T-Align, a web-based tool for comparison of multiple terminal restriction fragment length polymorphism profiles

Cindy J. Smith [a,b,1], Bret S. Danilowicz [a,2], Adrian K. Clear [c], Fintan J. Costello [c], Bryan Wilson [b], Wim G. Meijer [b,*]

[a] *Department of Zoology, University College Dublin, Dublin 4, Ireland*
[b] *Department of Industrial Microbiology, Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin 4, Ireland*
[c] *Department of Computer Science, University College Dublin, Dublin 4, Ireland*

Abstract

Terminal restriction fragment length polymorphism (tRFLP) is a potentially high-throughput method for the analysis of complex microbial communities. Comparison of multiple tRFLP profiles to identify shared and unique components of microbial communities however, is done manually, which is both time consuming and error prone. This paper describes a freely accessible web-based program, T-Align (http://inismor.ucd.ie/~talign/), which addresses this problem. Initially replicate profiles are compared and used to generate a single consensus profile containing only terminal restriction fragments that occur in all replicate profiles. Subsequently consensus profiles representing different communities are compared to produce a list showing whether a terminal restriction fragment (TRF) is present in a particular sample and its relative fluorescence intensity. The use of T-Align thus allows rapid comparison of numerous tRFLP profiles. T-Align is demonstrated by alignment of tRFLP profiles generated from bacterioplankton communities collected from the Irish and Celtic Seas in November 2000. Ubiquitous TRFs and site-specific TRFs were identified using T-Align.
© 2005 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

Keywords: tRFLP; Computer algorithm; Community analysis

1. Introduction

Since the first descriptions of terminal restriction fragment length polymorphism (tRFLP), its potential as a means to describe and monitor changes in microbial community structure was recognized [1–3]. As a result, tRFLP quickly became a widely used tool for analysis of complex natural microbial communities. tRFLP is based on the amplification of the 16S rRNA gene with a fluorescent label attached to the 5′ end of one or both primers followed by digestion of the PCR product with frequently cutting restriction enzymes. The sizes of the resulting terminal restriction fragments (TRFs) containing the fluorescent label are subsequently precisely determined using an automated fragment length analysis system.

tRFLP is a fast, cost-effective, high-resolution and reproducible method for examining microbial diversity. Data produced by tRFLP have been shown to be consistent with data from clone libraries [4] and are interchangeable with other PCR-based molecular techniques used to examine microbial diversity [4,5]. The method has been used to analyze bacterial, fungal and archaeal community structures from environments as diverse as fish intestines

---

* Corresponding author. Tel.: +353 1 7161364; fax: +353 1 7161183.
  E-mail address: wim.meijer@ucd.ie (W.G. Meijer).
[1] Present address: Department of Biological Sciences, University of Essex, Wivenhoe Park, Colchester C04 3SQ, United Kingdom.
[2] Present address: Allen E. Paulson College of Science and Technology, Georgia Southern University, P.O. Box 8044, Statesboro, GA 30460, USA.

[6], rat feces [7], soil [4,8,9], rice fields [10–12], marine sediments [13], marine bacterioplankton [5] and deep sea hydrothermal vents [14] and functional groups [15].

The aims of analyzing bacterial communities are often to determine their diversity, to identify species that are present and to compare communities separated in space and/or time. The number of peaks and peak area in a tRFLP profile immediately give insight into the richness and evenness of the population. In addition, software has been developed to obtain phylogenetic information from tRFLP profiles by correlation of empirical data with profiles obtained in silico using sequences deposited in 16S rRNA sequence databases [16–18]. These programs provide clues as to the identity of the species present in the community that is being analyzed. A recurring problem in tRFLP analysis is the comparison of profiles. This is first required when comparing duplicate tRFLP profiles in order to eliminate spurious peaks. In addition, tRFLP profiles need to be compared to determine whether a bacterial species is present in all communities analyzed or is unique to one. To date, comparison of T-RFLP profiles obtained from different samples to identify shared or unique TRFs is carried out manually, which is very time consuming, subjective and error prone.

The aim of this study was to develop a program that could be used for comparison of duplicate tRFLP profiles resulting in the generation of a consensus profile showing only the TRFs that occur in both profiles. In addition, the program would compare tRFLP profiles obtained from different communities resulting in a file that shows whether a TRF is present or absent in a particular profile. The resulting program, T-Align, is freely available on a dedicated website.

2. Methods

2.1. Sample collection and nucleic acid extraction

Water samples were taken from stations in the Irish and Celtic Seas in November 2000 (Fig. 1). A 7 l Niskin bottle was lowered to within 15 m of the sea floor (sample depths ranging from 16 to 117 m) from which water samples were taken. The water sample (5 l) was filtered through a Whatman GF/C membrane onto a 0.2 μm pore polycarbonate membrane (Millipore, MA, USA) by vacuum filtration. The filter was removed and frozen at −70 °C with 200 μl of TE (10 mM Tris–HCl pH 8.0, 1 mM EDTA) buffer. Nucleic acids were extracted from frozen samples using a modified version of the QIAamp mini DNA extraction kit (Qiagen, Crawley, UK). Lysozyme (20 mg/ml) was added to the filter and incubated for 30 min at 37 °C; after this initial step the Qiagen protocol was followed, but no vortexing of the sample took place.
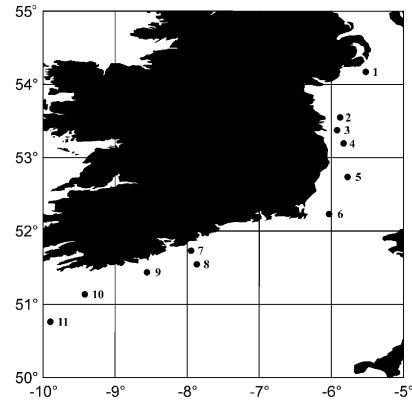

Fig. 1. Map of the positions of the stations surveyed along the Irish coast November 2001. At each site 5 l of water was collected from 15 m off the sea bed. The individual stations are indicated by black dots and numbers.

2.2. tRFLP

Bacterial 16S rRNA genes were amplified from total environmental DNA extracts, using the primer pair F63-D4 (5′-CAGGCCTAACACATGCAAGTC-3′) and R1389 (5′-AGCGGCGGTGTGTACAAG-3′) [19,20]. PCR reactions were carried out as recommended by the manufacturer (Promega, Madison, USA) except that primers were used at a concentration of 0.4 μM. The reaction mixture was incubated at 94 °C for 2 min followed by 27 cycles of 94 °C for 45 s, 55 °C for 1 min and 74 °C for 2 min and a 7-min extension step at 74 °C. The PCR products were purified with QIAquick PCR Purification Kit (Qiagen, Crawley, UK) according to the manufacturer's instructions. The PCR products were digested with MspI for 2 h at 37 °C according to the manufacturer's instructions (Roche, Sussex, UK). After digestion, 5 μl of the digest was desalted by ethanol precipitated in the presence of glycogen (5 μg). The samples were dried and dissolved in 5 μl of deionised formamide. DNA (1–2 ng) was subsequently added to 40 μl of deionised formamide and 0.5 μl of the 60–640 bp CEQ size standard (Beckman Coulter, CA, USA). The samples were analyzed using a Beckman Coulter CEQ 2000 XL DNA Analysis System for 60 min at 4.8 kV. The size of the fragments was determined using CEQ Fragment Analysis Software 2000XL, using a quartic polynomial to size the fragments and a percentage peak value of 1% and a slope threshold of 5.

3. Results and discussion

3.1. Accuracy of TRF size

Prior to comparison of tRFLP profiles it is essential that the error in calling duplicate TRF sizes in base pairs is known. To determine this, identical samples were ana-

lyzed in duplicate. For TRFs to be within 0.5 bp size of each other the standard deviation of specific TRF sizes generated in the duplicate comparisons must be less than 0.35. The average standard deviation of duplicate TRF size (ranging in size from 60.0 to 640.0 bp; $n = 300$) was 0.13, showing that the size of a TRF was determined within 0.5 bp. Subsequently, the variation in duplicate percentage peak areas was determined. It was shown that there was an average difference of only 2.7% in the percentage area of each peak between duplicate samples (data not shown). Previous literature cites errors between duplicate runs of up to 7% and 11% [4,20].

3.2. T-Align

Since tRFLP analysis is carried out using automated fragment length analysis systems, it is in principle a high-throughput method. However, although data can be generated rapidly, subsequent manual data analysis is cumbersome and subjective. An algorithm was therefore developed that compares tRFLP profiles in a statistically objective procedure, thus allowing processing of many tRFLP profiles without introducing human bias. The basis for the T-Align algorithm is detailed below after this brief overview of the program. T-Align initially generates a sample profile, which is constructed by comparison of replicate tRFLP profiles of the same sample. The sample profile only contains TRFs that occur in all replicate profiles (Fig. 2(a)), resulting in the removal of pseudo TRFs [21]. In a second step different sample profiles are compared and a file containing the TRFs of all sample profiles and their relative fluorescence intensity is produced (Fig. 2(b)). TRFs in different profiles that differ in size by 0.5 bp or less are considered the same and are aligned using the moving average algorithm. However, as this is instrument-dependent, this parameter can be changed by the user. T-Align can also be used for comparison of profiles that were not generated in duplicate. In this case the 'duplicate' stage of the program will be skipped when T-Align does not detect replicate profiles in the input file.

3.2.1. Alignment of replicated profiles using T-Align

Initially identical TRFs were identified in replicated tRFLP profiles using a moving average procedure. T-Align identified the smallest TRF present among all replicate profiles, and marked the tRFLP profile containing this fragment as 'used'. The remaining profiles were subsequently searched for TRFs that are up to 0.5 bp (TRFs within 0.5 bp are considered identical, see above) larger than the initiating TRF. Each profile can only contribute a single TRF to the overall alignment. The average size of all TRFs identified in this manner was determined and *only* profiles that did not contribute a TRF in the initial search were searched again for a TRF within +0.5 bp of the average TRF size. If a new

(a)

Replicate profile 1		Replicate profile 2		Sample profile	
TRF (bp)	Peak area	TRF (bp)	Peak area	Consensus TRF (bp)	% relative peak area
106.40	73466	106.54	74081	106.47	0.68
107.50	649493	107.49	476545	107.495	5.08
109.08	71621	108.79	49995	108.935	0.55
109.88	151150	109.77	117999	109.825	1.22
111.07	129978	110.96	97551	111.015	1.03
117.35	2497577	117.33	2192618	117.34	21.38
127.72	101538	127.69	74067	127.71	0.79
130.85	135695	142.80	109380	0	0.00
142.85	143291	156.99	144919	142.825	1.14
157.09	199959	253.57	107999	157.04	1.55
253.54	134655	403.39	251328	253.56	1.10
403.38	339749	405.39	309066	403.39	2.67
405.36	399007	407.07	2080181	405.38	3.20
407.05	2489027	409.15	2240136	407.06	20.77
409.15	2567410	432.82	213661	409.15	21.91
430.74	151399	454.65	252940	0	0.00
432.88	298002	456.69	419267	432.85	2.30
441.65	137266	459.33	166013	0	0.00
454.65	337476	461.13	163680	454.65	2.67
456.70	559639	508.79	151846	456.70	4.42
459.36	224748			459.35	1.76
461.21	203939			461.17	1.67
463.35	164393			0	0.00
508.83	234676			508.81	1.73

(b)

TRF	Sample Stations Profiles (% relative fluorescence)										
	1	2	3	4	5	6	7	8	9	10	11
106.70	0.00	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
107.43	2.72	5.08	0.00	0.00	2.67	0.00	0.00	0.00	2.21	0.00	2.12
108.94	0.00	0.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
109.72	0.58	1.22	0.85	0.00	1.35	0.00	0.00	0.00	0.00	0.00	1.28
110.97	0.00	1.03	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.74
117.38	19.01	21.38	11.27	35.09	33.45	15.98	19.94	14.28	22.09	24.12	25.88
127.61	0.87	0.79	2.10	0.00	1.98	1.23	0.00	5.19	1.33	1.20	1.83
129.25	0.00	0.00	0.46	0.00	0.00	0.00	0.00	9.66	1.51	0.00	0.00
130.86	0.92	0.00	1.87	0.00	0.00	0.00	0.00	4.68	0.00	0.00	0.98
137.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.47	0.00	0.00	0.00
141.81	0.95	1.14	2.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.11
157.04	0.00	1.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Fig. 2. Examples of input and output files of T-Align. (a) Two replicate tRFLP profiles from sample station 2 were compared and converted into a single, derived, consensus profile. The TRF size in the derived consensus profile contains the average size of the TRFs and the percentage fluorescence intensity of total fluorescence. For example, two TRFs with size 106.40 and 106.54 bp that are present in replicate profiles 1 and 2 are represented by TRF of 106.47 bp in the consensus profile, representing 0.68% of total fluorescence (box with solid line). Note that the number of TRFs in the sample profile is sum of all different TRFs present in the replicate profiles. Where a TRF is not present in all replicate profiles, the actual TRF size and fluorescence is given as '0'. For example, TRF of size 130.85 bp only occurs in replicate profile 1, and is therefore represented by '0' in the derived consensus profile (box with interrupted line). (b) Final output of T-Align. The program provides a list of TRF sizes and their relative fluorescense intensity. Any value higher than '0' indicates the presence of a TRF in sample profile. Not all data are shown in this figure due to large table generated when all TRFs are displayed.

TRF was identified, the process was repeated using a new average TRF size. All TRFs used in this average were marked as 'used'. When none was found, the entire process was repeated with the smallest 'unused' TRF among all profiles. This resulted in an aligned profile containing the average sizes of all TRFs found.

3.2.2. Construction of a consensus profile

All TRFs in all replicate profiles were again started as 'unused'. Starting with the smallest TRF, each TRF in all replicate profiles was checked against the first TRF in the aligned profile generated above. The smallest TRF in each profile within ±0.5 bp of the aligned profile TRF was found and marked as 'used'. Only if a matching TRF was found in every replicate profile was the TRF retained in the final consensus profile, and the peak fluorescence (abundance) of each matched replicate TRF was used to calculate average peak fluorescence for that consensus TRF. This resulted in a single consensus profile that only contained TRFs occurring in all replicate profiles with their corresponding average peak area. The peak areas of TRFs were subsequently normalized by representing each value as a percentage of total fluorescence (Fig. 2(a)).

3.2.3. Comparison of consensus profiles

Using the moving-average search method described above, all consensus profiles were used for creating a single master environmental profile. To align consensus profiles against the master environmental profile, an environmental matrix was constructed, with each column representing a final sample profile and each row a peak size from the master profile. Starting with the smallest TRF in this master environmental profile, the smallest TRF in each consensus profile within ±0.5 bp of this master TRF was found, marked as "used", and its associated peak area was placed into the corresponding sample/peak size location in the final environmental matrix. If no matching TRF could be found in a particular sample profile, a zero was placed into the corresponding location in the environmental matrix, representing an absence of that TRF in that consensus profile. The search then continued with the second TRF in the master file, and only 'unused' TRFs were searched in the consensus profiles. This resulted in a 'comparison' matrix, which contained all consensus profiles compared with all others, with each point containing either a zero in the absence of a TRF or the relative percentage fluorescence when the TRF present was present in a particular consensus profile (Fig. 2(b)). These profiles lend themselves to Bray–Curtis or other ordination statistics. These same profiles can be transformed to presence/absence matrixes, which can be compared using a binary matching statistic such as Jaccard's Correlation.

3.3. Program implementation and web interface

The T-ALIGN program is written as a stand-alone application. This program can be used in two ways, both of which are available from the T-Align web page at http://inismor.ucd.ie/~talign/. Intensive users can download the application from this webpage and compile and run it on their own machines. (The application is written in Java to allow maximum portability and ease of compilation.) It is expected, however, that most users will make use of the web interface to the program, also provided at the same T-Align web page. At this page users can upload an Excel spreadsheet that contains two columns; the first contains each TRF size in base pairs from the tRFLP profile and the second contains the corresponding TRFs peak area fluorescence. A duplicate tRFLP profile of the same sample should be in the two adjacent columns.
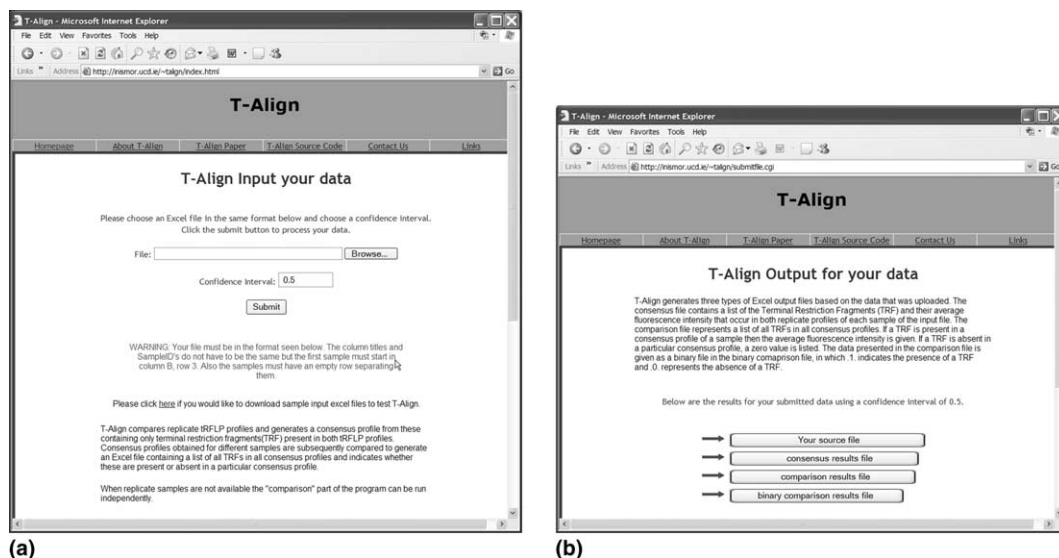

Fig. 3. Input and output screens of the T-Align web interface. (a) Data input screen. By clicking on the 'browse' button the user can select an Excel spreadsheet file located on a local drive. The default setting of the confidence interval is 0.5 bp, but can be changed by the user. By clicking the 'submit' button the consensus and comparison data are computed. These can be accessed from the output screen (b) by clicking on the relevant buttons.

An example of the correct format of the Excel spread sheet is provided on the web page. Duplicate tRFLP profiles of additional samples should be placed below the first sample and be separated by a single empty row. After uploading of the Excel file containing the tRFLP data to be compared, the user is presented with a page containing four buttons (Fig. 3). When these are selected they will provide an Excel file containing, respectively, the input file uploaded by the user, a derived consensus tRFLP profile containing only the TRFs present in both duplicate samples and their relative fluorescence intensity, an Excel file showing whether a TRF is present or absent in the individual samples as well as their relative fluorescence as a percentage of total fluorescence in that particular profile, and finally an Excel file showing simply whether a TRF is present or absent in the individual samples (without relative fluorescence). By default the T-ALIGN program uses a confidence interval of 0.5 bp when comparing TRFs, so that TRFs within 0.5 bp of each other are considered identical. When uploading their TRF and peak area fluorescence data, however, users have the option of selecting a different value for this confidence interval.

3.4. Example

In order to demonstrate the use of T-Align, bacterial populations present along the east coast of Ireland were compared. Water samples were collected from 11 sampling stations in the Irish and Celtic Seas 15 m from the sea bed (Fig. 1). Following DNA extraction, duplicate tRFLP profiles were generated from each sample and were compared in T-Align to generate a consensus profile for each of the 11 sampling stations that only contains the TRFs that occur in both duplicate profiles to eliminate spurious peaks. Individual TRFs that are within 0.5 bp are represented as the average size in bp; their peak area is given as a percentage of the total fluorescence (Fig. 2(a)). Subsequently, the consensus profiles generated for each of the 11 stations were compared using T-Align to show whether a particular TRF is present at a particular sampling station. The first column of the resulting Excel file lists all TRFs that were unambiguously identified in the 11 sampling stations. The subsequent 11 columns give the average percentage fluorescence for each of the TRFs. Absence of a TRF in a particular station is indicated by '0' fluorescence (Fig. 2(b)). The resulting file readily reveals whether a TRF is present or absent in any particular sample. For example, TRF 108.94 is only present in samples from station 2, whereas TRF 117.38 is present in all stations. Finally, the output from T-Align can easily be converted into a binary form, which can be used as input in for example a Jaccard similarity index (data not shown).

4. Conclusion

The main bottleneck in studies employing tRFLP to analyze complex microbial communities is not the generation of data, but the analysis of large data sets. Although this can be done manually, it is a very time consuming and error prone process. GeneMapper (ABI, Beckman Coulter) in conjunction with an Excel macro written by Rinehart [22] can be used to align amplified fragment length polymorphism data and to convert these into binary data sets. However, this software is not suitable for tRFLP data as it aligns whole numbers only. In addition, in contrast to the proprietary 'Genemapper' software, the web-based T-Align is freely available, and thus requires no investment on behalf of the researcher. The use of T-Align makes comparative studies of complex microbial communities feasible, by combination of high-throughput capabilities of automated fragment analysis systems and a fast and a statistically robust algorithm that can be adapted to all fragment analysis systems. The resulting output files can readily be used in statistical analysis such as Bray–Curtis ordination or be used to infer phylogenetic information.

References

[1] Avaniss-Aghajani, E., Jones, K., Chapman, D. and Brunk, C. (1994) A molecular technique for identification of bacteria using small subunit ribosomal RNA sequences. Biotechniques 17, 144–149.

[2] Avaniss-Aghajani, E., Jones, K., Holtzman, A., Aronson, T., Glover, N., Boian, M., Froman, S. and Brunk, C.F. (1996) Molecular technique for rapid identification of mycobacteria. J. Clin. Microbiol. 34, 98–102.

[3] Liu, W.T., Marsh, T.L., Cheng, H. and Forney, L.J. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. 63, 4516–4522.

[4] Dunbar, J., Ticknor, L.O. and Kuske, C.R. (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] Moeseneder, M.M., Arrieta, J.M., Muyzer, G., Winter, C. and Herndl, G.J. (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.

[6] van der Maarel, M.J.E.C., Artz, R.R.E., Haanstra, R. and Forney, L.J. (1998) Association of marine archaea with the digestive tracts of two marine fish species. Appl. Environ. Microbiol. 64, 2894–2898.

[7] Kaplan, C.W., Astaire, J.C., Sanders, M.E., Reddy, B.S. and Kitts, C.L. (2001) 16S ribosomal DNA terminal restriction fragment pattern analysis of bacterial communities in feces of rats fed *Lactobacillus acidophilus* NCFM. Appl. Environ. Microbiol. 67, 1935–1939.

[8] Edel-Hermann, V., Dreumont, C., Perez-Piqueres, A. and Steinberg, C. (2004) Terminal restriction fragment length polymor-

phism analysis of ribosomal RNA genes to assess changes in fungal community structure in soils. FEMS Microbiol. Ecol. 47, 397–404.

[9] Lukow, T., Dunfield, P.F. and Liesack, W. (2000) Use of the T-RFLP technique to assess spatial and temporal changes in the bacterial community structure within an agricultural soil planted with transgenic and non-transgenic potato plants. FEMS Microbiol. Ecol. 32, 241–247.

[10] Chin, K.J., Lukow, T. and Conrad, R. (1999) Effect of temperature on structure and function of the methanogenic archaeal community in an anoxic rice field soil. Appl. Environ. Microbiol. 65, 2341–2349.

[11] Derakshani, M., Lukow, T. and Liesack, W. (2001) Novel bacterial lineages at the (sub)division level as detected by signature nucleotide-targeted recovery of 16S rRNA genes from bulk soil and rice roots of flooded rice microcosms. Appl. Environ. Microbiol. 67, 623–631.

[12] Ludemann, H., Arth, I. and Liesack, W. (2000) Spatial changes in the bacterial community structure along a vertical oxygen gradient in flooded paddy soil cores. Appl. Environ. Microbiol. 66, 754–762.

[13] Scala, D.J. and Kerkhof, L.J. (2000) Horizontal heterogeneity of denitrifying bacterial communities in marine sediments by terminal restriction fragment length polymorphism analysis. Appl. Environ. Microbiol. 66, 1980–1986.

[14] Campbell, B.J. and Cary, S.C. (2001) Characterization of a novel spirochete associated with the hydrothermal vent polychaete annelid, *Alvinella pompejana*. Appl. Environ. Microbiol. 67, 110–117.

[15] Perez-Jimenez, J.R. and Kerkhof, L.J. (2005) Phylogeography of sulfate-reducing bacteria among disturbed sediments, disclosed by analysis of the dissimilatory sulfite reductase genes (dsrAB). Appl. Environ. Microbiol. 71, 1004–1011.

[16] Cole, J.R., Chai, B., Farris, R.J., Wang, Q., Kulam, S.A., McGarrell, D.M., Garrity, G.M. and Tiedje, J.M. (2005) The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res. 33, D294–D296.

[17] Kent, A.D., Smith, D.J., Benson, B.J. and Triplett, E.W. (2003) Web-based phylogenetic assignment tool for analysis of terminal restriction fragment length polymorphism profiles of microbial communities. Appl. Environ. Microbiol. 69, 6768–6776.

[18] Michel, F. and Sciarini, S. (2003) T-RFLP FRAGSORT: a Computer Program to Correlate Multiple 16S rRNA Gene T-RFLP Profiles with in silico Amplification and Digestions of Ribosomal Database Project II Alignments. Abstracts of the 103th Generall Meeting of the American Society for Microbiology.

[19] Marchesi, J.R., Sato, T., Weightman, A.J., Martin, T.A., Fry, J.C., Hiom, S.J., Dymock, D. and Wade, W.G. (1998) Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. Appl. Environ. Microbiol. 64, 795–799.

[20] Osborn, A.M., Moore, E.R.B. and Timmis, K.N. (2000) An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ. Microbiol. 2, 39–50.

[21] Egert, M. and Friedrich, M.W. (2003) Formation of pseudo-terminal restriction fragments, a PCR-related bias affecting terminal restriction fragment length polymorphism analysis of microbial community structure. Appl. Environ. Microbiol. 69, 2555–2562.

[22] Rinehart, T.A. (2004) AFLP analysis using GeneMapper® software and an Excel® macro that aligns and converts output to binary. Biotechniques 37, 188–194.