APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 1997, p. 4645–4650 Vol. 63, No. 12
0099-2240/97/$04.00+0

Frequency of Formation of Chimeric Molecules as a Consequence of PCR Coamplification of 16S rRNA Genes from Mixed Bacterial Genomes

GRACE C.-Y. WANG AND YUE WANG*

Microbial Collection and Screening Laboratory, Institute of Molecular and Cell Biology, National University of Singapore, Singapore 119076

Received 7 July 1997/Accepted 18 September 1997

PCR is routinely used in amplification and cloning of rRNA genes from environmental DNA samples for studies of microbial community structure and identification of novel organisms. There have been concerns about generation of chimeric sequences as a consequence of PCR coamplification of highly conserved genes, because such sequences may lead to reports of nonexistent organisms. To quantify the frequency of chimeric molecule formation, mixed genomic DNAs from eight actinomycete species whose 16S rRNA sequences had been determined were used for PCR coamplification of 16S rRNA genes. A large number of cloned 16S ribosomal DNAs were examined by sequence analysis, and chimeric molecules were identified by multiple-sequence alignment with reference species. Here, we report that the level of occurrence of chimeric sequences after 30 cycles of PCR amplification was 32%. We also show that PCR-induced chimeras were formed between different rRNA gene copies from the same organism. Because of the wide use of PCR for direct isolation of 16S rRNA sequences from environmental DNA to assess microbial diversity, the extent of chimeric molecule formation deserves serious attention.

DNA sequencing and phylogenetic analysis of rRNA genes obtained by PCR amplification of genomic DNAs extracted from environmental samples have led to reports of an immense diversity of undiscovered microorganisms (1–3, 6, 7, 12, 13, 29, 30). However, several studies have indicated that a significant fraction of the 16S ribosomal DNAs (rDNAs) could be chimeric as a consequence of PCR coamplification from mixed genomes (9, 16, 18, 26). Because of the explosive increase in the number of entries into public databases of rRNA sequences retrieved from environmental samples by PCR, a quantitative assessment of this problem is needed.

In a previous study (26), we investigated the contributions of several PCR parameters to the formation of chimeric molecules. We used a simple model in which pairs of cloned 16S rDNA fragments were mixed for PCR coamplification. However, we did not establish whether similar PCR conditions resulted in comparable or higher levels of chimeric molecule formation when 16S rDNAs were amplified from mixed genomes.

In this study, we employed an improved model in which 16S rRNA genes were amplified from mixed bacterial genomes instead of from pairs of cloned 16S rDNAs. We PCR amplified the 16S rRNA genes from a mixture of eight actinomycete genomes from species whose 16S rRNA sequences had been previously determined. We cloned the PCR fragments and obtained both 5′- and 3′-end sequences from a large number of randomly selected clones. The 5′- and 3′-end sequences, together with the corresponding sequences of the eight reference bacterial species, were used separately to generate two dendrograms. The cluster patterns of the two trees were compared, and a 16S rDNA clone was identified as chimeric if its two ends were placed into different species clusters. The complete nucleotide sequences of such clones were then determined, and the chimeric nature of the clones was confirmed when multiple-sequence alignments were analyzed.

Most bacteria have multiple copies of rRNA genes, and the sequences of these copies have commonly been found to be heterogeneous. The sequences of different copies of the small-subunit rRNA genes of an organism may differ by up to 6.5% (10, 28). During PCR amplification of the rRNA genes, the formation of chimeric sequences between different copies is also likely to occur. However, this problem has not been addressed before. In a previous study, the complete sequences of all four 16S rRNA copies of the actinomycete *Thermobispora bispora* were determined, and two distinct types of 16S rRNA genes were found, with each type consisting of two copies (28). In this study we used the genomic DNA of *T. bispora* to demonstrate and quantify the frequency of intercopy chimeric molecule formation.

MATERIALS AND METHODS

Organisms used in this study and culture conditions. The organisms used in this study were purchased from either the Institute for Fermentation (Osaka, Japan) or the Japan Collection of Microorganisms (Wako, Japan). The strains used were *Streptomyces griseus* subsp. *griseus* JCM4644, *Saccharopolyspora hordei* JCM8090, *Pilimelia terevasa* JCM3091, *Micromonospora chalcea* JCM3031, *Actinomadura madurae* IFO14623, *Microbispora bispora* JCM3082, and *Microtetraspora pusilla* IFO14684. The media and culture conditions used were the media and culture conditions described by the suppliers. A *Pseudonocardia* sp. strain was isolated by workers in our laboratory (unpublished data).

Preparation of genomic DNA. Mycelia of actinomycete strains were harvested by centrifugation at 3,600 × *g* for 5 min in a model 5410 C microcentrifuge (Eppendorf, Hamburg, Germany). Approximately 200 μl of wet mycelia was transferred to a 1.5-ml Eppendorf microcentrifuge tube. Mycelia were resuspended in 500 μl of lysozyme buffer containing 25 mM Tris-Cl (pH 8.0), 50 mM glucose, 10 mM EDTA, and 5 mg of lysozyme (Sigma) per ml and incubated at 37°C for 30 min. After lysozyme digestion, sodium dodecyl sulfate was added to a final concentration of 1%, and the tubes were incubated at 65°C for 30 min. The cell lysate was extracted twice with an equal volume of phenol and once with chloroform-isoamyl alcohol (24:1). Nucleic acid was precipitated with 1 volume of isopropanol at room temperature until a threadlike precipitate appeared. After centrifugation, the nucleic acid was resuspended in 100 μl of 10 mM Tris–1

* Corresponding author. Mailing address: Microbial Collection and Screening Laboratory, Institute of Molecular and Cell Biology, National University of Singapore, 15 Lower Kent Ridge Road, Singapore 119076. Phone: 65 8746861. Fax: 65 7791117. E-mail: mcbwangy@leonis.nus.sg.

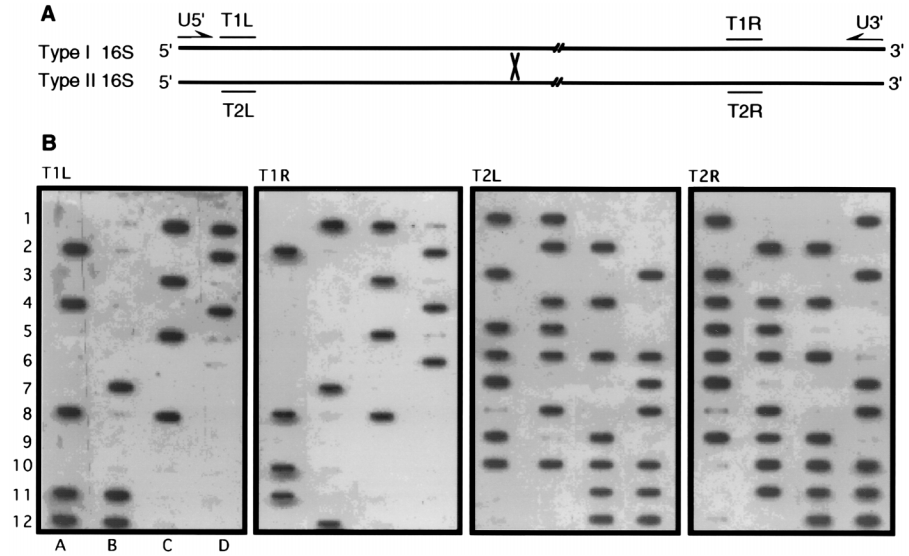

FIG. 1. Detection of chimeric molecule formation by hybridization. (A) Schematic illustration of the strategy used. Two pairs of oligonucleotide probes were designed for hybridization (positions are indicated by horizontal bars); T1L and T1R were specific for type I 16S rDNA, and T2L and T2R were specific for type II 16S rDNA. During PCR amplification of 16S rDNA with the two universal primers, U5′ and U3′ (arrows), any recombination between the two types of molecules occurring within the region confined by the hybridization probe's binding sites generated detectable chimeric molecules. See the text for the sequences of the oligonucleotides. (B) Example of hybridization results. Clones of the PCR-amplified 16S rDNA were slot blotted onto four duplicate membranes, and each membrane was hybridized with one of the four probes. The clones that showed hybridization with T1L and T2R or T2L and T1R were considered chimeric. The probe used for each membrane is shown at the top. The clones that did not hybridize with any of the four probes did not contain 16S rDNA.

mM EDTA (pH 7.5) containing 100 μg of DNase-free RNase A (Sigmal) per ml and incubated at 37°C for 30 min. Knowing that extensively fragmented DNA would also contribute to chimeric molecule formation (21), we carefully isolated the genomic DNAs to minimize mechanical shearing. The sizes of the genomic DNAs were confirmed to be greater than 50 kb (data not shown).

Oligonucleotides for PCR and sequencing. The two universal primers used for PCR amplification of nearly complete 16S rRNA genes were primer U5′ (5′TTA CCT GAT AGC GGC CGC AGA GTT TGA TCC TGG CTC AG3′; nucleotides 8 to 27 of the *Escherichia coli* 16S rRNA gene) (4) and primer U3′ (5′TAC AGG ATC CGC GGC CGC TAC GG[C/T] TAC CTT GTT ACG ACT T3′; nucleotides 1492 to 1513) (the underlined sequences are cleavage sites of restriction enzyme *Not*I) (Fig. 1). The nonspecific nucleotides on the 5′ side of the *Not*I site were included for efficient cleavage by *Not*I. The ends of cloned 16S rDNAs were sequenced by using the M13 forward sequencing primer 5′GTT TTC CCA GTC ACG AC3′ and reverse primer 5′GCG GAT AAC AAT TTC ACA CAG G3′. The internal regions were sequenced in both orientations by using eight oligonucleotide primers whose target sequences were derived either from previous reports (27) or from alignment of 16S rRNA gene sequences of many actinomycete species; these primers were 5′GCC TAT CAG CTT GTT GGT3′ (nucleotides 239 to 256), 5′CGT GCC AGC AGC CGC GGT3′ (nucleotides 514 to 531), 5′TAG ATA CCC TGG TAG TCC3′ (nucleotides 789 to 806), 5′CCG CAA CGA GCG CAA CCC3′ (nucleotides 1097 to 1114), 5′TAC CGC GGC (G/T)GC TGG CAC3′ (nucleotides 532 to 515), 5′GGA CTA CC(A/C/G) GGG TAT CTA AT3′ (nucleotides 806 to 787), 5′GGG TTG CGC TCG TTG CGG3′ (nucleotides 1114 to 1097), and 5′ACG GGC GGT GTG TAC (nucleotides 1406 to 1392). Oligonucleotides were synthesized by Oligos Etc., Inc.

PCR amplification of 16S rRNA genes. We chose PCR conditions similar to those used by most researchers for amplification of complete 16S rRNA genes (as described in references 2, 3, 9, 13, 18, and 24). Each PCR mixture contained 20 pmol of each of the two universal primers, 200 μM dATP, 200 μM dGTP, 200 μM dCTP, 200 μM dTTP, 2.5 U of *Taq* DNA polymerase (Amersham), and 1× buffer (the buffer provided with *Taq* DNA polymerase). The final volume was 100 μl. After a hot start at 95°C for 5 min, the PCR program included cycles consisting of 95°C for 30 s, 52°C for 30 s, and 72°C for 2 min. We mixed the genomic DNAs in equal amounts and used 10 ng as a template for 25 or 30 cycles of PCR amplification of the entire 16S rRNA gene preparation. The PCR products were purified with a QIAquick Spin PCR purification kit (QIAGEN) and were quantified by using the FastCheck nucleic acid quantification system (GIBCO BRL).

Cloning and sequencing of 16S rDNA. The purified PCR products of 16S rRNA genes were digested with *Not*I (Amersham) and cloned into plasmid Bluescript SK (Stratagene). *E. coli* DH5αF′ cells (GIBCO BRL) were prepared and transformed as described previously (19). Recombinant clones were picked and grown in 2 ml of Luria-Bertani medium overnight, and the recombinant plasmid DNA was prepared by using a miniplasmid preparation kit (QIAGEN). Cloned 16S rDNAs were sequenced in both orientations by using an Amersham Sequenase kit (version 2.1).

Slot blot hybridization. A slot blot hybridization experiment was carried out as described previously (18, 25). The nucleotide sequences of the hybridization probes used (Fig. 1A) are as follows: T1L, 5′CCG CAC CTT CGG GTG TGG3′ (nucleotides 76 to 93); T1R, 5′GGG TTA TGC CGG GCA CTC3′ (nucleotides 1122 to 1140); T2L, 5′CCG CAC TCT TTT GGG TGT3′ (nucleotides 76 to 93); and T2R, 5′ACG CCC CTT GTG GGG TGG3′ (nucleotides 1122 to 1140). The numbering of nucleotide positions is the numbering used for previously published sequences (28).

Sequence alignment and tree construction. Multiple alignment of sequences, calculation of sequence similarities, and construction of phylogenetic trees were carried out by using the DNASTAR program (DNASTAR Inc.).

Statistical analysis. The statistical significance of the difference between the expected and observed numbers of clones was determined as follows. The standard deviation (SD) of the expected probability (P) of clones from one actinomycete species or one type of 16S rRNA gene when a total of n clones was examined was equal to the square root of $P(1\text{-}P)/n$ (5), and the confidence limits for P were defined as $P \pm Z_{\alpha/2}\mathrm{SD}(P)$.

RESULTS

PCR coamplification of 16S rRNA genes from mixed bacterial genomes. We selected eight actinomycete species whose 16S rRNA sequences had been determined previously. The 16S rRNA sequences of most of these organisms, which were members of different genera, were less than 90% similar; the only exception was the *Micromonospora chalcea* and *P. teravasa* sequences, which were 94% similar (Table 1). The generally distant phylogenetic relationships should have allowed unambiguous identification of chimeric molecules consisting of segments from two different species. After the rDNAs amplified from the mixed genomes were cloned, 50 clones from each PCR experiment were randomly picked, and 300 nucleotides from both ends of each clone were determined. The 5′- and 3′-end sequences were aligned separately with the corresponding sequences of the eight actinomycetes to construct dendrograms. Figure 2 shows the two trees for the 16S rDNA clones from the 30-cycle PCR amplification experiment. These trees have no scale and are used only to reveal cluster patterns. Sixteen clones (32%) were identified as chimeric because their

TABLE 1. Levels of similarity of 16S rDNAs of eight actinomycetes

Taxon	% Similarity							
	Streptomyces griseus	*Saccharopolyspora hordei*	*Pilimelia terevasa*	*Micromonospora chalcea*	*Actinomadura madurae*	*Pseudonocardia* sp.	*Microbispora bispora*	*Microtetraspora pusilla*
Streptomyces griseus	100	81.2	85.6	85.7	84.0	84.0	76.9	84.4
Saccharopolyspora hordei		100	80.7	81.9	80.0	82.6	77.1	79.1
Pilimelia terevasa			100	94.0	86.3	86.4	77.8	86.1
Micromonospora chalcea				100	88.3	86.5	79.2	86.4
Actinomadura madurae					100	84.5	78.0	88.2
Pseudonocardia sp.						100	73.9	84.5
Microbispora bispora							100	85.4
Microtetraspora pusilla								100

two ends showed affiliations with different reference species. For example, clone 20 had a 5′ end that aggregated with *Streptomyces griseus* but a 3′ end that clustered with *P. terevasa*. The 5′-end sequences of clones 17 and 38 clustered with *Micromonospora chalcea* and *Microtetraspora pusilla*, respectively, but their 3′-end sequences were not clearly affiliated with any reference species. In a similar analysis of the clones from products of the 25-cycle PCR amplification, six clones (12%) were found to be chimeric (trees not shown). We found two clones and one clone that did not contain 16S rDNA inserts from the 50 random clones from the 25- and 30-cycle amplification experiments, respectively.

To confirm the chimeric nature of the clones, we determined the complete nucleotide sequences of 10 clones (clones 17, 28,

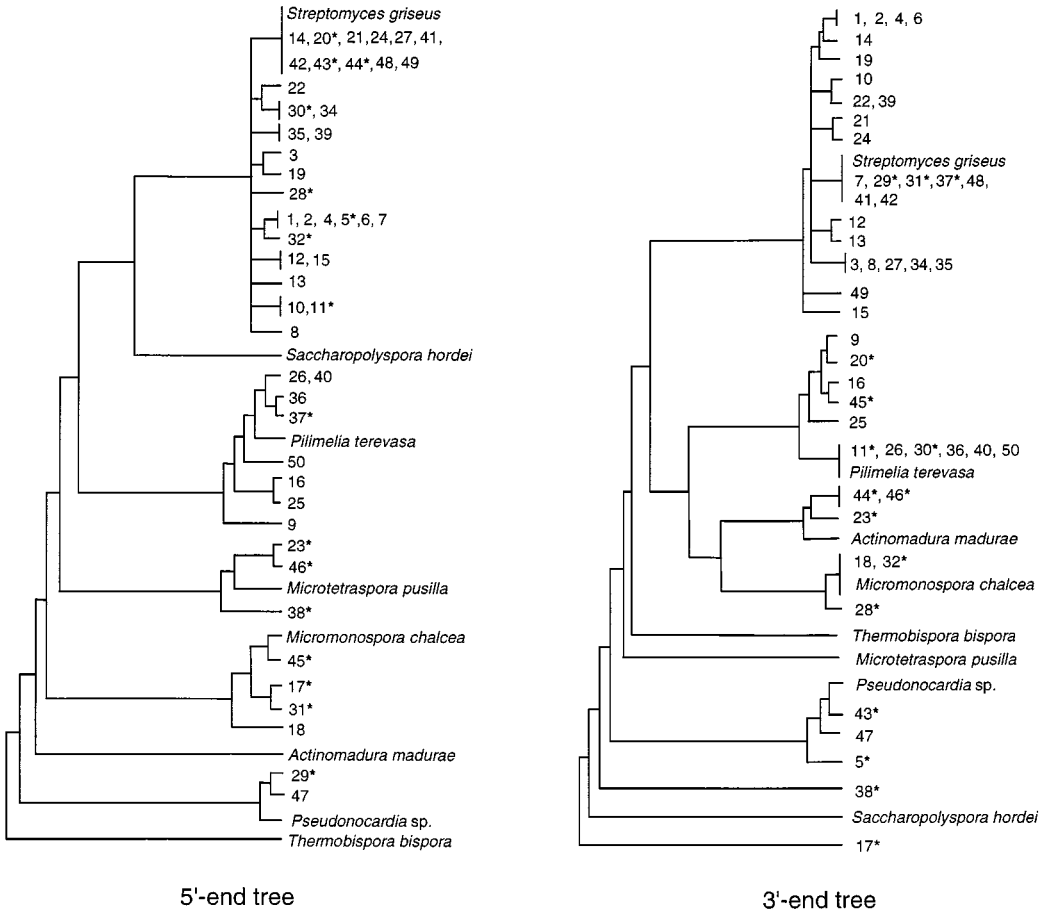

FIG. 2. Unscaled dendrograms of the 5′-end and 3′-end sequences of cloned 16S rDNAs. The asterisks indicate the clones whose two ends have different affinities with different reference species.

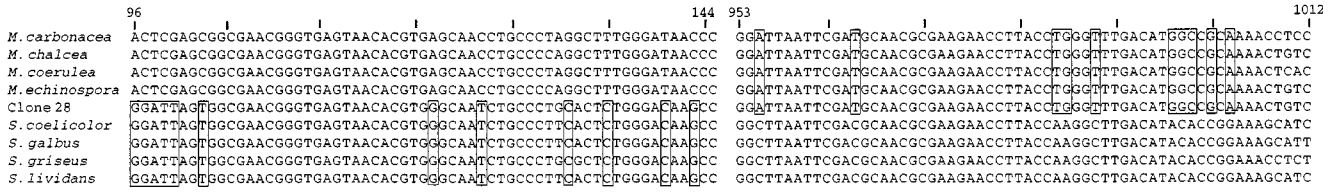

FIG. 3. Multiple-sequence alignment of a chimeric sequence (clone 28) with reference sequences. Two regions, nucleotides 96 to 145 and nucleotides 953 to 1012, are shown. Genus-specific nucleotides are enclosed in boxes to demonstrate the shifting of signatures along the chimeric sequence. Abbreviations: *M.*, *Micromonospora*; *S.*, *Streptomyces*.

30, 38, 44, 45, and 46 from the 30-cycle PCR and clones 30, 35, and 43 from the 25-cycle PCR). Each sequence was then aligned with several 16S rRNA sequences of species belonging to the two genera corresponding to the component segments of the chimeric molecule. The alignment was visually scanned to search for, along the chimeric molecule, shifting of nucleotides from signatures that characterize one genus to signatures that characterize another genus. Figure 3 shows an alignment of the sequences of two regions of chimeric clone 28 with the sequences of the actinomycete species specified by its two ends. From nucleotide 96 to nucleotide 144, clone 28 possessed all of the nucleotides that characterize *Streptomyces* species, while from nucleotide 953 to nucleotide 1012 it contained all of the bases specific to *Micromonospora* species. Such nucleotide signature shifting was clearly observed in the sequences of all 10 clones. The multiple-sequence alignment revealed that clone 17 consisted of segments from *Micromonospora chalcea* and *P. terevasa* and that clone 38 consisted of segments from *Microtetraspora pusilla* and *Saccharopolyspora hordei*. The ambiguous positions of the 3′-end sequences of the two clones in the 3′-end tree were found to be due to the occurrence of recombination within the last 300 bases of the 16S rDNA. In the other eight clones, the recombinations appeared to take place toward the middle of the fragments. However, the exact site of recombination could not be determined because the recombinations usually occurred within a region with very high sequence similarity.

The CHECK_CHIMERA program of the Ribosomal Database Project (17) identified 8 of the 10 clones as chimeric without ambiguity, because the component segments of these clones were from distantly related species. However, the program failed to identify clones 17 and 45 as chimeric because they contained segments from two closely related species, *Micromonospora chalcea* and *P. terevasa.*

We observed a markedly skewed ratio of 16S rDNA clones from different actinomycetes. Table 2 summarizes the numbers of 16S rDNA clones from different organisms. We found 52 clones (results of the 25- and 30-cycle amplification experiments combined) that contained sequences from *Streptomyces grisieus*, but no clone that contained 16S rDNAs from *T. bispora* and *Saccharopolyspora hordei*. Since the genomes of *Streptomyces griseus* and *T. bispora* are known to contain six and four copies of 16S rRNA genes, respectively (14, 28), the observed ratio of 52:0 for the numbers of 16S rDNA clones is statistically unlikely ($P < 0.01$) if different 16S rRNA genes were amplified equally. Before the genomic DNAs were mixed, all of the templates were tested separately and found to produce similar amounts of amplified 16S rDNA and comparable numbers of 16S rDNA clones (data not shown).

PCR coamplification of different 16S rRNA gene copies from a single genome. To demonstrate and quantify recombinations between different copies of the 16S rRNA gene in a single genome in the process of PCR amplification, we chose the actinomycete *T. bispora* as a model. *T. bispora* has two types of 16S rRNA genes which differ at 6.4% of the nucleotide positions, and there are two nearly identical copies of each type (28). Since only two types of molecules were studied, we used a hybridization-based method to detect chimeric molecules, as shown in Fig. 1A. Since we were limited by the selection of type-specific hybridization sites, we could quantify only the chimera formation occurring in an approximately 1-kb region. Figure 1B shows the hybridization results. Clones that hybridized with T1L and T2R (clones A4, B11, and D1) or with T2L and T1R (clones B1, A10, and D6) were identified as chimeric. These chimeric clones were subsequently confirmed by sequence analysis. Table 3 summarizes the frequencies of intercopy chimeric molecule formation as a result of 25, 30, and 35 cycles of PCR amplification; frequencies of 8.8, 11.1, and

TABLE 2. Numbers of chimeric and nonchimeric 16S rDNA clones after 25 and 30 amplification cycles

Source of 5′-end sequences	No. of clones with the following sources of 3′-end sequences after 25 amplification cycles (30 amplification cycles):							
	Streptomyces griseus	*Saccharopolyspora hordei*	*Pilimelia terevasa*	*Micromonospora chalcea*	*Actinomadura madurae*	*Pseudonocardia* sp.	*Thermobispora bispora*	*Microtetraspora pusilla*
Streptomyces griseus	28 (24)	0 (0)	0 (3)	0 (2)	0 (1)	0 (2)	1 (0)	0 (0)
Saccharopolyspora hordei	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Pilimelia terevasa	1 (1)	0 (0)	12 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Micromonospora chalcea	0 (1)	0 (0)	0 (2)	2 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Actinomadura madurae	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Pseudonocardia sp.	0 (1)	0 (0)	0 (0)	0 (0)	0 (0)	2 (1)	0 (0)	0 (0)
Thermobispora bispora	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Microtetraspora pusilla	0 (0)	0 (1)	0 (0)	0 (0)	1 (2)	0 (0)	0 (0)	1 (0)

TABLE 3. Occurrence of chimeric molecules

No. of PCR cycles	No. of clones examined	No. of clones hybridizing with the following oligonucleotide probes:				% Chimeric sequences
		T1L and T1R	T2L and T2R	T1L and T2R	T2L and T1R	
25	258	87 (118.5; $P < 0.01$)[a]	150 (118.5; $P < 0.01$)	17	4	8.8
30	190	42 (85.5; $P < 0.01$)	129 (85.5; $P < 0.01$)	13	6	11.1
35	228	37 (97.5; $P < 0.01$)	158 (97.5; $P < 0.01$)	17	16	15.5

[a] The numbers in parentheses are the expected numbers of clones if the PCR amplifications from the two types of 16S rRNA genes were equal, and the *P* values indicate the probability of the observed numbers of clones.

15.5% were observed, respectively. Table 3 also shows that the ratios of type II to type I 16S rDNA clones ranged from 2 to 4. A statistical analysis showed that it is highly unlikely that the skewed representation of clones with the two types of 16S rRNA genes could have resulted from unbiased amplification ($P < 0.01$).

DISCUSSION

The model used in this study is different from the previous model (26) in the following two respects. First, the template 16S rRNA genes were part of the entire genome; and second, the 16S rDNAs were amplified from a mixture of different genomes. This model provides new information, such as the frequency of formation of chimeras between different templates and biased amplification of genes in a single PCR. The results presented here have implications for studies in which PCRs are used to directly isolate 16S rDNAs from complex environmental samples.

A chimeric molecule is generated when a DNA fragment of one gene anneals with a homologous template to prime the next cycle of DNA synthesis. Prematurely terminated DNA strands, especially in later cycles of PCR, have been shown to be the main cause of chimera formation (26). We also demonstrated that when only two homologous sequences were coamplified, the frequency of chimeric molecule formation was much lower than the frequency of recombination because of frequent reannealing of DNA fragments with DNA strands from identical molecules. Thus, we hypothesized that during PCR coamplification of a large number of homologous DNA sequences in which each sequence accounts for only a minute fraction of the total number of sequences, a DNA fragment would rarely reanneal with strands from identical molecules and the recombination events would be maximally expressed as chimeric molecules. This hypothesis is strongly supported by two observations made in this study. First, in the products of the 30-cycle PCR coamplification from the mixed genomes, we detected 32% chimeric molecules. This frequency is much higher than the frequencies of 12.9 to 14.7% detected in the products of PCR coamplification from two templates that had approximately 85% sequence similarity (26). Second, we observed highly skewed ratios of clones from different 16S rDNAs, and it appeared that the fraction of chimeras in the less frequent 16S rDNAs was markedly higher than that in the abundant 16S rDNAs. For example, 8 of the 32 clones whose 5′ end was identified as a *Streptomyces griseus* sequence had a 3′ end belonging to another species. In striking contrast, all three clones with *Microtetraspora pusilla* 5′ ends had 3′ ends from other species, and three of four clones with *Micromonospora chalcea* 5′ ends had 3′ ends from different species. This result is in agreement with our previous experimental results which demonstrated that nearly all of the less abundant 16S rDNA molecules were involved in chimera formation when two templates were mixed at 1:5 and 1:10 ratios for 30 cycles of PCR coamplification.

The highly skewed ratios of 16S rDNA clones from different rRNA genes observed in our experiments were most likely a consequence of preferential PCR amplification of genes, which is a well-known phenomenon (8, 20, 22, 25). In extreme cases, amplification of the 16S rRNA gene from one genome could be completely inhibited by the presence of another genome (20, 22). Suzuki and Giovannoni (25) described one type of biased PCR amplification in which the ratio of two coamplified genes changed as a function of PCR cycles toward 1:1 in the final PCR products regardless of the initial proportions of the templates. However, in our study, this type of bias was not observed. Many factors could contribute to preferential amplification, such as different G+C contents, different folding of the template DNAs, and different ratios of templates (8, 11, 20, 25). Obviously, no PCR conditions can be optimized for amplification of all 16S rRNA genes in a complex mixture. However, one should be aware that biased amplification may lead to distorted representation of microbial community structures in the ecosystems under study and may increase the frequency of chimera formation occurring with the less amplified or less abundant genes, as described above.

16S rRNA genes are routinely retrieved by PCR from pure cultures of organisms for sequencing and subsequent phylogenetic analyses. An organism usually contains multiple copies of heterogeneous rRNA genes (10, 28). The phylogenetic position of an organism sometimes cannot be ascertained as the rDNA sequences of different copies may fall in different clades on the phylogenetic tree (10). The intercopy chimera formation observed in this study should certainly add to the ambiguity for taxonomic identification. We detected about 11.1 and 15.5% chimeric molecules as a result of 30 and 35 cycles of PCR coamplification of the two types of 16S rRNA genes of *T. bispora*. The relatively low frequencies can be explained by the short region over which chimera formation was detected. The extent of intercopy chimera formation in PCR amplification experiments is a function of the number of copies per genome and the levels of sequence heterogeneity between the different rRNA genes.

The use of PCR in the study of microbial diversity has successfully revealed the existence of many novel organisms that cannot be isolated by traditional culture-dependent techniques (1–3, 6, 7, 12, 24, 29, 30). However, the extent of chimeric molecule formation during PCR amplification of conserved genes from complex environmental DNA samples deserves serious attention. Many artifactual sequences may have been deposited in the public databases without being noticed, because there is no fail-safe way to detect all chimeric sequences (15, 16, 23, 26). Computer software programs, such as CHECK_CHIMERA (17), have been the main tools used for the identification of chimeric sequences. However, such programs may not detect chimeras when the component se-

quences of the chimeras are derived from closely related organisms (15, 23). The presence of chimeric sequences in the databases may further reduce the usefulness of a program which depends on comparison of an input sequence with all sequences in the databases. Accumulation of artifactual sequences may compromise the quality of the databases and lead to wrong phylogenies for organisms and misinterpretation of the extent of microbial diversity. To prevent potential damage to the databases, it would be wise to keep the sequences retrieved from environmental DNA samples by PCR in a separate category in the databases. Meanwhile, workers need to study factors that may cause chimeric sequences and biased amplification, to improve experimental conditions to minimize the problems of PCR, and to develop more sensitive computer programs for identification of chimeric molecules.

ACKNOWLEDGMENTS

This work was supported by the Institute of Molecular and Cell Biology.

We thank Soong Tuch Wah for critically reading the manuscript.

REFERENCES

1. **Amann, R., W. Ludwig, and K. H. Scheifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59:**143–169.
2. **Barns, S. M., R. E. Fundyga, M. W. Jeffries, and N. R. Pace.** 1994. Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. Proc. Natl. Acad. Sci. USA **91:**1609–1613.
3. **Barns, S. M., C. F. Delwiche, J. D. Palmer, and N. R. Pace.** 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. Proc. Natl. Acad. Sci. USA **93:**9188–9193.
4. **Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller.** 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. J. Mol. Biol. **148:**107–127.
5. **Brown, B. W., Jr., and M. Hollander.** 1977. Statistics: a biomedical introduction. John Wiley & Sons, New York, N.Y.
6. **Bruce, K. D., W. D. Hiorns, J. L. Hobman, A. M. Osborn, P. Strike, and D. A. Ritchie.** 1992. Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. Appl. Environ. Microbiol. **58:**3413–3416.
7. **Bull, A. T., M. Goodfellow, and J. H. Slater.** 1992. Biodiversity as a source of innovation in biotechnology. Annu. Rev. Microbiol. **46:**219–251.
8. **Chandler, D. P., J. K. Fredrickson, and F. J. Brockman.** 1997. Effect of PCR template concentration on the composition and distribution of total community 16S rDNA clone library. Mol. Ecol. **6:**475–482.
9. **Choi, B. K., B. J. Paster, F. E. Dewhirst, and U. B. Gobel.** 1994. Diversity of cultivable and uncultivable oral *Spirochetes* from a patient with severe destructive periodontitis. Infect. Immun. **62:**1889–1895.
10. **Clayton, R. A., G. Sutton, P. S. Hinkle, Jr., C. Bult, and C. Fields.** 1995. Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. Int. J. Syst. Bacteriol. **45:**595–599.
11. **Dutton, C. M., C. Paynton, and S. Sommer.** 1993. General method for amplifying regions of very high G + C content. Nucleic Acids Res. **21:**2953–2954.
12. **Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field.** 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature (London) **345:**60–63.
13. **Hershberger, K. L., S. M. Barns, A. L. Reysenbach, S. C. Dawson, and N. Pace.** 1996. Wide diversity of *Crenarchaeota*. Nature (London) **384:**420.
14. **Kim, E., H. Kim, S.-P. Hong, K. H. Kang, Y. H. Kho, and Y.-H. Park.** 1993. Gene organization and primary structure of a ribosomal RNA gene cluster from *Streptomyces griseus* subsp. *griseus*. Gene **132:**21–31.
15. **Komatsoulis, G. A., and M. S. Waterman.** 1997. A new computational method for detection of chimeric 16S rRNA artifacts generated by PCR amplification from mixed bacterial populations. Appl. Environ. Microbiol. **63:**2338–2346.
16. **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.
17. **Larsen, N., G. J. Olsen, B. L. Maidak, M. J. McCaughey, R. Overbeek, T. J. Macke, T. L. Marsh, and C. R. Woese.** 1993. The Ribosomal Database Project. Nucleic Acids Res. **21:**3021–3023.
18. **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.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Morrison, C., and F. Gannon.** 1994. The impact of PCR plateau on the quantitative PCR. Biochim. Biophys. Acta **1219:**493–498.
21. **Paabo, S., D. M. Irwin, and A. C. Wilson.** 1992. DNA damage promotes jumping between templates during enzymatic amplification. J. Biol. Chem. **265:**4718–4721.
22. **Reysenbach, A. L., 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.
23. **Robison-Cox, J. F., M. M. Bateson, and D. M. Ward.** 1995. Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. Appl. Environ. Microbiol. **61:**1240–1245.
24. **Suzuki, M. T., M. S. Rappe, Z. W. Haimberger, H. Winfield, N. Adair, J. Strobel, and S. J. Giovannoni.** 1997. Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. Appl. Environ. Microbiol. **63:**983–989.
25. **Suzuki, M. T., and S. J. Giovannoni.** 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. **62:**625–630.
26. **Wang, G. C.-Y., and Y. Wang.** 1996. The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. Microbiology **142:**1107–1117.
27. **Wang, Y., Z. S. Zhang, and J. S. Ruan.** 1996. A proposal to transfer *Microbispora bispora* (Lechevalier 1965) to a new genus *Thermobispora* gen. nov., as *Thermobispora bispora* comb. nov. Int. J. Syst. Bacteriol. **46:**933–938.
28. **Wang, Y., Z. S. Zhang, and R. Narendrakumar.** 1997. The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. J. Bacteriol. **179:**3270–3276.
29. **Ward, D. M., R. Weller, and M. M. Bateson.** 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. Nature (London) **345:**63–65.
30. **Weller, R., J. W. 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.