Rev. Brasil. Genet. 13, 4, 855-860 (1990)
(Brazil. J. Genetics)

METHODOLOGY

SIMILARITY COEFFICIENT OF JACCARD APPLIED TO STUDIES ON GENE REGULATION

Marion Schiengold and Helga Winge

ABSTRACT

An original form of utilization of the Similarity Coefficient of Jaccard is proposed using electrophoretic data obtained from different tissues and different developmental stages of the cleistogamous plant *Relbunium hypocarpium*. This approach allows us to estimate the effect of gene regulation on the variability observed among strains, that is, the degree in which the differences observed are due to differences in ontogenetic regulation (temporal regulation) and tissue differentiation (spatial regulation). This possibility of quantifying the effect of regulatory genes on the variability could also be applied to evolutionary studies.

INTRODUCTION

According to Korochkin (1981), the study of the modifications of protein and especially of isoenzyme patterns during ontogenesis and tissue differentiation is essential for research on gene activity and regulation in development.

The technique of zone electrophoresis in gel, followed by histochemical staining procedures, is very versatile and easily applied to studies on the genetic variability of organisms. The electrophoretic analysis of isoenzymes, besides allowing the study of intra- and interspecific genetic similarity, also permits us to evaluate the action of regulatory genes on tissue differentiation and on ontogenesis.

Departamento de Genética, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501 Porto Alegre, RS, Brasil.

The material studied in this work was the cleistogamous plant *Relbunium hypocarpium*, (Rubiaceae) an herbacious and perennial species widely distributed in the Neotropical region. As shown by Cavalli-Molina *et al.* (1989), *R. hypocarpium* is a species whose sibs always show identical isoenzymatic patterns when the same tissue at the same stage of maturation and the same ontogenetic stage is analyzed, which is evidence of an absence of structural and regulatory gene segregation. All results obtained up to this moment indicate that the species is autogamous, with consequent homozygosis.

The objective of this paper is to present an original way of applying the Similarity Coefficient of Jaccard to estimate the effect of gene regulation on the variability observed within and between populational samples, i.e., the degree in which the observed variation among individuals is due to differences in ontogenetic regulation and tissue differentiation.

METHODOLOGY AND DISCUSSION

The methodology described bellow was developed for the study of 25 strains of *R. hypocarpium*, five strains of each of five populations from different physiographic regions (Schiengold, 1985).

Isoesterase patterns of different tissues at two chosen stages of seedling development were analyzed, first stage (seedlings with expanded cotyledonal leaves) and second stage (seedlings with first leaves), as well as ripe and unripe fruits of adult plants. All tissues were analyzed separately at each stage studied.

The similarity among strains of each population and among the populations, at each developmental stage was estimated by the Similarity Coefficient of Jaccard (see Sneath and Sokal, 1973). The coefficients were always calculated, comparing the strains two by two, the numerator being the number of electromorphs common to the two strains, and the denominator the total number of electromorphs of these two strains. This coefficient (S_J) measures the similarity between strains, based on their active structural genes (electromorphs).

In order to estimate the effect of gene regulation during the ontogenetic development of the strains, we also calculated another Similarity Coefficient of Jaccard (S_J^*) adjusted to the observed regulatory differences between the same strains. In S_J^*, those electromorphs that were not detected in one of the two strains at the selected stage, but were active in any other developmental stage of that strain, were added to the numerator. In fact, both strains shared those electromorphs (and the alleles that control them) although they were detected in different ontogenetic stages.

Table I - Isoesterases of strains of a *Relbunium hypocarpium* population from Restinga, County of Porto Alegre, RS (only the most anodic bands are presented), and calculation of the Similarity Coefficient of Jaccard (S_J and S_J*) for seedlings at the 1st stage. (1st = first stage; 2nd = second stage; FR = fruits; POP = population; Rm = relative migration; Electromorphs: + = present; - = absent or not detected).

Esterase	Strains from Restinga - RS															
	HW-811B			HW-1468			HW-1470			HW-1471			HW-1472			
Rm	1st	2nd	FR	1st	2nd	FR	1st	2nd	FR	1st	2nd	FR	1st	2nd	FR	POP
1.10	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	+
1.06	-	-	-	-	-	-	-	-	-	+	-	+	+	-	+	+
1.03	+	-	+	+	+	-	+	-	+	+	+	+	+	-	+	+
1.00	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.96	-	-	+	+	-	+	-	-	-	+	-	+	+	-	+	+
0.93	+	-	+	-	-	+	-	-	+	-	-	-	+	-	-	+
0.91	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	+
0.86	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	+
0.84	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+
.																
.																
etc.																

$$S_J = \frac{\text{No. of bands common to the 2 strains (at 1st stage)}}{\text{Total No. of bands present in the 2 strains (at 1st stage)}}$$

$$S_J^* = \frac{\text{No. of common bands (1st stage) + No. of different bands (1st stage) due to temporal regulation}}{\text{Total No. of bands present in the 2 strains (at 1st stage)}}$$

As an example, Table I shows part of the results obtained for strains from Restinga, County of Porto Alegre, RS, Brazil.

As can be seen, the strains HW 1470 and HW 1471 at the first stage differ by the presence of bands 1.06, 0.96, 0.91 and 0.86; therefore, in the calculation of S_J (similarity based on structural genes), these bands are not included in the numerator but only in the denominator. However, as strain HW 1470 does present bands 0.91 and 0.86 in fruits, these bands are included in the numerator when S_J^* is calculated, since their absence at the first stage is simply due to temporal regulatory differences in relation to the strain HW 1471. Thus, considering all bands presented in Table I for the first stage, we verify that $S_J = 2/6$ but that $S_J^* = 4/6$, which means that the similarity between these two strains at the first stage passes from 33% to 66% when the ontogenetic regulatory differences are excluded.

To estimate the proportion of differences detected among the strains due to temporal regulation, we used the ratio $S_J^* - S_J / 1 - S_J$. the numerator estimates the fraction of the mean genetic similarity among the strains that is not detected when the similarity coefficient is based only on the bands present at one same stage. This fraction thus represents the dissimilarity among the strains at that stage due to ontogenetic regulation. The denominator, which is the complement of the total mean similarity, estimates the mean dissimilarity among the strains at that same stage.

The Coefficient of Jaccard adjusted to regulatory differences (S_J^*) can also be utilized to estimate the effect of gene regulation on tissue differentiation. In this case, the enzymatic patterns of different tissues at a single stage are taken into consideration.

Relbunium hypocarpium presents singular characteristics such as high variability in tissue differentiation among strains. The methodology of analysis described here made it possible to evaluate the intra- and interpopulational variability due to that kind of regulation (Schiengold, 1985). The strains were compared two by two, the number of common bands of a certain tissue being divided by the total number of bands that occur in the tissue in the two strains. Afterwards, the mean similarity of all tissues at that same stage was calculated. Exemplifying, if the mean similarity per tissue, between strains, at a determined stage is $S_J = 0.50$, and the mean similarity between these strains at that same stage, independently of the tissues, is $S_J = 0.60$, we can estimate the proportion of the similarity that would not be detected if only one tissue were analyzed. In this case, the average is 10% (0.60 - 0.50 = 0.10).

On the other hand, if we have the value of $\overline{S}_J^*$ for a certain stage (let us say, $\overline{S}_J^* = 0.80$), we can evaluate in the same way the genetic similarity between the strains that would be concealed due to the effect of temporal regulation. In the analysis per stage, averaging all tissues, this value would be 20% (0.80 - 0.60 = 0.20). We can also estimate the total effect of gene regulation, temporal plus spatial, on the observed isoenzymatic variability among strains; in our example it would be 30% (0.80 - 0.50 = 0.30). In this case, the spatial regulation would contribute with one third of the regulatory differences observed at this stage, and the temporal regulation with twice that amount.

Some final considerations can be made with respect to the methodology that is being proposed here.

It is expected that, by increasing the number of analyzed stages, we will be able to detect a greater proportion of regulatory differences between the strains.

The methodology described here for a new use of the Coefficient of Jaccard may also be utilized for evolutive studies, allowing the estimation of the degree of genotypic reformulation due to structural and to regulatory genes that occur during the speciation process.

The genetic identity among sib plants of *R. hypocarpium* made possible the accomplishment of the above described analyses. However, we believe that this methodology can eventually be applied to the study of allogamous plants. In this case, a previous knowledge of the genetic control of the electromorphs (loci/alleles) becomes indispensable, since heterozygosity does occur.

We also have to consider the possibility of seggregation of temporal regulatory genes in each progeny. For that reason, the number of individuals to be analyzed, of each progeny, at each ontogenetic stage, must be much larger than that analyzed in the case of autogamous plants. The utilization of clones of plants, obtained by *in vitro* culture of meristems, besides allowing the study of the same genotype along ontogenetic development, would permit us to obtain reliable data about tissue differentiation. Depending on the material, another possible alternative would be the electrophoretic analysis of tissues without endangering the seedling survival and development. As an example, we suggest the analysis of only one cotyledon, of leaves, or of lateral branches.

It is expected that these analyses will demonstrate that the major part of the intrapopulational variability in allogamous plants is due to polymorphism in structural genes, and that the variability in regulatory genes is low.

ACKNOWLEDGMENTS

We are grateful to Dr. Suzana Cavalli-Molina, Dept. Genetics - UFRGS, Porto Alegre, for reading the manuscript and for suggestions, to Prof. Dr. Antonio Rodrigues Cordeiro, Dept. Genetics - UFRJ, Rio de Janeiro, for encouragement during our work, and to Prof. Ilsa Simoni for the English translation.

The research is part of a project supported by PIG-CNPq/FINEP, CAPES, CNPq, FAPERGS.

RESUMO

A partir de dados eletroforéticos obtidos para a planta cleistógama *Relbunium hypocarpium*, em diferentes órgãos, em três estágios do desenvolvimento ontogenético, foi criada uma forma original de utilização do Coeficiente de Similaridade de Jaccard. Essa abordagem permite estimar o efeito da regulação gênica na variabilidade observada entre linhagens, ou seja, o grau em que as diferenças observadas são devidas a diferenças na regulação ontogenética (temporal) e diferenciação tissular (regulação espacial). Esta possibilidade de quantificação do efeito dos genes reguladores na variabilidade apresenta também utilidade em estudos evolutivos.

REFERENCES

Cavalli-Molina, S., Motta, V.E.P., Schiengold, M. and Winge, H. (1989). Identical isoenzyme patterns in sib plants of *Relbunium hypocarpium* (Rubiaceae). *Rev. Brasil. Genet. 12*: 361-368.

Korochkin, L.I. (1981). *Gene Interactions in Development.* Springer-Verlag, Berlin.

Schiengold, M. (1985). Efeito da regulação gênica na variabilidade isoesterásica de *Relbunium hypocarpium* (Rubiaceae). Master Thesis. Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

Sneath, P.H.A. and R.R. Sokal (1973). *Numerical Taxonomy.* W.H. Freeman, San Francisco, 573 pp.

(Received December 15, 1989)