Terminal Restriction Fragment Length Polymorphism for Soil Microbial Community Fingerprinting

Jane T. Aiken*
Univ. of Western Sydney
Penrith South DC
NSW 1797, Australia

Terminal restriction fragment length polymorphism (TRFLP) fingerprinting of 16S rRNA genes is a popular technique for analyzing bacterial communities. This review considers the technical aspects of deriving a community fingerprint with the goal of improving TRFLP as a tool for community ecologists, particularly those working in soil science. Analysis parameters are reviewed for samples run with GeneScan and the capillary system Peak Scanner software of ABI Prism genetic analysis units. Analysis parameters should be chosen using principles that preserve community composition and define inherent community variability, ensuring that methodology does not bias the distribution of an unknown population structure. Choices for fragment sizing include sizing algorithm, sizing accuracy, peak detection, split peak correction, peak smoothing options, minimum peak half width, baseline setting, and analysis range. Results indicate that a standard protocol should include choices of local southern sizing and use of a 35 relative fluorescence unit baseline with light smoothing. A size calling accuracy of 0.998 (for GeneScan gels) is recommended for each internal lane standard, based on the principle that the extent of linearity across the size-standard selection is the given limit for an analysis range. Analysis baseline is a choice that in combination with heavy peak smoothing can halve the number of fragments detected. Heavy smoothing will increase the number of replicate polymerase chain reaction samples necessary to represent all fragments of a soil community DNA and it is proposed that using heavy smoothing will produce a biased community profile and loss of population information. Choices during profile and alignment editing are also discussed. Generated profiles should be edited to ensure a minimum integrity of one base pair resolution between consecutive fragments. Results indicate that profile editing, for a specified analysis range of 80 to 827 base pairs, can decrease the linear relationship between total fluorescence and fragment number and thus increase independence between abundance and species richness.

Abbreviations: bp, base pair; rfu, relative fluorescence unit; TRFLP, terminal restriction fragment length polymorphism.
is the process of fragment size calling with the use of programs such as GeneScan or Peak Scanner to develop a community profile or fingerprint. Variables of the fragment sizing programs are size calling algorithm, sizing accuracy, peak detection, split peak correction, peak smoothing options, minimum peak half width, baseline setting, and analysis range. The third stage is a profile editing step, applied to ensure that the integrity between subsequent peaks along a fingerprint has a minimum of one base-pair resolution. The last stage, Stage 4, is called profile alignment, a step that requires an alignment of fingerprints for their comparison. This is a review of the protocols for the fragment sizing, profile editing, and profile alignment steps to develop consistent methodology for community analysis.

The TRFLP technique has been used to characterize microbial communities from a wide range of environments (Thiess, 2007). For soil, it is particularly useful because the method enables unknown populations, as communities, to be defined with a value of abundance, based on the digital information of peak area generated by the fragment analysis software. Thus, interpretations may include those of conventional indices and comparisons of community (Aiken, 2006; Ulrich and Becker, 2006; Blackwood et al., 2007). The technical merits, however, of diversity indices and community diversity derived by TRFLP (Blackwood et al., 2007; Bent et al., 2007), the merits of standardizing methodology in TRFLP to elucidate similarities between bacterial communities (Osborne et al., 2006), and the application of post statistical analyses (Grant and Ogilvie, 2003; Aiken, 2006; Culman et al., 2008) continue to be discussed in the literature. For a soil microbial community assay based on the analysis of a community DNA template, one important consideration is that the methodology should not bias the distribution of an unknown population structure. Only with unbiased community fingerprints should the various numerical methods of ecology be applied to make further inferences of ecological process.

Terminal fragments generated by the combination of fluorescently labeled PCR and restriction digestion are separated by electrophoresis, detected and sized, and then collated as a profile or fingerprint. The underlying premise of the TRFLP method, when applied to a community, must rely on accurate fragment size calling, which has as its basis the use of a standard curve of known fragment sizes. For researchers who run gels with Applied Biosystems equipment, a freeware program called GeneScan (Applied Biosystems, Foster City, CA) was provided to determine sample fragment sizes or, for TRFLP, terminal fragment lengths. To support the current capillary gel electrophoresis systems, Applied Biosystems provides a freeware program called Peak Scanner. Both programs allow the conversion of the gel data into a format for further analyses by, for example, Gene Mapper (Applied Biosystems). Although data presented here was derived using the GeneScan software, GeneScan was superseded in favor of the capillary systems; many researchers still have GeneScan data and many older laboratories may still run the older ABI Prism genetic analyzer units (Applied Biosystems). The current software program Peak Scanner still retains the original features developed for GeneScan. This review discusses the principles involved in choosing analysis parameters for a TRFLP assay, where choices will influence the structural characteristics of the community fingerprint profile, irrespective of the software used but does not compare TRFLP data derived from the different electrophoresis platforms. The necessity of an objective post-electrophoresis analysis is discussed with the goal of improving TRFLP as a tool for community ecologists, particularly those working in soil science.

MATERIALS AND METHODS

The experimental samples discussed here are bacterial community DNA amplified from selected soil and water samples including: (i) topsoil (0–10 cm) irrigated with effluent, (ii) subsoil (30–40 cm) irrigated with effluent, (iii) prepared saline soil (electrical conductivity = 10 dS m⁻¹), and (iv) sewage treatment plant effluent. Sewage treatment plant effluent was concentrated by freeze-drying before amplification for fingerprinting. Procedural details of community DNA extraction, PCR reactions, enzyme digestion, sample purification, and GeneScan electrophoresis running conditions were as described in Aiken (2006). All fingerprints were prepared by combining 3- by 100-μL PCR reactions before restriction digestion, except if analysis was for single 100-μL reactions specifically.

Tamra 2500 internal lane standards and experimental electrophoresis results files were opened in GeneScan Fragment Analysis 3.1 software (Applied Biosystems). A maximum analysis range between 80 and 827 base pairs (bp) was covered by using the first 17 fragments of 37, 94, 109, 116, 172, 186, 223, 238, 268, 286, 361, 470, 490, 536, 827, and 1115 bp sizes of the Tamra standard. For each sample, all these corresponding Tamra 2500 fragments had been calibrated to 0.998 accuracy. Minimum peak half-widths were calculated from a base-pair range between 94 and 536 bp (a range of 442 bp using the second and third last marker fragments). The number of data points between 94 and 536 bp for each gel lane were recorded and the number was divided by the number of base pairs in this range (442 bp), then halved to obtain the peak half width. These were calculated for each water and soil sample, which had been run on different gels at different times. Data of each sample’s internal lane standard (Tamra 2500) were then tested to determine the number of fragments detected using light or heavy smoothing. Fragment numbers across a set of 10 replicate, single, 100-μL PCR reactions derived from the topsoil bacterial community DNA were manipulated within the GeneScan software using heavy and light smoothing at 35, 45, 50, and 60 relative fluorescence unit (rfu) baselines. These fingerprints were also used to assess the effect of different baseline choices, analysis ranges, profile editing, and alignment principles using accumulation plots and regression analysis between total fluorescence and fragment number.

RESULTS

Comparisons among four Tamra 2500 internal lane standards of known fragment size were undertaken to demonstrate the inherent variation of the size calling marker and thus assess the need for preparing a calibrated size calling curve for each experimental sample. Table 1 shows the same standard run in
light smoothing was lower (Table 2). The comparison between the relative difference in the reduction of fragment number and total fluorescence, thus decreasing the linear dependence between total sample fluorescence and fragment number, when sub-band editing was undertaken. This was investigated further by comparison of linear $r^2$ values as a measure of dependence. For a different baseline rfu, the $r^2$ value reflects the strength of the relationship between fragment number and total sample fluorescence at $P < 0.05$ (Table 3). The comparison before profile editing to a 1-bp resolution and with no profile editing found that the selection of analysis range and dependence between fragment number and total fluorescence was significant. Alternatively, after profile editing the dependence was less, with lower $r^2$ values.

**DISCUSSION AND REVIEW**

**Stage 2: Fragment Size Calling**

To produce a fingerprint electropherogram or digital file, there are various computational choices to be made. In GeneScan and the Peak Scanner software, these choices of size calling, peak detection, and analysis range will...
influence the final characteristics of a fingerprint. Of the eight fragment sizing variables or choices in GeneScan, seven are used by Peak Scanner. Two variables, baseline setting and the choice of sizing algorithm, have been previously discussed for TRFLP fingerprinting microbial community analyses (Osborn et al., 2000; Kitts, 2001; Liesack and Dunfield, 2004). This review further investigates these and the other parameters that affect peak detection and size calling.

Output for all wavelengths detected within the Applied Biosystems fragment size analysis programs is contained within a single sample file for each gel lane. In Peak Scanner, sample files are formatted to an .fsa extension. This sample file contains all the information detected at each spectral wavelength, as tracked on the gel. In GeneScan, the results consist of five columns of “digital” data, which are (i) the peaks detected as the channel or wavelength detected (as B1, B2, B3, ... B45: blue 1–45), (ii) the time of detection, (iii) peak height, (iv) peak area, and (iv) data point number. Alternatively, the output of Peak Scanner will provide these and additional gel data. In both programs, the fragments detected and sized are referred to as peaks and the results are visualized as electropherograms and may be exported to spreadsheet files.

**Choice of Sizing Algorithm**

GeneScan version 3.1 and Peak Scanner version 1.0 provide sizing algorithms including second-order least squares, third-order least squares, global southern, local southern, and cubic spline. Each algorithm will produce different peak sizes from one sample file, so consistency is necessary. Osborn et al. (2000) reported that the local southern method was the most appropriate algorithm of those available. Technical advice from Applied Biosystems, the protocol published by Liesack and Dunfield (2004), and Thies (2007) support this decision. Indicative of the local southern sizing algorithm is that three points are used for two curves averaged, the first using two known points below the unknown and one above and the second curve using one known point below and two above the unknown (Southern, 1979). It is recommended that only points between defined size standards be used.
A fundamental requirement of the TRFLP method is control of the accuracy and precision of fragment sizing. The ability to accurately call the sizes of fragments is related to designating the known fragment sizes in base pairs of a molecular marker that is mixed with the sample. In GeneScan, the linear accuracy of size standards is derived for each sample file on the basis of manually prepared standard curves. The recommended accuracy for all size calling standard curves prepared from known fragments is an \( r^2 \) of 0.998 (A. Maslem, Applied Biosystems, personal communication, 2001). This was prepared by manually defining the known fragment sizes of each sample internal lane standard. In this manner, the unknown fragments on any gel lane could be sized from a size-calling curve relative to its sample lane electrophoresis condition. For large-scale comparisons, the between-gel biases could be eliminated and samples run on different gels objectively compared when standard curves are prepared to a similar quality. Thus, a size calling curve, prepared to an equivalent accuracy for each sample (e.g., 0.998), ensures that all unknown fragments for all samples have been sized relative to their own gel running conditions.

Alternatively, the newer software uses the distance between the peaks in the size standard definition to work out peak identity and performs an automatic calculation (L. Cullen, Applied Biosystems, personal communication, 2009). Subsequent quality of the linearity across the size standard definition is then defined for a specified analysis range and a Size Quality Invalidated (SQI) value is derived. The SQI value represents a pass or fail for the program to continue. This aspect for TRFLP analysis is worth considering for any comparative microbial studies with many samples. Generally, the accuracy of size calling was not included in protocols reporting any slab-gel electrophoresis studies, but now Peak Scanner requires all marker fragment sizes to be included or the program will not proceed, nor will the program proceed if linearity is below the desired settings.

Fig. 1. (cont.) Fragment accumulation curves of replicate polymerase chain reaction (PCR) reactions using (A) a 35 or 45 relative fluorescence unit (rfu) baseline and (B) a 50 or 60 rfu baseline, heavy or light smoothing, and four analysis ranges of base pairs (bp).
Table 3. The relationship of determined fragment number to total profile fluorescence under baseline conditions from 30 to 80 relative fluorescence units (rfu) without profile editing in an undefined analysis range of 0 to 1115 base pairs (bp) or with profile editing in a defined analysis range of 80 to 827 bp. A reduced linear dependence was demonstrated after subband editing for community profiles.

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Relationship of total fluorescence to fragment number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No profile editing</td>
</tr>
<tr>
<td></td>
<td>$r^2$</td>
</tr>
<tr>
<td>rfu</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.72</td>
</tr>
<tr>
<td>35</td>
<td>0.68</td>
</tr>
<tr>
<td>50</td>
<td>0.52</td>
</tr>
<tr>
<td>80</td>
<td>0.63</td>
</tr>
</tbody>
</table>

**Peak Detection**

The following reviews the main choices to detect peaks. These choices are made after the choice for size calling, which included the algorithm and accuracy. They are included to highlight the distinct possibilities that subtle differences to a community profile can be produced by choices made during software analysis.

**Multicomponent Spectral Definition and “Baselining.”** In Applied Biosystems software, peaks are detected with respect to the intensity of blue, green, yellow, or red color channels. For peak detection as defined by GeneScan, it was necessary to choose the multicomponent feature to adjust for spectral overlap of the fluorescent dyes used. In conjunction with baselining, these two adjustments provided a better comparison of relative signal intensity between each wavelength (Applied Biosystems, 1997) and need to be activated within an analysis. In Peak Scanner, these analysis parameters are implemented automatically as program defaults.

**Split-Peak Correction.** Split-peak correction is associated with the use of an internal lane standard made from double-stranded DNA and can be used to correct for denatured digest products. Without using a split-peak correction in GeneScan, the program might resolve a difference for the two strands of the same internal standard peak, and more fragments may be detected for the marker or sample. Although this does not directly affect the fragment size called, to reduce artifact peaks a correction choice would be a standard setting for a comparative assay. Alternatively, Peak Scanner includes split-peak correction without necessitating a choice.

**Peak Smoothing Options.** Peak smoothing changes the configuration of the peak shoulders and, as such, can change the number of fragments present within a fingerprint. GeneScan and Peak Scanner provide this option to help reduce the number of false peaks (Applied Biosystems, 1997). The choices are none, heavy, and light and are provided to compensate for factors of the running conditions. For slab gels run in GeneScan mode, the none option should be selected if the data have very sharp narrow peaks of interest. Light smoothing provides the best results for normal data. Heavy smoothing is applied to data from slower runs that have very broad peaks. In early TRFLP assays, both Clement et al. (1998) and Kaplan et al. (2001) used the heavy smoothing option. Clement et al. (1998) ran a 16-h gel but Kaplan et al. (2001) did not specify the gel running time. Liesack and Dunfield (2004) recommended the light smoothing option but the choice for peak smoothing still remains arbitrary and has not been discussed in terms of a community assay. The data presented here also support the use of the light smoothing option. Peak Scanner data should be run without smoothing.

The fingerprint profiles presented here are shown as examples of manipulation on the basis of choice of peak-smoothing options. Peak-smoothing options produced different community population structures (Fig. 1). This raises several issues. First, there is a need for consistency when undertaking comparative analyses. Second, there can be no assumption other than the one of objectivity. For example, accumulation curves using heavy smoothing may mask detected fragments. For community analyses, this means that the community structure (richness) can be manipulated. Further work would be needed to determine if a comparative statistical difference would be computed on the basis of peak smoothing options. Thus, it is recommended to define the smoothing option for any reported community fingerprint.

**Minimum Peak Half Width**

Peak half width is a peak detection choice that determines the number of data points to constitute a peak or the size of a single base pair. The setting of the peak half width for each sample lane on the gel is the actual relationship between the number of data points of laser detection, the length of the gel, and the number of base pairs along that length; this may vary for different gel runs. In GeneScan, it is possible to determine a peak width value for each sample file, a practice also possible in Peak Scanner, so that samples are analyzed independently and there is no introduced lack of precision across multiple gel comparisons.

Peak half width is an aspect of size calling in a GeneScan analysis that has only been noted once, in a technical paper by Troth et al. (2002), concerning the use of multicapillary electrophoresis for microbial community analysis. Peak half width is also associated with normalizing any effect of gel running conditions. Different lanes or capillaries may run faster or slower and thus the distance along the gel as defined by its data points may vary across a corresponding number of base pairs. A data point/base pair ratio is defined for each sample. Although generally a straightforward ratio, defining this parameter is not straightforward because the assumption within the software is that the peak half width is an integer, not a decimal.

If the data point/base pair ratio is 10, then the minimum peak half width would be 5. If the ratio is 9 then the peak half width would be 4.5. Thus, the values for half widths need to be rounded up or down. Theoretically, if raised to the higher integer, fewer peaks would be detected than if rounded down. Thus, by rounding down and subsequently conducting “profile editing,” as discussed below, to ensure that each peak cannot be smaller than one base pair, fragments would not be underdetected. In addition, it is known that different fragment sizes may run at differ-
ent rates during electrophoresis and the data point ratio may be different along the gel. While the capillary technique produces more consistent running conditions, there remains no easy way to produce differential peak half widths using either GeneScan or Peak Scanner. Thus, the analysis assumption remains that all fragment sizes at any location along the gel have the same base pair/data point ratio. This may be assumed in parallel with the linearity of size standards.

When conducting a comparative assessment of bacterial community fingerprints, uniformity can be maximized by ensuring that, for each gel lane, the gel running conditions have been independently and consistently defined. The provisions for independently defined samples as presented here included minimum peak half widths (Table 1), as a data point/base pair ratio and to ensure the quality of size calling accuracy. In this respect, to attain comparative integrity, the accuracy of one internal lane standard to represent all gel samples should not be relied on.

Additional fragments may be impurities as vagrant bands in a marker. For example (Table 1), fragments in a Tamra 2500 standard across a 94- to 536-bp range may include up to 24 bands rather than the 14 used to set up a size standard curve across this range. Such contaminant fragments are readily distinguished from fragments of known size when each of the known fragments is “defined” manually on the basis of the base pair/data point ratio for the gel lane or accurately defined automatically by the software. What this implies is that if a peak is detected, it will show in the fingerprint. Therefore, it is important that fingerprints are derived systematically and objectively for both the known and unknown gel lanes.

Baseline Setting

Baseline choice is an important factor needing careful consideration for the detection of fragments by the analysis software. A baseline corresponds directly to the height of a peak as a Gaussian curve expressed in rfu. The baseline choice determines a minimum height for fragments to be included within a fingerprint. Generally, baseline setting is an arbitrary determination. By changing the baseline setting at the peak detection stage of fragment analysis, the reproducibility of a TRFLP profile can be manipulated to contribute to sample reproducibility (Osborn et al., 2000). As Kitts (2001) noted, the robust sample reproducibility reported by Osborn et al. (2000) was primarily a function of the baseline setting of 100 rfu, an observation that was valid because at 25 rfu in the study of soil TRFLP profiles by Dunbar et al. (2000), sample reproducibility was low. Gonzalez et al. (2000) also found high variability between replicate samples of bacterioplankton, but when the smaller peaks, at a level less than 4% of the total sample fluorescence, were taken out, their average coefficient of variation (by peak area) decreased from 70 to 13%.

Baseline choice concerns issues of reproducibility between profiles and fragment numbers and thus the determination of richness values. Liesack and Dunfield (2004) published a standardized protocol for TRFLP with 35 rfu as the baseline threshold, but the baseline threshold is rarely noted as relevant (e.g., Singh et al., 2006). A choice of 35 rfu used by all researchers for microbial community assays would provide an unbiased technical protocol for comparative TRFLP fingerprints. The adoption of a standard 35 rfu baseline reduces any options for theoretical choices. There is also evidence of dependence between the choice of analysis range and baseline, and between total fluorescence and baseline. Results of the accumulation curves (Fig. 1) with different baselines support the assumption that a baseline at 35 rfu will reduce potential bias in community population structure. The data show that increasing the baseline has more effect on community structure and population reproducibility if coupled with the heavy smoothing option because fragment numbers can be reduced by half. Thus, light smoothing and a low baseline (i.e., 35 rfu) provides a consistent basis for an assessment of community structure.

Analysis Range

The total fluorescence values calculated from a GeneScan fingerprint profile result from the choice of analysis range. Consequently, this is an issue for comparative studies because all fragments within the sample should accurately represent the population of interest. Unfortunately, some fragments can be an artifact of the GeneScan analysis parameters or derived from the technical process of fluorescent labeling of DNA primers. The latter is shown by the phenomenon of a 35- to 37-bp size fragment frequently representing a considerable proportion of the total fluorescent units in a sample (a result of not cleaning the sample after the restriction digestion of the original PCR product before electrophoresis). Marsh (1999) proposed that an effective analysis range was from 50 to 550 bp.

Choosing an analysis range is important when abundance values expressed as proportional frequencies are derived from the GeneScan rfu values for each sample. Consequently, if an analysis range is short, then the analysis range can also be a selective process against potential diversity in community composition. Technically, the analysis range is related to maintaining linearity across the size standard selection.

Stage 3: Profile Editing

Profile editing as proposed here, establishes an objective rather than a subjective process to define the integrity of consecutive fragments within a fingerprint. Profile editing is not a GeneScan process. Its scope covers the definition of fingerprint resolution to 1 bp and determination of the final size of a fragment (peak area or height as a relative abundance value) and occurs as an additional step.

The process of profile editing needs to be undertaken within a tabulated data set in a spreadsheet rather than manually from an electropherogram within a fragment analysis program. Essentially, profile editing is based on an ordinary assumption that a nucleotide within a DNA sequence should not be <1 bp. Profile editing ensures that, from the smaller to the larger fragments along the fingerprint, there is a continuous 1-bp resolution between those fragments. In this manner, subsequent fragments.
of the same size can be combined by adding their peak heights and areas to produce a single fragment.

One reason editing is required is that fragment size data are calculated to two decimal places in both GeneScan and Peak Scanner, with peaks sized at a fraction of one base pair. Theoretically a decision needs to be consistent, but resolution should not be < 1 bp. Rounding up using a spreadsheet computation does not solve this aspect because it is still possible to have two fragments called at the same size. Previous researchers called this editing process **binning** (Marsh, 1999), made manually while aligning in a spreadsheet layout (Dunbar et al., 2001); however, the process of profile editing is different than a decision to bin fragments. Ultimately, a method of profile editing needs to be simple and reproducible. The application of a standard method is particularly relevant for a large-scale comparative analysis of communities.

**Binning** is a grouping or classing process and was raised as an issue by Kitts (2001). It also makes an assumption that consecutive fragments are not equivalent along a gel. However, if data point/base pair ratios are checked and the sizing accuracy of each set of standards for each gel lane is consistent, then the earlier proposals for binning are not needed. Binning also contributes to subjectivity. Dunbar et al. (2001) reported a method that gathered aligned peaks in a 0.5-bp bin, thus placing one sample in a 1-bp bin when aligning fragment sizes. This is not practical because it is subjective and it relies on change done as part of the process to align same fragments sizes across multiple samples in a spreadsheet.

Determination of a fingerprint’s integrity by profile editing, before alignment across a spreadsheet as proposed here, ensures that the fragment analysis software has not called the same size for two consecutive bands. It also identifies stutter bands produced during the PCR reaction (A. Maslam, Applied Biosystems, personal communication, 2002), which occur as artifacts of PCR extension cycles (Speksnijder et al., 2001). By profile editing along fingerprint gel data, the value of total sample fluorescence remains unchanged and relative abundance values remain accurately derived. For example, typical stutter bands can be identified as a large peak with a smaller peak on either side. An example of stutter bands would be consecutive fragments of sizes 104.56, 105.45, and 105.90, which would be one fragment of 105.45 bp. If not identified within the fingerprint and rounded up, there would be two bands, 105 and 106 bp, and thus an inaccuracy would be introduced.

This type of editing along a profile, termed **profile editing**, was developed here to define fingerprint integrity for comparative data sets and is used to ensure a minimum resolution of 1 bp between each fragment along the fingerprint. The resulting fingerprint data can then be called as integers, for example 189 and 190 from values that have been computed as 189.45 and 190.6, irrespective of decimal place figures. The method also solves the issue of subjective profile alignments, creating a basis for an objective TRFLP methodology. The logic of this step is in the application of consistency and creation of a method of objectivity that can be applied for any sample from any gel analysis. Examples (Table 2) demonstrate that the choice of light or heavy smoothing does not eliminate the need to objectively define consecutive fragments to 1-bp resolution. The results indicate that the total number of fragments will differ with and without profile editing. In the case of profile editing for the light smoothing choice, there were more bands differing by < 1 bp requiring editing and more bands overall. Profile editing resulted in a consistent number of peaks whether the profiles were subjected to heavy or light smoothing.

Profile editing also reduces the linear dependence between sample fluorescence and fragment number. For example, the relationship between total sample fluorescence and fragment number was significant (Table 3) based on different baseline values for the 10 replicate PCR reactions. Additional editing for sub-bands to ensure that consecutive profile fragments had at least 1 bp difference achieved nonlinearity ($R^2 < 0.5$) between fragment number and sample fluorescence when coupled with the use of a restricted base-pair range of 80 to 827 bp. Consequently, profile editing to ensure that peaks were a minimum of 1-bp size difference provides an objective method for consistency and reduces the dependence between fragment numbers and cumulative sample fluorescence. In terms of community analysis, this means that richness and abundance could be evaluated independently.

### Stage 4: Profile Alignment and Community Data

Stage 4 in a TRFLP assay is the step that aligns the digital data of multiple fingerprint profiles in a spreadsheet matrix for computation. Profile alignment is the ordering in rows and columns of the peak size, height, or area for comparison with other fingerprint profiles. This is necessary for comparative analysis of multiple samples and is best done by manual spreadsheet row and column manipulations. Data matrices are either binary or abundance. For binary as either presence or absence, the comparative matrix will be filled with zeros and ones. For abundance comparisons, the relative fluorescence values are required. In both cases, to achieve an objective method for profile alignment, several aspects need consideration.

For a large-scale comparison among many fingerprints with many fragments at 1-bp resolution, making decisions about whether fragments are the same or not is subjective. Subjective decisions can lead to a loss of analytical integrity or unique community composition, with researchers finding they have to make decisions implying that the analysis parameters set within the fragment analysis program had called the fragment size incorrectly. Decisions about peak alignment were reviewed by Kitts (2001) to address the issue of fractional peak lengths. If fingerprints were previously checked by profile editing, however, there is no subsequent confusion. The use of integer values without reference to the decimal fraction does not rely on artificially determined sizes or lead to variability when comparative multiple-gel analyses are required. Trotha et al. (2002) confirmed by empirical testing that precision based on the integer and not the decimal values is an appropriate choice.
Community Data

The other step necessary when abundance values and quantitative comparisons are to be undertaken is normalization of sample fluorescence, because there is no guarantee that the amounts loaded for electrophoresis are equivalent (Marsh, 1999) even when using capillary electrophoresis.

The process of normalizing the fluorescent signal can be complicated (Dunbar et al., 1999) or as simple as using a proportional frequency value (Lukow et al., 2000). What is important is the process of normalization of the total signal that enables each sample to be independent within any multiple sample comparison. Essentially, the aim of normalization for comparative profiles is to control variability and contribute to reproducibility. Several methods have been used. The first method, by Kaplan et al. (2001), multiplied the sum of the ratio of peak area by the smallest possible peak (a value of 580 area units) detected by GeneScan at a 50 rfu baseline to create a new minimum threshold value. They did not, however, provide a rationale for the choice of a 50 rfu baseline.

The second method, by Dunbar et al. (1999), normalized all samples on the gel to the lowest value of fluorescence to reduce variability. By adjusting all peak heights accordingly, this method results in all sample fluorescence peak heights equal to the sample with the lowest total fluorescence. This strategy is not practical for true comparative analyses because one cannot add or remove samples within the comparison without recalculating values for each fragment, nor is it practical for large data sets or for anything other than a single gel investigation.

The third method derives a proportional relative peak area for each peak (Lüdemann et al., 2000) calculated from the total fluorescence. This method is useful for high-throughput analysis.

A fourth method uses the actual value of fluorescence of each peak without normalization (Fey and Conrad, 2000). This latter choice also leads into the debate concerning absolute or relative values of peak heights or peak area (Suzuki et al., 1998; Kitts, 2001), a matter seemingly resolved because recent publications have been consistently reporting percentage peak areas (e.g., Nagashima et al., 2003). Liesack and Dunfield (2004) did not advocate any choice and some recent studies have still adopted the method for normalization of Dunbar et al. (1999) (e.g., Heath and Saunders, 2006). For all example data within this review, the method of Lüdemann et al. (2000) of proportional relative fluorescence units for peak area was used.

For abundance computations, the choice of either peak height or peak area is necessary. The issues for either choice have been discussed (Kitts, 2001; Osborn et al., 2000). Some researchers use peak height (such as Dunbar et al., 1999), others use peak area (Lukow et al., 2000). Alternatively, some still follow the method of Dunbar et al. (1999), such as Sessitsch et al. (2001). As Kitts (2001) rightly noted, however, a small peak height can be associated with a large peak area and is omitted if the baseline is higher than the relative peak height. Thus, for determinations using peak height, it is necessary to understand that height is related to the baseline and, when varied, this will affect the relative abundance values. In this review, the effect of different baselines was demonstrated.

CONCLUSIONS

To ensure that valid assumptions of ecological inference can be attributed to a biological process, the methodology for obtaining the values of abundance needs to be understood and documented. This review has noted aspects of the GeneScan fragment size calling software and the more recent Peak Scanner fragment size calling software for capillary separations, explaining the implications of analysis parameters for microbial community comparisons.

The GeneScan and Peak Scanner software for fragment analysis include a suite of analysis parameters that contribute to fingerprint composition. Consistency is needed for the choices of baseline, peak smoothing, and analysis range. In addition, there are further manipulations, such as profile editing and fingerprint alignment, which will contribute to the final population structure of a microbial community. A standard protocol would include those previously published choices of local southern sizing and the use of a 35 rfu baseline with light smoothing in conjunction with the determination of a consistent peak minimum half width and size calling accuracy for each sample file. For GeneScan gels, size calling accuracy is recommended to be 0.998 for each internal lane standard. Finally, the choice of an analysis range should be consistent with the extent of linearity across the size standard selection, with the resulting generated profiles or fingerprints edited to ensure a minimum integrity of 1-bp resolution. These attributes are recommended as a consistent methodology for TRFLP fingerprinting, particularly when assessing inherently variable populations of a soil microbial community.

An investigation of the effect on fingerprint results of heavy or light peak-smoothing options showed that for replicate PCR samples, the heavy peak-smoothing option reduced but did not remove the necessity to edit out sub-bands below 1-bp difference. Heavy smoothing may actually increase the number of replicate PCR samples necessary to represent all fragments of a soil community DNA and may produce a biased community profile and possible loss of population information. Light or no smoothing options are recommended.

Profile editing along each fingerprint ensures an objective integrity for each community to a minimum, 1-bp resolution. It is one way of reducing between-sample variability in artifact fragments and is an objective strategy for further comparative community analyses. Its use will influence the number and types of peaks detected, but it provides a logical basis for comparison that has been derived objectively for all samples. Evidence also indicates that profile editing for a specified analysis range also eliminates artifact primer peaks and decreases the linear relationship between total fluorescence and total fragment number. Thus, profile editing and defining an analysis range will improve the independent integrity for population studies assessing species richness and evenness. For microbial community assays, population distribution and diversity will differ in different environments,
and it is important to establish a basis for sample comparisons that are consistent for any sample from any environment.

ACKNOWLEDGEMENTS

The author extends a grateful thank you to the manuscript reviewers, and especially to the Associate Editor Jeanette Norton, Professor of Soil Microbiology, Utah State University.

REFERENCES


