The Concept and Measurement of Phenotypic Diversity in *Puccinia graminis* on Wheat

J. V. Groth and A. P. Roelfs

Professor, Department of Plant Pathology, and research plant pathologist and professor, Cereal Rust Laboratory, USDA, University of Minnesota, St. Paul 55108.

Paper 15,033, Scientific Journal Series, Minnesota Agricultural Experiment Station, St. Paul 55108. Supported in part by Competitive Grant USDA 58-5759-6-1.

Accepted for publication 31 March 1987 (submitted for electronic processing).

ABSTRACT

Groth, J. V., and Roelfs, A. P. 1987. The concept and measurement of phenotypic diversity in *Puccinia graminis* on wheat. Phytopathology 77:1395-1399.

Four different diversity indexes were applied to race survey samples of wheat stem rust (*Puccinia graminis* f. sp. *tritici*) made annually during 1918–1982 rate survey samples of wheat and screened on the Stakman differential wheat line set for the years 1918–1982. Several kinds of diversity were measured by these indexes, including phenotype number, evenness of frequency, and temporal, summed change in frequency of each phenotype. The trend for all four indexes was downward through the years. Multiple regression analysis established that the contributions of two aspects of diversity, phenotype number and evenness, and a third variable, sample size, to the Gleason, Shannon, and Simpson indexes were different, so that, depending on the nature of the population being sampled, one or other of the diversity indexes should prove to be more appropriate. The ease of computation of these indexes and the complementary nature of the Rogers index to the other three suggest that more than one of them can be applied to more fully describe diversity of a plant pathogen.

The concept of phenotypic or genetic diversity is central to the field of ecological genetics (3). Diversity, however, has several meanings, resulting in imprecision and confusion in communication. Plant pathogenic microorganisms are often categorized below the species level into "physiological races," which are phenotypes that differ from one another solely on the basis of their visible disease reactions on a set of host genotypes called the differential set. In plant pathology, when it is said that a pathogenic organism is phenotypically diverse, one or more of at least four distinct meanings can be conveyed. These shall be referred to by number in the following presentation. Meaning 1) is a relatively large number of phenotypes (races) for a given number of isolates; 2) is an even distribution of phenotypes, the low diversity alternative being the case where a small number of phenotypes dominates the population (referred to as "dominance"), all others being rare; 3) is relatively large numbers of differences in virulence or other genetic attributes between phenotypes; and 4) is a relatively high rate of temporal or spatial change in some or all of the first three meanings. Spatial change is usually best considered as an aspect of the scale of sampling of a population, e.g., regional diversity might be expected to be greater than local diversity. As will be illustrated by the diversity measures discussed later, these concepts can be understood and used separately, but they often occur in combination.

Plant pathologists have expressed meaning 1) in its simplest form, as a ratio of phenotypes (races) to isolates sampled (1,4). This measure is extremely sensitive to sample size, since the rarer the phenotype, the larger the sample that is needed to detect it once. An improvement on the measure is provided in the Gleason index (14),

$$H_g = (r - 1)/\ln(N),$$

where r = the number of distinct phenotypes detected and N = the number of isolates in the sample. This index is less sensitive to sample size than the ratio r/N because as r increases more slowly with increasing sample size (because the more common phenotypes have been found), the increase in sample size is correspondingly diminished in its logarithmic form. The Gleason index is simple in that it directly includes only meaning 1) of diversity and in that it is easy to calculate.

Two indexes of diversity widely used in ecology to characterize species diversity (8,9,14,16) have been advocated in plant pathology (5,6) to characterize intraspecific diversity of plant pathogens. One is the Shannon index (variously referred to as the Shannon-Weaver or Shannon-Wiener index),

$$H_w = -\sum_i p_i \ln(p_i),$$

where p_i = the frequency of the ith phenotype. A correction factor is desirable because the Shannon index as calculated above is a biased estimate of the actual diversity because frequencies are only estimated from samples (10, p. 392). For samples as large as most of those included in this study, the bias is generally of the order of +0.01. We have thus chosen to use the uncorrected Shannon index. If samples are small, say less than 100, the correction should be made. The complement of Simpson's index of concentration, which we shall refer to as the Simpson index of diversity, is the other popular diversity index:

$$H_s = 1 - \sum_i \left[n_i(n_i - 1)/N(N - 1) \right],$$

where n_i = the number of isolates of the ith phenotype and N = the sample size. Sometimes the reciprocal of Simpson's index is used.

Both the Shannon and Simpson indexes, which are similarly derived (8, p. 9), directly include meanings 1), number of phenotypes, and 2), evenness of distribution, and therefore are somewhat more complex than is the Gleason index.

Meaning 3), phenotypic or genetic distance, is brought into diversity indexes in various ways. Some indexes primarily measure genetic distance between populations, requiring that identity of alleles be known (7). A more widely applicable approach that can incorporate a degree of phenotypic distance into the Shannon index is the index of hierarchical diversity (8, pp. 17-18).

$$H'_w = \Sigma_{i=1}^{g} \Sigma_{j=1}^{h} q_i p_{ij} \ln(q_i p_{ij}),$$

where g = the number of arbitrarily or previously determined groupings of phenotypes, h = the number of phenotypes in the ith grouping, q_i = the proportion of all isolates that belong to the ith group, and p_{ij} = the proportion of isolates of the ith group that are of the jth phenotype.

This index weights the Shannon index higher according to the amount or proportion of intergroup, as opposed to intragroup separation, and, as such, it accounts for phenotype distance in a discrete and limited way. It can be extended to include more levels of hierarchy. An example of groupings is the race clusters of stem rust of wheat (12). In general, there has been no consensus among plant pathologists in how to account for meaning 3) in describing populations. Genetic distance measures (7) are more appropriate for genetic, as opposed to phenotypic, analyses.

One way of measuring meaning 4), spatial or temporal change in diversity, is by using the index of proportional overlap (14), sometimes referred to as the Rogers index (2,13):

$$H_r = 0.5 \Sigma_{i=1}^{m} {}_{i=1}^{m} \mid p_{i1} - p_{i2} \mid ,$$

where m = the total number of phenotypes in both populations, p_{i1} = the frequency of the ith phenotype in the first population and p_{i2} = the frequency of the ith phenotype in the second population.

Like the Simpson index, this index has limits of $0 <= H_r <= 1$. For all phenotypes present in at least one of the two populations being compared, it specifically accounts for the amount of change in frequency between them. Because they are calculated on single populations, none of the other indexes above directly includes this important aspect of phenotypic diversity.

Lebeda (6) has reviewed some of the above as well as other useful indexes of diversity as applied either to race numbers and frequency or to frequencies of selected individual virulence in lettuce downy mildew. The present applications correspond to one used by Lebeda where overall virulence phenotype of each isolate is used to define the isolate as distinct. In this way, our application is similar to both the ecologists' interspecific use of the indexes and to the plant pathologists' definition and use of physiological race. Depending on the circumstances, both uses have value. For analysis of historical data, the application to individual virulence frequency is not usually possible because single-gene differential host lines were not used.

The objectives of this study were to apply the above nonhierarchical diversity indexes to annual data from 1918 to 1982 U. S. collections of *Puccinia graminis* Pers. f. sp. *tritici* Eriks. & Henn. and to examine the changes in indexes to assess trends in diversity with changes in agronomic practice and other influences on stem rust populations. Direct comparisons of the influences of three components on three of the indexes were also made using multiple regression on these data with the view that the analysis might help others in choosing which index to use for their specific case.

MATERIALS AND METHODS

The USDA Cereal Rust Laboratory has conducted an annual race (virulence phenotype) survey of the population of *P. g.* f. sp. *tritici* in the United States since 1918. The data have been published almost annually since 1939, and data from earlier years were retained at the Laboratory. Collection procedures have varied, but they have been relatively standard for the past 30 yr (11). Collections of diseased host material were sent to the laboratory where uredospores were removed and used to inoculate a susceptible host (originally cultivar Little Club, but more recently McNair 701). Uredospores from these plants were then used to inoculate a series of 12 different host cultivars chosen by Stakman and Piemeisel in 1917 (15). These cultivars are now known to possess a number of different genes for resistance (11), some of which have been used widely in commercial bread and durum wheats over the years. Other genes have probably never occurred in commercial wheats. The differential cultivars contain various numbers of resistance genes, i.e., they are not single-gene differential lines.

Several BASIC programs were developed on the Apple II+ computer to facilitate the calculation of the many indexes of diversity, using actual numbers of isolates and races. Stepwise multiple regression analysis was performed using Apple Computer software, with the index for each year as the dependent variable and the three components that affect diversity as the independent variables. The three components are race number, evenness of frequencies, and sample size. While it is not a normally distributed variable, we chose a simple standard deviation of frequencies of races as an easily calculated measure of evenness that could be applied to all three indexes equally. The Shannon index is the only one of the three, to our knowledge, for which a standard deviation can be calculated directly (10).

RESULTS

Table 1 presents some of the data used in calculating the diversity indexes and used in regression analyses.

Annual values of the Gleason index are plotted in Figure 1 and those of similar Shannon and Simpson indexes in Figure 2. For all three, trends are downward based on visual inspection from 1918 to the present, but not to the same degree or equally uniformly for all three. In particular, the Shannon index decreases much less than the other two. The Gleason and Shannon indexes do not appear to decrease on average during the approximate period 1944–1958, mainly because the number of races detected was somewhat higher during that period than it was just before or after (Table 1). Also contributing, however, was the slightly greater evenness especially during 1954–1959. The extreme high value of the Gleason index in 1952 is due to the fact that 38 races were detected that year, about 10–20 more than were normal for that period.

The Rogers index is plotted in Figure 3, where, except for 1918, the comparison is always with the previous year. Again, the trend is downward, indicating reduced temporal variability in race identity and frequency in recent years. The extreme high value for the index in 1934 represents the rise of race 56 during the first year of the major stem rust epidemic of that period. Other large fluctuations are not readily explained by single race frequency changes. Figure 4 presents the percentage of the Rogers index that is accounted for by changes in frequency of the three most prevalent races in either of the years that are compared. The trend appears to be upward from 1918 to about 1948, reflecting the decrease in number of phenotypes during the early part of this period (up to about 1938) followed by an increase in dominance of the predominant three phenotypes from about 1938 to 1950. From about 1950 to 1970 there was enough collective change in less prevalent phenotypes to cause a reduced if erratic contribution of the most prevalent single race each year of approximately 50%. Finally, the period from 1972 to 1980 reflects the lower number of phenotypes, resulting in a greater contribution from the most prevalent three.

Table 2 presents the contribution of two components of diversity, race number and frequency distribution or evenness, and the variable sample size to three of the diversity indexes as determined by the coefficients of determination in multiple linear regression. The associations between the various independent variables and diversity indexes were approximately linear by inspection of scatter diagrams of individual indexes plotted over the components singly.

DISCUSSION

The general trend in diversity of annually sampled stem rust populations, regardless of index used, is downward from 1918 to the present. Beyond this agreement in a gross trend, the different indexes have little else in common and represent real choices in how best to characterize diversity of a population. The choice must depend on the objectives of the worker and the properties of the population(s) of pathogens to be characterized or compared. For example, if one wished to compare two populations that have about equal numbers of phenotypes, in order to uncover more subtle differences, an index that is more sensitive to variation in evenness would be most desirable.

Pielou (9, p. 308) has stated that the explicit contributions of phenotype number and evenness cannot be separated in general. Our regression analysis only presents approximate contributions for specific data. There are probably characteristics of these data that render any conclusions drawn to be less applicable to populations or samples that do not share these characteristics. One such characteristic is the amount of dominance of the most prevalent phenotype. In many years, particularly after 1926, when the effects of sexual reproduction on diversity had been obviated, stem rust populations were characterized by a single dominant phenotype, which sometimes accounted for 70% or more of the sample. The large magnitude of the effect of evenness on all three

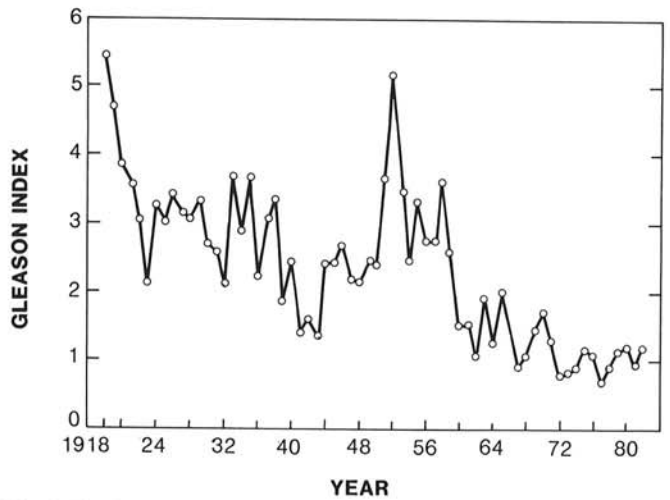

Fig. 1. The Gleason index of diversity for samples of stem rust of wheat collected annually in the United States from 1918 to 1982 and phenotypically analyzed using the 12 Stakman differential wheat lines.

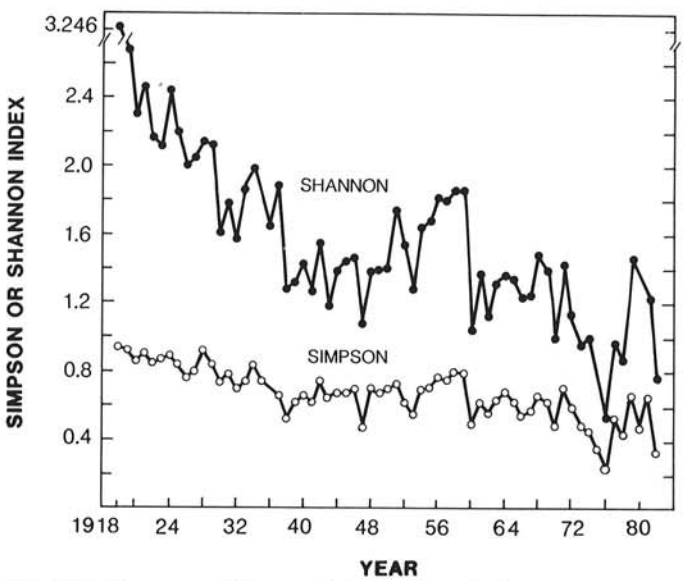

Fig. 2. The Simpson and Shannon indexes of diversity for samples of stem rust of wheat collected annually in the United States from 1918 to 1982 and phenotypically analyzed using the 12 Stakman differential wheat lines.

TABLE 1. Diversity characteristics of wheat stem rust samples collected annually in the United States from 1918 to 1982 and screened on the Stakman differential host set[a]

Year	Sample size	Phenotypes (no.)	Standard deviation of frequencies	Dominant phenotype(s)[a] (%)	
				(1)	(3)
1918	142	28	0.032	15	32
1919	162	25	0.037	12	54
1920	80	18	0.073	28	59
1921	90	17	0.056	19	48
1922	68	14	0.082	25	63
1923	70	10	0.066	21	54
1924	185	28	0.058	24	45
1925	144	16	0.086	35	58
1926	194	19	0.101	44	69
1927	406	20	0.087	30	67
1928	344	17	0.078	28	66
1929	720	23	0.073	26	62
1930	371	17	0.112	36	84
1931	498	17	0.097	27	74
1932	444	14	0.132	46	83
1933	369	23	0.096	37	76
1934	737	20	0.083	28	70
1935	1,493	28	0.088	42	78
1936	825	15	0.126	47	81
1937	1,341	23	0.116	56	73
1938	1,290	25	0.133	66	88
1939	1,063	14	0.154	56	89
1940	1,662	19	0.122	42	86
1941	1,269	11	0.168	51	90
1942	1,014	12	0.124	31	86
1943	1,455	11	0.162	49	97
1944	1,740	18	0.120	43	90
1945	1,035	13	0.123	42	89
1946	1,211	20	0.114	38	90
1947	601	15	0.177	69	93
1948	1,745	17	0.120	33	93
1949	669	17	0.127	45	92
1950	1,180	18	0.122	44	90
1951	950	26	0.094	41	76
1952	1,279	38	0.098	59	84
1953	1,025	25	0.130	63	89
1954	1,481	19	0.115	49	78
1955	755	23	0.105	48	80
1956	1,005	20	0.072	31	75
1957	1,059	20	0.096	34	75
1958	775	25	0.084	29	71
1959	700	18	0.092	32	72
1960	721	11	0.199	61	92
1961	1,347	12	0.162	57	80
1962	657	8	0.208	61	94
1963	1,535	15	0.144	46	88
1964	1,286	10	0.151	40	87
1965	1,808	16	0.141	55	88
1966	281	9	0.206	65	84
1967	706	7	0.213	61	87
1968	1,731	9	0.162	52	80
1969	2,016	12	0.161	57	82
1970	2,040	14	0.184	68	93
1971	1,514	16	0.142	39	88
1972	1,990	7	0.206	53	98
1973	1,277	7	0.244	67	94
1974	2,410	8	0.238	72	95
1975	2,415	10	0.248	77	97
1976	1,703	8	0.298	86	99
1977	1,207	6	0.242	61	97
1978	790	7	0.260	71	97
1979	420	8	0.173	49	86
1980	134	7	0.232	51	87
1981	272	7	0.186	44	93
1982	345	8	0.276	80	95

[a] Percentages of the sample accounted for by the dominant single and three phenotypes are shown in the last two columns.

TABLE 2. Coefficients of determination (r^2 or R^2) between parameters or components of diversity and three indexes of diversity applied to samples of stem rust of wheat collected from 1918 to 1982 in the United States and analyzed using the Stakman differential host set[a]

Diversity index	Component of diversity							
	Sample size (S)	Race number (R)	Standard deviation of frequency (E)	S+R	S+E	R+E	S+R+E	Intercorrelation
Gleason	0.160	0.828	0.686	0.904	0.692	0.884	0.920	0.618
Shannon	0.337	0.337	0.794	0.591	0.853	0.798	0.856	0.835
Simpson	0.240	0.249	0.856	0.429	0.873	0.927	0.931	0.419[b]
Intercorrelation	0.020NS	0.539	0.162[c]					

[a] All coefficients are significant at $\alpha = 0.05$, except where indicated by NS.
[b] Between the Gleason and Simpson indexes.
[c] Between sample size and standard deviation of frequency, all data.

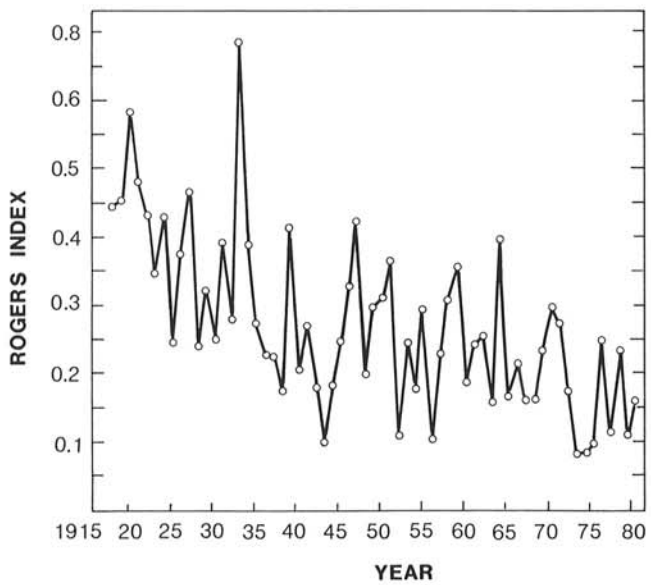

Fig. 3. The Rogers index of phenotypic change for samples of stem rust of wheat collected annually in the United States from 1918 to 1982 and phenotypically analyzed using the 12 Stakman wheat differential lines. Each plotted value is a comparison with the previous year except for the first year, 1918.

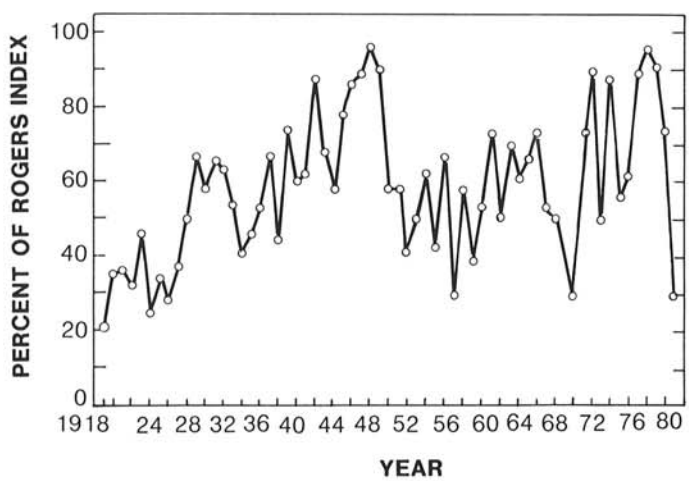

Fig. 4. The percentage of the Rogers index of phenotypic change accounted for by change in frequency of the most prevalent three phenotypes (races) found in either the plotted year or the previous year with which it is compared. Data are from samples of stem rust of wheat collected annually in the United States from 1918 to 1982 and phenotypically analyzed using the 12 Stakman wheat differential lines.

indexes was probably strongly influenced by dominance. The effect was indirect in the application of the Gleason index because as dominance became more extreme in later years, a corresponding decrease in number of phenotypes was observed. Although this would naturally result in fewer phenotypes being detected for a given sample size, it cannot be totally explained this way because sample sizes were large most years. In other words, the small number of phenotypes and the large degree of dominance were both a function of the nature of the asexual stem rust populations, rather than either being a function of the other.

The populations of pathogens that were being sampled in these studies are probably not typical of most plant pathogens. Phenotype numbers are low, especially in later years. Dominance is high. Sample sizes are larger than will generally be possible with smaller surveys. However, if populations of plant pathogens are being characterized that generally resemble stem rust populations in numbers of phenotypes, degree of dominance, and size of samples that can be handled, the following observations taken from the data in Table 1 are probably useful with respect to whether the Gleason, Shannon, or Simpson index should prove a better choice. The Gleason index is by far the most responsive to number of phenotypes. If frequency data are not to be included it is the only index of those discussed here that can be used, and represents an improved method of describing the race/collection ratio. If frequency data are used, the Gleason index should prove superior in comparing populations that vary primarily in number of phenotypes, but not in degree of dominance or sample size, to both of which it is least responsive of the three. If populations are to be compared that differ primarily in degree of dominance, the Simpson index is slightly more responsive than is the Shannon index, but both are probably acceptable. The greater responsiveness of the Simpson than the Shannon index to dominance has been demonstrated earlier (14), and the difference in the two with our data is smaller than might have been expected. As stated above, the Gleason index is only indirectly responsive to dominance as it influences number of phenotypes. Finally, if populations are being compared for which sample sizes are variable, but which do not appear to differ greatly in degree of dominance or number of phenotypes, the Shannon index might be a relatively poor choice because it was most responsive to this variable. Intuitively and ideally, sample size should not be permitted to influence the index used. The overall magnitude of sensitivity to sample size was low for all three indexes. The greater sensitivity of the Shannon index to sample size was surprising to us because Sanders (14) found that this index was least sensitive to sample size of the three, particularly when sample sizes were over 200. This further suggests that the characteristics of the populations can influence one another and the dynamics of the various indexes.

Combined (two and three independent variable) coefficients of determination in Table 2 suggest which indexes might be most useful where two of the three, or all three, parameters are apparently important.

Finally, the use of more than one index to describe more fully the populations is desirable if several of the meanings of diversity are of interest. In particular, either the Simpson or Shannon index might be calculated for each of two populations to assess absolute magnitude of diversity. Change in diversity from one population to the other might, however, be better expressed by the Rogers index, which incorporated phenotype identity as well as (in a less precise form) number and frequency.

LITERATURE CITED

1. Fleischmann, G. 1965. Variability in the physiologic race populations of oat crown rust isolated from aecia and uredia. Plant Dis. Rep. 49:132-133.
2. Futuyma, D. J. 1979. Evolutionary Biology. Sinauer, Sunderland, MA. 565 pp.
3. Grassel, S. F., Patil, G. P., Smith, W., and Taille, C. 1979. Ecological Diversity in Theory and Practice. International Cooperative Publishing House, Fairland, MD.
4. Green, G. J. 1971. Physiologic races of wheat stem rust in Canada from 1919 to 1969. Can. J. Bot. 49:1575-1588.
5. Groth, J. V., and Roelfs, A. P. 1982. Effect of sexual and asexual reproduction on race abundance in cereal rust fungus populations. Phytopathology 72:1503-1507.
6. Lebeda, A. 1982. Measurement of genetic diversity of virulence in populations of phytopathogenic fungi. Z. Pflanzenkrankr. Pflanzenschutz 89:88-95.
7. Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283-292.
8. Pielou, E. C. 1975. Ecological Diversity. John Wiley & Sons, New York. 165 pp.
9. Pielou, E. C. 1977. Mathematical Ecology. John Wiley & Sons, New York. 385 pp.
10. Poole, R. W. 1974. An Introduction to Quantitative Ecology. McGraw-Hill Book Co., New York. 532 pp.
11. Roelfs, A. P. 1984. Race specificity and methods of study. Pages 131-164 in: The Cereal Rusts. Vol. I. Origins, Specificity, Structure, and Physiology. W. R. Bushnell and A P. Roelfs, eds. Academic Press, Orlando, FL.
12. Roelfs, A. P., and Groth, J. V. 1980. A comparison of virulence phenotypes in wheat stem rust populations reproducing sexually and asexually. Phytopathology 70:855-862.
13. Rogers, J. S. 1972. Measures of genetic similarity and genetic distance. Pages 145-153 in: Studies in Genetics. University of Texas, Austin.
14. Sanders, H. L. 1968. Marine benthic diversity: A comparative study. Am. Nat. 102:243-282.
15. Stakman, E. C., and Levine, M. N. 1922. The determination of biologic forms of *Puccinia graminis* on *Triticum* spp. Minn. Agric. Exp. Stn. Tech. Bull. 8:1-10.
16. Whittaker, R. H. 1972. Evolution and measurement of species diversity. Taxon 21:213-251.