Analysis of Genotypic Diversity Data for Populations of Microorganisms

Niklaus J. Grünwald, Stephen B. Goodwin, Michael G. Milgroom, and William E. Fry

First author: U.S. Department of Agriculture-Agricultural Research Service (USDA-ARS), 24106 N. Bunn Rd., Prosser, WA 99350; second author: USDA-ARS, Department of Botany and Plant Pathology, 915 West State Street, Purdue University, West Lafayette, IN 47907; and third and fourth authors: Department of Plant Pathology, 334 Plant Science Building, Cornell University, Ithaca, NY 14853.

Accepted for publication 12 January 2003.

ABSTRACT

Grünwald, N. J., Goodwin, S. B., Milgroom, M. G., and Fry, W. E. 2003. Analysis of genotypic diversity data for populations of microorganisms. Phytopathology 93:738-746.

Estimation of genotypic diversity is an important component of the analysis of the genetic structure of plant pathogen and microbial populations. Estimates of genotypic diversity are a function of both the number of genotypes observed in a sample (genotype richness) and the evenness of distribution of genotypes within the sample. Currently used measures of genotypic diversity have inherent problems that could lead to incorrect conclusions, particularly when diversity is low or sample sizes differ. The number of genotypes observed in a sample depends on the technique used to assay for genetic variation; each technique will affect the maximum number of genotypes that can be detected. We developed an approach to analysis of genotypic diversity in plant pathology that makes specific reference to the techniques used for identifying genotypes. Preferably, populations that are being compared should be very similar in sample size. In this case, the number of genotypes observed can be used directly for comparing richness. In most cases, sample sizes differ and use of the rarefaction method to calculate richness is more appropriate. In all cases, scaling either Stoddart and Taylor's G or Shannon and Wiener's H' by sample size should be avoided. Under those circumstances where it might be important to distinguish whether richness or evenness contribute more to diversity, a bootstrapping approach, where confidence intervals are calculated for indices of diversity and evenness, is recommended.

Additional keywords: epidemiology, microbial ecology, population genetics.

Genotypic diversity is one of several components estimated during analysis of the genetic structure of populations of microorganisms. Two indices of genotypic diversity, Stoddart and Taylor's G (45) and Shannon-Wiener's H' (41), have been used most frequently in the estimation of genotypic diversity of plant pathogen populations (3,4,6–8,12–14,23,30,32,46). Indices of diversity have been used in plant pathology to measure phenotypic diversity—for example, for race (or pathotype) structure of rust fungi (15) and oomycetes (2)—as well as genotypic diversity for allozymes (14) or other molecular markers (7).

Currently used measures of genotypic diversity have inherent problems that could lead to incorrect conclusions, particularly when diversity is low and sample sizes differ. Stoddart and Taylor's G statistic (45) has a minimum value of 1 and a maximum equal to the sample size. The relationship between the maximum value of G (G_{max}) and sample size (Fig. 1A) makes comparison of genotypic diversity values among populations with unequal sample sizes difficult, particularly when diversity is high; the population with the highest G most likely will be the one with the largest sample size, regardless of whether it actually has the highest diversity. To remove the sample size bias, Chen et al. (7) recommended using a normalized statistic obtained by dividing G by the sample size. This solution works when comparing populations that all have high diversity (i.e., near G_{max}) (Fig. 1B). However, it fails badly when diversity is low (Fig. 1C). As an example, consider a population composed of a single clone, as is the case for many populations of the potato late blight pathogen *Phytophthora infestans* in the United States (11). Genotypic diversity should be zero in populations containing a single clone, regardless of how it is estimated. However, for the normalized G statistic, the minimum diversity equals $1/n$, where n is the sample size. Therefore, for samples of 10 and 100 individuals, the normalized G would be 0.1 and 0.01, respectively. This implies that the first population has a higher diversity than the second, even though the actual genotypic diversity of both populations is zero.

A second problem can occur with the normalized G statistic when the maximum possible diversity, given the amount of allelic variation present in a sample (or in a species), is less than the number of individuals sampled (Fig. 1D). For example, suppose we have amplified fragment length polymorphism (AFLP) or random amplified polymorphic DNA (RAPD) data for 30 loci in two haploid populations, but only two loci are polymorphic. If both loci have two alleles (1 and 0, for presence and absence of bands) then only four multilocus genotypes are possible (11, 10, 01, and 00) and the G_{max} will be 4 regardless of sample size. If we have unequal samples from populations with maximum diversity (i.e., all four genotypes are present at equal frequencies), then dividing by sample size will result in the population with the smallest sample size having the highest estimated diversity, even though the true diversity in all samples is the same. This will occur any time the sample size is larger than the number of possible genotypes that can be generated given the observed level of allelic variation. This was not a problem for Chen et al. (7) because the amount of variation was very high in each population they assayed. However, this problem could be significant in populations with low diversity.

Corresponding author: N. J. Grünwald
E-mail address: ngrunwald@pars.ars.usda.gov

*The *e*-Xtra logo stands for "electronic extra" and indicates that the online article contains supplemental material not included in the print edition. Two additional tables contain the raw data sets obtained from the literature on which simulations were based and a validation of the algorithm for calculation of rarefaction curves.

Publication no. P-2003-0421-02R

The normalized Shannon-Weaver index suffers from the same problem as Stoddart and Taylor's *G* index when diversity is low. For this statistic, the natural logarithm of the sample size is used as the denominator in normalization. In this case, there is not a problem with zero diversity, because the normalized Shannon information statistic ranges from 0 to 1 (Fig. 1B and C). However, the normalized Shannon statistic does have the same problem as the normalized *G* statistic when diversity is greater than zero but still low (Fig. 1D). Attempting to compare genotypic diversity among populations with small and unequal sample size is common in plant pathology, particularly for pathogens that are rare or obligate pathogens that may be difficult to sample and assay. There is a need for an improved understanding of the measurement of genotypic diversity within populations of plant pathogens and for the development of methods that allow comparison of populations with unequal sample sizes.

Indices of genotypic diversity are different from indices of gene diversity commonly used in analysis of populations of microorganisms (33,35). Gene diversity is calculated from allele frequencies, whereas genotypic diversity is calculated from genotype frequencies based on individuals. Gene diversity also is referred to as average heterozygosity when restricted to the case of a diploid, randomly mating organism for a population in Hardy-Weinberg equilibrium (35,47). Gene diversity can be applied to haploid, diploid, and polyploid organisms as well as for different means of reproduction, including asexual reproduction, random mating, or selfing. For populations of microorganisms, particularly those that have both sexual and asexual reproduction, both gene and genotypic diversity are needed to estimate genetic diversity; this approach commonly is observed in plant pathology, as shown for populations of *Mycosphaerella graminicola* or *Cryphonectria parasitica* (30,31).

The number of genotypes observed in a sample depends on the technique used to reveal genetic variation. The most commonly studied markers include virulence, isozymes, RAPD, AFLP, and restriction fragment length polymorphism (RFLP) (22). Each technique will affect the maximum number of genotypes that can be detected (5). Yet, to date, the technique used for genotyping has not been considered when analyzing genotypic diversity.

The goal of this study was to reassess the commonly used methods for analysis of genotypic and phenotypic diversity to determine which approaches are most suitable for populations of microorganisms. One specific objective was to analyze the shortcomings of currently used statistics to identify conditions under which they might give misleading or incorrect results. A secondary objective was to analyze how different molecular techniques (with specific reference to isozyme, RAPD, AFLP, and RFLP analysis) affect analysis of genotypic diversity and how they can be incorporated into statistical measures. The final objective was to develop approaches for comparing indices of diversity among populations by estimating confidence intervals through a bootstrap approach when sample sizes differ, and to illustrate this with specific examples from the literature.

THEORY AND APPROACHES

Background on diversity. Most of the theory and application of diversity indices was developed for community ecology to ana-

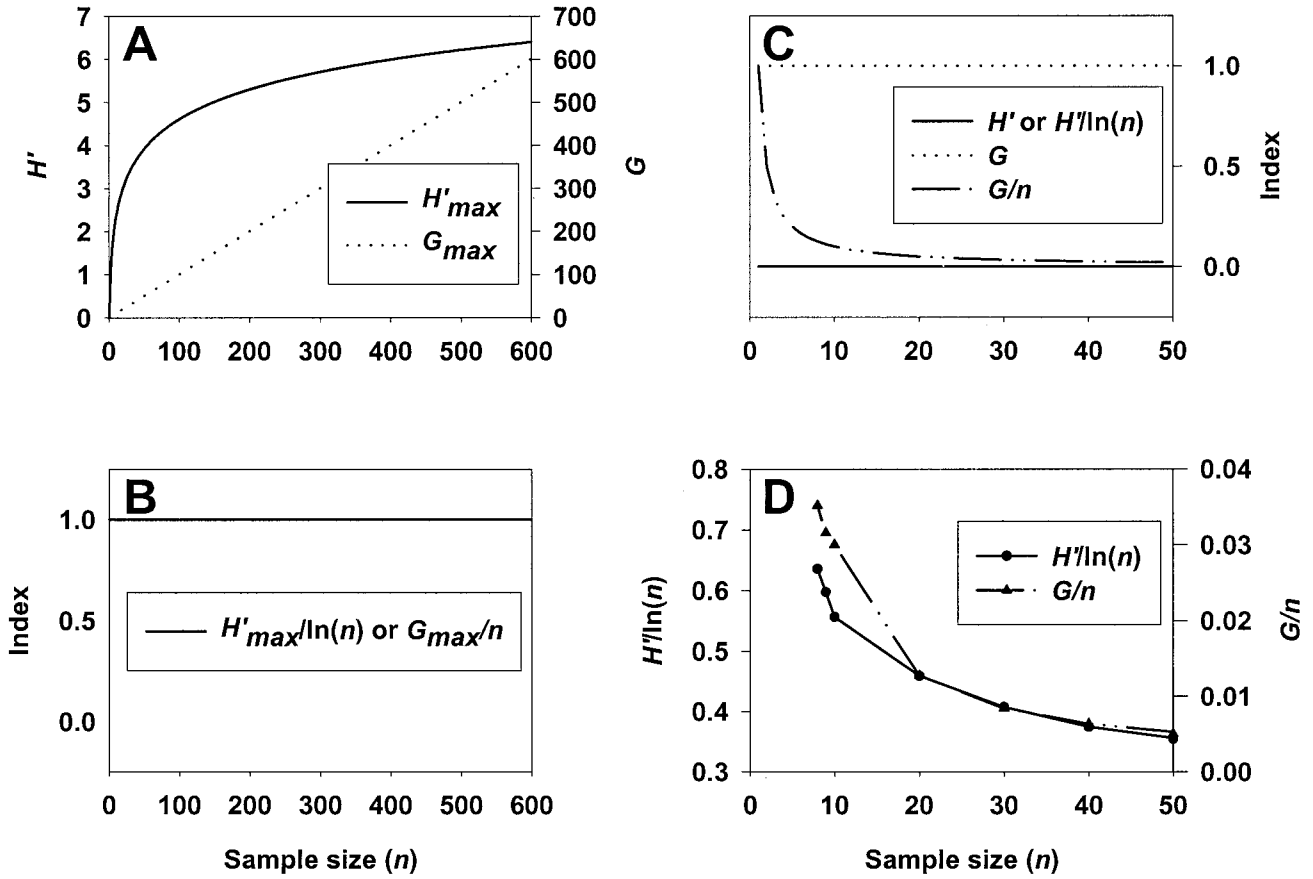

Fig. 1. Behavior of Stoddart and Taylor's *G* (45) and Shannon-Wiener's *H′* (41) as a function of sample size under different scenarios of genotypic diversity. **A,** Behavior of G_{max} and H'_{max} when each genotype is unique. **B,** Scaling G_{max} by *n* and H'_{max} by ln(*n*) when each genotype is unique. **C,** Behavior of *G* and *H′* and scaled indices (G_{max}/n and $H'_{max}/\ln[n]$) when only one genotype is observed (i.e., a single clone). **D,** Effect of scaling G_{max} by *n* and H'_{max} by ln(*n*) when the population consists of only four genotypes at equal frequency (e.g., two random amplified polymorphic DNA or amplified fragment length polymorphism loci with two alleles each) as sample size increases.

lyze patterns of species diversity. Thus, statistics referred to herein as indices of genotypic diversity are discussed as indices of species diversity in the ecological literature (21,25,26,37).

Diversity is composed of two aspects: richness and evenness (21,25,26,37). Richness is the number of genotypes contained in a population; intuitively, diversity increases with increasing richness. Evenness measures how genotypes are distributed within a population. If a small number of genotypes dominate the population, evenness is low and, intuitively, so is diversity. If each genotype occurs at an equal frequency then evenness (and diversity) is maximal. Most indices of diversity combine both richness and evenness. Thus, indices like Stoddart and Taylor's G (45) and Shannon and Wiener's H' (41) increase as richness (more genotypes in the population) or evenness (less domination of the population by one or a few genotypes) increase.

Richness. Genotypic richness is an estimate of the number of genotypes contained in a population. The simplest estimate of richness is the number of unique genotypes observed within a sample (g). Because g tends to increase with sample size, particularly when sample sizes are small, g is not a valid statistic for comparing richness of different populations unless sample sizes are equal.

To compare richness when sample sizes differ, ecologists have used rarefaction curves (16,20,21,25,26). Rarefaction curves yield the number of genotypes expected in a sample corresponding to the smallest sample size n of all populations being compared. This method assumes that the number of expected genotypes $E(g_n)$ in a random sample of n individuals out of a total sample of N individuals, where n_i corresponds to the number of individuals per genotype, is

$$E(g_n) = \sum_{i=1}^{g}\left[1-\left(\frac{\binom{N-n_i}{N}}{\binom{N}{n}}\right)\right]$$

$E(g_n)$ follows a hypergeometric distribution (18,43) and the expectation is based on the sum of the probabilities that each genotype will be included in the sample.

Use of rarefaction to estimate richness is appropriate if several restrictions are kept in mind (21,43). First, the sampling method needs to be consistent across populations; the same sampling design and technique must be employed in each population. Second, the rarefaction algorithm cannot be used to extrapolate to sample sizes larger than N in any particular sample. Third, if a certain genotype in a population is spatially aggregated, then the rarefaction method tends to overestimate genotype richness (10,43). This is particularly important for pathogens with both a vegetative and a sexual cycle, as is the case for *Phytophthora infestans* in central Mexico. For example, aggregation of one genotype can be observed if several isolates are sampled from one infected plant or one infected plot.

The method to assess genotypic richness using rarefaction curves was implemented in C using the function *bico* and associated routines as described by Press et al. (39). This function returns a binomial coefficient using logarithms of factorials. Use of logarithms of factorials avoids overflow of binomial coefficients for situations in which large factorials have to be calculated, but constitutes a trade-off in accuracy. The algorithm <Rarefac.c> calculates $E(g_n)$ for each sample size n = 1 to N and was validated using published data sets (19,21,24-26). <Rarefac.c> is available upon request from N. J. Grünwald.

Indices of genotypic diversity. The ideal index of diversity should take into consideration both richness and evenness. The index should be largest when each individual sampled has a unique genotype (i.e., g is maximal). Additionally, given the same richness, the index should be greater if genotypes are distributed more evenly within the sample.

Hill (17) developed a conceptual framework for analysis of diversity indices, and derived a family of diversity statistics given by

$$N_a = \left(\sum_{i=1}^{g} p_i^a\right)^{1/(1-a)}$$

in which p_i is the frequency of the ith genotype. Setting a = 0, 1, or 2 results in three of the most widely used diversity indices. For a = 0, N_0 is the number of genotypes observed (g), which is actually a measure of genotypic richness. For a = 1, we obtain

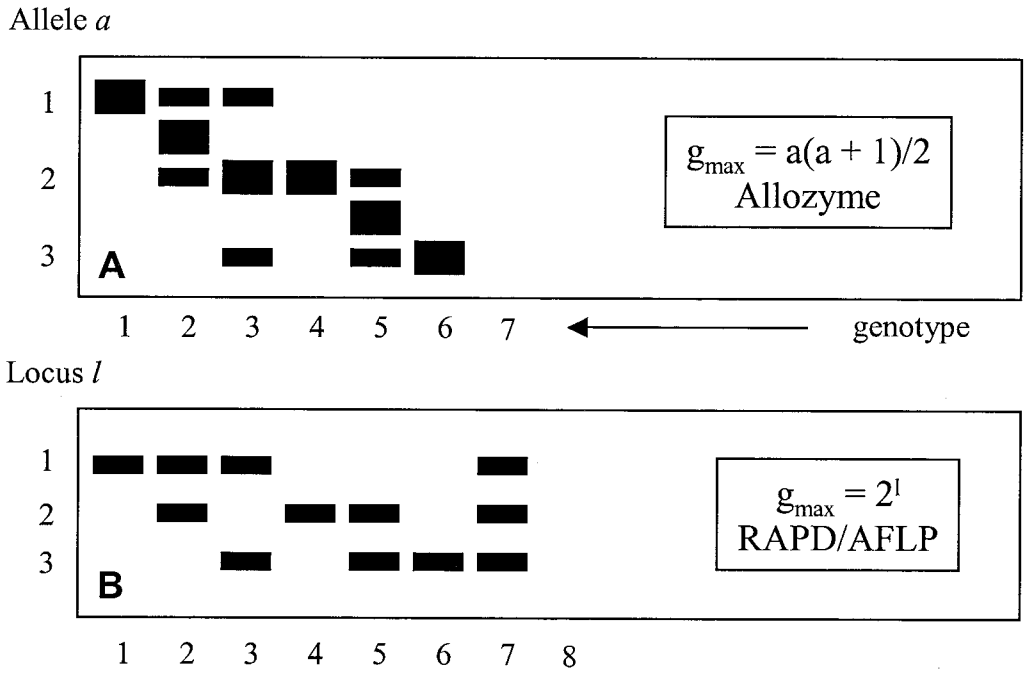

Fig. 2. Comparison of different molecular techniques for evaluation of the maximally expected number of genotypes (g_{max}) for a diploid organism. **A,** Allozyme analysis (one locus with three alleles for a dimeric enzyme), and **B,** random amplified polymorphic DNA or amplified fragment length polymorphism analysis (three loci each with two alleles, presence and absence of a band); a = allele; l = locus.

$N_l = e^{H'}$, where H' refers to Shannon-Wiener's $H' = \{-\Sigma_i[p_i \times \ln(p_i)]\}$ (41). Because H' is proportional to the logarithm of n, it can be expressed as $N_1 = e^{H'}$, which is proportional to the number of genotypes. N_1 then represents the number of equally common genotypes which would produce the same diversity as H' (17,25). For $a = 2$ we get $N_2 = 1/\lambda$, where λ is Simpson's index (44):

$$\lambda = \sum_{i=1}^{g} p_i^2$$

N_2 corresponds to the genotypic diversity index presented in Stoddart and Taylor (45): $G = 1/\Sigma p_i^2$. Stoddart and Taylor's index can also be calculated as $G = 1/\{\Sigma[(f_x)(x/n)^2]\}$, where n is the sample size, and f_x is the number of genotypes observed x times. G and N_2 are identical and will be referred to as Stoddart and Taylor's G because this has become common usage in plant pathology.

N_1 and G measure how effectively proportional abundances are distributed among the different genotypes (17). G weighs the number of abundant genotypes more strongly, whereas N_1 weighs rarer genotypes more strongly. N_1 generally falls between N_0 (i.e., g) and G.

Evenness. Indices of evenness indicate how genotypes are distributed within a sample. Usually, these indices are calculated by scaling a diversity index by the maximum number of expected genotypes (21,37). A desirable property of an index of evenness is that it should be 0 for a population composed of a single genotype and equal to 1 when all genotypes occur at the same frequency, regardless of richness.

The most common index of evenness, E_1 (36), scales the Shannon-Wiener index by the maximally expected number of genotypes g_{max}:

$$E_1 = \frac{H'}{\ln(g_{max})} = \frac{\ln(N_1)}{\ln(N_0)}$$

The mathematical derivation for why $H'_{max} = \ln(g_{max})$ is given in the Appendix. Note that the maximally expected value for H' (H'_{max}) is not equal to $\ln(n)$ and that H' cannot be scaled correctly by $\ln(n)$ (Appendix).

Sheldon (42) proposed scaling the diversity index N_1 by the number of genotypes observed in a population (Appendix provides derivation of appropriate scaling factor) such that

$$E_2 = \frac{e^{H'}}{g} = \frac{N_1}{N_0}$$

Another index, E_5 (25), actually consists of the ratio of G and N_1 calculated as

$$E_5 = \frac{(1/\lambda) - 1}{e^{H'} - 1} = \frac{G - 1}{N_1 - 1}$$

E_5 is preferable to E_1 and E_2 for several reasons (1,25). Essentially, E_5 is the ratio of the number of abundant genotypes to the number of rarer genotypes. E_5 is less dependent on the number of genotypes in a sample, because it is a ratio of two indices of diversity, which cancels out the effect of sample size. The value of E_5 is shifted by subtracting 1 from N_1 and G to make E_5 converge to 0 rather than 1 as a single genotype becomes more and more dominant.

Effect of genotyping technique on diversity estimates. Estimates of evenness depend on appropriate estimates of richness. Richness, in turn, depends on the technique used to assay genetic variation and the number of loci assayed. For example, with allozyme data for a diploid organism with three co-dominant alleles at one locus, the maximum possible number of genotypes is six (Fig. 2A). In general, the maximally expected number of genotypes at a locus for a diploid is given by $g_{max} = a(a + 1)/2$ (co-dominant) and $g_{max} = a(a - 1)/2$ (dominant) for co-dominant and dominant alleles, respectively. For a haploid, the maximum number of genotypes equals the number of alleles a at that locus. In the case of RAPD or AFLP loci with only two alleles at each locus, the maximum expected number of genotypes is given by $g_{max} = 2^l$ (Fig. 2B), where l is the number of loci, because only two phenotypes (plus or minus) can be scored at each locus for dominant markers.

To estimate the maximum number of multilocus genotypes, one calculates the product of the number of possible genotypes at each locus over all loci assayed. For instance, in a study on population structure of *P. infestans* in central Mexico, mating type and allozyme pattern at the *Gpi* and *Pep* locus were combined to define multilocus genotypes (16). For this pathogen, mating type can either be *A1* or *A2*. In this particular study, four alleles were observed for *Pep*, $g_{max} = (4)(4 + 1)/2 = 10$, and six for *Gpi*, $g_{max} = (6)(6 + 1)/2 = 21$, allozyme loci. All possible combinations of g_{max}

TABLE 1. Calculation of indices of richness, evenness, and diversity for three previously published data sets from *Cryphonectria parasitica*, *Phytophthora infestans*, and *Colletotrichum graminicola*

	Cryphonectria parasitica[a]			*Phytophthora infestans*[b]			*Colletotrichum graminicola*[c]	
Statistic	Teano	Finzel	Bartow	1988–89	1997–98	Combined	G92	G93
Sample size								
n	194	54	50	179	401	580	98	208
Indices of richness								
g_{obs}	4	23	22	32	48	54	5	8
$E(g_n)$	3.99	22.78	21.80	31.92	47.95	53.97	4.99	7.99
$E(g_n)$ for smallest n	2.96	22.20	21.80	31.92	34.61	34.17	4.99	6.24
g_{max}	4	128	128	200	420	420	144	1,152
Indices of diversity								
H'	0.614 (0.39–0.84)[d]	2.825 (2.57–3.08)	2.832 (2.59–3.07)	2.815 (2.67–2.96)	3.089 (2.94–3.24)	3.089 (2.94–3.24)	0.735 (0.54–0.93)	0.790 (0.56–1.02)
N_1	1.848 (1.45–2.25)	16.855 (13.6–20.2)	16.977 (13.8–20.1)	16.689 (14.4–18.9)	21.964 (19.1–24.8)	21.953 (19.0–24.9)	2.086 (1.68–2.49)	2.204 (1.72–2.69)
G	1.537 (1.21–1.87)	12.678 (9.21–16.2)	13.158 (9.68–16.6)	11.596 (9.53–13.7)	15.371 (13.0–17.8)	15.097 (12.8–17.5)	1.571 (1.30–1.85)	1.553 (1.28–1.83)
Indices of evenness								
E_1 [$=H'/ln(g_{obs})$]	0.443	0.901	0.916	0.812	0.798	0.774	0.457	0.380
G/g_{obs}	0.384	0.551	0.598	0.362	0.320	0.280	0.314	0.194
E_5	0.633 (0.47–0.79)	0.737 (0.62–0.85)	0.761 (0.64–0.88)	0.675 (0.61–0.75)	0.686 (0.63–0.74)	0.673 (0.62–0.73)	0.526 (0.44–0.61)	0.459 (0.38–0.54)

[a] Data from Milgroom and Cortesi (31).
[b] Data from Grünwald et al. (16).
[c] Data from Rosewich et al. (40).
[d] Numbers in parentheses indicate confidence intervals calculated by the bootstrapping approach for the common sample size of the smallest population.

for the three loci results in a g_{max} (multilocus genotype) of 2 × 10 × 21 = 420 for the maximum expected number of multilocus genotypes.

Evaluation of scaling methods. Several authors have normalized indices of diversity to correct for sample size bias, thus facilitating comparisons among populations when sample sizes differ. The Shannon-Wiener index often is used as an index of diversity scaled by the logarithm of sample size *n,* such that $H'_{scaled} = H'/\ln(n)$ (3,4,12–14,23,46). Similarly, Stoddart and Taylor's *G* sometimes is scaled by dividing it by the sample size *n* (7,8,30). The scaling factor depends on the index calculated. For instance, for the Shannon-Wiener diversity index *H′,* the correct scaling factor is the maximally expected number of genotypes $H' = \ln(g_{max})$ (Appendix). In the case of Stoddart and Taylor's *G*, the correct scaling factor is g_{max} (Appendix).

It is not clear how to choose g_{max} when estimating evenness. The easiest way is to set g_{max} equal to the maximal number of genotypes observed in any sample. However, this method is not appropriate when sample sizes differ. For example, increasing sample size from 179 to 401 in a population of *P. infestans* sampled at the same location resulted in an increase of observed *g* from 32 to 48 (Table 1). Another way of estimating g_{max} is by using the rarefaction method to calculate $E(g_n)$. This is particularly appropriate when sample sizes differ. However, $E(g_n)$ often may be different from the g_{max} calculated as all possible combinations of the alleles present in a sample. For the *P. infestans* example in Table 1, all possible combinations of alleles at the mating type *Gpi* and *Pep* loci generate 420 possible multilocus genotypes. However, only 54 genotypes were detected in a sample of 580 (Table 1). In this scenario, it is not clear whether a g_{max} of 420 or 54 is the appropriate scaling factor. Under conditions in which all alleles are present in a population, it would suffice to use g_{max} calculated according to the technique used for genotyping (Appendix), provided it is smaller than the sample size. If the potential number of genotypes is larger than the sample size, then the sample size is the appropriate scaling factor, because it would be impossible to detect more than *n* genotypes in a sample of *n* individuals. Given the difficulties with the choice of g_{max} as an appropriate scaling factor, we suggest the use of bootstrapping to compare diversity in populations with differing sample sizes. Bootstraps can be run on any sample size.

Bootstrapping. Bootstrapping is a resampling technique based on randomly drawing a sample with replacement from an original sample (9,28). The underlying idea is that, in the absence of any other knowledge about a population, the distribution of values found in a random sample of size *n* from the population is the best

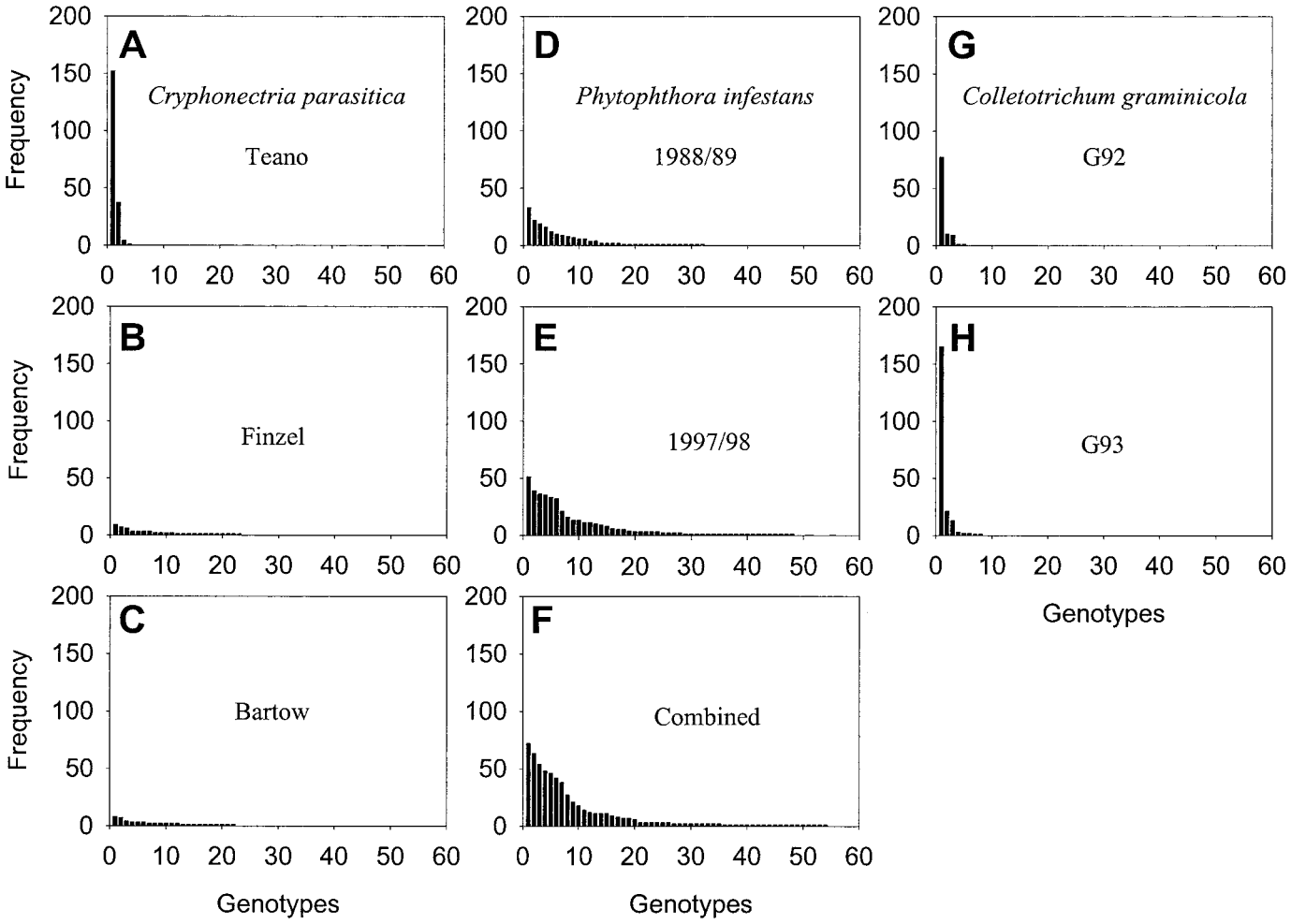

Fig. 3. Frequency distribution of published data sets used to contrast different scenarios of genotypic diversity analysis. The first set of data comes from **A,** one European and **B and C,** two American populations of the haploid chestnut blight pathogen, *Cryphonectria parasitica* (31). Multilocus genotypes are based on vegetative incompatibility polymorphisms at six unlinked loci, each with two alleles. The second set of data comes from **D and E,** two samples of a sexual population of the diploid potato late blight pathogen, *Phytophthora infestans*, that were **F,** combined into a larger data set (16). Genotypes are defined as multilocus genotypes (mating type: two alleles; *Gpi* and *Pep* isozyme pattern with six and four alleles each, respectively). The third set of data comes from **G and H,** two populations of the haploid pathogen *Colletotrichum graminicola* (40). Genotypes are based on restriction fragment length polymorphisms (RFLPs) from seven probes; restriction size variants detected by each probe were treated as alleles at a single RFLP locus. Alleles at different loci were combined to form multilocus genotypes.

guide to the distribution in the population (28). Boostrapping is particularly powerful because it allows calculation of confidence intervals. Confidence intervals contain the population mean with a fixed probability determined by the confidence coefficient, often chosen to be 95%. Bootstrap tests of significance have not been as well studied as bootstrap confidence intervals (28).

Bootstrapping was conducted using the SAS macro <jackboot.sas> (available online at no cost from the website of the SAS Institute, Cary, NC) modified to calculate indices of diversity and evenness. Bootstrapping was conducted using 2,000 resamples at a confidence interval of 95% using the accelerated bootstrap procedure (BCa method) (9).

Simulated data sets. Artificial data were constructed to evaluate the behavior of indices of diversity across gradients of evenness and richness. One group of data sets differing only in number of genotypes, but with constant evenness, was used to evaluate dependence of diversity indices on richness. The data sets consisted of sample sizes n = 2, 4, 6, 10, 20, 40, 60, 80, and 100 isolates, where each genotype occurred twice, resulting in numbers of genotypes g = 1, 2, 3, 5, 10, 20, 30, 40, and 50, respectively. All genotypes were distributed evenly; therefore, evenness E_5 = 1.0. Another group of data sets differing only in evenness, but with richness and sample size constant, was used to evaluate dependence of diversity indices on evenness. Data sets were constructed to have the same sample size n = 1,000 and genotypic richness g = 200, but differed in distribution of genotypes within a data set. Evenness ranged from completely equal genotype frequencies to one genotype dominating 80% of the population (the rest of the genotypes occurred at equal frequencies). The range of evenness was quantified using E_5, intermediate in response to either E_1 or E_2.

Sample data sets. Several published data sets differing in richness, evenness, and sample size were selected to illustrate the importance and difficulties of determining richness, evenness, and diversity (Fig. 3). The first set of data comes from one European and two American populations of the haploid chestnut blight pathogen, *C. parasitica* (31). Multilocus genotypes are based on vegetative incompatibility polymorphisms at six unlinked loci, each with two alleles. The frequency of detection of genotypes is shown in descending order (Fig. 3). The second set of data comes from two samples of a sexual population of the diploid potato late blight pathogen, *P. infestans* (16). Genotypes are defined as multilocus genotypes (mating type: two alleles; the allozyme loci *Gpi* and *Pep* with six and four alleles each, respectively). The third data set comes from two populations of the haploid pathogen *Colletotrichum graminicola* (40). Genotypes are based on RFLPs using seven probes. Restriction size variants detected by each probe were treated as alleles at single RFLP loci. Alleles at different loci were combined to form multilocus genotypes.

RESULTS

Effect of richness and evenness on diversity indices. Most indices of diversity increase with increasing genotypic richness and sample size (Fig. 4). N_1 and G (= N_2) increase linearly as the number of genotypes g increase (Fig. 4A). In contrast, λ decreases and H' increases nonlinearly as g increases (Fig. 4B). The linear increase of a diversity index with increasing richness is more intuitive.

With increasing evenness, but constant richness, N_1 and G (=N_2) increase (Fig. 5A). H' increases almost linearly, and λ decreases as evenness of genotypes within a sample increases (Fig. 5B). For maximum evenness (that is, all n_i/N are equal), $N_1 = G = g$ (Fig. 5A). Thus, with a uniform distribution of genotypes, both diversity indices N_1 and G result in the number of observed genotypes. In contrast, H' = 5.298 and λ = 0.005 when evenness E_5 = 1.0 (Fig. 5B). These values of H' are abstract and not amenable to direct interpretation. The linear behaviors with changing richness and the intuitive interpretation when evenness is maximal make N_1 and G more desirable indices of diversity.

Evaluation of scaling factors. Scaling N_1, G, and H' by g or ln(g) has the desired effect of scaling indices between 1/g and 1 for N_1 and G and 0 and 1 for H' (Fig. 5C), while scaling by sample size does not (Fig. 5D). Scaling indices of diversity by some measure of sample size or richness makes these indices directly dependent on sample size. Scaling by n is not the same as scaling by g. If $g \leq n$, which is normally the case in most populations of plant pathogens sampled (Fig. 3), scaling by n will have a stronger effect than scaling by g.

Rarefaction curves. The algorithm <Rarefac.c> performs well when contrasted to values reported in the literature. The algorithm was evaluated using published data sets (19,21,24–26). Values for $E(g_n)$ calculated by <Rarefac.c> deviated from published results by only one decimal place. Because the algorithm uses a logarithmic approach for calculation of binomial coefficients to avoid overflow (39), the results are expected to differ somewhat from those published in the literature.

An example of the application of rarefaction curves is presented with data from Milgroom and Cortesi (31) (Fig. 6). Given any data set, the smallest sample size present among all populations is chosen for comparison. This sample size n then is used to determine the expected number of genotypes for all samples. For example, the smallest sample size in the three populations presented in Figure 3 is n = 50. At n = 50, we would expect the number of genotypes in our sample to be $E(g_n)$ = 2.96, 22.20, and 21.80 for the Teano, Finzel, and Bartow populations, respectively (Table 1; Fig. 6).

Application to published data. In an analysis of eight published data sets, in which sample sizes and richness vary considerably, it is not always clear whether differences in diversity are due

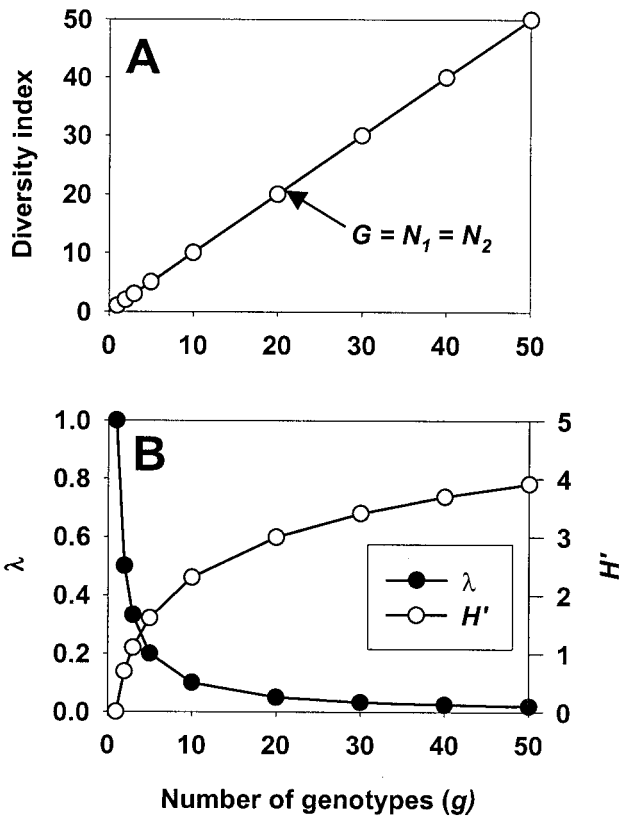

Fig. 4. Evaluation of genotypic diversity indices with artificial data sets constructed to vary in richness and sample size, but with constant evenness. **A,** Hill's index N_1 and Stoddart and Taylor's G (=N_2), and **B,** Simpson's λ and Shannon-Wiener's H'. The data sets consisted of sample sizes n = 2, 4, 6, 10, 20, 40, 60, 80, and 100 isolates, where each genotype occurred twice, resulting in numbers of genotypes g = 1, 2, 3, 5, 10, 20, 30, 40, and 50, respectively. All genotypes were distributed uniformly; therefore, evenness E_5 = 1.0.

to changes of richness or of evenness and whether indices are different (Table 1). The Bartow and Finzel data sets for *Cryphonectria parasitica* have higher diversity (contrast H', N_1, and G) than the Teano data set. Bootstrapped confidence intervals of diversity do not overlap, whereas those for evenness do (Table 1). Thus, the differences in diversity must be due to differences in richness, which is larger in the Bartow and Finzel data sets. If we contrast the 1988–89 and the combined data sets from *P. infestans*, it is not clear whether we detect more genotypes (32 and 54, respectively) in the second population due to a higher richness or simply because the sample size increased from 179 to 580 (Table 1). When an estimate of evenness is combined with the rarefaction method to calculate richness (g is not an appropriate measure of richness because the sample sizes differ), it becomes clear that the higher diversity in the combined data set is due mostly to higher richness because evenness actually decreases and bootstrapped confidence intervals for E_5 are almost identical (Table 1). Finally, the two *Colletotrichum graminicola* data sets exemplify a case where, while there are no significant differences in diversity, diversity measured by different indices could result in different conclusions. The two indices H' and G show opposite trends: a diversity index biased toward rare genotypes (H' or N_1) results in higher diversity in the *G93* population compared with the *G92* population, whereas an index that weighs common genotypes more (G) results in lower diversity in population *G93* compared with *G92* (Table 1). This difference in diversity is due to higher richness and lower evenness in the *G93* sample.

DISCUSSION

Our analysis confirmed that richness, evenness, and diversity (including both richness and evenness) are different qualities of population genotypic diversity. Indices of diversity clearly depend on both the number of genotypes in the sample (Fig. 4) and on how those genotypes are distributed within a sample (Fig. 5). Differences in an index of diversity can reflect variation in richness, evenness, or both; therefore, our analysis suggests separating the two components of diversity. A combined approach, in which comparisons are at a common sample size or richness is calculated by using the rarefaction method and diversity and evenness are calculated with corresponding confidence intervals using bootstrapping, allows inferences about the relative importance of richness and evenness in a genotypic diversity analysis. This is particularly important when sample sizes differ between populations. Richness can be reported as either the number of genotypes observed (g), in cases where sample size is nearly equal for all populations to be compared, or can rely on the rarefaction method. Rarefaction curves estimate the number of genotypes expected in each sample if all samples are of a standard size and is the method of choice when sample sizes differ (25,26).

Not all indices of diversity contrasted in this study perform equally well. Hill's indices N_1 and Stoddart and Taylor's G have several favorable qualities. Most importantly, they carry units of "effective numbers of genotypes" which can be interpreted intuitively. Both indices increase linearly with increasing richness and have a maximal value equal to richness g when evenness is maximal. H' and λ cannot be interpreted intuitively and result in abstract numbers even under conditions where evenness is equal to unity. In addition, λ behaves inversely to richness, which is counterintuitive. In light of these advantages, we recommend use of N_1 and G over H' and λ.

Indices of diversity can be more or less sensitive to changes in the number of rare genotypes (25). The Shannon-Wiener index H' is most sensitive to changes in rare genotypes, while Simpson's

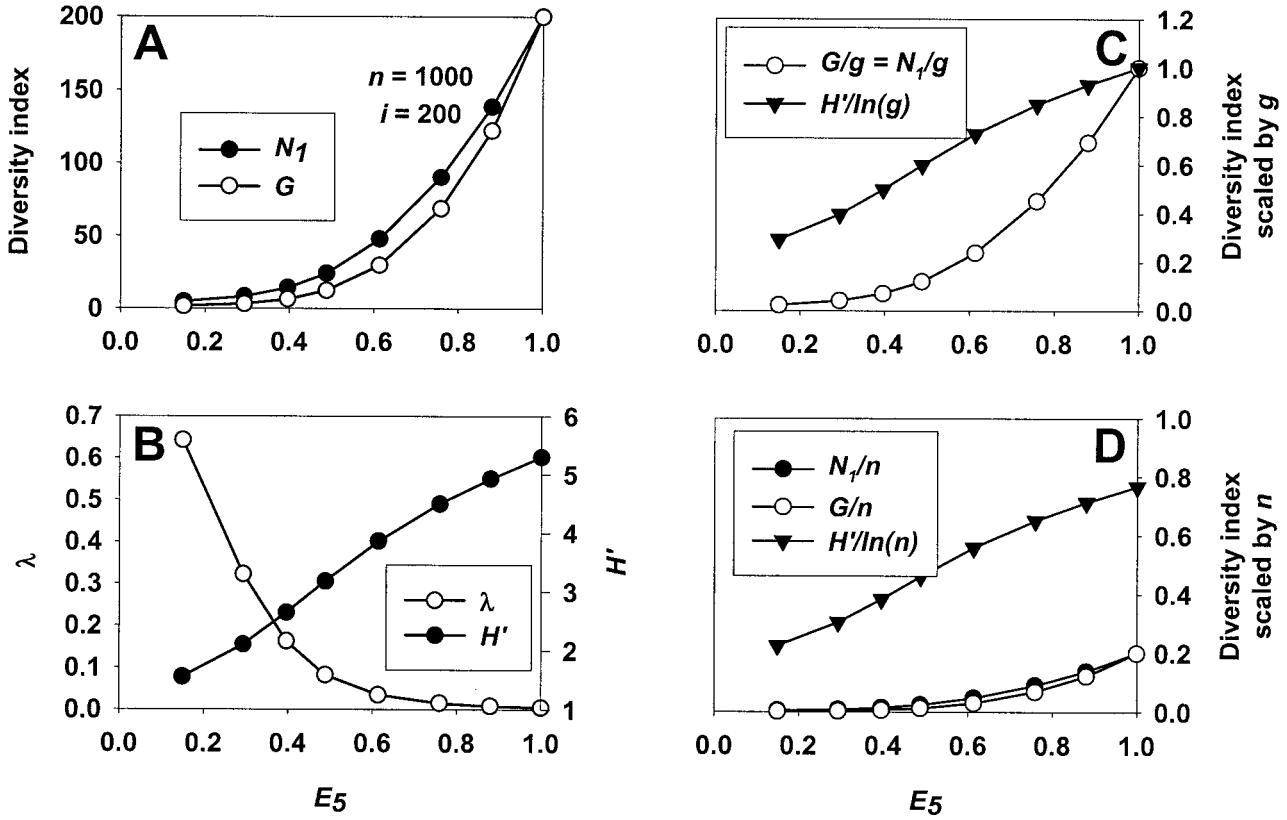

Fig. 5. Evaluation of genotypic diversity indices with artificial data sets constructed to vary in evenness of genotype frequency, but with constant richness. **A,** Hill's index N_1 and Stoddart and Taylor's G ($=N_2$), and **B,** Simpson's λ and Shannon-Wiener's H'. **C and D,** Effect of scaling by g and n, respectively. Data sets were constructed to have the same sample size n = 1,000 and genotypic richness g = 200, but differed in distribution of genotypes within a data set. Evenness (E_5) ranged from completely equal in distribution to one genotype dominating 80% of the population.

index G is most sensitive to changes in the abundant genotypes. Similarly, N_1 is more sensitive to rare and G to abundant genotypes. Analysis of genotypic diversity in plant pathology should take into account whether rare or common genotypes should be given more emphasis.

Several recent publications have scaled indices of diversity by sample size or the logarithm of sample size. Several researchers have corrected G by dividing it by n (6–8,30). Similarly, the H' is often used as an index of diversity scaled by the logarithm of sample size n, such that $H'_{scaled} = H'/\ln(n)$ (3,4,12–14,46). We showed that scaling by $\ln(n)$ or n is both mathematically and conceptually incorrect, particularly when diversity is low. Most importantly, the scaled indices are strongly affected by sample size. Our analysis shows that a data set with the same diversity would result in decreasing values of genotypic diversity scaled by either $\ln(n)$ or n, once n is greater than the number of genotypes that can be detected in a sample of that size with the technique used. The effect is strongest with sample sizes typically used in studies of plant pathogen populations (e.g., $n < 200$).

A more appropriate scaling factor would be the maximal expected number of genotypes. For the Shannon-Wiener diversity index, the scaling factor can be calculated as $\ln(g_{max})$, where g_{max} is the maximum number of genotypes possible depending on the type of marker, observed level of allelic diversity, and ploidy of the organism. Scaling H' as $E_1 = H'/\ln(g)$ is more correctly considered a measure of evenness, because diversity is scaled by the number of genotypes (25,26,37,38) and, thus, the index reflects how uniformly genotypes are distributed within a population. The evenness index E_5 is preferable to E_1 because it is not affected by sample richness (25). E_5 describes the ratio of the effective numbers of very abundant to abundant genotypes, scaled to approach 0 when one genotype becomes more and more abundant by subtracting 1 from N_2 and N_1.

The equations for calculating g_{max} also may be useful in deciding how many markers are needed to analyze a population. McDonald (29) recommends 6 to 12 unlinked marker loci to obtain accurate measurements of population genetic structure. The equations for g_{max} permit calculating exactly how many genotypes can be detected for analyses of genotypic diversity. For example, if we use six biallelic RAPD or AFLP loci, we can detect only 64 genotypes ($g_{max} = 2^6$), which is sufficient for a clonal population but insufficient to study population structure of sexual populations. With 12 loci, the number of possible genotypes would be 4,096, which should be more appropriate for studying sexual populations, assuming each locus is polymorphic.

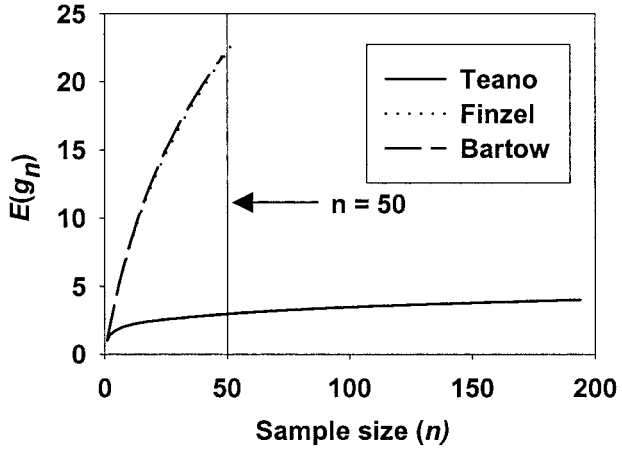

Fig. 6. Richness estimated using the rarefaction method, $E(g_n)$, for three populations of the chestnut blight pathogen *Cryphonectria parasitica* (31). The smallest sample size $n = 50$ (vertical reference line) of the three populations assayed is used to compare richness estimates between populations. The Teano population is much less rich than either Finzel or Bartow, which have the same degree of richness.

Analysis of genotypic diversity, as discussed in this article, can be complemented by analysis of gene diversity (33,34), which is based on frequencies of alleles, not individuals or genotypes. Correction for sample size bias has been described for gene diversity and unbiased estimators are available (35). Attempts to integrate genotypic and gene diversity into a single index have been made and applied to populations of *Puccinia recondita* f. sp. *tritici* (20,27).

Our analysis of indices used to estimate genotypic diversity results in several practical recommendations. Preferably, sample size should be very similar for all populations being compared. In this case, the number of genotypes observed (g) can be used directly as a measure of richness. When sample sizes differ, estimation of richness by rarefaction is more appropriate. In all cases, scaling either Stoddart and Taylor's G or Shannon and Wiener's H' by sample size should be avoided. If scaling is necessary, it should be done with g or $\ln(g)$. Finally, under those circumstances in which it might be important to distinguish whether richness or evenness contributes more to diversity, a bootstrapping approach is recommended to calculate confidence intervals for indices of diversity and evenness.

APPENDIX

Derivation of appropriate scaling factors for diversity indices. The maximal value for the Shannon-Wiener index H'_{max} occurs when each $p_i = 1/g$ (i.e., when every individual sampled has a unique genotype). Thus we get

$$H'_{max} = -g\left(\frac{1}{g}\ln\frac{1}{g}\right) = \ln(g)$$

It is observed that, in the case where $p_i = 1/g$, we also get $p_i = 1/n$, because $n = g$. However, when diversity is low, then n may be much greater than g and $\ln(n) > \ln(g)$.

The maximal value for Sheldon's index $N_1 = e^{H'}$ is given by $N_{1max} = e^{H'max} = e^{\ln(g)} = g$. In this case using $\ln(n)$ as a scaling factor will underestimate the true value of H', and the magnitude of this bias will increase as n gets larger. The appropriate scaling factor N_{2max} for $N_2 = 1/\lambda$, where λ is Simpson's index

$$\lambda = \sum_{i=1}^{g} p_i^2$$

is obtained in the case where $p_i = 1/g$. It follows that $\lambda_{max} = g \times (1/g)^2 = 1/g$ and $N_{2max} = g$.

ACKNOWLEDGMENTS

This work was funded by the CEEM (Cornell-Eastern Europe-Mexico) Potato Late Blight Project, PICTIPAPA (Programa Internacional Cooperativo del Tizón Tardío de la Papa), and USDA CRIS projects 3602-22000-009-00 and 5354-21220-009-00. We thank L. R. Gale for the opportunity of corresponding with her about interpretation of her published data included in this study.

LITERATURE CITED

1. Alatalo, R. V. 1981. Problems in the measurement of evenness in ecology. Oikos 37:199-204.
2. Andrivon, D. 1994. Race structure and dynamics in populations of *Phytophthora infestans*. Can. J. Bot. 72:1681-1687.
3. Borchart, D. S., Welz, H. G., and Geiger, H. H. 1998. Genetic structure of *Setosphaeria turcica* populations in tropical and temperate climates. Phytopathology 88:322-329.
4. Borchart, D. S., Welz, H. G., and Geiger, H. H. 1998. Spatial and temporal variation of genetic marker patterns in *Setosphaeria turcica* populations from Kenya. J. Phytopathol. 146:451-457.
5. Brown, J. K. M. 1996. The choice of molecular marker methods for population genetic studies of plant pathogens. New Phytol. 133:183-195.
6. Caffier, V., Brändle, U. E., and Wolfe, M. S. 1999. Genotypic diversity in barley powdery mildew populations in northern France. Plant Pathol.

48:582-587.
7. Chen, R. S., Boeger, J. M., and McDonald, B. A. 1994. Genetic stability in a population of a plant pathogenic fungus over time. Mol. Ecol. 3:209-218.
8. Chen, R. S., and McDonald, B. A. 1996. Sexual reproduction plays a major role in the genetic structure of populations of the fungus *Mycosphaerella graminicola*. Genetics 142:1119-1127.
9. Dixon, P. M. 1993. The bootstrap and the jackknife: Describing the precision of ecological indices. Pages 290-318 in: Design and Analysis of Ecological Experiments. Chapman and Hall, New York.
10. Fager, E. W. 1972. Diversity: A sampling study. Am. Nat. 106:293-310.
11. Forbes, G. A., Escobar, X. C., Ayala, C. C., Revelo, J., Ordonez, M. E., Fry, B. A., Doucett, K., and Fry, W. E. 1997. Population genetic structure of *Phytophthora infestans* in Ecuador. Phytopathology 87:375-380.
12. Goodwin, S. B., Cohen, B. A., and Fry, W. E. 1994. Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. Proc. Natl. Acad. Sci. USA 91:11591-11595.
13. Goodwin, S. B., Drenth, A., and Fry, W. E. 1992. Cloning and genetic analyses of two highly polymorphic, moderately repetitive nuclear DNAs from *Phytophthora infestans*. Curr. Genet. 22:107-115.
14. Goodwin, S. B., Saghai Maroof, M. A., Allard, R. W., and Webster, R. K. 1993. Isozyme variation within and among populations of *Rhynchosporium secalis* in Europe, Australia and the United States. Mycol. Res. 97:49-58.
15. Groth, J. V., and Roelfs, A. P. 1989. The analysis of genetic variation in populations of rust fungi. Pages 318-339 in: Plant Disease Epidemiology. Vol. 2: Genetics, Resistance, and Management. McGraw-Hill Publishing Company, New York.
16. Grünwald, N. J., Flier, W. G., Sturbaum, A. K., Garay-Serrano, E., van den Bosch, T. B. M., Smart, C. D., Matuszak, J. M., Lozoya-Saldaña, H., Turkensteen, L. J., and Fry, W. E. 2001. Population structure of *Phytophthora infestans* in the Toluca Valley region of Central Mexico. Phytopathology 91:882-890.
17. Hill, M. O. 1973. Diversity and evenness: A unifying notation and its consequences. Ecology 54:427-432.
18. Hurlbert, S. H. 1971. The nonconcept of species diversity: A critique and alternative parameters. Ecology 52:577-586.
19. James, F. C., and Rathbun, S. 1981. Rarefaction, relative abundance, and diversity of avian communities. Auk 98:785-800.
20. Kosman, E. 1996. Difference and diversity of plant pathogen populations: A new approach for measuring. Phytopathology 86:1152-1155.
21. Krebs, C. J. 1989. Ecological Methodology. HarperCollins Publishers, New York.
22. Leung, H., Nelson, R. J., and Leach, J. E. 1993. Population structure of plant pathogenic fungi and bacteria. Adv. Plant Pathol. 10:157-205.
23. Liu, Y.-C., Cortesi, P., Double, M. L., Macdonald, W. L., and Milgroom, M G. 1996. Diversity and multilocus genetic structure in populations of *Cryphonectria parasitica*. Phytopathology 86:1344-1351.
24. Livingston, R. J. 1976. Diurnal and seasonal fluctuations of organisms in a north Florida estuary. Estuarine Coastal Marine Sci. 4:373-400.
25. Ludwig, J. A., and Reynolds, J. F. 1988. Statistical Ecology: A Primer on Methods and Computing. John Wiley & Sons, New York.
26. Magurran, A. E. 1988. Ecological Diversity and Its Measurement. Princeton University Press, Princeton, NJ.
27. Manisterski, J., Eyal, Z., Ben-Yehuda, P., and Kosman, E. 2000. Comparative analysis of indices in the study of virulence diversity between and within populations of *Puccinia recondita* f. sp. *tritici* in Israel. Phytopathology 90:601-607.
28. Manly, B. F. J. 1997. Randomization, Bootstrap and Monte Carlo Methods in Biology. 2nd ed. Chapman and Hall, New York.
29. McDonald, B. A. 1997. The population genetics of fungi: Tools and techniques. Phytopathology 87:448-453.
30. McDonald, B. A., Mundt, C. C., and Chen, R.-S. 1996. The role of selection on the genetic structure of pathogen populations: Evidence from field experiments with *Mycosphaerella graminicola* on wheat. Euphytica 92:73-80.
31. Milgroom, M G., and Cortesi, P. 1999. Analysis of population structure of the chestnut blight fungus based on vegetative incompatibility genotypes. Proc. Natl. Acad. Sci. USA 96:10518-10523.
32. Milgroom, M. G., Lipari, S. E., and Powell, W. A. 1992. DNA fingerprinting and analysis of population structure in the chestnut blight fungus, *Cryphonectria parasitica*. Genetics 131:297-306.
33. Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70:3321-3323.
34. Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.
35. Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.
36. Pielou, E. C. 1966. The measurement of diversity in different types of biological collections. J. Theor. Biol. 13:131-144.
37. Pielou, E. C. 1975. Ecological Diversity. Wiley, New York.
38. Pielou, E. C. 1977. Mathematical Ecology. Wiley, New York.
39. Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. 1992. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, Cambridge, MA.
40. Rosewich, U. L., Pettway, R. E., McDonald, B. A., Duncan, R. R., and Frederiksen, R. A. 1998. Genetic structure and temporal dynamics of a *Colletotrichum graminicola* population in a sorghum disease nursery. Phytopathology 88:1087-1093.
41. Shannon, C. E., and Weaver, W. 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL.
42. Sheldon, A. L. 1969. Equitability indices: Dependence on the species count. Ecology 50:466-467.
43. Simberloff, D. 1979. Rarefaction as a distribution-free method of expressing and estimating diversity. Pages 159-176 in: Ecological Diversity in Theory and Practice. International Co-operative Publishing House, Fairland, MD.
44. Simpson, E. H. 1949. Measurement of diversity. Nature 163:688.
45. Stoddart, J. A., and Taylor, J. F. 1988. Genotypic diversity: Estimation and prediction in samples. Genetics 118:705-711.
46. Sujkowski, L. S., Goodwin, S. B., Dyer, A. T., and Fry, W. E. 1994. Increased genotypic diversity via migration and possible occurrence of sexual reproduction of *Phytophthora infestans* in Poland. Phytopathology 84:201-207.
47. Weir, B. S. 1996. Genetic Data Analysis. Sinauer Associates Inc., Sunderland, MA.