

# Comparing Morphological With Genetic Distances Between Populations: A New Method and its Application to the *Prosopis chilensis* – *P. flexuosa* complex

By A. VERGA<sup>1)</sup> and H.-R. GREGORIUS<sup>2)</sup>

(Received 18<sup>th</sup> May 2005)

## Abstract

The biological units that are the object of management, preservation and improvement for the development of sustainable productive systems in natural areas, need to be differentiated and analyzed. Attending to this need, a new morphological distance is presented in this work. This distance is based on qualitative criteria and is applied to numerical taxonomy studies. The characteristics of this trait allow its comparison with the genetic distance of GREGORIUS (1974). Both parameters are essential tools in basic studies of native species populations. The morphological distance is applied to reveal genetically differentiated units in a swarm of hybrids between closely related species, and this result is compared with the results obtained from the application of traditional methods of numerical taxonomy.

*Key words:* genetic distance, morphological distance, taxonomic distance, *Prosopis* species.

## Introduction

In order to achieve the development of sustainable productive systems based on the management of native flora resources, it is essential to know the species that compose them. In order to gain knowledge of this, the main subject of research starts with the study of the populations and the relationships between the species that compose the flora, which is the basis of these productive systems.

These studies on natural populations and native species should start with a genetic and phenotypic analysis of the individuals because a central concern is to distinguish parts of the natural distribution which are reproductively isolated against others. Many studies are specifically aimed at differentiating biological units within species complexes in order to contribute to efficient breeding as well as utilization and maintenance of genetic resources (CHIN-SUNG CHANG and HUI KIM, 2003; PETIT, R. et. al., 2003; NES, R. et. al., 2002; CHIA-SZU WEN and JU-YING HSIAO, 1999; BUCK, G. and BIDLACK, J. 1998). These studies typically apply methods of classification and differentiation with special reference to reproductively coherent units.

However, many traditional methods which rely on clustering according to spatial or geographic criteria fail to reveal distinct biological units when the object of study are sympatrically occurring taxonomic species

that form species complexes, or when hybrid swarms involving closely related species are to be analyzed. In such cases it is meaningful to use morphological criteria at a first step to look for structure within geographically definable units. At a second step the morphological characterization (when leading to distinguishable clusters) is checked for consistency with its associated genetic features. Only after this process yields units of sufficient genetic and morphological homogeneity, other criteria (such as physiological, ecological or relating to productivity) can be used for further characterization, or the units may serve as basic material for breeding and genetic conservation. Furthermore, if the combined morphological and genetic method of analysis suggests the existence of distinct biological units within a species complex or a hybrid swarm, the microevolutionary processes that led to the differentiation of these units can be studied more efficiently. Traditional studies, which rely on primarily geographic and morphological criteria are unlikely to allow for such conclusions.

From a genetic point of view, there are many efficient tools to obtain and analyze the data. On the other hand, for the phenotypic analysis, it is essential to have a methodology for data analysis in agreement with the genetic methods. Comparisons of phenotypic with genetic distributions between populations are particularly relevant in studies of evolutionary adaptive differentiation. In such studies phenotypic differences between populations are to be analyzed as to their adaptational significance. This includes differences up to the level of speciation. The detection of tendencies at this level requires observation of phenotypic traits that can be argued to be adaptationally relevant or to have diagnostic value for species distinction. These observations have to be complemented by studies of genetic differentiation among the populations in order to corroborate the hypothesis that the adaptational differences have an evolutionary basis.

This article presents a new morphological distance that satisfies the need of a phenotypic classification parameter that can be compared to the genetic parameters which result from population genetic studies. Since these parameters compare frequencies of trait states irrespective of special trait characteristics, our approach will concentrate on the measurement of differences in frequency distributions of quantitative traits between populations. Therefore, other approaches, which focus on statistical moments such as expectations or variances, or which consider more general differences in trait states (see GREGORIUS et al., 2003 for a comprehensive discussion), will not be treated here.

<sup>1)</sup> IFFIVE-INTA. Camino a 60 Cuadras, km 5.5. 5119, Córdoba, Argentina. E-Mail: [arverga@yahoo.com.ar](mailto:arverga@yahoo.com.ar).

<sup>2)</sup> Institut für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen, Büsgenweg 2, 37077 Göttingen, Germany.

The newly developed measure of phenotypic distance is applied to the results of a study on a number of *Prosopis* populations, which are considered to belong to different species.

*The morphological distance  $d_m$*

The concept of a morphological distance will be developed in the following with reference to the idea underlying the genetic distance  $d_o$  as was introduced by GREGORIUS (1974). Since this distance compares genetic structures and is thus based on qualitative features, the question to be answered here is as to how the idea can be transferred to quantitative morphological characters.

According to GREGORIUS (1974) the genetic distance between two units of individuals (OTUs, "Operational Taxonomic Unit" (SNEATH and SOKAL, 1973)) X and Y is defined as:

$$d_o = \frac{1}{2} \sum_{i=1}^k |x_i - y_i|$$

Where  $x_i$  and  $y_i$  are the relative frequencies of the genetic type  $i$  in the OTUs X and Y, respectively,  $k$  is the number of genetic types, and

$$\sum_i x_i = \sum_i y_i = 1$$

Consistent application of this concept to distributions of quantitative characters yields a morphological distance between two OTUs defined as

$$d_m = \frac{1}{2} \int_{-\infty}^{\infty} |f(x) - g(x)| dx \tag{1}$$

where  $f(x)$  and  $g(x)$  represent the probability density functions for the character  $x$  in the two OTUs  $F$  and  $G$ . Since both  $d_m$  and  $d_o$  measure differences between frequency distributions and assume their maximum value of 1 exactly if two populations have no character in common, they are commensurate and thus directly comparable. Yet, while  $d_o$  can be easily estimated from sample distributions, this is more intricate with  $d_m$  because of the necessity to estimate probability density functions for quantitative characters. The main reason is that the number of observations represented in a sample and the measuring accuracy are always finite. This difficulty is usually circumvented by any of the following two methods:

a) A normal distribution of the character or some other model of character distribution is accepted a priori and fitted to the sample data. The thus obtained estimate of the probability density function is then characterized solely by its parameters (such as the expectation, variance or other moments).

b) The continuously distributed character values are assigned to a finite number of classes in order to obtain a discrete character that summarizes as much as possible of the originally continuous information. This character change requires establishment of an interval a priori, to define the classes.

Both methods have considerable disadvantages. In method a) an estimate of the population distribution is

obtained by accepting a particular model. Thus, estimates are model dependent (indirect) and are therefore meaningless if the model is incorrect. Model testing must therefore precede estimation, and a model dependent estimate is acceptable only if the model is not rejected (see e.g. GREGORIUS, 1998). In method b), even if the frequency distribution considered for the character covers all sample observations, the result of the comparison between discrete frequency distributions depends on the particular decomposition into classes.

To avoid these shortcomings, a morphological distance  $d_m$  is required that measures differences between continuous frequency distributions which neither rests on estimates of parameters from model distributions nor on arbitrarily definable discrete classifications of characters. It rather should be aimed at taking account of the totality of data on frequency distributions. In the first place this requires a transformation of the data that reflects their underlying density functions.

In order to arrive at such a transformation consider the cumulative sample distribution. Because the accuracy (resolution) of measurement is always limited and samples are finite, cumulative sample distributions consist of steps whose width cannot be smaller than the unit of measurement and whose height may be greater or equal than 1 (repeated occurrence of the same measurement). The cumulative sample distribution can be used to estimate the cumulative distribution function of the character in the population. For this purpose consider a polygon which connects half the height of two successive steps in the cumulative sample distribution (see top of *Figure 1*). Since polygons are piecewise differentiable and continuous functions, they are suitable for the estimation of probability density functions. Recall that the derivative of the cumulative distribution function  $F(x)$  yields the probability density function  $f(x)$ :

$$\frac{dF(x)}{dx} = f(x) \tag{2}$$

Returning to the sample distribution, note that a value  $x_i$  can appear  $n_i \gg 1$  times because of the limited accuracy of measurement. A sample thus consists of a finite number  $N$  of observations, which can be arranged in an ordered sequence of pairs

$$(x_i, n_i) \quad i = 1, \dots, k, \quad x_i < x_{(i+1)}, \quad \sum_i n_i = N,$$

where  $k$  is the number of the different values observed in the sample. Note that  $n_i = 1$  for  $i = 1, \dots, k$  if each measure is unique.

The infinitesimal terms  $d_x$  in (2) now become a sequence of disjoint character intervals of length  $x_{(i+1)} - x_i$  with  $i = 1, \dots, k-1$ . As a consequence of the polygonal approximation, the slope of the transformed cumulative sample distribution function is constant in each of these intervals and equals  $0.5(n_{(i+1)} + n_i) / (x_{(i+1)} - x_i)$ . In order to guarantee that the cumulative distribution is normalized to 1, the transformed frequencies  $0.5(n_{(i+1)} + n_i)$  must be divided by their sum. This sum can be written as  $N - (n_1 + n_k) / 2$ , so that we obtain for the slope in the  $i$ -th character interval:

$$\frac{1}{2} \frac{n_{i+1} + n_i}{x_{i+1} - x_i}$$

$$N - \frac{n_1 + n_k}{2}$$

This slope equals the probability density function  $f^*$  corresponding to the transformed cumulative sample distribution function, and we therefore arrive at the representation

$$f_{(x)}^* = \frac{1}{2} \frac{n_{i+1} + n_i}{x_{i+1} - x_i} \text{ for } x_i \leq x < x_{i+1}$$

$$N - \frac{n_1 + n_k}{2}$$

This frequency distribution can be represented graphically by plotting the  $x_i$  on the abscissa and the  $f_{(x)}^*$  on the ordinate such that each character interval is constantly assigned the  $f^*$  value corresponding to the interval (see middle of the left panel of Figure 1). In the thus resulting frequency histogram, the probability to obtain a value in the  $i$ -th character interval is given by the area of the rectangle (column) above this interval. The size of the total area below the  $f^*$  graph is equals 1.

Now consider the estimates  $f^*$  and  $g^*$  of the probability density functions from two samples taken from populations  $F$  and  $G$ . It then follows directly from the above graphical representation that the estimate  $d_m$  of the morphological distance between the two populations is given by half the size of the complement of the intersec-

tion of the areas below the graphs  $f^*$  and  $g^*$  (see bottom in the left and right panel of Figure 1).

Formally,  $d_m$  is obtained from considering the intersection between a character interval from the  $F$ -distribution and from the  $G$ -distribution, multiply the size of the intersection with the absolute difference between the  $f^*$  and  $g^*$  value in the intersection, sum these products over all combinations of character intervals of the two populations and divide the result by 2. Since the intersection of many of these pairs of intervals is empty, they do not contribute to the morphological distance.

It should be noted that quantitative traits that are decomposed into distinct classes correspond to qualitative traits, so that  $d_m$  directly reduces to  $d_0$  for such traits.

The morphological distance  $d_m$  can only take values that range between 0 and 1. This contrasts with traditional measures of taxonomic difference, which are based on assumptions on the form of the character distribution (like normality) and which use estimates of distribution parameters (like the expectation).  $d_m$ , on the other hand, requires no assumptions on underlying distribution characteristics and compares whole frequency distributions of populations rather than individual distribution parameters. As a result all individual data affect the estimation of differences between populations, and no information is lost due to data condensation.

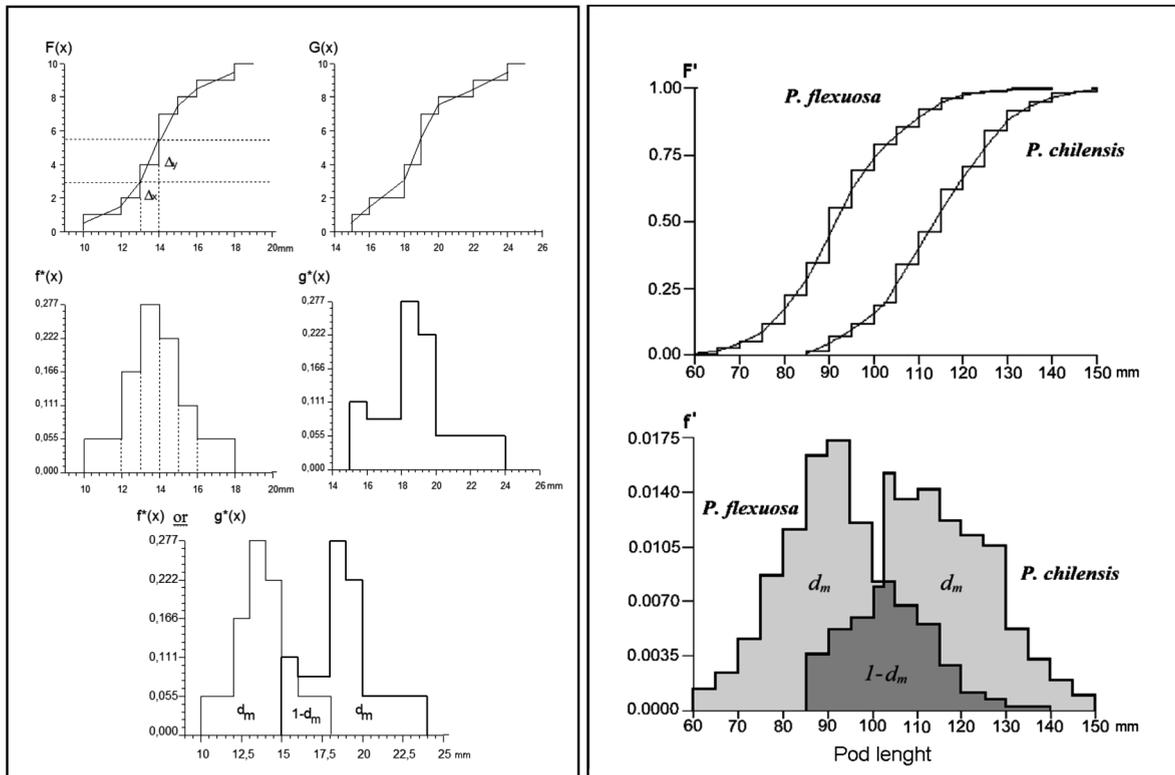


Figure 1. – Graphic representation of the computation of  $d_m$ . Above left: Cumulative distributions  $F(x)$  and  $G(x)$  for 10 measurements of length of foliolule in two hypothetical OTUs. In the left distribution the computation of the height in an interval is shown. The slope in that interval ( $\Delta y/\Delta x$ ) corresponds to the height of the estimated probability density function  $f(x_i)$  in the interval. In the center (left) of the figure, the standardized distribution functions  $f^*(x)$  y  $g^*(x)$  are represented. Below: Both distributions are represented in the same coordinate axis. The area of superposition between both distributions equals  $1-d_m$ . Right: Example of  $d_m$  computation by Pod length between *P. chilensis* and *P. flexuosa*.

There is also a more basic conceptual difference between the two methods: the distance  $d_m$  assumes its maximum value of 1 only when there is no overlapping between the distributions. This is a situation that cannot be detected by comparisons on the basis of the average values (expectations) of the distributions. This fact reflects the empirical concept that populations which share no characteristics cannot differ to various degrees.

In the same way,  $d_m$  can take the value 0 only when there is total equality between the distributions that is when each trait state appears in each population at the same frequency. This is another main difference from the traditional taxonomic methods, in which the distance is 0 when the population means are equal.

Two species, which are classified as taxonomically different may show morphological differences smaller than 1, so that both species have morphological characters in common. This could support the idea that the two species belong to a higher level of taxonomical organization.

*Comparison between taxonomic, morphologic and genetic distances. An application to the distinction of Prosopis species*

The following demonstrations are based on unpublished data from a study of Aníbal Verga, Jacqueline Joseau, Carlos Carranza and Marcela Ledesma. The study is based on observations on 81 trees from 3 allopatric and 3 sympatric populations of *Prosopis chilensis* and *P. flexuosa* in the arid Chaco of southern Argentina.

In the sympatric populations, where both taxonomic species exist in close neighborhood, only trees were sampled that are phenotypically classified as hybrid or intermediate according to traditional taxonomic standards (denoted as "I" trees in the present paper). Additionally, trees were sampled, which are unambiguously classified as *Prosopis chilensis* according to traditional taxonomic standards, and which were scattered among *P. flexuosa* trees (denoted as "H" trees). In the allopatric populations, individuals of both species ("F" for *P. flexuosa* and "C" for *P. chilensis*) were sampled at random.

A herbarium sample was taken from each tree for numerical taxonomic analyses. A total of 14 quantitative and qualitative characters were scored for 10 leaves and 5 fruits from each tree.

In order to apply one of the most widely used taxonomic distances, the averages were obtained for each character within each tree (individual trees are considered as OTUs in the following). Numerical data analysis was performed with the help of the NTSYS-software (Applied Biostatistics, Inc. 1990). The individual measurements were standardized with the help of the STAND module of NTSYS in the usual way by

$$y' = \frac{(y - \bar{y})}{S_{(y)}}$$

where  $y$  denotes the (average) character value of an OTU,  $\bar{y}$  denotes the average over all OTUs, and  $S_{(y)}$  denotes the standard deviation. The resulting quadratic

matrix of distances between OTUs was computed with the help of the SIMINT module of NTSYS, where the distances were computed as the "average taxonomic distance" (DST)

$$d_{ij} = \sqrt{\frac{1}{n} \sum_{k=1}^n (y'_{ki} - y'_{kj})^2}$$

In this expression  $d_{ij}$  is the distance between OTU  $i$  and OTU  $j$ ,  $n$  is the number of characters,  $y'_{ki}$  and  $y'_{kj}$  are the standardized values of the  $k$ -th character of OTU  $i$  and OTU  $j$ , respectively. UPGMA (unweighted pair-group method of arithmetic averages) served as the clustering method.

Figure 2, shows two dendrograms obtained from the matrix of taxonomic distances and the matrix of morphologic distances between trees. In this context, each tree is considered as representing a distribution of leaf and fruit characteristics, so that morphological distances can be obtained. These distance matrices have been calculated from ten measurements on 14 qualitative and quantitative characters of leaf and fruit for each tree classified.

The dendrogram obtained for the taxonomic distances shows 4 distinct classes, whereas in the dendrogram obtained for morphological distances three classes can be clearly distinguished.

In order to detect possible genetic differences between the morphological classes, allele frequencies were determined for the enzyme locus ADH-A in each of the classes. This locus was earlier shown by VERGA (1995) to be diagnostic for the differentiation between both species in the Argentinean arid Chaco.

In order to analyze the genetic significance of both classification criteria, taxonomic and morphologic, we performed Mantel tests for correlation between morphological and genetic distances in the morphological classification as well as tests for correlation between taxonomical and genetic distances in the taxonomic classification (Table 1). The correlation in the morphologic classification was distinctly stronger than in the taxonomic classification, and both correlations were significant.

The two species actually form a species complex in the sense that they hybridize and show various degrees of introgression including asymmetric introgression (VERGA, 1995). The above results clearly distinguish between the four parts of this complex when adult individuals are considered. To see whether these results are consistent across developmental stages, the morphological distance approach was also applied to offspring from the four state classes I, H, F and C. Within class H only those offspring were considered, which were heterozygous for the species diagnostic ADH-A locus. In this way it was possible to identify the hybrids. The offspring were grown in a nursery.

In Figure 3 the morphological distances ( $d_m$ ) between the phenotypic classes C, F, H and I as well as the pertinent UPGMA dendrogram are given. It is clear that the smallest distance is realized between the classes H and I even though the seed of class H come from taxonomically well classified *P. chilensis* trees. Recall that these

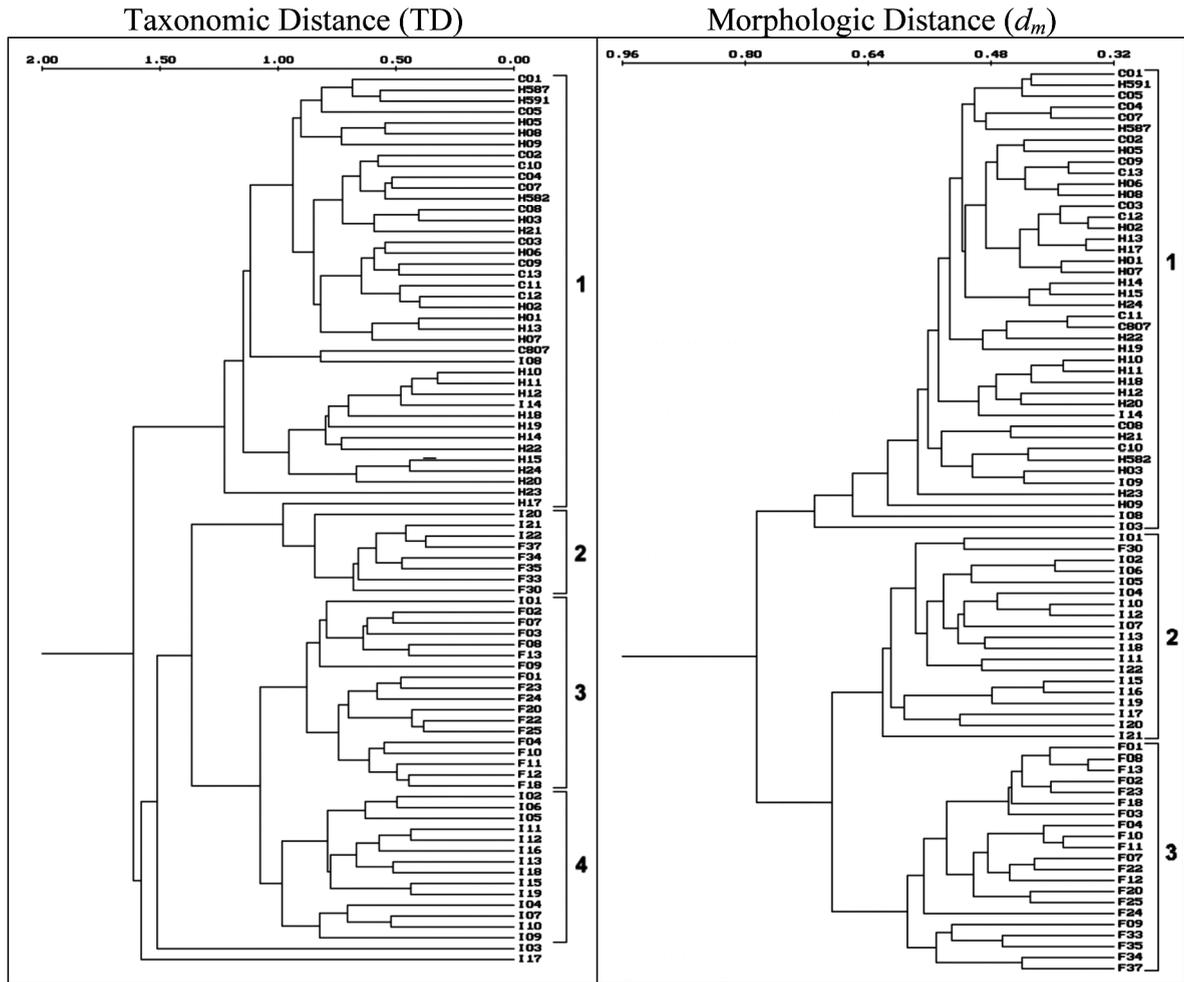


Figure 2. – UPGMA dendrograms obtained from the matrix of taxonomic distance and the matrix of morphologic distance.

Table 1. – Allele frequencies for the enzyme locus ADH-A in each of the classes and correlation coefficients ( $r^2$ ) between Morphologic Distance ( $d_m$ ) and Genetic Distance ( $d_o$ ) matrixes by morphological classification, and between Taxonomic Distance and Genetic Distance ( $d_o$ ) matrixes by taxonomic classification.

Allele	Allele frequencies in Morphological classes			Allele frequencies in Taxonomic classes			
	1	2	3	1	2	3	4
ADH-A2	0.038	0.447	0.857	0.038	0.611	0.853	0.464
ADH-A3	0.963	0.553	0.143	0.963	0.389	0.147	0.536

Allelic frequency of each group identified from the dendrogram

	Morphologic Distance $d_m$ between Morphological classes			Genetic Distance $d_o$ between Morphological classes			
	1	2	3	1	2	3	
1	0.000			1	0.000		
2	0.519	0.000		2	0.410	0.000	
3	0.711	0.473	0.000	3	0.820	0.410	0.000

Correlation coefficient  $r^2 = 0.967$

	Taxonomical Distance between Taxonomic classes				Genetic Distance $d_o$ between Taxonomic classes				
	1	2	3	4	1	2	3	4	
1	0.0000				1	0.0000			
2	0.1252	0.0000			2	0.5736	0.0000		
3	0.1564	0.1013	0.0000		3	0.8154	0.2418	0.0000	
4	0.0820	0.1095	0.1126	0.0000	4	0.4268	0.1468	0.3887	0.0000

Correlation coefficient  $r^2 = 0.529$ .

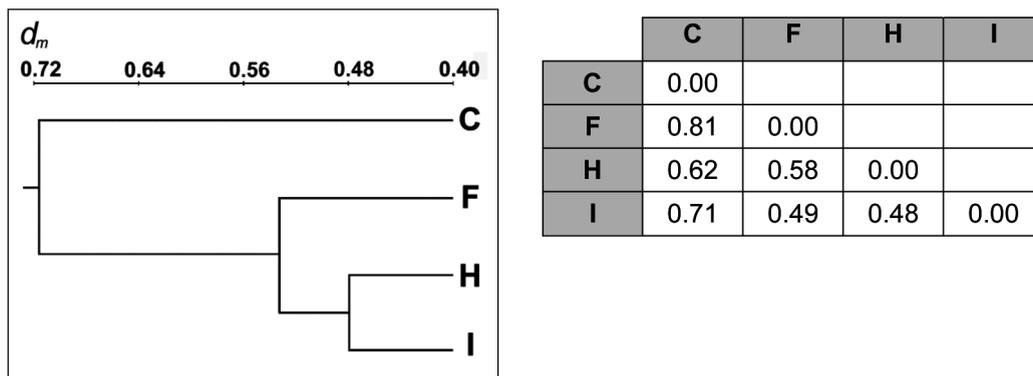


Figure 3. – UPGMA dendrogram and matrix of morphological distances between the offspring from the four state classes I, H, F and C.

seed trees are isolated from other *P. chilensis* and *P. flexuosa* trees as well as from the hybrid stand. Also recall that class H offspring result only from interspecific matings as was confirmed with the help of ADH-A analysis. Hence, class H seed trees are *P. chilensis* and class H offspring are “putative” hybrids. Moreover, the morphological distance between classes H and F ( $d_m(H,F)=0.58$ ) is smaller than between classes H and C ( $d_m(H,C)=0.62$ ). Thus the ranking of the differences of F and C from H is the same as from I ( $d_m(I,F)=0.49$  and  $d_m(I,C)=0.71$ ).

It thus turns out that the morphological distance approach yields developmentally stable classifications.

### Discussion and Conclusion

In the process of classification and differentiation it is essential to define the groups, at least throughout qualitative characters, in order to decide if an individual does or does not belong to a special group. This process of differentiation is in essence a qualitative process. The contradiction appears when we have, besides, quantitative characters, such as group of morphological characters of the units that will compose the groups. The main difference between the traditional taxonomic distance and  $d_m$  is that the first one is in essence a quantitative methodology, as long as  $d_m$  is based on qualitative criteria.

The differences here are not defined by means of distances between the average (mean) expressions of the metric character, but by means of the comparison of the frequency distributions of the characters within each population. This allows to distinguish populations up to the possibility of complete disjunction (separation). The idea of complete phenotypic separation of populations has no meaning in methods that rely on statistical moments of random variables.

The morphological method takes account of each individual data collected for the comparison between populations. In this way, not even one piece of information is missed during the data analysis. In contrast, the taxonomic distance methodology the differentiation is reached throughout the comparison of the average values of each character. This does not provide any information about elements in common between the OTUs that are compared and, all the information generated

from the intra-OTU variability is missed. Even when there is a possibility to include, as a new character, a dispersion measure for each character in order to conserve these data, both, mean and variance, assume, before hand, a normal distribution of the characters that is not always true in the reality.

Both qualities of the  $d_m$  have allowed the differentiation between intermediate groups, allowing the identification of interspecific hybrids that were impossible to separate by means of taxonomic distance.

In the differentiation method there are two initial situations that can be identified when the groups that will be compared are formed:

- *A priori* groups of individuals can be formed according to a criterion that can be: geographic location, age classes, degree of resistance to diseases, etc. Then, these groups can be compared with other characters.
- Groups can be formed according to a set of characters of the individuals, without prioritize any of them.

In the first case, the formation of the groups is made according to a characteristic set *a priori*. It implies that the chosen characteristic is already considered an essential element for the classification. Once the groups have been formed this way, they can be studied using other set of characters. The finding of differences between the formed sets compared with the new characters can or can not justify the original formation of the groups, but any way, the groups formed this way will always show subjective slant, since its formation depended on which character they were grouped. We must not forget the objective that these studies follow, which is the identification of the groups that behave as truly biological units. These units will be used then as units for management, conservation and improvement. Only when the units have been identified (delimited), it is possible to work on them and to know the interaction processes with other units. The election, *a priori*, of a character we suppose will define the groups, implies the knowledge of the groups that we want to identify indeed. This contradiction appears when we apply the first concept to the group formation mentioned below.

We can say then that for the formation of true groups, in the sense that they have a biological significance and they have not simply capricious groupings, we must begin from the individual, because our work consists in

grouping types of individuals. The classification will not be on the basis of predetermined characters (whether these are genetic, or morphologic, or physiological, or geographic, or ecological, etc); on the other hand, the characters capable to differentiate the individuals which are the object of the study, must come from the same process of classification.

### Acknowledgements

The authors appreciate the comments of two reviewers, which helped considerably in improving the presentation.

### Literature

- BUCK, G. and J. BIDLACK (1998): Identification of *Quercus* and *Celtis* species using morphological and electrophoretic data. *Proc. Okla. Acad. Sci.* 1998; **78**: 23–33.
- CHIA-SZU WEN and JU-YING HSIAO (1999): Genetic differentiation of *Lilium longiflorum* Thunb. var. *scabrum* Masam. (*Liliaceae*) in Taiwan using Random Amplified Polymorphic DNA and morphological characters. *Bot. Bull. Acad. Sin.* (1999).
- CHIN-SUNG CHANG and HUI KIM (2003): Analysis of morphological variation of the *Acer tshonoskii* complex in eastern Asia: implications of inflorescence size and number of flowers within sect. *Macrantha*. *The Linnean Society of London, Botanical Journal of the Linnean Society*, 2003, **143**, 29–42.

- GREGORIUS, H.-R. (1974): Genetischer Abstand zwischen Populationen. *Silvae Genetica* **23**: 22–27.
- GREGORIUS, H.-R. (1998): The system analytical approach to the study of hypotheses. URL <http://www.uni-forst.gwdg.de/forst/fg/index.htm>
- GREGORIUS, H.-R., E. M. GILLET and M. ZIEHE (2003): Measuring differences of trait distributions between populations. *Biometrical Journal* **45**(8): 959–973.
- JONES, R., D. STEANE, B. POTTS and R. VAILLANCOURT (2002): Microsatellite and morphological analysis of *Eucalyptus globulus* populations. *Can. J. For. Res.* **32**: 59–66.
- KLEINSCMIT, J., R. BACILIERI, A. KREMER and A. ROLOFF (1995): Comparison of morphological and genetic traits of pedunculata oak (*Quercus robur* L.) and sessile oak (*Q. petraea* Matt, Liebl.). *Silvae Genetica*. 1995, **44**: 5–6, 256–269.
- PETIT, R., C. BODENES, A. DUCOUSSO, G. ROUSSEL and A. KREMER (2003): Hybridization as a mechanism of invasion in oaks. *New Phytologist*. **161**: 151–164.
- SNEATH, P. H. A. and R. R. SOKAL (1973): Numerical Taxonomy. *The Principles and Practices of Numerical Classification*, Freeman, San Francisco, Ca., XV, 573 pp.
- VERGA, A. R. (1995): Genetische Untersuchungen an *Prosopis chilensis* und *P. flexuosa* (*Mimosaceae*) im trockenen Chaco Argentinien. *Göttingen Research Notes in Forest Genetics* 19. Abteilung für Forstgenetik und Forstpflanzenzüchtung der Universität Göttingen.

## The Genetic Structure of *Sorocea bonplandii* in Southern Brazilian Forest Fragments: AFLP Diversity

By A. R. RUSCHEL<sup>1),2),\*</sup>, R. O. NODARI<sup>2)</sup> and B. M. MOERSCHBACHER<sup>1)</sup>

(Received 28<sup>th</sup> July 2005)

### Abstract

To analyse possible genetic erosion due to fragmentation in the Brazilian Atlantic Forest ecosystem, we investigated the genetic diversity within and among one large and six small populations of one of the key species of this ecosystem, *Sorocea bonplandii*, using AFLP analyses of 468 plants. Eight primer pairs yielded 299 polymorphic fragments for analysis. *S. bonplandii* was characterized by an unusually high genetic diversity within the species and also within individual populations, around 94% of the total genetic diversity occurred within populations. Genetic distances between popula-

tions were low in spite of extensive fragmentation. Genetic distance was significantly correlated with geographical distances between fragments, but these differences may have existed before fragmentation. Our results have direct implications for sustainable management of *S. bonplandii*, indicating that conservation strategies might be based on a random sample of trees taken throughout the Atlantic forest. However, the minimum population size required for maintaining the huge genetic diversity of this species is unknown. In order to establish a sustainable management plan for the species, further ecological studies are needed.

*Key words:* dioecious plant, forest fragmentation, genetic diversity, Moraceae, Subtropical Atlantic Forest.

### Introduction

The Atlantic forest, classified fifth among the hotspots of global biodiversity containing endemic species (MEYER

<sup>1)</sup> Department of Plant Biochemistry and Biotechnology, University of Münster, Hindenburgplatz 55, 48143 Münster, Germany.

<sup>2)</sup> Departamento de Fitotecnia, Universidade Federal de Santa Catarina, C. P. 476, Florianópolis CEP 88040-900 SC, Brazil.

\*) Corresponding author: ADEMIR R. RUSCHEL, Phone: +55-9132041234, Fax: +55-32772888. E-mail: [arruschel@yahoo.com.br](mailto:arruschel@yahoo.com.br).