

**Classification of Phenolic Compounds by Chemical Structural Indicators and Its
Relation to Antioxidant Properties of *Posidonia Oceanica* (L.) Delile**

Gloria Castellano*

C tedra Energesis de Tecnolog a Interdisciplinar, Universidad Cat lica de
Valencia San Vicente M rtir, Guillem de Castro-94, E-46003 Valencia,
Spain

gloria.castellano@ucv.es

Jos  Tena

Instituto de Medioambiente y Ciencia Marina, Universidad Cat lica de
Valencia San Vicente M rtir, Guillem de Castro-94, E-46003 Valencia,
Spain

josetena@ucv.es

Francisco Torrens

Institut Universitari de Ci ncia Molecular, Universitat de Val ncia, Edifici
d'Instituts de Paterna, Spain.

francisco.torrens@uv.es

(Received March 16, 2011)

ABSTRACT

A total of 33 phenolic compounds are classified into a periodic table by using an algorithm based on the theory of information entropy. Five features in hierarchical order are used to classify structurally the phenols. From these features, the first three mark the group or column, while the last two mark the row or period in the table. Those phenols in the same group and period are suggested to have maximum similarity in properties. Furthermore, those with only the same group will present moderate similarity. The Mediterranean endemic seagrass *Posidonia oceanica* is constituted by cells rich in phenolic compounds. In this report, we related the phenolic components of this phanerogam to compounds of the table whose bioactivity and antioxidant properties have been previously published.

INTRODUCTION

Polyphenols are among the most widespread class of secondary metabolites in nature. Most polyphenols arise from a common origin: the amino acids phenylalanine or tyrosine. These amino acids are deaminated to cinnamic acids, which enter the phenylpropanoid pathway. A key step in this biosynthetic route is the introduction of one or more

hydroxyl groups into the phenyl ring. Biosynthesis, produces the large variety of plant phenols such as cinnamic acids (C6-C3) and benzoic acids (C6-C1) (Figure 1) [1]. Despite their origins, these molecules present antioxidant, antimutagenic, antiviral, antibacterial (bactericidal, bacteriostatic), algicidal, antifungal, insecticidal, estrogenic and keratolytic activities that may serve to protect the organism from competing ones in their biological environment [2-6].

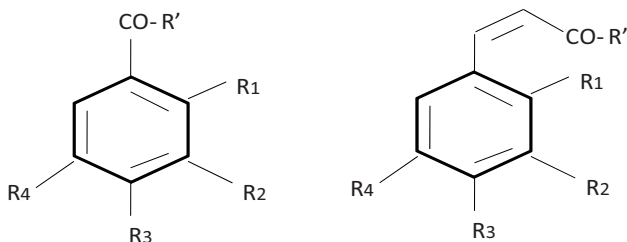


Figure 1. Molecular structures of benzoic and cinnamic acid derivatives

Phenolic hydroxyl groups are good H-donating antioxidants, which scavenge reactive oxygen species and breaks the cycle of generation of new radicals. The radical scavenging antioxidants inhibit free-radical-mediated oxidation of lipids, proteins and DNA, which is implicated in illness. Phenolics act as antioxidants by inhibiting enzymes involved in radical generation [7-9].

Among the several methods designed to measure the activity compounds, such as the lipid peroxidation inhibition capacity (LPIC) assay [10], 1,1-diphenyl-2-picrylhydrazyl (DPPH)-scavenging [11], cyclic voltammetry [12] and using the induction period method [13]. The antioxidants act synergistically together with other antioxidants *in vivo* [14].

With regard to the aromatic substitution, the increase in the number of methoxy groups or the catechol structure substantially increased the antioxidant activity of the compounds by further stabilizing the phenoxyl radical [15].

P. oceanica (L.) Delile is a marine seagrass endemic to the Mediterranean Sea where it forms vast meadows. *P. oceanica* is used as a bio-indicator of ecosystem. Phenolics are formed to protect the plant from photosynthetic stress, reactive oxygen, anthropogenic pressures and inter-specific competition [16, 17]. *P. oceanica* meadows suffer a great regression, especially from last century. Its lost is severe in the surroundings of big urban and port areas [18-20]. Nowadays, it is protected by the EU (Managerial Habitat 92/43 EEC and Community Board 97/62/EEC).

A total of 23 compounds were detected in *P. oceanica* samples [16, 17]. Of them, 20 are phenolic and cinnamic acid derivatives; of these, acetosyringone and ferulic acid are the main compounds. Other compounds exhibit much lower average concentrations: syringaldehyde, vanillic acid, acetovanillone, 4- hydroxybenzoic acid, p-anisic acid, protocatechualdehyde and protocatechuic acid. The remaining compounds represent less than 2%, including caffeic and gallic acid as traces.

In the present report 33 phenolic compounds (Table 1) have been studied. A lot of them are present in *P. oceanica*.

Table 1. Chemical structures of phenolic acids given as R', R₁, R₂, R₃ and R₄ substitutions of the benzoic and cinnamic scaffolds in Figure 1.

<i>Benzoic derivative</i>	<i>R'</i>	<i>R₁</i>	<i>R₂</i>	<i>R₃</i>	<i>R₄</i>	<i>Cinnamic derivative</i>
	H	H	H	H	H	Cinnamaldehyde
4-Hydroxybenzaldehyde*	H	H	H	OH	H	
Protocatechualdehyde*	H	H	OH	OH	H	
Vanillin*	H	H	OCH ₃	OH	H	
Syringaldehyde*	H	H	OCH ₃	OH	OCH ₃	
Veriatricaldehyde	H	H	OCH ₃	OCH ₃	H	
Piccol*	CH ₃	H	H	OH	H	
Acetovanillone*	CH ₃	H	OCH ₃	OH	H	
Acetosyringone*	CH ₃	H	OCH ₃	OH	OCH ₃	
Methylvanillate	OCH ₃	H	OCH ₃	OH	H	
Methylsyringate	OCH ₃	H	OCH ₃	OH	OCH ₃	
	OCH ₃	H	H	OH	H	Methylcoumarate
	OCH ₃	H	OCH ₃	OH	H	Methylferulate
Benzoic acid*	OH	H	H	H	H	Cinnamic acid*
p-Hydroxybenzoic acid*	OH	H	H	OH	H	p-Coumaric acid*
Salicylic acid	OH	OH	H	H	H	
3-Hydroxybenzoic acid	OH	H	OH	H	H	
α-Resorcylic acid	OH	H	OH	H	OH	
β-Resorcylic acid	OH	OH	H	OH	H	
Protocatechuic acid*	OH	H	OH	OH	H	Caffeic acid*
Pyrocatechuic acid	OH	OH	OH	H	H	
Gentisic acid*	OH	OH	H	H	OH	
Gallic acid*	OH	H	OH	OH	OH	
Vanillic acid*	OH	H	OCH ₃	OH	H	Ferulic acid*
Isovanillic acid	OH	H	OH	OCH ₃	H	
p-Anisic acid*	OH	H	H	OCH ₃	H	
Syringic acid*	OH	H	OCH ₃	OH	OCH ₃	Sinapic acid*
Veratric acid	OH	H	OCH ₃	OCH ₃	H	

*Acid derivatives presents in *Posidonia oceanica*

The aim of this article is to develop the learning potentialities of our classification program of molecular structures. The molecules with similar chemical structures are described. We have expanded the study of general approaches to the processing of structural information. The following section presents the computational method. The next section describes the classification algorithm. Following that, the equipartition conjecture of entropy production is exposed. Next, a section analyzes the learning procedure. Then, a section describes program MolClas. A new section explains program GraphCor. Then the calculation results are presented and discussed. Finally, the conclusions are summarized.

COMPUTATIONAL METHOD

The key problem in classification studies is to define similarity indices when several criteria of comparison are involved. The first step in quantifying the concept of similarity for molecules of phenolic compounds is to list the most important moieties with respect to antioxidant activity of such molecules. Furthermore, the vector of properties $\vec{i} = \langle i_1, i_2, \dots, i_k, \dots \rangle$ should be associated to each phenolic compound i , whose components correspond to different characteristic groups of the molecule, in a hierarchical order according to the expected importance of their anti-oxidant potency. The components i_k are "1" or "0" according to experimental conclusions of antioxidant power to structural variations in the phenolic compound: index $i_1 = 1$ denotes an acid group $-(CH=CH)_n-COOH$, $i_2 = 1$ signifies one phenolic hydroxyl group (OH) on either o- or p- position, $i_3 = 1$ indicates the presence of an -OR group on o-position with regard to the OH above, $i_4 = 1$ signifies the presence of methoxyl groups on o- or p- with respect to the OH above and $i_5 = 1$ means that the molecule is a cinnamic acid derivative. In phenolics without OH group as veratricaldehyde $i_3 = i_4 = 0$. Table 2 contains the vectors associated to 33 phenolic compounds.

Let us denote by r_{ij} ($0 \leq r_{ij} \leq 1$) the similarity index of two phenolic compounds associated to the \vec{i} and \vec{j} vectors, respectively. The relation of similitude is characterized by a similarity matrix $\overline{R} = [r_{ij}]$. The similarity index between two phenolic compounds $\vec{i} = \langle i_1, i_2, \dots, i_k, \dots \rangle$ and $\vec{j} = \langle j_1, j_2, \dots, j_k, \dots \rangle$ is defined as:

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

where $0 \leq a_k \leq 1$ and $t_k = 1$ if $i_k = j_k$, but $t_k = 0$ if $i_k \neq j_k$. This definition assigns a weight $(a_k)^k$ to any property involved in the description of molecules i or j . The hierarchical order of the

five structural features is expressed in their corresponding weights. For instance, for all $\alpha_k = 0.5$ these weights are 0.5, 0.25, 0.125, 0.0625, 0.03125.

Table 2. Vector properties of phenolic compounds.

1. 4-Hydroxybenzaldehyde*	<01000>
2. Protocatechualdehyde*	<01100>
3. Vanillin*	<01110>
4. Syringaldehyde*	<01110>
5. Veriatricaldehyde	<00000>
6. Piccol*	<01000>
7. Cinnamicaldehyde	<00001>
8. Acetovanillone*	<01110>
9. Acetosyringone*	<01110>
10. Methylvanillate	<01110>
11. Methylsyringate	<01110>
12. Methylcoumarate	<01001>
13. Methylferulate	<01111>
14. Benzoic acid*	<10000>
15. p-Hydroxybenzoic acid*	<11000>
16. Salicylic acid	<11000>
17. 3-Hydroxybenzoic acid	<11000>
18. α -Resorcylic acid	<11000>
19. β -Resorcylic acid	<11000>
20. Protocatechuic acid*	<11100>
21. Pyrocatechuic acid	<11100>
22. Gentisic acid*	<11000>
23. Gallic acid*	<11100>
24. Vanillic acid*	<11110>
25. Isovanillic acid	<00001>
26. p-Anisic acid*	<10000>
27. Syringic acid*	<11110>
28. Veratric acid	<10000>
29. Cinnamic acid*	<10001>
30. p-Coumaric acid*	<11001>
31. Caffeic acid*	<11101>
32. Ferulic acid*	<11111>
33. Sinapic acid*	<11111>

*Acid derivatives presents in *Posidonia oceanica*

CLASSIFICATION ALGORITHM

The grouping algorithm uses the stabilized similarity matrix, obtained by applying the max-min composition rule defined by:

$$(\bar{R} \circ \bar{S})_{ij} = \max_k [\min_k (r_{ik}, s_{kj})] \quad (2)$$

Where $\bar{R} = [r_{ij}]$ and $\bar{S} = [s_{ij}]$ are two similarity matrices, and $(\bar{R} \circ \bar{S})_{ij}$ is the (i,j) -th element of the matrix $\bar{R} \circ \bar{S}$ [21]. It can be shown that when applying this rule iteratively so that

$\bar{R}^{(n+1)} = \bar{R}^{(n)} \circ \bar{R}$, there is an integer n such that: $\bar{R}^{(n)} = \bar{R}^{(n+1)} = \dots$. The resulting matrix $\bar{R}^{(n)}$ is called the stabilized similarity matrix. The importance of this matrix lies in the fact that in the classification process, it will generate a partition into disjoint classes. From now on, it is understood that the stabilized matrix is used and designated by $\bar{R}^{(n)} = [r_{ij}^{(n)}]$. The grouping rule is the following: i and j are assigned to the same class if $r_{ij}^{(n)} \geq b$, where b is the grouping level of the classification. The class of i noted \hat{i} is the set of species j that satisfies the rule $r_{ij}^{(n)} \geq b$. The matrix of classes is:

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

where s stands for any index of a species belonging to the class \hat{i} (similarly for t and \hat{j}). Rule (3) means finding the largest similarity index between species of two different classes.

In information theory, the information entropy h measures the surprise that the source emitting the sequences can give [22-23]. For a single event occurring with probability p , the degree of surprise is proportional to $-\ln p$. Generalizing the result to a random variable X (which can take N possible values x_1, \dots, x_N with probabilities p_1, \dots, p_N), the average surprise received on learning the value of X is $-\sum p_i \ln p_i$. The information entropy associated with the matrix of similarity \bar{R} is:

$$h(\bar{R}) = -\sum_{i,j} r_{ij} \ln r_{ij} - \sum_{i,j} (1 - r_{ij}) \ln (1 - r_{ij}) \quad (4)$$

Denote also by C_b the set of classes and by \bar{R}_b the matrix of similarity at the grouping level b . The information entropy satisfies the following properties.

1. $h(\bar{R}) = 0$ if $r_{ij} = 0$ or $r_{ij} = 1$. $h(\bar{R})$ is maximum if $r_{ij} = 0.5$, i.e., when the imprecision is maximum.
2. $h(\bar{R}_b) \leq h(\bar{R})$ for any b , i.e., classification leads to a loss of entropy.
3. $h(\bar{R}_{b_1}) \leq h(\bar{R}_{b_2})$ if $b_1 < b_2$, i.e., the entropy is a monotone function of the grouping level b .

THE EQUIPARTITION CONJECTURE OF ENTROPY PRODUCTION

In the classification algorithm, each hierarchical tree corresponds to a dependence of entropy on the grouping level, and thus an h - b diagram can be obtained [25]. The Tondeur and Kvaalen equipartition conjecture of entropy production is proposed as a selection criterion among different variants resulting from classification among hierarchical trees. According to this conjecture, the best configuration of a flow sheet is the one in which entropy production is most uniformly distributed, *i.e.*, closest to a kind of equipartition. One proceeds here by analogy using information entropy instead of thermodynamic entropy. Equipartition implies a linear dependence, which is a constant production of entropy along the b scale, so that the equipartition line is described by:

$$h_{\text{eqp}} = h_{\text{max}} b \quad (5)$$

Indeed, since the classification is discrete, a realistic way of expressing equipartition would be a regular staircase function. The best variant is chosen to be that minimizing the sum of squares of the deviations:

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

Initially, b was introduced as a limiting value; in Equation (5) b is used as a factor; finally, in Equation (6) b is discretized.

LEARNING PROCEDURE

Learning procedures similar to those encountered in stochastic methods are implemented as follows [24]. Consider a given partition into classes as good or ideal from practical or empirical observations. This corresponds to a reference similarity matrix $\bar{\bar{S}} = [s_{ij}]$ obtained for equal weights $a_1 = a_2 = \dots = a$ and for an arbitrary number of fictitious properties. Next, consider the same set of species as in the good classification and the actual properties. The similarity degree r_{ij} is then computed with Equation (1) giving the matrix $\bar{\bar{R}}$. The number of properties for $\bar{\bar{R}}$ and $\bar{\bar{S}}$ may differ. The learning procedure consists in trying to find classification results for $\bar{\bar{R}}$ as close as possible to the good classification. The first weight a_1 is taken constant and only the following weights a_2, a_3, \dots are subjected to random variations. A new similarity matrix is obtained using Equation (1) and the new weights. The distance between the partitions in classes characterized by $\bar{\bar{R}}$ and $\bar{\bar{S}}$ is given by:

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

The definition was suggested by that introduced in information theory by Kullback to measure the distance between to probability distribution [25].

The result of the algorithm is a set of weights allowing adequate classification. Such a procedure has been applied in the synthesis of complex flow-sheets using information entropy [26].

DESCRIPTION OF CALCULATION PROGRAM MOLCLAS

We have written program MolClas for molecular classification based on the equipartition conjecture of entropy production. The MolClas allows reading the coefficients from input and optimizing them. MolClas allows either calculating the correlation matrix or reading it from input. An option has been included to transform negative correlation coefficients to zero. The program punches both similarity and difference matrices, as well the latter in format NEXUS (.NEX) for programs PAUP, MacClade and SplitsTree. MolClas performs both single- and complete- linkage hierarchical cluster analyses (CAs) of the compounds by using the IMSL subroutine CLINK [27].

DESCRIPTION OF GRAPHIC PROGRAM GRAPHCOR

We have written graphic program GraphCor for performing the graphical correlation diagram. GraphCor can present high partial correlations ($|r| \geq 0.75$) in red, medium partial correlations ($0.50 \leq |r| < 0.75$) in orange, low partial correlations ($0.25 \leq |r| < 0.50$) in yellow and “zero” partial correlations ($|r| < 0.25$) in white. An option has been included to transform negative correlation coefficients to zero. MolClas and GraphCor are available from the authors at Internet (francisco.torrens@uv.es) and are free for academics.

CALCULATION RESULTS AND DISCUSSION

The matrix of Pearson correlation coefficients between each pair of vector properties $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ of the 33 phenolic compounds has been calculated. The Pearson intercorrelations are illustrated in the partial correlation diagram, which contains high partial correlations ($r \geq 0.75$), medium partial correlations ($0.50 \leq r < 0.75$), low partial correlations ($0.25 \leq r < 0.50$) and zero partial correlations ($r < 0.25$). Pairs of phenolic compounds with

high partial correlations show a similar vector property. However, the results should be taken with care, because the two compounds with constant $\langle 11111 \rangle$ vector (Entries 32 and 33 in Table 1) show null standard deviation, causing high partial correlations ($r = 1$) with any phenolic compounds, which is an artifact. With the equipartition conjecture, the intercorrelations are illustrated in the partial correlation diagram, which contains 171 high, 97 medium and 175 low partial correlations (Figure 2). Notice that 36 of the 64 ($2 \times 18/32$) high partial correlations of Entries 32 and 33 have been corrected; *e.g.*, for Entry 32, ferulic acid, the correlations with Entries 14, 25, 26, 28 and 29 are medium, its correlations with Entries 1, 2, 3, 4, 6, 8, 9, 10, 11, 12 and 13 are low partial correlations, and its correlations with Entries 5 and 7 are *zero* partial correlations.

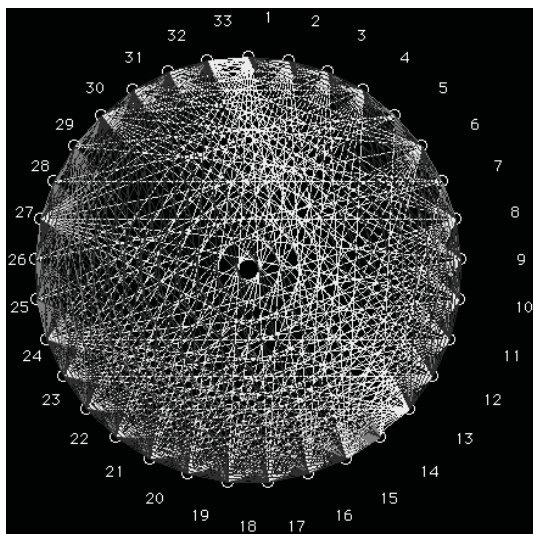


Figure 2. Partial correlation diagram: High, medium and low correlations in grayscale

The grouping rule in the case with equal weights $a_k = 0.5$, for the classification level $0.93 \leq b_1 \leq 0.91$ allows the classes:

$$C_{b_1} = (1,6,12) (2) (3,4,8-11,13) (5,7) (14,25,26,28,29) (15-19,22,30) (20,21,23,31) \\ (24,27,32,33)$$

The eight classes are obtained with the associated entropy $h(\mathbf{R}_{b_1}) = 31.12$. Both dendrogram (binary tree) and radial tree [28, 29] matching to $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ and C_{b_1} are calculated [30.] The former is illustrated (Figure 3).



Figure 3. Dendrogram for the studied phenols

A comparative analysis of the set containing 1–33 classes is summarized in Table 3. The C_{b1} separates the same eight classes in agreement with partial correlation diagram, dendrogram and radial tree (Figure 2 and 3).

The principal component analysis (PCA) [31] F_2 - F_1 scores plot (cf. Fig 4) phenolic compounds with the same vector property appear superimposed. The same eight classes of polyphenols are clearly distinguished in concordance with the reference clustering: class 1 (compounds 1, 6, 12, $F_1 < F_2$, centre-left), class 2 (compound 2, $F_1 > F_2$, bottom), etc. The classification is in qualitative agreement with the partial correlation diagram, dendrogram and radial tree (Fig 2-3).

Table 3. Classification level, number of classes and information entropy for the phenolic compounds.

Classification level b	Number of classes	Entropy h
1.00	33	521.20
0.96	15	109.22
0.93	8	31.12
0.87	6	17.71
0.75	4	7.80
0.68	3	4.05
0.50	2	1.66
0.18	1	0.14

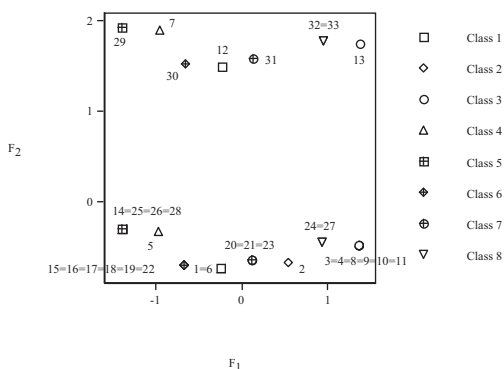


Figure 4. Principal component analysis F_2 vs. F_1 scores plot for phenolic compounds

The recommended format for the periodic table (PT) of phenolic compounds, *cf.* Table 4, shows that they are classified first by i_1 , then by i_2 , i_3 , i_4 and, finally, by i_5 . However, instead of 2^5 now six classes with common bit patterns are of interest (Table 4). Periods of six units are assumed; *e.g.*, group $\underline{g}000$ stands for $\langle i_1, i_2, i_3 \rangle = \langle 000 \rangle$, *viz.* $\langle 00000 \rangle$ (veriatricaldehyde) and $\langle 00001 \rangle$ (cinnamaldehyde), *etc.* The polyphenols in the same column appear close in partial correlation diagram, dendrogram, radial tree and PCA (Figures 2–4).

Table 4 shows our results of classification of the studied phenolic compounds. Generally, groups of compounds with similar structures follow similar trends.

Table 4. Table of periodic properties of the phenolic compounds corresponding to the vector $\langle i_1, i_2, i_3, i_4, i_5 \rangle$

	g000	g010	g011	g100	g110	g111
p00	vanilinaldehyde	4-hydroxybenzaldehyde* piccol*	probenzothioaldehyde*	benzoic acid* isovanillic acid p-anisic acid* veratric acid	p-hydroxybenzoic acid* salicylic acid 3-hydroxybenzoic acid α -resorcylic acid β -resorcylic acid gentisic acid*	protocatechuic acid* pyrocatechuic acid* gallic acid*
p01	cinnamicaldehyde	methylcoumarate		cinnamic acid*		caffeic acid*
p10			Vanillin* syringaldehyde* acetovanillone* acetosyringone* methylvanilate methylsyringate			vanillic acid* syringic acid*
p11			methylferulate			ferulic acid* synapic acid*

*Acid derivatives presents in *Posidonia oceanica*

Table 5 includes data of antioxidant power taken from the literature [32, 33] Antioxidant character increases across the periods, i.e. for period p00: g000 < g010 < g011 < g100 < g110 < g111. In most cases it also increases when descending in a group, i.e. for group p111: p00 < p01 < p10 < p11. This order is the same as that found in Table 4.

The antioxidant activity of phenolic acids is related to the number and position of hydroxyl groups in the molecule. The antioxidant efficiency of mono-phenols is strongly enhanced by the introduction of a second methoxyl or hydroxyl group at the *o*- or *p*-position with respect to the hydroxyl above, i.e. Table 4 show that, according to publications [31, 32] the introduction of a second hydroxyl group in the *o*- position (caffeic acid) or *p*-position (protocatechuic acid), g111 group, enhances the antioxidant activity, making these phenolic acids more efficient than their respective monofenols *p*-hydroxybenzoic acid, g010, and *p*-coumaric acid, g100. This is consistent with the electron withdrawing potential of the single carboxyl functional group on the phenol ring affecting the *o*- and *p*-positions. Ferulic and sinapic acid (hydroxycinnamic acids) in p11-g111 position are the most antioxidant compounds. This is in agreement with the literature, in which hydroxycinnamic acids were found to be more effective than their hydroxybenzoic acid counterparts, possibly because of the aryloxy-radical stabilizing effect of the $-\text{CH}=\text{CH}-\text{COOH}$ linked to the phenyl ring by resonance [33].

Moreover, we predict that the extract of the intermediate leaves of *P. oceanica* has an important antioxidant character, since it contains a great number of phenolic acid derivatives (benzoics and cinnamics), especially ferulic acid and acetosyringone, which are the main compounds, as well as other phenolic acids that exhibit lower concentrations: vanillic acid, acetovanillone, 4-hydroxybenzoic acid, *p*-anisic acid, protocatechuic acid and protocatechualdehyde. On the other hand, caffeic acid and gallic acid are present merely as trace compounds. However, in situations of stress, the composition of phenols increases in *P. oceanica* because these are formed to protect the plant from reactive oxygen species, anthropogenic pressures and inter-specific competition.

Table 5. Antioxidant activity of phenolic compounds

	g000	g010	g011	g100	g110	g111
P00						
(AE)	vanillic aldehyde			benzoic acid		protocatechuic acid
TEAC	-0,08			0,24		0,36
				veratric acid		1,19
				0,16		
(AE)						
P01						
(AE)	cinnamic aldehyde	methylcoumarate			p-coumaric acid	caffeic acid
TEAC	-0,08	0,09			0,24	0,47
						1,26
P10						
TEAC						vanillic acid
						1,43
(AE)						syngaldehyde
						0,12
(AE)						methylvanillate
						0,17
(AE)						methylsyngalate
						0,22
P11						
(AE)						ferulic acid
TEAC						0,61
						1,90

AE: Antioxidant Effect phenolic compounds after 2 hours of incubation with linoleic acid
TEAC: TROLOX

Figure 5 exhibits the variation of vector property P of vector $\langle i_1, i_2, i_3, i_4, i_5 \rangle$, expressed in the decimal system $P=10^4i_1 + 10^3i_2 + 10^2i_3 + 10i_4 + i_5$ as a function of structural parameters $\{i_1, i_2, i_3, i_4, i_5\}$ for the polyphenols. The property was not used in the development of the PT (Table 4) and serves to validate it. The results agree with a PT of properties with vertical groups defined by $\{i_1, i_2, i_3\}$ and horizontal periods described by $\{i_4, i_5\}$.

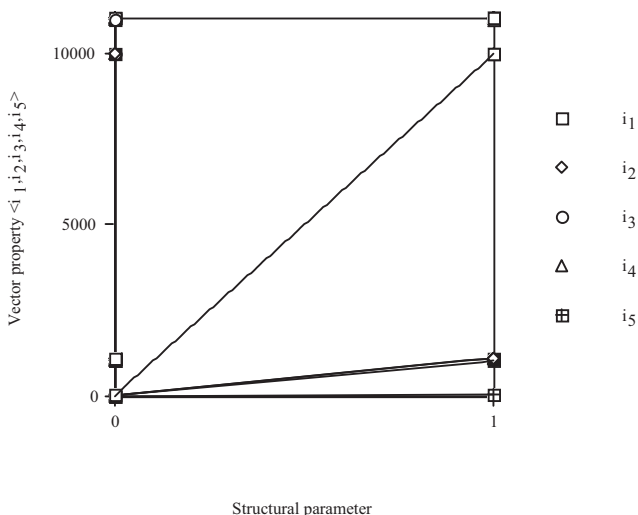


Figure 5. Variation of vector property of polyphenols in base 10 vs. counts $\{i_1, i_2, i_3, i_4, i_5\}$

The variation of property P of vector $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ in base 10 as a function of the number of the group in PT for the phenols, *cf.* Figure 6, reveals minima corresponding to phenols with $\langle i_1, i_2, i_3 \rangle$ close to $\langle 000 \rangle$ (group g000). Periods p00, p01, p10, p11 represent rows 1, 2, 3 and 4 in Table 4, respectively. The corresponding function $P(i_1, i_2, i_3, i_4, i_5)$ reveals a series of periodic waves clearly limited by maxima or minima, which suggest a periodic behavior that recalls the form of a trigonometric function. For $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ a maximum is clearly shown. Some points appear superimposed, especially the p11 line is on p10. The distance in $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ units between each pair of consecutive maxima is six, which coincides with the phenolic sets belonging to the same group in PT and in the successive periods. The maxima occupy analogous positions in the curve and are in phase. The representative points in phase should correspond to the elements of the same group in PT. For $\langle i_1, i_2, i_3, i_4, i_5 \rangle$

maxima there is coherence between the two representations; however, this coherence is not complete. The comparison of the waves shows only periods p00 and p01 are complete or almost complete. It should be remarked that the results for period p00 and p01 are similar and periods p10 and p11 with lesser number of data could present similar trends.

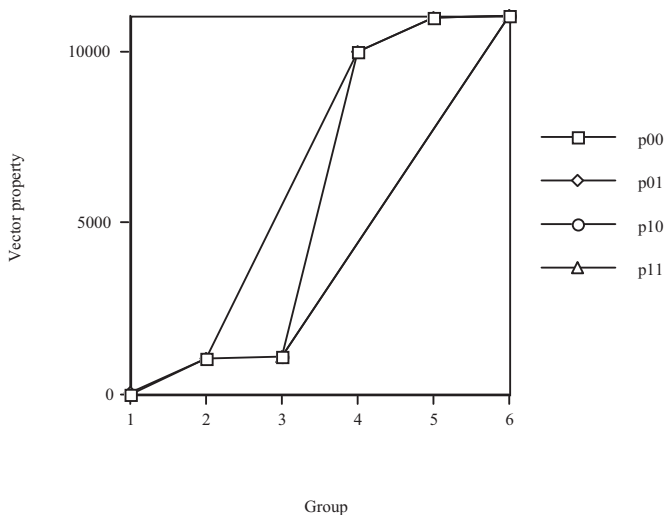


Figure 6. Variation of vector property of phenolic compounds in base 10 vs. group number.

An empirical function $P(p)$ reproduces the different $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ values. A minimum of $P(p)$ has meaning only if it is compared with the former $P(p-1)$ and later $P(p+1)$ points, needing to fulfill:

$$\begin{aligned} P_{\min}(p) &< P(p-1) \\ P_{\min}(p) &< P(p+1) \end{aligned} \quad (8)$$

Order relations (8) should repeat at determined intervals equal to the period size and are equivalent to:

$$\begin{aligned} P_{\min}(p) - P(p-1) &< 0 \\ P(p+1) - P_{\min}(p) &> 0 \end{aligned} \quad (9)$$

As relations (9) are valid only for minima more general others are desired for all the values of p . The differences $Z(p)$ are calculated assigning each of their values to phenolic compound p :

$$Z(p) = P(p+1) - P(p) \quad (10)$$

Instead of $Z(p)$ the $R(p) = P(p+1)/P(p)$ values can be taken assigning them to phenolic compound p . If periodic low (PL) were general, the elements in the same group in analogous positions in the different waves would satisfy:

$$\text{either } Z(p) > 0 \text{ or } Z(p) < 0 \quad (11)$$

$$\text{either } R(p) > 1 \text{ or } R(p) < 1 \quad (12)$$

Figure 7 shows variation of $Z(p)$ vs. grow number. The results, in general, are as expected, $Z(p) > 0$ for all periods according to the periodic low, as expected. For group 6 the result $Z(p) < 0$ should be taken with care because $Z(p)$ is calculated from the following position in the periodic table in the subsequent period.

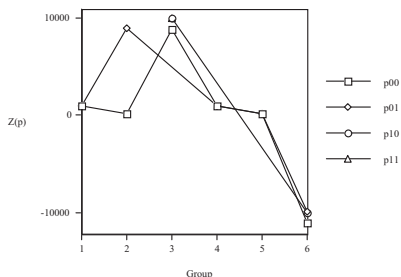


Figure 7. Variation of $Z(p) = P(p+1) - P(p)$ vs. grow number for phenolic compounds

Moreover Figure 8a shows the change of $R(p)$ vs. grow number, pointing up that $R(p) > 1$. Figure 8b details that groups 2-5 show $R(p) > 1$ according to the periodic low. The limitation of group 6 is explained in the previous paragraph.

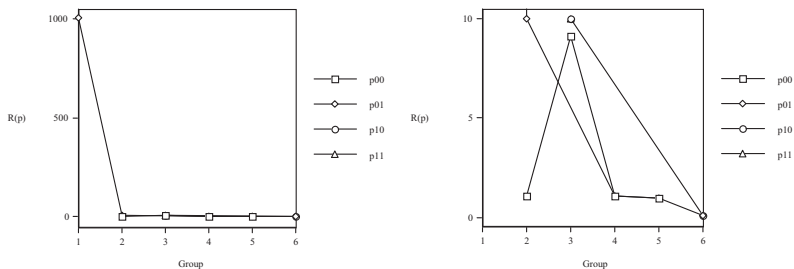


Figure 8.(a). Variation of $R(p) = P(p+1) / P(p)$ vs. grow number for phenolic compounds. (b) Detail.

CONCLUSION

Information entropy and principal component analysis allow classifying the phenolic compounds and agree. The cinnamic and benzoic acid derivatives are grouped into different classes. Most polyphenols are present in *P. oceanica*. The final classification is shown precise. MolClas has been written not only to analyze the equipartition conjecture of entropy production, but also to explore the world of molecular classification.

This report present results of classification of the 33 studied phenolic compounds. The analysis includes such chemical compounds that fit the following general scheme: presence of a carboxylic acid group type $-(CH=CH)_n-COOH$ ($n=0, 1$), hydroxyl group on either *o*- or *p*-position, -OR group on *o*-position with regard to the hydroxyl above, methoxyl group on *o*- or *p*- with respect to the hydroxyl and the molecule is a cinnamic acid derivative.

Generally, groups of compounds with similar structure follow the same trend. The antioxidant character increases through the periods. In most cases it also augments when descending in a group. This order is the same as that found in literature.

REFERENCES

- [1] D. M. Pereira, P. Valentão, J. A. Pereira, P. B. Andrade, Phenolics: From chemistry to biology, *Molecules* **14** (2010) 2202-2211.
- [2] L. Lacikova, M. Jancova, J. Muselik, I. Masterova, D. Grancai, M. Fickova, Antiproliferative, cytotoxic, antioxidant activity and polyphenols contents in leaves of four *Staphylea* L. species, *Molecules* **14** (2009) 3259-3267.
- [3] S. Khadem, R. J. Marles, Monocyclic phenolic acids; hydroxy- and polyhydroxybenzoic acids: Occurrence and recent bioactivity studies, *Molecules* **15** (2010) 7985-8005.
- [4] K. P. Chong, S. Rossall, M. Atong, In vitro antimicrobial activity and fungitoxicity of syringic acid, caffeic acid and 4-hydroxybenzoic acid in *Ganoderma Boninense*, *J. Agr. Sci.* **1** (2010) 15-20.
- [5] K. D. Croft, The chemistry and biological effects of flavonoids and phenolic acids, *Ann. New York Acad. Sci.* **854** (1998) 435-442.
- [6] C. S. Yang, J. M. Landau, M. Huang, H. L. Newmark, Inhibition of carcinogenesis by dietary polyphenolic compounds, *Ann. Rev. Nutr.* **21** (2001) 381-406.

- [7] L. R. Fukumoto, G. Mazza, Assessing antioxidant and prooxidant activities of phenolic compounds. *J. Agric. Food Chem.* **48** (2000), 3597-3604.
- [8] E. Anouar, P. Kosinová, D. Kozłowski, R. Mokrini, J. L. Duroux, P. Trouillas, New aspects of the antioxidant properties of phenolic acids: A combined theoretical and experimental approach, *Phys. Chem. Phys.* **11** (2009) 7659-7668.
- [9] P. M. Dewick, *Medicinal Natural Products: A Biosynthetic Approach*, Wiley, New York, 2002.
- [10] J. Zhang, R. A. Stanley, L. D. Melton, M. A. Skinner, Inhibition of lipid oxidation by phenolic antioxidants in relation to their physicochemical properties. *Pharmacology on-line* **12** (2007) 1496-1547.
- [11] W. Brand-Williams, M. E. Cuvelier, C. Berset, Use of free radical method to evaluate antioxidant activity, *Food Sci. Technol.* **28** (1995) 25-30.
- [12] A. Simić, D. Manojlović, D. Segan, M. Todorović, Electrochemical behavior and antioxidant and prooxidant activity of natural phenolics, *Molecules* **12** (2007) 2327-2340.
- [13] Y. Kadoma, S. Fujisawa, A comparative study of the radical-scavenging activity of the phenolcarboxylic acids caffeic acid, p-coumaric acid, chlorogenic acid and ferulic acid, with or without 2'-mercaptoethanol, a thiol, using the induction period method, *Molecules* **13** (2008) 2488-2499.
- [14] Y. Omata, Y. Saito, Y. Yoshida, E. Niki, Simple assessment of radical scavenging capacity of beverages, *J. Agric. Food Chem.* **56** (2008) 3386-3390.
- [15] F. Natella, M. Nardini, M. Di Felice, C. Scaccini, Benzoic and cinnamic acid derivatives as antioxidants: Structure-activity relation, *J. Agric. Food Chem.* **47** (1999) 1453-1459.
- [16] S. Awstini, J. Dmomm, G. Perwnt, Distribution of phenolic compounds in the seagrass *Posidonia oceanica*, *Phytochemistry* **48** (1998) 611-617.
- [17] O. Dumay, J. Costa, J. Desjobert, G. Pergent, Variations in the concentration of phenolic compounds in the seagrass *Posidonia oceanica* under conditions of competition, *Phytochemistry* **65** (2004) 3211-3220.
- [18] C. F. Boudouresque, A. Meinesz, Découverte de l'herbier de Posidonies, *Cahier Parc nation.* **4** (1982) 1-7.

- [19] A. Leriche, V. Pasqualini, C. F. Boudouresque, G. Bernard, P. Bonhomme, P. Clabaut, J. Denis, Spatial, temporal and structural variations of a *Posidonia oceanica* seagrass meadow facing human activities. *Aquat. Bot.* **84** (2006) 287–293.
- [20] M. Montefalcone, G. Albertelli, C. Morri, C.N. Bianchi, Urban seagrass: Status of *Posidonia oceanica* facing the Genoa city waterfront (Italy) and implications for management, *Mar. Pollut. Bull.* **54** (2007) 206–213.
- [21] A. Kaufmann, *Introduction à la Théorie des Sous-ensembles Flous*, Masson, Paris, 1975.
- [22] C. E. Shannon, A mathematical theory of communication: Part I, discrete noiseless systems, *Bell Syst. Tech. J.* **27** (1948) 379-423.
- [23] C. E. Shannon, A mathematical theory of communication: Part II, the discrete channel with noise, *Bell Syst. Tech. J.* **27** (1948) 623-656.
- [24] H. White, Neural network learning and statistics, *AI Expert* **4** (1989) 48-52.
- [25] S. Kullback, *Information Theory and Statistics*, Wiley, New York, 1959.
- [26] O. Iordache, J. P. Corriou, L. Garrido-Sánchez, C. Fonteix, D. Tondeur, Neural network frames. Application to biochemical kinetic diagnosis, *Comput. Chem. Eng.* **17** (1993) 1101-1113.
- [27] IMSL, *Integrated Mathematical Statistical Library (IMSL)*, IMSL, Houston, 1989.
- [28] R. C. Tryon, A multivariate analysis of the risk of coronary heart disease in Framingham, *J. Chronic Dis.* **20** (1939) 511-524.
- [29] R. A. Jarvis, E. A. Patrick, Clustering using a similarity measure based on shared nearest neighbors, *IEEE Trans. Comput.* **C22** (1973) 1025-1034.
- [30] R. D. M. Page, *Program TreeView*, University of Glasgow, 2000.
- [31] P. J. A. Shaw, *Multivariate Statistics for the Environmental Sciences*, Hodder-Arnold, New York, 2003.
- [32] D. Szwajgier, J. Pielecki, Z. Targoriski, Antioxidant activities of cinnamic and benzoic acid derivatives, *Acta Sci. Pol. Technol. Aliment.* **4** (2005) 129-142.
- [33] C. A. Rice-Evans, N. J. Miller, G. Paganga, Structure-antioxidant activity relationships of flavonoids and phenolic acids, *Free Radic. Biol. Med.* **20** (1996) 933-956.