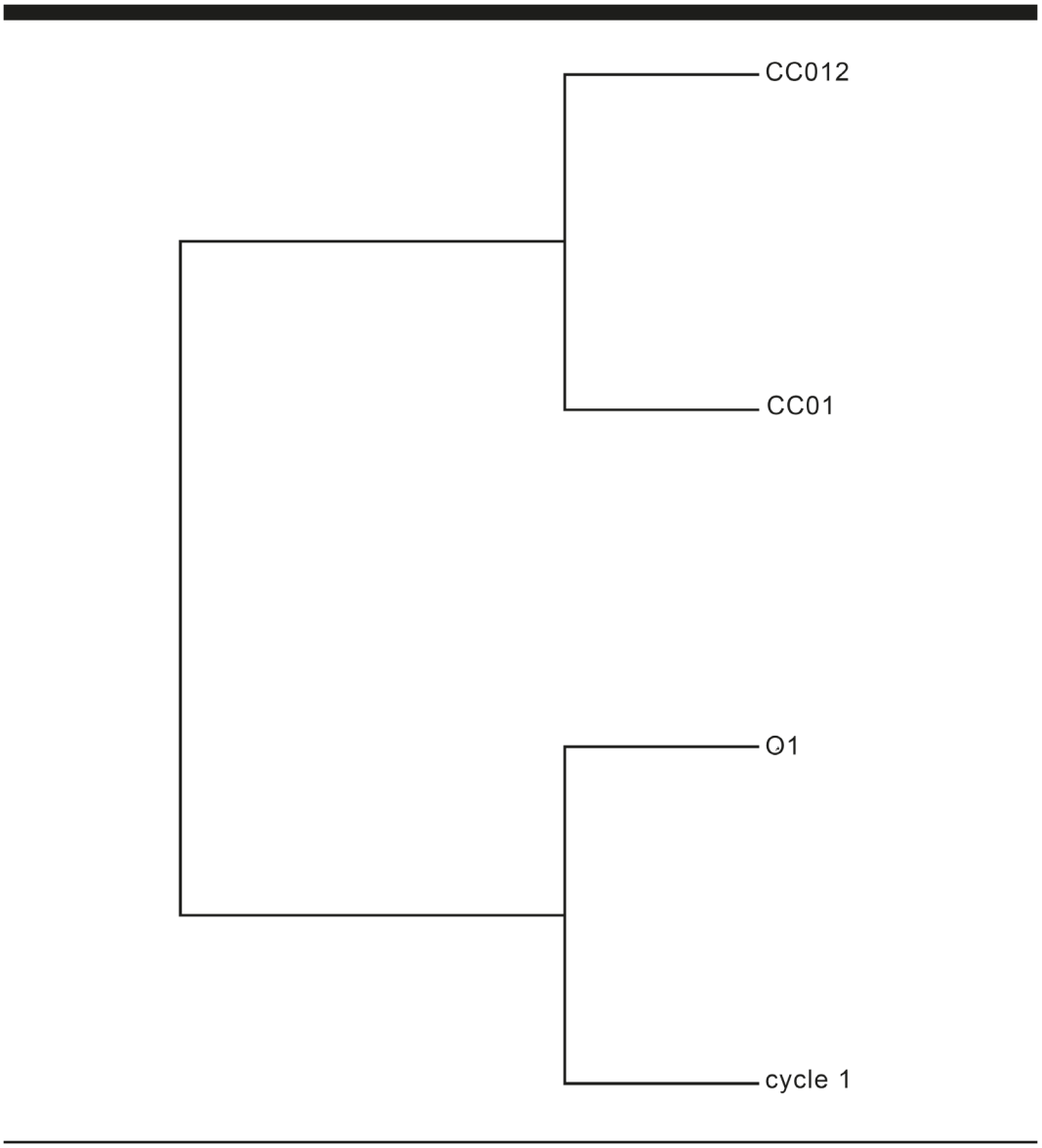

Fig. 8. Binary tree for the properties matching the five components of essential oils.

Radial tree (*cf*. fig. 9) separates equal groupings as loadings plot and binary tree (figs. 7 and 8).

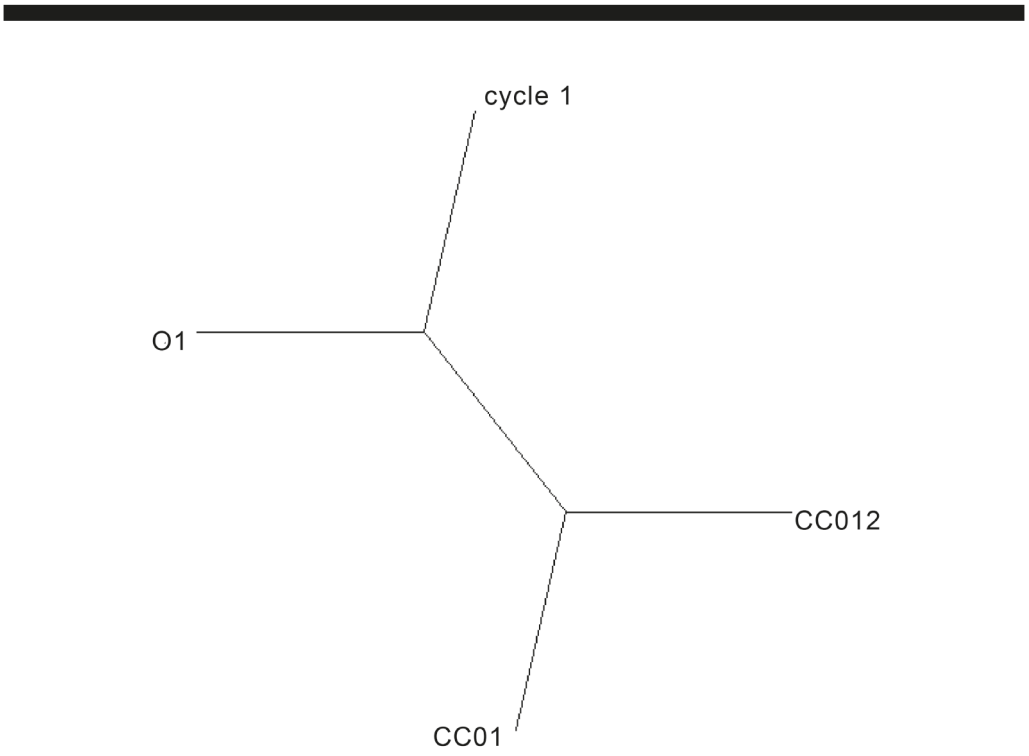

Fig. 9. Radial tree for the properties matching the five components of essential oils.

Properties splits graph (*cf*. fig. 10) displays a contradictory association among constituents of classes 1 and 2 owing to interdependences. It points to a false association resulting from base-composition consequences. It results in accord with PCA loadings plot, as well as binary/radial trees (figs. 7-9).

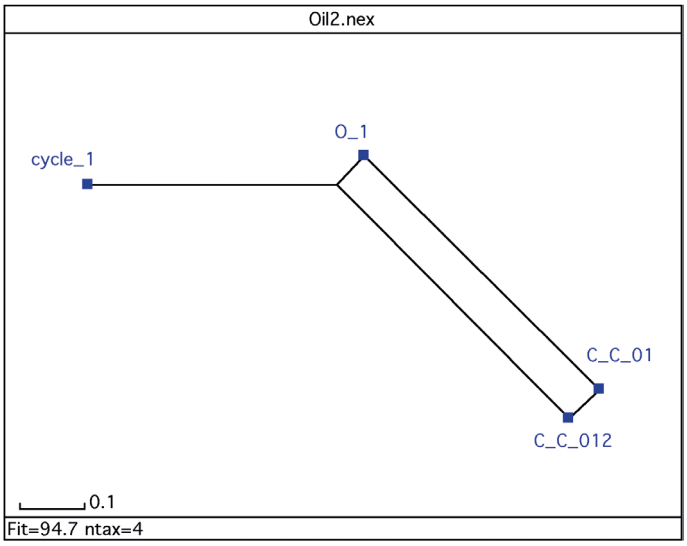

Fig. 10. Splits graph for the properties matching the five components of essential oils.

A PCA was carried out for the properties. The first factor F_1 represents 37 % of the explained variance (63 % of the error), the two first factors (F_1, F_2), 73 % of the explained variance (27 % of the error), the three first factors (F_1, F_2, F_3), 90 % of the explained variance (10 % of the error) and so on. In the F_2–F_1 scores plot of PCA, the same two clusters of properties are differentiated: grouping 1 {C=C_{012},C=C_{01}} ($F_1 > F_2$, *cf*. fig. 11, *bottom left*) and class 2 {O_1,cyc_1} ($F_1 < F_2$, *top right*), in accord with the loadings plot of PCA, binary and radial trees, as well as splits graph (figs. 7-10).

In proposed arrangement of periodic classification (periodic table, PT, *cf*. Table 6), EO components are categorized first by i_1, then by i_2, i_3 and, finally, by i_4. Vertical columns (*groups*) are described by $<i_1,i_2,i_3>$ and horizontal rows (*periods*), by $<i_4>$. For EOs constituents in table 1 separation of table 6, periods of four out of six components are considered. For instance, group g000 indicates $<i_1,i_2,i_3>$ = <000>: <0001> (C=C_3, O_0, cyc_1) and so on. Compounds in the same column result near in partial correlation chart, binary/radial trees, splits graph, as well as scores plot of PCA (figs. 2-6).

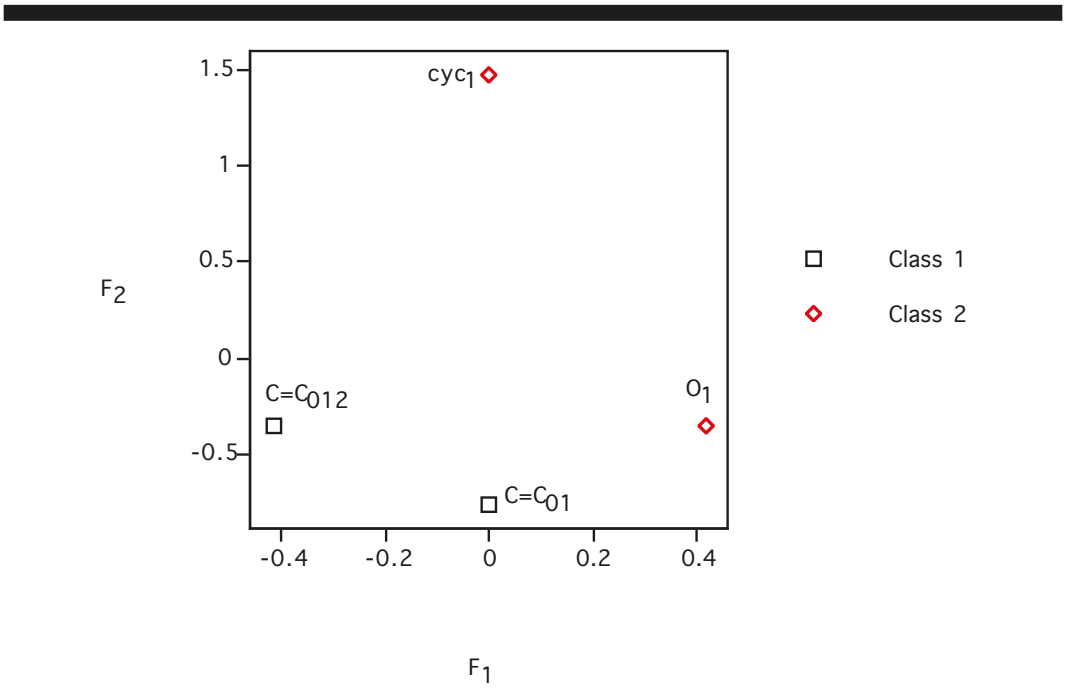

Fig. 11. F_2 *vs*. F_1 scores plot of PCA for the properties matching the five components of oils.

Table 6. Periodic properties of components of *E. camaldulensis* and *M. pulegium* oils and acetylsalicylic acid (ASA).

Pr.	g000	g001	g100	g101	g110	g111
p0	myrcene	*cis*-nerolidol	linalyl acetate β-caryophyllene	linalool geraniol	α-thujene α-pinene β-pinene δ-2-carene aromadendrene *allo*-aromadendrene	octan-3-ol 1,8-cineole fenchol *trans*-pinocarveol pinocarvone verbenone spathulenol globulol γ-eudesmol α-eudesmol
p1	*p*-cymene ASA	cuminaldehyde thymol	α-phellandrene limonene γ-terpinene *p*-mentha-3,8-diene	*trans*-carveol carvone eucarvone		3-methyl*cyclo*hexanone isopulegol menthone menthol terpinen-4-ol α-terpineol dihydrocarvone (*R*)-(+)-pulegone piperitone

DISCUSSION

It is tempting to speculate that molecular heterogeneity could explain the diversity of compositions, properties, functions, taste, odour, flavour, stability, volatility, storage and food of aromatic EOs. Oil effects infer that such natural volatilized EOs provide not only flavour but also possess biopharmacological activities. Having draft the molecular profile of EO components should give the basic elements, which cause/accompany composition and provide the necessary information for defining EOs and compositional subtypes in a rational way. What was so far mainly defined by compositional observation could also be approached hopefully better with full molecular characterization, which information maybe not replace but it certainly complements their constitution. The question is beyond the intellectual exercise of ordering observations: conceptually, it provides an EO compositional map, and facilitates reason/drawing hypotheses that will lead to understanding the nature of the constituents, activities and bioprinciples that govern it; in practice, it presents enormous implications for the activities and atlas of EO compositional types. The success of these attempts to define new molecular classification of EOs should not hide the fact that clustering data remains a challenging task, from a methodological viewpoint. In particular, many parameters affect the categorization obtained by grouping methods, for instance, the features and metric used to compare samples, clustering algorithm and procedure to select the quantity of groupings. Despite these limitations, the new molecular classifications of EOs obtained by automatic data categorization revolutionized the way one apprehends EO heterogeneity. As larger samples collections are analyzed, it is likely that finer classifications into well-specified and robust subtypes will emerge from clustering methods, and allow a more precise stratification of constituents and activities into subcategories, which would not be captured by only compositional parameters. As different subgroups can present different uses or activities, a more precise and robust categorization of constituents and powers can improve food and pharmacological uses. Classification improvement is its extension to other components, EOs, analgesics, anti-inflammatories, antipyretics, local anaesthetics, alcohols, ice and so on for situations being difficult to arrange *a priori*. Possible applications are revealed, for instance, the use of EOs instead of synthetic medicines to circumvent pathogens raising resistance. They could not only be used for infectious-diseases therapy but, also, as preservatives in the agro-food manufacturing. Use of EOs and their bioactive components in agro-food applications undergoes more than a few main disadvantages, *e. g*., short stability and bioavailability, poison for non-target organisms or sensory alterations in agro-food uses. Some EOs constituents present antimicrobial activity, *e. g*., thymol, thymoquinone, eugenol, carvacrol, diallyl disulphide, allyl isothiocyanate, cinnamaldehyde, etc. As the constituents are volatile, usually they evaporate previous to show antimicrobial effect. It is interesting to let them go in a maintained way to decrease volatility and extend bioeffect. Oils are good antimicrobial agents; however, their high volatility and reactivity limit their application as food preservatives [61]. A further possibility is, *e. g*., EO application in skin products to care for and prevent dermal infections. The dominant components of the particularized antimicrobial active EOs was complicated. Studies were performed concerning EO antimicrobial activity. However, some breeds were reported to present resistance to EOs. Oils possess strong antimicrobial activities *vs*. a number of microorganisms. Incomplete constancy and elevated volatility result EO disadvantages that make difficult the *in vitro* assays, as well as storage and application. Microbial resistance was informed because of natural re-

sistance or coming out tolerance by habituation. The antimicrobial potency cannot be assigned to one specific compound because essential oils are based on more than a few ingredients. Additional assays are required to study more concerning the connections among the individual substances of the intricate multicomponent combinations. Oil activities result of major components and minor-constituents synergism. Possible EO applications, as antimicrobials, include the chemical alteration of constituents of essential oils to increase antimicrobial potency or utilizing essential oils to care for infections by medicine-resistant bacterium breeds; however, additional work is advocated.

CONCLUSION

As of the discussion of the current results, the next conclusions can be sketched.

1. Several criteria to decrease the examination to a convenient number of oils components passed on *structural elements* that are *ranked*: less than 3 or 2 C=C > one O > one cycle. Pulegone ($C=C_1$, O_1, cyc_1) <1111> was selected as *reference*. A lot of categorization procedures are founded on *information entropy*. Intended for collections of reasonable dimension, a extreme quantity of outcomes result matching the information and suffer a combinatorial upsurge; notwithstanding, following the *equipartition conjecture*, one gets a criterion for selection coming out from categorization among hierarchical dendrograms, consistent with which the most excellent pattern is that in which entropy production is mainly consistently delivered. The procedure keeps away from the question of range variables since for <1111>, the zero standard deviation produces a Pearson correlation coefficient of unit. The MolClas code is an easy, dependable, effective and quick algorithm for molecular categorization. Information entropy allows clustering constituents of oils and is in agreement with principal component analyses. Periodic classification of components of essential oils indicates that the ones in the same column present similar features; constituents as well in the same row, utmost similarity.

2. Fungal strains resistant to chemical treatments, based on synthetic fungicides during apples-preservation period, push safety alternatives research. Work was devoted to composition antifungal activity of *Eucalyptus camaldulensis* and *Mentha pulegium* oils, on mycelium growth of *Alternaria alternata* and *Penicillium expansum* fungi responsible for apples deterioration. Constituents vary on region because of differences in geography. Oils inhibit fungi growth. Antifungal activity is because of components, which show synergism. Results deserve to be deepened by *in vivo* tests to develop a biofight milieu based on natural substances *vs*. fungi causing apples rot; further studies concerning isolation/identification of oil active components and testing for pharmacological activities are useful.

3. Program MolClas allows classifying oils and anti-inflammatory drugs for hard situations that are difficult to arrange *a priori* (*e. g*., oils, analgesics, anti-inflammatory agents, antipyretics, local anaesthetics, alcohols, ice). Although anaesthetics and oil components perform in ache management, ice shows benefits in simple of service, quick action and economy. Periodic table of constituents was extended to oils, analgesics, anti-inflammatories, procaine, benzyl alcohol and ice. Forthcoming work will deal with more *Eucalyptus* and *Mentha* spp. Further study will classify edible virgin olive, sunflower, high-oleic sunflower and soya bean oils.

ACKNOWLEDGEMENT

The author thanks support from Fundación Universidad Católica de Valencia San Vicente Mártir (Project No. 2019-217-001UCV).

LITERATURE CITED

[1] Zhang CQ, Liu YH, Zhu GN. Detection and characterization of benzimidazole resistance on *Botrytis cinerea* in greenhouse vegetables. Eur J Plant Pathol. 2010;126:509-15.
[2] Moleyar V, Narasimhan P. Antifungal activity of some essential oil components. Food Microbiol. 1986;3:331-6.
[3] Soliman KM, Badeaa RI. Effect of oil extracted from some medicinal plants on different mycotoxigenic fungi. Food Chem Toxicol. 2002;40:1669-75.
[4] Jacet Dongmo PM, Tatsadjieu LN, Tchinda Sonwa E, Kuate J, Amvam Zollo PH, Menut C. Essential oils of *Citrus aurantifolia* from Cameroon and their antifungal activity against *Phaeoramularia angolensis*. African J Agric Res. 2009;4:354-8.
[5] Bourkhiss M, Hnach M, Bourkhiss B, Ouhssine M, Chaouch A. Composition chimique et propriétés antimicrobiennes de l'huile essentielle extraite des feuilles de *Tetraclinis articulata* (Vahl) du Maroc. Afrique Sci. 2007;3:232-42.
[6] Magina MDA, Dalmarco EM, Wisniewski A, Simionatto EL, Dalmarco JB, Pizzolatti MG, Brighente IMC. Chemical composition and antibacterial activity of essential oils of *Eugenia* species. J Nat Med. 2009;63:345-50.
[7] Bouzouita N, Kachouri F, Ben Halima M, Chaabouni MM. Composition chimique et activités antioxydante, antimicrobienne et insecticide de l'huile essentielle de *Juniperus phoenicea*. J Soc Chim Tunis. 2008;10:119-25.
[8] Erler F, Ulug I, Yalcinkaya, B. Repellent activity of five essential oils against *Culex pipiens*. Fitoterapia. 2006;77:491-4.
[9] Tang GW, Yang CJ, Xie LD. Extraction of *Trigonella foenum-gracum* L. by supercritical fluid CO_2 and its contact toxicity to *Rhyzopertha dominica* (Fabricius) (Coleoptera: Bostrichidae). J Pest Sci. 2007;80:151-7.
[10] Cheng S, Huang C, Chen Y, Yu J, Chen W, Chang S. Chemical compositions and larvicidal activities of leaf essential oils from two eucalyptus species. Bioresour Technol. 2009;100:452-6.
[11] Bhaskara Reddy MV, Angers P, Gosselin A, Arul J. Characterization and use of essential oil from *Thymus vulgaris* against *Botrytis cinerea* and *Rhizopus stolonifer* in strawberry fruits. Phytochemistry. 1997;47:1515-20.
[12] Nielsen PV, Rios R. Inhibition of fungal growth on bread by volatile components from species and herbs and the possible application in active packaging, with special emphasis on mustard essential oil. Int J Food Microbiol. 2000;60:219-29.
[13] Tzortzakis NG. Maintaining postharvest quality of fresh produce with volatile compounds. *Innovat Food Sci Emerg Technol*. 2006;8:111-6.

[14] Amarti F, Satrani, B, Ghanmi M, Farah A, Aafi A, Aarab L, El Ajjouri M, Chaouch A. Composition chimique et activité antimicrobienne des huiles essetieles de *Thymus algeriensis* Boiss. and Reut et *Thymus ciliatus* (Desf) Benth. du Maroc. Biotechnol Agron Soc Environ. 2010;14:141-8.
[15] Bachir Raho G, Benali M. Antibacterial activity of leaf essential oils of *Eucalyptus globulus* and *Eucalyptus comaldulensis*. African J Pharm Pharmacol. 2008;2:211-5.
[16] Batish DR, Singh HP, Kohli RK, Kaur S. Eucalyptus essential oil as a natural pesticide. For Ecol Manage. 2008;256:2166-74.
[17] Mahboubi M, Haghi G. Antimicrobial activity and chemical composition of *Mentha pulegium* L. essential oil. J Ethnopharmacol. 2008;119:325-7.
[18] Snoussi M, Hajlaoui H, Noumi E, Usai D, Sechi LA, Zanetti S, Bakhrouf A. *In vitro* anti-*vibrio* spp. activity and chemical composition of some Tunisian plants. World J Microbiol Biotechnol. 2008;24:3071-6.
[19] Ouraini D, Agoumi A, Ismaili-Alaoui M, Alaoui K, Cherrah Y, Alaoui MA, Belabbas MA. Activité antifongique de l'acide oléique et des huiles essentielles de *Thymus saturejoides* L. et *Mentha pulegium* L., comparée aux antifongiques dans les dermatoses mycosiques. Phytotothérapie, 2007;1:6-14.
[20] Hajlaoui H, Trabelsi N, Noumi E, Snoussi M, Fallah H, Ksouri R, Bakhrouf A. Biological activities of the essential oils and methanol extract of two cultivated mint species (*Mentha longifolia* and *Mentha pulegium*) used in the Tunisian folkloric medicine. World J Microbiol Biotechnol. 2009;25:2227-38.
[21] Verdeguer M, Blazquez MA, Boira H. Phytotoxic effects of *Lantana camara*, *Eucalyptus comaldulensis* and *Eriocephalus africanus* essential oils in weeds of Mediterranean summer crops. Biochem Syst Ecol. 2009;37:362-9.
[22] Akin M, Aktumsek A, Nostro A. Antibacterial activity and composition of the essential oils of *Eucalyptus comaldulensis* Dehn. and *Myrtus communis* L. growing in Northern Cyprus. African J Biotechnol. 2010;9:531-5.
[23] Hmiri S, Rahouti M, Habib Z, Satrani B, Ghanmi M, El Ajjouri M. Évaluation du potentiel antifongique des huiles essentielles de *Mentha pulegium* et d'*Eucalyptus camaldulensis* dans la lutte biologique contre les champignons responsables de la détérioration des pommes en conservation. Bull Soc R Sci Liège. 2011;80:824-36.
[24] Baser KHC, Buchbauer G, editors. Handbook of essential oils: Science, technology, and applications. Boca Raton (FL): Taylor and Francis; 2010.
[25] Lang G, Buchbauer G. A review on recent research results (2008-2010) on essential oils as antimicrobials and antifungals. A review. Flavour Fragrance J. 2012;27:13-39.
[26] Ibáñez Jaime MD, Blázquez Ferrer MA. Post-emergent herbicidal activity of *Eucalyptus globulus* Labill. essential oil. Nereis. 2018;2018(10):25-36.
[27] Ibáñez MD, Blázquez MA. Essential oils: Quality indicators of spices in supermarkets. Nereis. 2019;2019(11):39-50.
[28] Ibáñez MD, Blázquez, MA. Phytotoxic effects of commercial *Eucalyptus citriodora*, *Lavandula angustifolia*, and *Pinus sylvestris* essential oils on weeds, crops, and invasive species. Molecules. 2019;24:2847–1-15.

[29] Varmuza K. Pattern recognition in chemistry. New York: Springer; 1980.
[30] Benzecri JP. L'analyse des données. Paris: Dunod; 1984. Vol. 1.
[31] Tondeur D, Kvaalen E. Equipartition of entropy production. An optimality criterion for transfer and separation processes. Ind Eng Chem Fundam. 1987;26:50-6.
[32] Castellano G, Tena J, Torrens, F. Classification of polyphenolic compounds by chemical structural indicators and its relation to antioxidant properties of *Posidonia oceanica* (L.) Delile. MATCH Commun Math Comput Chem. 2012;67:231-50.
[33] Castellano G, González-Santander JL, Lara A, Torrens F. Classification of flavonoid compounds by using entropy of information theory. Phytochemistry. 2013;93:182-91.
[34] Torrens F, Castellano G. From Asia to Mediterranean: Soya bean, Spanish legumes and commercial *soya bean* principal component, cluster and meta-analyses. J Nutr Food Sci. 2014;4(5):98-8.
[35] Castellano G, Lara A, Torrens F. Classification of stilbenoid compounds by entropy of artificial intelligence. Phytochemistry. 2014;97:62-9.
[36] Castellano G, Torrens F. Information entropy-based classification of triterpenoids and steroids from *Ganoderma*. Phytochemistry. 2015;116:305-13.
[37] Castellano G, Torrens F. Quantitative structure–antioxidant activity models of isoflavonoids: A theoretical study. Int J Mol Sci. 2015;16:12891-906.
[38] Castellano G, Redondo L, Torrens F. QSAR of natural sesquiterpene lactones as inhibitors of Myb-dependent gene expression. Curr Top Med Chem. 2017;17:3256-68.
[39] Torrens F, Castellano G. Structure-activity relationships of cytotoxic lactones as inhibitors and mechanisms of action. Curr Drug Discov Technol. 2020;17:166-182.
[40] Torrens F, Redondo L, Castellano G. Artemisinin: Tentative mechanism of action and resistance. Pharmaceuticals, 2017;10:20-4-4.
[41] Kaufmann A. Introduction à la théorie des sous-ensembles flous. Paris: Masson; 1975. Vol. 3.
[42] Cox E. The fuzzy systems handbook. New York: Academic; 1994.
[43] Kundu S. The min-max composition rule and its superiority over the usual max-min composition rule. Fuzzy Sets Sys. 1998;93:319-29.
[44] Lambert-Torres G, Pereira Pinto JO, Borges da Silva LE. Minmax techniques. In: Wiley encyclopedia of electrical and electronics engineering. New York: Wiley; 1999.
[45] Shannon CE. A mathematical theory of communication: Part I, discrete noiseless systems. Bell Syst Tech J. 1948;27:379-423.
[46] Shannon CE. A mathematical theory of communication: Part II, the discrete channel with noise. Bell Syst Tech J. 1948;27:623-56.
[47] White H. Neural network learning and statistics. AI Expert. 1989;4(12):48-52.
[48] Kullback S. Information theory and statistics. New York: Wiley; 1959.
[49] Iordache O, Corriou JP, Garrido-Sánchez L, Fonteix C, Tondeur D. Neural network frames. application to biochemical kinetic diagnosis. Comput Chem Eng. 1993;17:1101-13.
[50] Iordache O. Modeling multi-level systems. Berlin: Springer; 2011.
[51] Tryon RC. A multivariate analysis of the risk of coronary heart disease in Framingham. J Chronic Dis. 1939;20:511-24.
[52] Jarvis RA, Patrick EA. Clustering using a similarity measure based on shared nearest neighbors. IEEE Trans Comput. 1973;C22:1025-34.

[53] Page RDM. Program TreeView. Glasgow (UK): Universiy of Glasgow; 2000.
[54] Huson DH. SplitsTree: Analizing and visualizing evolutionary data. Bioinformatics. 1998;14:68-73.
[55] Hotelling H. Analysis of a complex of statistical variables into principal components. J Educ Psychol. 1933;24:417-41.
[56] Kramer R. Chemometric techniques for quantitative analysis. New York: Marcel Dekker; 1998.
[57] Patra SK, Mandal AK, Pal MK. State of aggregation of bilirubin in aqueous solution: Principal component analysis approach. J Photochem Photobiol Sect A. 1999;122:23-31.
[58] Jolliffe IT. Principal component analysis. New York: Springer; 2002.
[59] Xu J, Hagler A. Chemoinformatics and drug discovery. Molecules, 2002;7:566-600.
[60] Shaw PJA. Multivariate statistics for the environmental sciences. New York: Hodder-Arnold; 2003.
[61] Kloucek P, Smid J, Flesar J, Havlik J, Titera D, Rada V, Drabek O, Kokoska L. *In vitro* inhibitory activity of essential oil vapors against *Ascophaera apis*. *Nat Prod Commun*. 2012;7:253-6.