Interpreting Ecological Diversity Indices Applied to Terminal Restriction Fragment Length Polymorphism Data: Insights from Simulated Microbial Communities[▽]

Christopher B. Blackwood,[1,3,4]* Deborah Hudleston,[2] Donald R. Zak,[2,3] and Jeffrey S. Buyer[4]

Department of Biological Sciences, Kent State University, Kent, Ohio 44242[1]; Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109-1048[2]; School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michigan 48109-1041[3]; and Sustainable Agricultural Systems Laboratory, USDA-ARS, Beltsville Agricultural Research Center, Beltsville, Maryland 20705[4]

Ecological diversity indices are frequently applied to molecular profiling methods, such as terminal restriction fragment length polymorphism (T-RFLP), in order to compare diversity among microbial communities. We performed simulations to determine whether diversity indices calculated from T-RFLP profiles could reflect the true diversity of the underlying communities despite potential analytical artifacts. These include multiple taxa generating the same terminal restriction fragment (TRF) and rare TRFs being excluded by a relative abundance (fluorescence) threshold. True community diversity was simulated using the lognormal species abundance distribution. Simulated T-RFLP profiles were generated by assigning each species a TRF size based on an empirical or modeled TRF size distribution. With a typical threshold (1%), the only consistently useful relationship was between Smith and Wilson evenness applied to T-RFLP data (TRF-E_{var}) and true Shannon diversity (H'), with correlations between 0.71 and 0.81. TRF-H' and true H' were well correlated in the simulations using the lowest number of species, but this correlation declined substantially in simulations using greater numbers of species, to the point where TRF-H' cannot be considered a useful statistic. The relationships between TRF diversity indices and true indices were sensitive to the relative abundance threshold, with greatly improved correlations observed using a 0.1% threshold, which was investigated for comparative purposes but is not possible to consistently achieve with current technology. In general, the use of diversity indices on T-RFLP data provides inaccurate estimates of true diversity in microbial communities (with the possible exception of TRF-E_{var}). We suggest that, where significant differences in T-RFLP diversity indices were found in previous work, these should be reinterpreted as a reflection of differences in community composition rather than a true difference in community diversity.

Microbial ecologists deal with arguably the most diverse biological communities on Earth (3, 7) and with organisms which are among the most difficult to study in their natural environments. Molecular profiling methods based on the heterogeneity of a specific gene resulting in differences in electrophoretic mobility are popular approaches for characterizing microbial community composition. Diversity indices originally adopted for macroorganisms are frequently used on microbial community profile data (12, 23), including data generated by terminal restriction fragment length polymorphism (T-RFLP) (see reference 24 for details on T-RFLP). This is appealing because diversity is at the center of a large body of ecological theory (27) and is a concept that the general public appreciates. Univariate diversity indices (e.g., Table 1) may also be an elegant summary of a complex community.

Despite their use in microbial ecology, the application of diversity indices to T-RFLP data has not been analytically validated, and several aspects common to all electrophoresis-based community profiling methods should be considered before this practice is accepted. For example, molecular profiling methods normally characterize only "dominant" organisms (e.g., >1% of the community) due to detection limits that arise whenever many markers are quantified simultaneously. Hence, rare species can never be detected, but these species often make up the vast majority of the diversity in microbial communities (13, 33). In addition, terminal restriction fragments (TRFs) of the same size can be generated from multiple taxa, which can be disparately related (1, 5, 8). The diversity of even those dominant taxa that are detected will therefore be underestimated.

Given the aforementioned properties of microbial community profiles, what can comparison of diversity indices applied to them tell us about the underlying diversity of the communities? Here, we make use of database information to answer this question for simulated communities with various levels of diversity. We examine T-RFLP of bacterial ribosomal genes because it is a popular, high-throughput method which we hope will not be abused, and databases and bioinformatics tools are available to support our simulations. The utilities of several diversity indices, including species richness, evenness, and integrated diversity indices, are evaluated.

MATERIALS AND METHODS

The simulations required assumptions about two distributions: a species abundance distribution, which determines the "true" diversity of the community in terms of the number of species present and their relative abundances, and a TRF

TABLE 1. Indices used to quantify the diversity of simulated bacterial communities and their associated community profiles[a]

Index	Formula
Richness (S)	S
Shannon index (H')	$H' = \sum_{i=1}^{S} p_i \ln(p_i)$
Shannon effective no. of species ($e^{H'}$)	$e^{H'} = \exp(H')$
Simpson index ($1/D$)	$1/D = 1 / \sum_{i=1}^{S} p_i^2$
Berger-Parker index ($1/d$)	$1/d = 1/\max(p_i)$
Shannon evenness (J')	$J' = \dfrac{H'}{\ln(S)}$
Simpson evenness ($E_{1/D}$)	$E_{1/D} = \dfrac{1/D}{S}$
Smith and Wilson evenness (E_{var})	$E_{var} = 1 - \dfrac{2}{\pi} \arctan\left\{ \sum_{i=1}^{S}\left(\ln(p_i) - \sum_{j=1}^{S} \ln(p_j)/S \right)^2 \Big/ S \right\}$

[a] S is the number of species in the community or the number of biomarkers present; p_i is the relative abundance of species or biomarker i. For further information, see Jost (21), Magurran (29), and Smith and Wilson (37).

size distribution, which describes how the sizes of TRFs vary among species (Fig. 1). The simulations were performed by sampling a TRF size distribution to determine the T-RFLP profile for a given species abundance distribution. Species within the bacteria are difficult to define, and we use the term here to be consistent with the literature on diversity indices.

Species abundance distributions. Community species abundance distributions were simulated using the lognormal distribution, which has previously been applied to soil microbial communities (3, 6, 13). Equations defining the lognormal distribution were obtained from Dunbar et al. (6). The distribution is defined by the following equation:
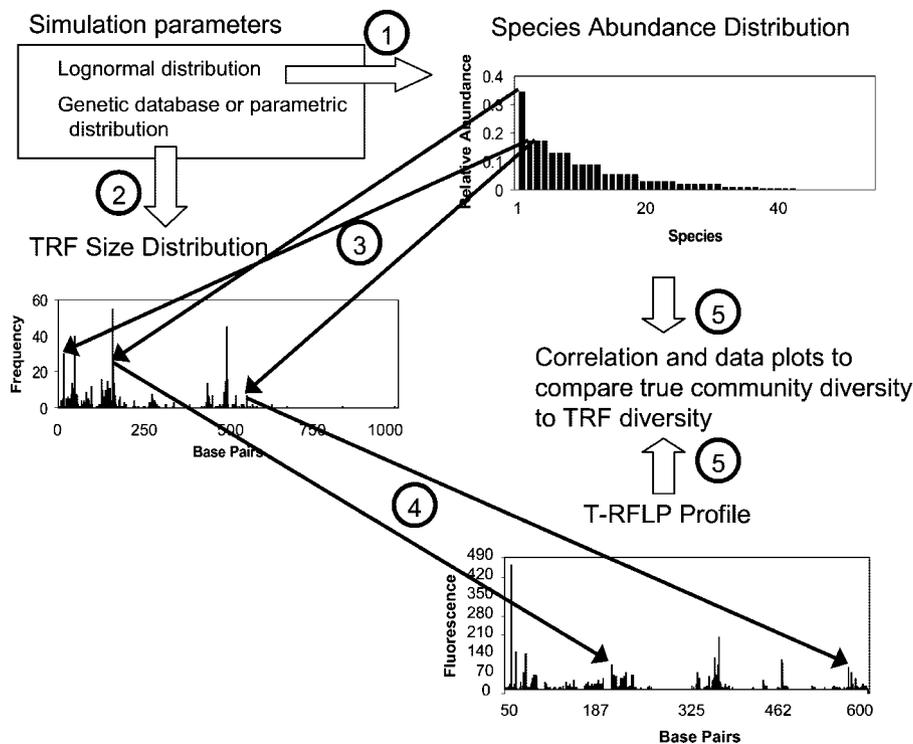

FIG. 1. Illustration of how the species abundance distribution and TRF distribution interact to form a simulated T-RFLP profile. Illustrated steps include calculating relative abundances in a lognormal species abundance distribution based on systematically iterated distribution parameters (1), defining a TRF size probability distribution (one of the four empirical restriction enzyme databases was used for OTU sampling and sequence sampling simulations, whereas a new TRF size distribution was generated for each iteration in parametric sampling simulations) (2), randomly assigning each species in the simulated community to a TRF size according to the TRF size probability distribution (3), summing signals from TRFs and applying a relative abundance (fluorescence) threshold and upper and lower fragment size cutoffs (4), and comparing diversity statistics calculated on the underlying community (species abundance distribution) and the simulated T-RFLP profile by calculating correlation coefficients and plotting data density graphs (5).

TABLE 2. Database and simulation information for restriction enzymes used to create empirical TRF size distributions

Enzyme	No. in database		Maximum no. of species[a] for:	
	OTUs	Sequences	OTU sampling	Sequence sampling
HaeIII	1,552	4,836	1,200	4,000
HhaI	1,512	4,679	1,200	4,000
MspI	1,532	4,724	1,200	4,000
RsaI	1,118	3,759	900	3,000

[a] Maximum number of species simulated in community abundance distributions.

$$S(R) = \frac{S_T a}{\sqrt{\pi}} \exp(-a^2 R^2)$$

where R is the $\log_2$ species abundance octave, $S(R)$ is the number of species in octave R, S_T is the total number of species in the community, and a is the dispersion constant which incorporates the variance of the distribution [$a = (0.5/\sigma^2)^{0.5}$]. The abundance of each species in octave $R(N_R)$ equals $N_0 2^R$. The focus of our simulations was relative abundance, so the abundance of each species was found relative to N_0, and then the total arbitrarily scaled population size of the community (from $N_{-R_{max}}$ to $N_{R_{max}}$) was determined so that relative abundances could be calculated.

The simulations consisted of systematically varying initial parameters which determined the diversity and shape of the distribution, including a, S_T, and $S(R_{max})$. The parameter $S(R_{max})$ is the number of species in the octave with the most abundant species. The parameter a was varied from 0.15 to 0.3 by intervals of 0.05, and $S(R_{max})$ was varied from 1 to 7 by intervals of 2. While it is often assumed that $S(R_{max})$ equals 1, we did not see an a priori reason to assume this when considering species with highly diverse ecological niches, as is the case for soil bacteria. S_T was varied from 10 species to different maxima, depending on the analysis (Table 2), by intervals of 10 species. For a given set of these parameters, $S(0)$ was calculated as $S_T \times a/\pi^{0.5}$, and then R_{max} was determined iteratively. The number of species in each octave, its relative abundance, and S_T for the resulting distribution were then calculated. To accommodate the fixed value of a and the discrete lognormal distribution, values of S_T and $S(R_{max})$ were allowed to vary slightly below and above their initial values, respectively, after the number of octaves in the distribution was determined. Combinations of parameters which did not result in meaningful distributions were discarded (e.g., distributions with a values of 0.2, $S(R_{max})$ values of 5, and 40 or fewer species were not used, because at least 42 species are required for this combination of parameters).

Empirical TRF size distributions. A database of TRF sizes was generated using the ARB 2004 small-subunit ribosomal database (28) and the ARB tool TRF-CUT (36). A total of 5,133 sequences which matched the primers Bac8F and Univ1492R underwent in silico digestions using the restriction enzymes HhaI, RsaI, MspI, and HaeIII, assuming that Bac8F was the labeled primer. A sequence was excluded from analysis for a particular enzyme if it contained missing nucleotide data before the first restriction site. This empirical database was sampled in two ways. In "sequence sampling," sequences were randomly sampled without replacement and assigned to each species in a species abundance distribution. In "OTU sampling," sequences were first categorized into operational taxonomic units (OTUs) by using the following approach. Genetic distances between the primed gene fragments were calculated using the Olsen distance correction factor recommended by ARB. Sequences were then clustered into OTUs by using a 3% distance cutoff and the average distance method in SAS (SAS Institute, Cary, NC). OTU sampling consisted of randomly choosing one OTU without replacement and then randomly choosing one sequence within that OTU to represent it. OTU sampling therefore avoided bias due to overrepresentation of certain OTUs within the sequence database, whereas sequence sampling was prone to this bias.

Simulations of community diversity and T-RFLP profiles. Simulations were performed using a macro written in SAS Proc IML. For a given simulation, true community diversity was varied by altering the parameters of the species abundance distribution as described above, and 10 replicate T-RFLP profiles were generated for each species abundance distribution. Replicate T-RFLP profiles were generated by repeating the random assignment of sequences to species in the distribution. The TRF signal in the simulated T-RFLP profile (representing height or area) was assumed to be proportional to the relative abundance of the

species. TRF signals were pooled when two species had the same TRF size in base pairs. TRFs smaller than 50 bp and greater than 600 bp were deleted, and TRFs were expressed as relative abundances of the total signal detected in the T-RFLP profile. TRFs below a threshold (0.1, 1, or 4% of the total profile signal) were also deleted, and relative abundances of remaining TRFs were recalculated. Diversity statistics (Table 1) were calculated on both the T-RFLP profile (TRF indices) and the species abundance distribution that it was derived from (true indices).

The utilities of TRF indices were assessed by calculating their correlation coefficients with true indices. However, testing the statistical significance of the correlation coefficients is not appropriate, because the very large number of data points makes all correlations significant, even those very close to zero and obviously not useful as an analytical tool. Therefore, we used graphical displays of the data cloud (data density plots) to further interpret correlation coefficients. These were constructed by plotting the mean values and percentile ranges of a true index over each 5-percentile interval of a TRF index.

Theoretical TRF size distributions and simulations. "Parametric sampling" consisted of using mathematical approximations of TRF size distributions, allowing the total number of species in abundance distributions to be increased beyond the number of empirical sequences available in the database. Candidate models for the general form of the parametric TRF size distribution were evaluated for their abilities to fit the HhaI, RsaI, MspI, and HaeIII TRF size distributions. Models were initially screened by one of two methods. The first involved visually comparing empirical TRF frequency plots arranged by size (bp) to plots generated using combinations of normal and uniform probability distributions. The second screening method consisted of fitting nonlinear equations supplied with the software GOSA (Bio-Log, Ramonville, France) (e.g., exponential decay, two-phase exponential decay, lognormal, normal, power, Pareto, and polynomial, etc.) to empirical TRF frequencies arranged as rank abundance data. We were modeling TRF size distributions for the purpose of evaluating TRF diversity indices by simulation, so our main criterion was the ability of the models to generate correlations similar to those obtained in empirical simulations when performed over the same range of S_T. After the initial screen, simulations were performed to 1,200 and 4,000 species by using selected models, and correlations were compared to results from OTU sampling and sequence sampling, respectively. Simulations were performed as described above, except that assignment of TRF size to each species was according to a probability distribution defined by the model in question (the parametric TRF size distribution) rather than an empirical TRF size distribution. Also, variability between empirical TRF size distributions was incorporated into the simulations by creating a new parametric TRF size distribution for each iteration within a simulation. This was done by allowing random variability in model parameters based on the range of parameter values obtained during fitting of the model to the four empirical TRF size distributions. Correlation matrices obtained using different parametric TRF size distribution models were compared to OTU sampling and sequence sampling correlation matrices by conducting Mantel tests and by calculating the mean absolute deviation between correlation matrices.

Our goal was to simulate communities with up to 10,000 species. Through the above-described process, we chose to use the following model in further simulations: an assemblage of six normal curves and one uniform random distribution. Each normal curve had a randomly generated probability (P_O) that a species would have a TRF size derived from that curve. P_O values for all the normal curves summed to 90%, while the underlying random-uniform-distribution P_O was 10%. Each normal curve was defined by a randomly generated mean (25 to 650 bp) and standard deviation (SD; 2 to 20 bp). The uniform random distribution ranged from 1 to 1,400 bp. Parametric TRF size distributions were sampled by randomly assigning each species in an abundance distribution to one of the six normal curves or the uniform distribution (according to their P_O values) and then randomly assigning a specific TRF size within the assigned curve.

RESULTS

OTU sampling. OTU sampling was conducted using species abundance distributions with 10 to 1,200 species for most restriction enzymes (Table 2). Table 3 shows correlation results for a 1% threshold (the most commonly used threshold in the literature) for some of the indices tested. The highest correlation between a true index and its corresponding TRF index was between H' and TRF-H' (0.71). However, TRF-E_{var} had the highest correlations with true indices overall, with the correla-

TABLE 3. Representative correlation coefficients from OTU sampling simulations of communities with up to 1,200 species, using a TRF relative abundance threshold of 1%[a]

Index	Avg (SD) for indicated true index			
	S	H'	$1/D$	J'
TRF-S	0.15 (0.04)	0.58 (0.06)	0.45 (0.07)	0.54 (0.04)
TRF-H'	0.23 (0.03)	0.71 (0.05)	0.56 (0.06)	0.67 (0.03)
TRF-$e^{H'}$	0.21 (0.04)	0.67 (0.06)	0.57 (0.08)	0.63 (0.04)
TRF-$1/D$	0.22 (0.04)	0.68 (0.06)	0.61 (0.09)	0.64 (0.05)
TRF-$1/d$	0.16 (0.05)	0.55 (0.09)	0.52 (0.13)	0.56 (0.07)
TRF-J'	0.26 (0.02)	0.72 (0.04)	0.59 (0.05)	0.70 (0.03)
TRF-$E_{1/D}$	0.24 (0.04)	0.66 (0.07)	0.64 (0.10)	0.66 (0.06)
TRF-E_{var}	**0.36 (0.02)**	**0.81 (0.04)**	**0.72 (0.04)**	**0.71 (0.03)**

[a] The highest coefficients for each true index are shown in bold. Values are means (standard deviations) of results for four restriction enzymes.

tion between H' and TRF-E_{var} the highest (0.81). Data density plots (Fig. 2A and B) show the extent to which TRF-H' and TRF-E_{var} may be useful over this range of abundance distributions. Compared to the data in Fig. 2A, the decreased slope and the smaller range of true H' for each category in Fig. 2B show the greater resolution that TRF-E_{var} has in reflecting true H'. This is particularly true when H' is greater than 4.5, which is the case for approximately half the data. Species richness (S) and the evenness indices E_{var} and $E_{1/D}$ are not predicted well by any TRF index ($r < 0.4$). This is illustrated for the relationship between S and TRF-S in Fig. 2C. Other indices tested (J', $1/D$, $1/d$, and $e^{H'}$) were moderately well correlated with TRF indices ($r < 0.73$).

When a 4% threshold was used, the correlations between most true and TRF indices declined substantially (data not shown), even for those TRF indices with the highest correlations with true indices (Table 4). However, the relationship between TRF-E_{var} and H' may still be useful, with a correlation coefficient of 0.66 (SD = 0.03). This was the highest correlation observed between true and TRF indices when a threshold of 4% was used. The correlation between true H' and TRF-H' dropped to -0.54 (SD = 0.03).

With a 0.1% threshold, true indices were generally well correlated with their corresponding TRF indices (see examples in Table 4).

Sequence sampling. When species richness was extended to 4,000 species by using sequence sampling (Table 2) and a 1% peak threshold, almost all correlations between true and TRF indices declined compared to those described above for simulations using up to 1,200 species and OTU sampling (data not shown). The highest correlation remained between H' and TRF-E_{var}, which declined from 0.81 to 0.71 (SD = 0.03) (Table 4), whereas the correlation between H' and TRF-H' declined from 0.71 to 0.51 (SD = 0.08). Differences between results for sequence sampling and OTU sampling for 4% or 0.1% thresholds were similar to the differences described for the 1% threshold. Correlations remained basically the same or declined slightly (data not shown).

Parametric sampling for complex communities. A variety of parametric TRF size distributions were compared to the four empirical size distributions obtained from TRF-CUT. The comparisons were based on shapes of the size distributions and on how well correlations between true and TRF indices ob-

tained with communities of up to 1,200 or 4,000 species matched results obtained using empirical TRF size distributions over the same range of S_T. According to both Mantel tests and the mean absolute deviations, assemblages of six
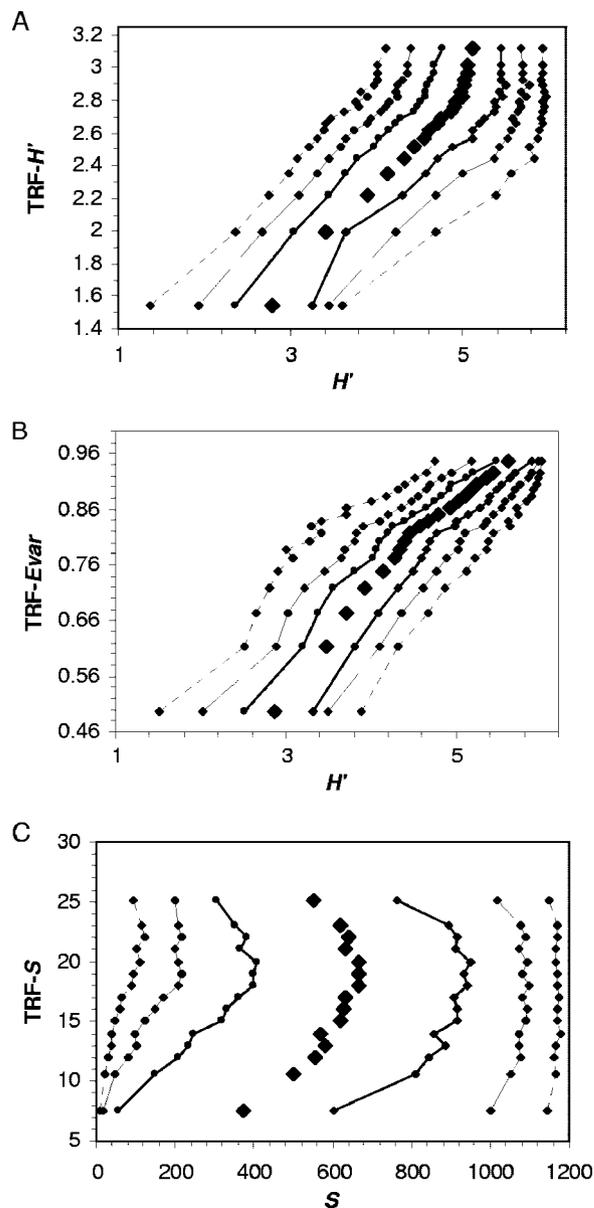

FIG. 2. Data density plots for OTU sampling simulations of communities with up to 1,200 species, using a relative abundance threshold of 1% and the HhaI TRF distribution. These plots summarize a data cloud. Lines can be interpreted as contour lines of data point density around a "peak" showing the central relationship between the two variables. Overlapping percentile ranges for two values of the TRF index indicate the potential for the true indices to be the same for the two communities. Therefore, the slope of the central relationship, as well as the width of the percentile ranges, is important. The TRF diversity index data were broken up into five percentile groups. The mean value for the TRF index in each group was plotted against the following true community index values: mean (diamond), 25th and 75th percentiles (bold line), 10th and 90th percentiles (thin line), and 2.5th and 97.5th percentiles (dashed line). In panels A and B, n was 940 for each group. In panel C, n was 638 to 1,998 due to TRF-S ties.

TABLE 4. TRF indices with the highest correlation with select true indices[a]

Threshold (%)	TRF size distribution group	Correlation with indicated true index			
		S	H'	$1/D$	J'
1	OTU	0.36 (TRF-E_{var})	0.81 (TRF-E_{var})	0.72 (TRF-E_{var})	0.71 (TRF-E_{var})
	Sequence	0.30 (TRF-E_{var})	0.71 (TRF-E_{var})	0.51 (TRF-E_{var})	0.63 (TRF-E_{var})
	Parametric	0.30 (TRF-E_{var})	0.73 (TRF-E_{var})	0.57 (TRF-$E_{1/D}$)	0.69 (TRF-$E_{1/D}$)
4	OTU	0.23 (TRF-$E_{1/D}$)[b]	0.66 (TRF-E_{var})	0.55 (TRF-$E_{1/D}$)	0.60 (TRF-E_{var})
	Sequence	0.24 (TRF-$E_{1/D}$)	0.60 (TRF-$E_{1/D}$)	0.44 (TRF-$E_{1/D}$)	0.56 (TRF-$E_{1/D}$)
	Parametric	0.15 (TRF-$E_{1/D}$)[b]	0.56 (TRF-$E_{1/D}$)[b]	0.33 (TRF-$E_{1/D}$)	0.54 (TRF-$E_{1/D}$)
0.1	OTU	0.68 (TRF-S)	0.97 (TRF-H')	0.92 (TRF-$1/D$)	0.80 (TRF-J')
	Sequence	0.59 (TRF-S)	0.96 (TRF-S)	0.83 (TRF-S)	0.73 (TRF-J')
	Parametric	0.37 (TRF-S)	0.76 (TRF-H')	0.65 (TRF-$1/d$)	0.68 (TRF-J')

[a] Values show the best correlations from among the eight TRF indices evaluated for each of the true indices. OTU sampling included communities with up to 1,200 species, sequence sampling up to 4,000 species, and parametric sampling up to 10,000 species. Values shown for OTU sampling and sequence sampling are averages from analyses of four restriction enzymes (standard deviations were less than or equal to 0.04). Values shown for parametric sampling incorporate variability due to TRF size distribution by using a different randomly generated size distribution for each iteration.

[b] TRF-$E_{1/D}$ and TRF-E_{var} correlations round to the same two-decimal number.

normal curves plus an underlying random uniform distribution were the most similar to results obtained using the empirical distributions (Table 5). This distribution performed better by these criteria than a two-phase exponential-decay model, which had the best fit of built-in nonlinear equations according to analysis with the software GOSA. When the simulations were limited to 1,200 species (1% threshold), the correlation of H' and TRF-H' declined to 0.64 for the parametric distribution, whereas the correlation of H' and TRF-E_{var} was 0.82;

these values were the closest to those obtained by OTU sampling for all models tested.

When species abundance distributions were extended up to 10,000 species and a threshold of 1% was used, the correlation between H' and TRF-E_{var} remained high (0.73), whereas the correlation between H' and TRF-H' dropped to 0.15 (Fig. 3). Other TRF evenness indices also correlated well with H' and J', but the TRF species richness and diversity indices did not correlate well with any true indices (Table 6). The results were

TABLE 5. Evaluation of different models in creating correlation matrices similar to average empirical matrices[a]

Distribution type	Value for indicated threshold and group			
	1%		4%	0.1%
	OTU	Sequence	OTU	OTU
Uniform plus six normal curves[b]	0.98	0.92	>0.99	0.99
Two-phase exponential decay[c]	0.91	0.68	>0.99	0.99
Power[d]	0.80	0.70	ND	ND
Uniform[e]	0.65	0.54	ND	ND
Contrast[f]	**	**	**	*

[a] Values shown are Mantel statistics which reflect the congruence of (i) the average correlation matrix created using empirical TRF size distributions with (ii) the correlation matrix created using the model shown, performed over the same range in S_T (1,200 for OTU sampling and 4,000 for sequence sampling). ND, not determined.

[b] Distribution described in text.

[c] Within the analyzed size range (50 to 600 bp), TRF size distribution followed $y = a_1 \exp(-k_1 \times x) + a_2 \exp(-k_2 \times x) + b$, where x is TRF size minus 49. Based on fitted empirical data, a_1 was allowed to vary from 0 to 14, a_2 from 17 to 111, k_1 from 0.0025 to 0.03373, k_2 from 0.093 to 0.7513, and b from 0.12 to 1. The probability that a species had TRF in the analyzed size range varied between 0.2 and 0.34.

[d] Same as two-phase exponential-decay distribution, except that TRF size distribution followed $y = ax^b$; a varied from 90 to 120, and b varied from -0.87 to -1.

[e] The species was assigned a random TRF size between 1 and 1,000 bp; the analyzed size range remained 50 to 600 bp.

[f] Significance of difference between the uniform-plus-six-normal-curve and two-phase exponential-decay distributions in terms of absolute deviations from OTU sampling and sequence sampling correlations (two-tailed pairwise t test). **, $P < 0.01$; *, $P < 0.1$. The mean absolute deviation was smaller for the uniform-plus-six-normal-curve distribution for each comparison.
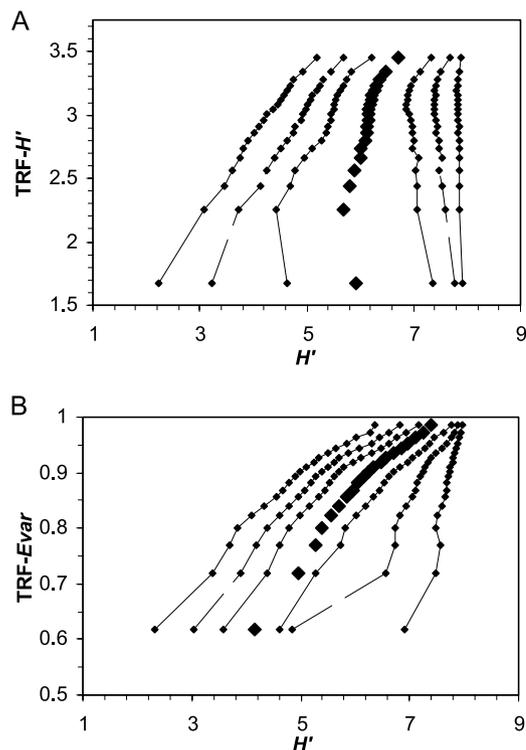
FIG. 3. Data density plots for parametric sampling simulations of communities with up to 10,000 species, using a relative abundance threshold of 1%. See the legend to Fig. 2 for explanation. n was 7,973 for each group.

TABLE 6. Representative correlation coefficients from parametric sampling simulations of communities with up to 10,000 species, using a TRF relative abundance threshold of 1%

Index	Value for indicated true index[a]			
	S	H'	$1/D$	J'
TRF-S	−0.07	0.02	−0.08	0.07
TRF-H'	−0.01	0.15	0.01	0.19
TRF-$e^{H'}$	0.00	0.19	0.06	0.23
TRF-$1/D$	0.05	0.30	0.16	0.33
TRF-$1/d$	0.11	0.40	0.31	0.42
TRF-J'	0.25	0.68	0.47	0.66
TRF-$E_{1/D}$	0.27	0.71	**0.57**	**0.69**
TRF-E_{var}	**0.30**	**0.73**	0.53	0.66

[a] The highest coefficients for each true index are shown in bold. Values shown incorporate variability due to TRF size distribution by using a different randomly generated size distribution for each iteration.

found to be essentially the same when a 4% threshold was used (correlation between H' and TRF-E_{var} of 0.56) (Table 4). In contrast, when a 0.1% threshold was used, all true indices correlated well with their corresponding TRF indices, with the best correlations between H' and TRF-H' (0.76) and between J' and TRF-J' (0.68) (e.g., Table 4).

DISCUSSION

Estimates of species richness in soil bacterial communities range from thousands to millions of species, whether based on sequencing of ribosomal genes (3, 31) or DNA reassociation kinetics (13, 39). Here, we asked whether differences in diversity between communities can be meaningfully portrayed by common diversity indices applied to T-RFLP profiles derived from bacterial communities. Our results demonstrate a large disparity between true community diversity and diversity estimates derived from T-RFLP data, indicating that diversity estimates based on T-RFLP inaccurately portray the true underlying diversity of bacterial communities.

When typical peak height or area thresholds (~1%) were used, richness and diversity indices applied to T-RFLP profiles had limited capabilities to discriminate between levels of diversity in the underlying community. The number of bands in T-RFLP profiles (TRF-S) did a very poor job at predicting the number of taxa in the community (S) in all simulations. Loisel et al. (26) found that the number of bands did not correspond with S for other community profiling methods (single-stranded conformation polymorphism and denaturing gradient electrophoresis profiles) and proposed that the Curtis estimator could be used to evaluate the diversity of the underlying community. The Curtis estimator is based on the reciprocal of the Berger-Parker index ($1/d$), which we did not find to be particularly useful here, as demonstrated by low correlations between TRF-$1/d$ and true community indices.

H' is a well-known and widely used diversity index integrating both species richness and evenness, although interpretation of the index itself is not without some conceptual difficulties (19, 21). TRF-H' did correlate well with the true community value of H' in the simulations using the lowest number of species (10 to 1,200). However, this correlation declined substantially in simulations using greater numbers of species, to the point where TRF-H' cannot be considered a useful metric

of diversity. We were surprised to find that one community evenness index applied to T-RFLP profiles, TRF-E_{var}, consistently had the highest correlations with the true community value of H' in all simulations using 1% or 4% thresholds (correlations above 0.71 to 0.81 and 0.56 to 0.66, respectively). The fact that different taxa can generate the same TRF (1, 8) may in part explain why evenness of T-RFLP profiles (reflected in TRF-E_{var}) is affected by true community richness as well as true evenness (reflected in true H'). An increase in species richness will likely increase the average number of taxa contributing to each TRF, which will reduce the impact of differences in abundance and increase the evenness of the profile. Additionally, our results may indicate that evenness within the most abundant taxa is often affected by the overall species richness of a community.

The factors captured in our simulations which cause T-RFLP diversity indices to underestimate, and differ independently from, true community diversity include (i) TRFs of the same size generated from multiple taxa, (ii) exclusion of TRFs if they are outside the size range reliably resolved during electrophoresis, and (iii) TRFs not being detected if they fall below a fluorescence threshold. This fluorescence threshold is translated to a relative abundance threshold by the standardization processes that have been recommended to reduce effects of analytical variability in overall profile strength (2, 5, 32). While there are some similarities between the threshold and the lognormal distribution "veil line" of Preston (34), they differ in that the veil line is a consequence of random sampling, whereas the threshold specifically excludes rare TRFs. Hence, there is no reason to expect the well-known relationship between diversity and sampling effort (14) to hold for T-RFLP profiles, if sampling effort is increased by repeating the assay on the same DNA extract. The relationship may or may not hold if sampling effort is increased by repeating the analysis on replicate environmental samples, depending on the spatial scale of the sampling relative to the spatial variability of the community (35).

Sampling effort can also be increased by enhancing the sensitivity of detection of rare TRFs. We found that an order-of-magnitude increase in the simulated sensitivity of T-RFLP, from a 1% threshold to 0.1%, completely alters the relationships between true diversity indices and TRF indices. Evenness indices are replaced by richness or integrated indices as the best predictors of true S, H', and $1/D$ for a 0.1% threshold. To our knowledge, current technology cannot be used to consistently differentiate peaks from background noise at this low level, so we present the analysis for comparative purposes. This sensitivity to fluorescence threshold highlights the importance of choosing a common threshold to be applied uniformly across profiles.

All diversity indices tested here, other than TRF-S, are based on comparisons of peaks within a single profile and the assumption that these comparisons reflect differences in relative abundances of taxa. It is worth noting that this assumption is also tenuous due to complex interactions between taxa in ribosomal copy number, PCR bias, and DNA extraction bias (11). It is important to note that analysis of profiles by ordination is based on comparisons between rather than within samples and on the more realistic assumption that biases affecting taxa are the same across samples. The simulations are

also conservative (in favor of diversity indices) because TRFs were assumed to be sized with single-base-pair resolution, which often cannot be achieved (22, 30).

While we have demonstrated that diversity indices applied to T-RFLP profiles do not tell us much about the underlying community diversity, empirical studies using the indices often find significant differences between communities from different environments. In other words, the indices do not vary randomly in the field. Significant differences could, however, arise from changes in the identities or relative abundances of the taxa present rather than species richness and evenness. Calculation of diversity indices results in an inevitable loss of information, as the community data are reduced to a single value. Indeed, we are unaware of any studies in which diversity indices indicated a difference between communities and multivariate methods (an ordination or cluster analysis) did not, whereas there are numerous studies where both methods of analysis indicated a difference between communities (9, 17, 20, 25) or where multivariate methods indicated a difference but diversity indices did not (4, 10, 15, 16, 40). It has been suggested that the greater sensitivity of multivariate methods for detecting community differences is true even across methods which seem optimal for each type of analysis (18).

From our analyses, we are led to conclude that T-RFLP-based estimates of diversity provide inaccurate insights into the actual diversity of bacterial communities, even as a comparative method. We suggest that previously published work using diversity indices applied to T-RFLP data should be reinterpreted. For example, Fierer and Jackson (9) found that diversity indices (TRF-S and TRF-H') and community ordinations of bacterial T-RFLP profiles were correlated with soil pH. Based on conclusions reached here, both of these analyses indicate that community composition responds to pH, but no reliable statements can be made regarding the diversity (richness or evenness) of those communities. In addition, there is little reason to apply diversity indices to T-RFLP data to detect differences in community composition, because multivariate methods are more sensitive and better suited for this task. If it is necessary to measure bacterial diversity using T-RFLP, we recommend using TRF-E_{var} as a measure of diversity because it was the index that best represented the true underlying community diversity.

REFERENCES

1. **Blackwood, C. B., and J. S. Buyer.** 2007. Evaluating the physical capture method of terminal restriction fragment length polymorphism for comparison of soil microbial communities. Soil Biol. Biochem. **39:**590–599.
2. **Blackwood, C. B., T. Marsh, S.-H. Kim, and E. A. Paul.** 2003. Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. Appl. Environ. Microbiol. **69:**926–932.
3. **Curtis, T. P., W. T. Sloan, and J. W. Scannell.** 2002. Estimating prokaryotic diversity and its limits. Proc. Natl. Acad. Sci. USA **99:**10494–10499.
4. **Dunbar, J., L. O. Ticknor, and C. R. Kuske.** 2000. Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal restriction fragment analysis. Appl. Environ. Microbiol. **66:**2943–2950.
5. **Dunbar, J., L. O. Ticknor, and C. R. Kuske.** 2001. Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities. Appl. Environ. Microbiol. **67:**190–197.
6. **Dunbar, J., S. M. Barns, L. O. Ticknor, and C. R. Kuske.** 2002. Empirical and theoretical bacterial diversity in four Arizona soils. Appl. Environ. Microbiol. **68:**3035–3045.
7. **Dykhuizen, D. E.** 1998. Santa Rosalia revisited: why are there so many species of bacteria? Antonie Leeuwenhoek **73:**25–33.
8. **Engebretson, J. J., and C. L. Moyer.** 2003. Fidelity of select restriction endonucleases in determining microbial diversity by terminal-restriction fragment length polymorphism. Appl. Environ. Microbiol. **69:**4823–4829.
9. **Fierer, N., and R. B. Jackson.** 2006. The diversity and biogeography of soil bacterial communities. Proc. Natl. Acad. Sci. USA **103:**626–631.
10. **Fierer, N., J. P. Schimel, and P. A. Holden.** 2003. Influence of drying-rewetting frequency on soil bacterial community structure. Microb. Ecol. **45:**63–71.
11. **Frey, J. C., E. R. Angert, and A. N. Pell.** 2006. Assessment of biases associated with profiling simple, model communities using terminal-restriction fragment length polymorphism-based analyses. J. Microbiol. Methods **67:**9–19.
12. **Fromin, N., J. Hamelin, S. Tarnawski, D. Roesti, K. Jourdain-Miserez, N. Forestier, S. Teyssier-Cuvelle, F. Gillet, M. Aragno, and P. Rossi.** 2002. Statistical analysis of denaturing gel electrophoresis (DGE) fingerprinting patterns. Environ. Microbiol. **4:**634–643.
13. **Gans, J., M. Wolinsky, and J. Dunbar.** 2005. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science **309:**1387–1390.
14. **Gotelli, N. J., and R. K. Colwell.** 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. **4:**379–391.
15. **Graff, A., and R. Conrad.** 2005. Impact of flooding on soil bacterial communities associated with poplar (*Populus* sp.) trees. FEMS Microbiol. Ecol. **53:**401–415.
16. **Grüter, D., B. Schmid, and H. Brandl.** 2006. Influence of plant diversity and elevated carbon dioxide levels on belowground bacterial diversity. BMC Microbiol. **6:**68.
17. **Hackl, E., S. Zechmeister-Boltenstern, L. Bodrossy, and A. Sessitsch.** 2004. Comparison of diversities and compositions of bacterial populations inhabiting natural forest soils. Appl. Environ. Microbiol. **70:**5057–5065.
18. **Hartmann, M., and F. Widmer.** 2006. Community structure analyses are more sensitive to differences in soil bacterial communities than anonymous diversity indices. Appl. Environ. Microbiol. **72:**7804–7812.
19. **Hill, T. C. J., K. A. Walsh, J. A. Harris, and B. F. Moffett.** 2003. Using ecological diversity measures with bacterial communities. FEMS Microbiol. Ecol. **43:**1–11.
20. **Janus, L. R., N. L. Angeloni, J. McCormack, S. T. Rier, N. C. Tuchman, and J. J. Kelly.** 2005. Elevated atmospheric CO$_2$ alters soil microbial communities associated with trembling aspen (*Populus tremuloides*) roots. Microb. Ecol. **50:**102–109.
21. **Jost, L.** 2006. Entropy and diversity. Oikos **113:**363–375.
22. **Kaplan, C. W., and C. L. Kitts.** 2003. Variation between observed and true terminal restriction fragment length is dependent on true TRF length and purine content. J. Microbiol. Methods **54:**121–125.
23. **Kent, A. G., and E. W. Triplett.** 2002. Microbial communities and their interactions in soil and rhizosphere ecosystems. Annu. Rev. Microbiol. **56:**211–236.
24. **Kim, S. H., and T. L. Marsh.** 2004. The analysis of microbial communities with terminal restriction fragment length polymorphism (T-RFLP), p. 789–808. *In* G. A. Kowalchuk, F. J. de Bruijn, I. M. Head, A. D. Akkermans, and J. D. van Elsas (ed.), Molecular microbial ecology manual, 2nd ed. Springer, New York, NY.
25. **LaMontagne, M. G., J. P. Schimel, and P. A. Holden.** 2003. Comparison of subsurface and surface soil bacterial communities in California grassland as assessed by terminal restriction fragment length polymorphisms of PCR-amplified 16S rRNA genes. Microb. Ecol. **46:**216–227.
26. **Loisel, P., J. Harmand, O. Zemb, E. Latrille, C. Lobry, J. Delgenès, and J. Godon.** 2006. Denaturing gradient gel electrophoresis (DGE) and single-strand conformation polymorphism (SSCP) molecular fingerprintings revisited by simulation and used as a tool to measure microbial diversity. Environ. Microbiol. **8:**720–731.
27. **Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle.** 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. Science **294:**804–808.
28. **Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A. W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. Konig, T. Liss, R. Lussmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K. H. Schleifer.** 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32:**1363–1371.
29. **Magurran, A. E.** 2004. Measuring biological diversity. Blackwell Publishing, Malden, MA.

30. **Moeseneder, M. M., J. M. Arrieta, G. Muyzer, C. Winter, and G. J. Herndl.** 1999. Optimization of terminal restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. **65:** 3518–3525.

31. **Neufeld, J. D., and W. W. Mohn.** 2005. Unexpectedly high bacterial diversity in arctic tundra relative to boreal forest soils, revealed by serial analysis of ribosomal sequence tags. Appl. Environ. Microbiol. **71:**5710–5718.

32. **Osborne, C. A., G. N. Rees, Y. Bernstein, and P. H. Janssen.** 2006. New threshold and confidence estimates for terminal restriction fragment length polymorphism analysis of complex bacterial communities. Appl. Environ. Microbiol. **72:**1270–1278.

33. **Pedrós-Alió, C.** 2006. Marine microbial diversity: can it be determined? Trends Microbiol. **14:**257–263.

34. **Preston, F. W.** 1948. The commonness, and rarity, of species. Ecology **29:** 254–283.

35. **Ranjard, L., D. P. H. Lejon, C. Mougel, L. Schehrer, D. Merdinoglu, and R. Chaussod.** 2003. Sampling strategy in molecular microbial ecology: influence of soil sample size on DNA fingerprinting analysis of fungal and bacterial communities. Environ. Microbiol. **5:**1111–1120.

36. **Ricke, P., S. Kolb, and G. Braker.** 2005. Application of a newly developed ARB software-integrated tool for in silico terminal restriction fragment length polymorphism analysis reveals the dominance of a novel *pmoA* cluster in a forest soil. Appl. Environ. Microbiol. **71:**1671–1673.

37. **Smith, B., and J. B. Wilson.** 1996. A consumer's guide to evenness indices. Oikos **76:**70–82.

38. Reference deleted.

39. **Torsvik, V., J. Goksoyr, and F. L. Daae.** 1990. High diversity of DNA of soil bacteria. Appl. Environ. Microbiol. **56:**782–787.

40. **Väisänen, R. K., M. S. Roberts, J. L. Garland, S. D. Frey, and L. A. Dawson.** 2005. Physiological and molecular characterisation of microbial communities associated with different water-stable aggregate size classes. Soil Biol. Biochem. **37:**2007–2016.