APPLIED AND ENVIRONMENTAL MICROBIOLOGY, July 1995, p. 2798–2801 Vol. 61, No. 7
0099-2240/95/$04.00+0

Effect of Genome Size and *rrn* Gene Copy Number on PCR Amplification of 16S rRNA Genes from a Mixture of Bacterial Species

VINCENT FARRELLY, FREDERICK A. RAINEY, AND ERKO STACKEBRANDT*

DSM-German Collection of Microorganisms and Cell Cultures GmbH, 38124 Braunschweig, Germany

Received 21 October 1994/Accepted 11 April 1995

In order to assess the effect of genome size and number of 16S rRNA genes (rDNAs) on the quantities of PCR-generated partial 16S rDNA fragments, equimolar amounts of DNA from pairs of different species for which these parameters are known were subjected to gene amplification. The experimentally determined ratio of PCR products obtained, as determined by image analysis of SYBR-Green I-stained amplification products, corresponded well with the predicted ratio calculated from the number of *rrn* genes per equimolar amounts of DNA in mixtures of *Escherichia coli* and "*Thermus thermophilus*" and of *Pseudomonas aeruginosa* and "*T. thermophilus*." The values for the pair of *Bacillus subtilis* and "*T. thermophilus*" showed greater deviations from the predicted value. The dependence of the amount of 16S rDNA amplification product on these two parameters makes it impossible to quantify the number of species represented in 16S rDNA clone libraries of environmental samples as long as these two parameters are unknown for the species present.

As the quantification of microbial populations is the ultimate goal of molecular ecology studies, it appears important to place more emphasis on the identification of factors that might influence individual steps in the process. However, the present lack of knowledge regarding the influence of all parameters involved in such analysis excludes quantitative conclusions. Each step in community analysis is open to error or bias, including cell lysis (17), DNA extraction and purification (10, 15, 29), choice of primers (19), PCR conditions (2, 12, 21, 32), and cloning (19). Some of these restrictions, such as the formation of chimeric PCR products (8), can be detected, while others can be minimized by improved PCR conditions (21). Besides errors related to methodology, which may be overcome with improved techniques, the genomic properties of a bacterial cell, such as the size and number of genomes within a cell as well as the organization and number of rRNA genes (rDNAs), may constitute a previously uninvestigated source of bias. Genome size, genome number, and organization of *rrn* operons vary widely among prokaryotic species. The size of the genome may range from 600 to 13,000 kb (3, 22). Even the number of genomes may vary; e.g., *Brucella melitensis* contains two genomes of 2,100 and 1,150 kb (13). The number of genes coding for 16S rRNA can vary from 1 (3) to 14 (34). Sequence heterogeneity within 16S rRNA genes of a given species has been detected (16), but the extent to which these sequence deviations occur is not well studied. In most cases genes coding for rRNA species are linked in *rrn* operons, but exceptions to this rule have been described (5, 9, 28). Copies of *rrn* operons are mostly located on the same chromosome, but location on two different genomes has been reported (3). Differences in genome size (27, 34) and in numbers (33) and types (11) of *rrn* operons have been detected for very closely related species and even for strains of a single species.

For phylogenetic studies of pure cultures, the effect of these genomic properties on the PCR amplification of 16S rDNA has not been considered important as long as an unambiguous sequence is obtained. However, such genomic properties need to be taken into account when drawing conclusions from molecular ecological studies involving PCR amplification of 16S rRNA by using DNA that has been isolated either directly from the environment or from isolated cells. The question of whether the percentage distribution of clones determined to have identical or very similar sequence can be correlated with the quantitative abundance of species from which the DNA was extracted has not yet been addressed. It was therefore the goal of this study to determine if the genome size and the number of *rrn* genes influence the amounts of PCR amplification products by comparing the predicted amounts of products with the amounts of products obtained.

Bacillus subtilis DSM 402, *Escherichia coli* DSM 498, *Pseudomonas aeruginosa* DSM 1707, and "*Thermus thermophilus*" DSM 579 were grown on media as indicated in the DSM catalog of strains. DNA was extracted from bacterial colonies or pellets as described previously (18). The DNA was further purified with a Prep-a-Gene kit (Bio-Rad, Hercules, Calif.) according to the manufacturer's instructions, with the exception that the DNA was eluted in 45 μl of sterile distilled H_2O. The DNA yield and purity were determined spectrophotometrically by the A_{260}/A_{280} ratio.

The 16S rDNA primers used were 10-30f (5′GAGTTTGATCCTGGCTCAG3′) and 530r (5′G[A/T]ATTACCGCGGCGGCTG3′) (26) synthesized by Pharmacia Biotech, Roosendaal, The Netherlands. The target sites of these oligonucleotide primers are conserved in the majority of organisms of the domain *Bacteria* (31). PCR amplification was performed in a total volume of 100 μl containing 0.5 μg of each primer, 10 μl of 10× reaction buffer (Boehringer Mannheim, Mannheim, Germany), 1 mM (each) dATP, dGTP, dTTP, and dCTP, and a total amount of 200 ng of DNA, as indicated in the tables. The reaction mixtures were incubated in a thermal cycler (Perkin-Elmer thermal cycler 480) for 3 min at 98°C. Two units of *Taq* polymerase (Boehringer Mannheim) was added, and the reaction was performed for 28 cycles of 55°C for 1 min, 72°C for 2 min, and 93°C for 1 min. A final extension step of 72°C for

* Corresponding author. Mailing address: DSM—German Collection of Microorganisms and Cell Cultures GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Germany. Phone: 49 531 2616 352. Fax: 49 531 2616 418. Electronic mail address: stackebrandt@venus.gbf-braunschweig.d400.de.

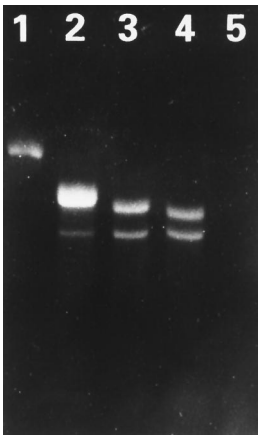

FIG. 1. SYBR-Green I-stained 3% Metaphor agarose gel of PCR products obtained with primers 10-30f and 530r targeting 16S rDNA in mixtures containing equal concentrations of DNA. Lane 1, molecular weight marker III (Boehringer Mannheim); lane 2, *B. subtilis* and "*T. thermophilus*"; lane 3, *E. coli* and "*T. thermophilus*"; lane 4, *P. aeruginosa* and "*T. thermophilus*"; lane 5, negative control. The lower bands indicate the amplification product of the 16S rDNA from "*T. thermophilus*."

5 min was performed. The annealing temperature and amount of DNA used were optimized for the PCR. An annealing temperature of 55°C was found to give the most specific PCR products for all four bacterial strains. As acetamide has previously been shown to affect the outcome of PCRs with DNA isolated from mixtures of archaeal and eukaryotic organisms (21), the effect of 5% acetamide was investigated, but no difference was found under the conditions described above.

Initially, temperature gradient gel electrophoresis was investigated as a method of separating PCR products of similar size. However, a large number of bands were detected, rather than the two expected bands (data not shown). The occurrence of multiple bands could be due to the use of PCR primers lacking a GC clamp. As the use of this method would involve the redesign of the universal primers most commonly used in molecular ecology studies and introduce possible bias, the separation of fragments by temperature gradient gel electrophoresis (14) was not continued. Agarose gel electrophoresis (1.5%) did not resolve the fragments satisfactorily; therefore, Metaphor agarose (3%) was used for the visualization of PCR products. Five-microliter quantities of the PCR solutions were loaded onto a vertical gel consisting of 3% (wt/vol) Metaphor agarose (FMC, Rockland, Maine) in Tris-borate-EDTA prepared according to the manufacturer's instructions. Gels were run at 5 V/cm for 14 to 16 h at 4°C and subsequently were stained with SYBR-Green I (Molecular Probes, Eugene, Oreg.) according to the manufacturer's instructions (Fig. 1). Gel images were converted to digitized files by using a Variocam video camera and Centronics software (Biotec-Fischer, Reiskirchen, Germany). The images were analyzed with Image Quant 3.3 (Molecular Dynamics), in which the pixel densities of the bands were transformed into pixel intensity ratios. The use of the stain required correction of the experimentally determined values, as larger fragments bind more stain than smaller fragments. Therefore, correction values were determined in order to quantify the amount of DNA present in any given band. The correction factor was determined by staining the molecular weight markers III and VIII (Boehringer Mannheim), which had been separated on Methaphor agarose and stained with SYBR-Green I. These factors were used to correct the predicted values in order to compare the theoretical values with experimentally determined values (see Table 1).

In order to determine the effect of genome size and *rrn* operon copy number on the amounts of PCR amplification products of 16S rRNA genes from mixtures of bacterial strains, a simplified model of a bacterial community was constructed. The organisms used and their relevant properties were as follows: (i) *B. subtilis* DSM 402, with a 4,165-kb genome, 10 *rrn* copies (6, 7), and, with the primer pair targeting conserved regions of the 16S rDNA, a PCR product of 541 bp; (ii) *E. coli* DSM 498, with a 4,595-kb genome, 7 *rrn* copies (25), and a PCR product of 534 bp; (iii) *P. aeruginosa* DSM 1707, with a 5,900-kb genome, 4 *rrn* copies (20, 23), and a PCR product of 513 bp; and (iv) "*T. thermophilus*" DSM 579, with a 1,740-kb genome, 2 *rrn* copies (1, 30), and a PCR product of 502 bp. *B. subtilis*, *E. coli*, and *P. aeruginosa* were analyzed in combination with "*T. thermophilus*." Other combinations were shown to yield unsatisfying separations of amplification products, even when the optimal separation protocol was used.

The predicted ratios of PCR products and the corresponding experimentally determined values from eight independent PCR amplifications with DNAs of the three pairs of organisms are shown in Table 1. In the mixtures of *P. aeruginosa* with "*T. thermophilus*" and of *E. coli* with "*T. thermophilus*," the deviations from the predicted fragment ratio were 31 and 23%, respectively. The degree of variation appears small considering the possible introduction of errors in the calculation of DNA

TABLE 1. Effect of genome size and copy number of 16S rRNA genes on the amounts of PCR amplification products of partial 16S rDNA fragments[a]

Pair of organisms	Amt of DNA (ng) per 100-μl PCR mixture	Genome size (kb)	Genome size ratio	16S rDNA copy ratio	16S rDNA copies per genome	16S rDNA copy ratio per 100 ng of DNA of each bacterial genome	Predicted band intensity (%)	Size of 16S rDNA PCR fragments (bp)	% Error due to SYBR-Green I binding to larger fragments	Adjustment to predicted band intensity	Predicted band ratio	Exptl determined ratio of PCR products	SD	Deviation from predicted value (%)
P. aeruginosa/"*T. thermophilus*"	100/100	5,900/1,740	1/3.39	4/2	2/1	1/1.69	37.1/62.9	513/502	0.8	37.9/62.1	1/1.64	1/1.23	0.497	31
E. coli/"*T. thermophilus*"	100/100	4,595/1,740	1/2.64	7/2	3.5/1	1/0.75	57.1/42.9	534/502	2.5	59.6/40.4	1/0.68	1/0.521	0.214	23
B. subtilis/"*T. thermophilus*"	100/100	4,165/1,740	1/2.39	10/2	5/1	1/0.48	67.6/32.4	541/502	3.3	70.9/29.1	1/0.41	1/0.11	0.03	73

[a] Results are averages from eight experiments.

TABLE 2. Effect of equal 16S rRNA gene copy ratio on the amounts of PCR amplification products of partial 16S rDNA fragments

Pair of organisms	16S rDNA copy ratio per 100 ng of each bacterial genome	DNA (ng) required for equal 16S rDNA copy no. ratio for 200 ng per PCR	Predicted band intensity (%)	Adjustment to predicted ratio according to SYBR-Green I binding (%)	Predicted band ratio	Exptl determined ratio of PCR	Deviation from predicted value (%)
P. aeruginosa/"*T. thermophilus*"	1/1.69	125.6/74.3	50/50	50.8/49.2	1/0.968	1/0.818	15
E. coli/"*T. thermophilus*"	1/0.75	85.7/114.3	50/50	52.5/47.5	1/0.904	1/0.651	28
B. subtilis/"*T. thermophilus*"	1/0.48	64.8/135.2	50/50	53.3/46.7	1/0.876	1/0.225	71

concentrations and/or in pipetting. In contrast, the experimental value determined for the pair of *B. subtilis* and "*T. thermophilus*" is 73%. It seems unlikely that this rather high value is due solely to experimental errors. The values were cross-checked by adjusting the DNA concentration in each mixture to give an equal copy ratio of *rrn* genes (Table 2). For the three mixtures the deviation factors were 15, 28, and 71%, respectively, which are of the same order of magnitude as found in the main experiment. The effect of equal genome ratio for the pair of *E. coli* and "*T. thermophilus*" was determined (Table 3), and the degree of variation from the predicted value was 13%.

In two of the DNA combinations, i.e., DNAs obtained from *E. coli* and "*T. thermophilus*" and from *P. aeruginosa* and "*T. thermophilus*," the yields of experimentally determined PCR products match the predicted values very closely. It can thus be concluded that the relative ratio of DNAs of these two partners can be determined on the basis of the relative ratio of PCR products of 16S rDNA. In order to apply this to cell numbers, other factors, such as the absence of free DNA, recovery and lysis of cells, DNA extraction, and DNA quality, need to be considered. This study has shown that knowledge of the genome size and the copy number of *rrn* genes is necessary but that even having such information does not enable the exact prediction of ratios, as shown in the experiment with *B. subtilis* and "*T. thermophilus*." The higher deviation from the predicted value in that case cannot be explained. A possible source of error could be the difference in base composition of the amplified stretch of 16S rDNA. The G+C content of 16S rDNA ranges between 53 to 55% for the mesophilic strains used in this study, while "*T. thermophilus*" has a higher value of 65.7%. It can be concluded from the values obtained with "*T. thermophilus*" in combination with *E. coli* and *P. aeruginosa*, for which the predicted and the experimental values were similar, that the high G+C content of "*T. thermophilus*" has no influence on the PCR product ratio in combination with *B. subtilis*. It appears rather that under the conditions applied, PCR amplification of DNAs from *B. subtilis* and "*T. thermophilus*" is influenced by intrinsic but unknown features of the *B. subtilis* genome. Possible factors include 16S rDNA proportions and the locations of *rrn* operons on the genome, as follows. (i) Of the three partner species of "*T. thermophilus*," *B. subtilis* has a slightly smaller genome but the highest number of rRNA genes. In *B. subtilis* 0.36% of the genome is composed of 16S rRNA genes; this value is significantly lower for *P. aeruginosa* (0.1%) and slightly lower for *E. coli* (0.22%). (ii) In contrast to *E. coli* (4) and *P. aeruginosa* (23), in which *rrn* operons are more evenly distributed on the genome, 7 of the 10 *rrn* operons of *B. subtilis* are closely linked, with one pair and one triplicate being directly adjacent (24). The density of *rrn* operons could possibly increase the efficiency of PCR amplification, but no experimental data are currently available.

Recent studies on 16S rDNAs amplified directly from nucleic acids that have been extracted from natural environments have demonstrated that the vast majority of bacterial species are uncultured. Information on the two parameters whose effect on the amount of PCR amplification products of 16S rDNA has been investigated in this study is completely lacking for the vast majority of described species and for all of the uncultured microbial diversity. Since this information is required before attempts to quantify the most simple community structures (i.e., mixtures of two species) can be made, it can be concluded that with the methods currently available, quantification of microbial communities is not possible from analysis of 16S rDNA clone libraries.

TABLE 3. Effect of equal genome ratio on the amounts of PCR amplification products of partial 16S rDNA fragments

Parameter	Value for *E. coli*/"*T. thermophilus*" pair
Genome size ratio	1/2.64
DNA (ng) required from each strain for equal genome no. ratio (200 ng)	145/55
Copy ratio of 16S rDNAs for equal genome ratio	1/0.286
Predicted band intensity (%)	77.8/22.2
Adjustment to predicted ratio according to SYBR-Green I binding (%)	80.3/19.7
Predicted band ratio	1/0.245
Exptl determined ratio	1/0.212
Deviation from predicted value (%)	13

Thanks are expressed to Bert Ingelmann (Biologische Bundesanstalt für Landwirtschaft und Forsten) for work on temperature gradient electrophoresis, to Ester Wellens (Gesellschaft für Biotechnologische Forschung [GBF]) for help with image analysis, and to Edward Moore (GBF) for helpful comments.

REFERENCES

1. **Borges, K. M., and P. L. Bergquist.** 1993. Genomic restriction map of the extremely thermophilic bacterium *Thermus thermophilus* HB8. J. Bacteriol. **175:**103–110.
2. **Cariello, N. F., W. G. Thilly, J. A. Swenberg, and T. R. Skopek.** 1991. Deletion mutagenesis during polymerase chain reaction: dependence on DNA polymerase. Gene **99:**105–108.
3. **Cole, S. T., and I. S. Girons.** 1994. Bacterial genomics. FEMS Microbiol. Rev. **14:**139–160.
4. **Condon, C., J. Philips, Z. Y. Fu, C. Squires, and C. L. Squires.** 1992.

Comparison of the expression of the seven ribosomal RNA operons in *Escherichia coli*. EMBO J. **11:**4175–4185.
5. **Fukunaga, M., and I. Mifuci.** 1990. Unique organization of *Leptospira interrogans* rRNA genes. J. Bacteriol. **171:**5763–5767.
6. **Green, C. J., G. C. Stewart, M. A. Hollis, B. S. Vold, and K. F. Bott.** 1985. Nucleotide sequence of *Bacillus subtilis* ribosomal RNA operon, *rrn* B. Gene **37:**261–266.
7. **Itaya, M., and T. Tanaka.** 1991. Complete physical map of the *Bacillus subtilis* 168 chromosome constructed by a gene-directed mutagenesis method. J. Mol. Biol. **220:**631–648.
8. **Kopczynski, E. D., M. M. Bateson, and D. M. Ward.** 1994. Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. Appl. Environ. Microbiol. **60:**746–748.
9. **Liesack, W., and E. Stackebrandt.** 1989. Evidence for unlinked *rrn* operons in the planctomycete *Pirellula marina*. J. Bacteriol. **171:**5025–5030.
10. **Liesack, W., H. Weyland, and E. Stackebrandt.** 1991. Potential risks of gene amplification by PCR as determined by 16S rDNA analysis of a mixed-culture of strict barophilic bacteria. Microb. Ecol. **21:**191–198.
11. **Linton, D., F. E. Dewhirst, J. P. Clewley, R. J. Owen, A. P. Burnens, and J. Stanley.** 1994. Two types of 16S rRNA gene are found in *Campylobacter helveticus*: analysis, applications and characterization of the intervening sequence found in some strains. Microbiology **140:**847–855.
12. **Meyerhans, A., J. Vartanian, and S. Wain-Hobson.** 1990. DNA recombination during PCR. Nucleic Acids Res. **18:**1687.
13. **Michaux, S., J. Paillisson, M. J. Carles-Nurit, G. Bourg, A. Allardet-Servent, and M. Ramuz.** 1993. Presence of two independent chromosomes in the *Brucella melitensis* 16M genome. J. Bacteriol. **175:**701–705.
14. **Muyzer, G., E. C. De Waal, and A. G. Uitterlinden.** 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. **59:**695–700.
15. **Ogram, A. V., G. S. Sayler, and T. T. Barkay.** 1988. The extraction and purification of microbial DNA from sediments. J. Microbiol. Methods **7:**57–66.
16. **Pettersson, B., K. E. Johansson, and M. Uhlen.** 1994. Sequence analysis of 16S rRNA from mycoplasmas by direct solid-phase DNA sequencing. Appl. Environ. Microbiol. **60:**2456–2461.
17. **Picard, C., C. Ponsonnet, E. Paget, X. Nesme, and P. Simonet.** 1992. Detection and enumeration of bacteria in soil by direct DNA extraction and polymerase chain reaction. Appl. Environ. Microbiol. **58:**2717–2722.
18. **Rainey, F. A., M. Dorsch, H. W. Morgan, and E. Stackebrandt.** 1992. 16S rDNA analysis of *Spirochaeta thermophila*: its phylogenetic position and implications for the systematics of the order *Spirochaetales*. Syst. Appl. Microbiol. **15:**197–202.
19. **Rainey, F. A., N. Ward, L. I. Sly, and E. Stackebrandt.** 1994. Dependence on the taxon composition of clone libraries for PCR amplified, naturally occurring 16S rDNA, on the primer pair and the cloning system used. Experientia **50:**796–797.
20. **Ratnaningsih, E., S. Dharmsthiti, V. Krishnapillai, A. Morgan, M. Sinclair, and B. W. Holloway.** 1990. A combined physical and genetic map of *Pseudomonas aeruginosa* PAO. J. Gen. Microbiol. **136:**2351–2357.
21. **Reysenbach, A., L. J. Giver, G. S. Wickham, and N. R. Pace.** 1992. Differential amplification of rRNA genes by polymerase chain reaction. Appl. Environ. Microbiol. **58:**3417–3418.
22. **Riley, M., and S. Krawiec.** 1987. Genome organization, p. 967–981. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
23. **Römling, U., D. Grothues, W. Bautsch, and B. Tymmler.** 1989. A physical genome map of *Pseudomonas aeruginosa* PAO. EMBO J. **8:**4081–4089.
24. **Rudner, R., A. Chevrestt, S. R. Buchholz, B. Studamire, A. White, and E. D. Jarvis.** 1993. Two tRNA gene clusters associated with rRNA operons *rrnD* and *rrnE* in *Bacillus subtilis*. J. Bacteriol. **175:**503–509.
25. **Smith, C. L., J. G. Econome, A. Schutt, S. Klco, and C. R. Cantor.** 1987. A physical map of the *Escherichia coli* K12 genome. Science **236:**1448–1453.
26. **Stackebrandt, E., and O. Charfreitag.** 1990. Partial 16S rRNA primary structure of five *Actinomyces* species: phylogenetic implications and development of an *Actinomyces israelii*-specific oligonucleotide probe. J. Gen. Microbiol. **136:**37–43.
27. **Stackebrandt, E., and W. Liesack.** 1993. Nucleic acids and classification, p. 151–194. *In* M. Goodfellow and A. O'Donnell (ed.), Handbook of new bacterial systematics. Academic Press, London.
28. **Taylor, D. E., M. Eaton, W. Yan, and N. Chang.** 1992. Genome maps of *Campylobacter jejuni* and *Campylobacter coli*. J. Bacteriol. **174:**2332–2337.
29. **Tebbe, C. C., and W. Vahjen.** 1993. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. Appl. Environ. Microbiol. **59:**2657–2665.
30. **Ulbrich, N., I. Kumagai, and V. A. Erdmann.** 1984. The number of ribosomal RNA genes in *Thermus thermophilus* HB8. Nucleic Acids Res. **7:**2055–2060.
31. **Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane.** 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. **173:**697–703.
32. **Weller, R., J. Walsh Weller, and D. M. Ward.** 1991. 16S rRNA sequences of uncultivated hot spring cyanobacterial mat inhabitants retrieved as randomly primed cDNA. Appl. Environ. Microbiol. **57:**1146–1151.
33. **Widom, R. L., E. D. Jarvis, G. Lafauci, and R. Rudner.** 1988. Instability of rRNA operons in *Bacillus subtilis*. J. Bacteriol. **170:**605–610.
34. **Young, M., and S. T. Cole.** 1993. *Clostridium*, p. 35–52. *In* A. L. Sonenshein,, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive organisms. American Society for Microbiology, Washington, D.C.