Pak. J. Agri. Sci., Vol. 52(2), 285-291; 2015
ISSN (Print) 0552-9034, ISSN (Online) 2076-0906
http://www.pakjas.com.pk

AN IMPROVED METHOD FOR BACTERIAL IDENTIFICATION THROUGH 16S rDNA SEQUENCING USING DOUBLE SET OF PRIMERS: A CASE STUDY

Amanat Ali[1,2,3], Sohail Hameed[1,*], Ivan J. Oresnik[2], Muhammad Shahid[1,4], Mohsin Tariq[1,4] and Naveed Ahmad[1]

[1]National Institute for Biotechnology and Genetic Engineering (NIBGE), Jhang Road, P.O. Box 577, Faisalabad, Punjab, Pakistan; [2]Department of Microbiology, University of Manitoba, Winnipeg, MB, Canada R3T 2N2; [3]Soil Science Division, Nuclear Institute of Agriculture (NIA), Tandojam, Sindh, Pakistan; [4]Department of Bio-informatics and Biotechnology, Government College University, Faisalabad, Pakistan.
[*]Corresponding author's e-mail: shameed58@hotmail.com

16S rDNA sequencing is an established technique for molecular identification and phylogenetic studies of bacteria. In most of the previous studies, single set of primers (16S forward and 16S reverse) was unlikely to amplify and sequence full fragment of 16S gene and the product size remained shorter. Thus, an improved method was adopted to sequence the full-length 16S rDNA fragment employing an additional set of internal primers (Henk 16S and Hoff 16S) along with external primers (Loff 16SF and Loff 16SR) in a single run to enhance the precision in bacterial identification by getting a sequence length of over 1300 bp. Moreover, we presented another way to tackle the missing sequences and some faulty readings in the raw sequence data that appear to be another impediment causing imprecise identification of bacteria. Therefore, peaks cross-matching was done while editing the raw sequence data prior to alignment through NCBI BLAST. Finally, CAP3 sequence assembly online program was employed for precise assembling of the sequence data obtained from each of the four primers to obtain a single contiguous sequence for greater precision in bacterial identification. This approach will certainly help taxonomists to describe novel bacterial species.

Keywords: Soil microbes, bacterial identification, 16S rDNA sequencing, PCR primers, CAP3 assembly program

INTRODUCTION

Microbes are indispensable in clinical and agricultural research hence, their precise identification is inevitable (Ramadhar *et al.,* 2013). Isolation and culture-based studies of bacterial phenotypic characteristics are traditional methods of identifying bacteria. This includes Gram staining, colony and cell morphologies, culture requirements and biochemical reactions. Since anaerobic bacteria cannot tolerate oxygen, these methods are difficult to be performed (Woo *et al*., 2003). It is also difficult to identify non-cultivable or slow growing bacteria, former of which occur 99% in nature (Amann *et al.,* 1995). Bacteria which exhibit deviant biochemical characteristics are also difficult to identify. Thus, phenotypic methods generally used in this context, are not feasible as they are time-consuming and often result in false identification in common environmental or specific clinical situations (Mignard and Flandrois, 2006). Thus, correct identification of microbes in clinical research facilitates the diagnosis of a patient's condition and gives an immediate indication of appropriate antibiotic therapy to be applied. The same is essential in agricultural systems in exploring the potential of novel strains of microbes for beneficial plant-microbe interaction studies and ultimately for inoculum production (Coombs and Franco, 2003).

For taxonomic implications, rRNA is highly conserved gene in all cells. Genes encoding rRNA have extensively been used for taxonomic classification, phylogenetic studies and estimations of species divergence rates among bacteria (Yarza *et al.,* 2008). Thus, the comparison of 16S rDNA sequence (one of the representatives of rRNA genes) can exhibit evolutionary relatedness among microorganisms. Carl Woese was the pioneer of this work who proposed the three domain system of classification (Archaea, Bacteria, and Eucarya) based on sequence information of 16S rDNA (Woese *et al*., 1990). The analysis of cloned 16S rDNA provides information mainly of taxonomic value of bacteria and has become a regular method for identifying bacteria. The method is considered to overcome the disadvantages of bacterial culture, time consumption and uncertainty in the bacterial identification (Simmon *et al*., 2008). Drancourt *et al*. (2000) used 16S rDNA with a large collection of phenotypically unidentifiable isolates for the first time, and reported this approach efficient in the majority of cases, with 88.7% of isolates identified to the genus level and 76.3% to the species level.

The universal primers (Loff 16S F and Loff 16S R) are generally used for amplification and sequencing of 16S rRNA gene. However currently used universal primers have been reported to target only a portion of total diversity of bacterial species (Rajendhran and Gunasekaran, 2011) and

additional primers have been recommended for perfect identification (Klindworth *et al*., 2012). Recently, Starke *et al.* (2014) concluded the sequencing studies that two or more set of primers should be used to overcome PCR studies problems and to cover the bacterial diversity completely. The problem in using single set of primers is that the middle portion of 16S rRNA gene may remain unsequenced as the forward and reverse primers sequences terminates before they reach their full length resulting in shorter length of 16S rRNA gene sequence which can lead to incorrect identification of microbes. However, the entire 16S rRNA gene length (1500 bp) is necessary to distinguish between particular taxa or strains (Sacchi *et al*., 2002a). So, the current case study was focused on sequencing the middle portion of 16S rDNA and an improved method was described for full-length gene sequencing resulting in precise bacterial identification and phylogenetic studies.

MATERIALS AND METHODS

***Bacterial culture, media and growth conditions*:** Soybean root-nodule rhizobia were isolated by serial dilution method as described earlier (Somasegaran and Hoben, 1994) after sterilizing them by immersing in ethanol for 1 min and then transferring to 5% sodium hypochlorite solution for 2-5 min and subsequent rinsing in five changes of sterile water. Rhizobia were cultured in congo red YEM (CR-YEM) agar plates and incubated at 28±2°C for 36 h. Well isolated bacterial colonies showing different morphologies were selected and sub-cultured until the purity was achieved. Four isolates (SR-3, SR-5, SR-10 and SR-11) were selected for identification purpose.

***DNA Extraction*:** DNA from one loopful of pure bacterial morphotypes was extracted using MO BIO's Power Soil™ DNA isolation kit according to the manufacturer's protocol with the modification of using pure bacterial culture instead of soil. . Isolated DNA was re-suspended in 100µL of sterile dH_2O and confirmed on 1% agarose gel.

***PCR primers*:** PCR amplification of 16S rDNA was carried out using standard primers (Loff 16S forward and Loff 16S reverse) The primers Loff 16S F (5`-AGAGTTTGATCCTGGCTCAG-3`) and Loff 16S R (5`-AAGGAGGTGATCCAGCC- 3`), which correspond to *Escherichia coli* K12 16S rRNA gene positions 8-27 and 1541-1522, respectively, were used for the PCR amplification of the 16S rRNA gene fragment of bacterial isolates as described. In addition to these universal primers, two internal primers Henk 16S (5`-GTGCCAGCAGCCGCGGTAA-3` (that binds to 515-533 in *E. Coli* K12) and Hoff 16S (5`-GTAAGGTTCTTCGCGTT-3`) (that binds to 984-968 in *E. Coli* K12) were used during 16S rRNA gene sequencing of the isolates (Henckel *et al*., 1999; Hoffmann *et al*., 2002)

***PCR amplification*:** Gene fragments were amplified in a PCR thermocycler with a reaction mixture volume of 50 µL containing about 50 ng DNA or 3 µL of colony mixture, 0.4 µM of each primer, 200 µM dNTPs (Fermentas), 1.5 mM $MgCl_2$, 0.1% Dimethyl sulfoxide (DMSO), and 1.25 U *Taq* DNA polymerase (Invitrogen). The reaction was performed in 1 X PCR buffer. PCR reaction was performed as 5 min denaturation at 95°C, then 35 cycles of 30 sec at 95°C, 30 sec at 58°C and 90 sec at 72°C and final extension at 95°C for 10 min.

***Electrophoresis*:** The amplified PCR products were loaded onto 1% agarose gel in 0.5 X TBE (Trisborate-EDTA) buffer containing 2 µL ethidium bromide (20 mg mL^{-1}). 1 kb and 100bp DNA ladders (Gene Ruler, Fermentas) were used as size markers. The gels were viewed under UV light and photographed using Vilber Lourmat gel documentation system.

***Purification of PCR product*:** Amplified PCR products of 16S rRNA genes were purified by cutting the specific band from the gel and DNA was isolated by melting the band gel in Sodium iodide (NaI) at 60°C, collecting it in silica milk at 0°C, washing with NEET wash solution (Zhuo-hua *et al.*, 2000) to remove residues and then releasing DNA from silica milk by melting at 60°C.

***Cycle sequencing and cleanup* reaction:** The cycle sequence PCR with four primers i.e. Loff 16S F, Loff 16S R, Henk 16S and Hoff 16S was carried out (each in separate tube) so that full-length 16S rRNA gene could be sequenced. The two strands of the DNA were sequenced separately, generating four sequences (one with each primer). After that a cleanup reaction with 0.125M EDTA and Ethanol (95% and 70%, respectively) was carried out to remove the contaminants.

***16S rRNA gene* sequencing:** The purified PCR products were directly sequenced with an ABI Prism 3110 Genetic Analyzer (Hitachi, Japan) using Big Dye Terminator V3.0 Cycle Sequencing chemistry.

***Editing of raw sequence data*:** The missing sequences and wrong readings present in the raw sequence data as appeared in electropherogram were corrected by matching the nitrogen bases sequences with electropherogram peaks accordingly.

***Assembling edited sequence data*:** The sequences for each primer of the same isolate (4 for each as 4 primers were used) were needed to get assembled as they may contain the overlapping sequences. For this purpose, CAP3 an online program (http://doua.prabi.fr/software/cap3) was used. This program is efficient in assembling the sequences and it shows the assembly details along with overlapping sequences and single sequences and there is no need to complementate the reverse sequence data before assembling.

***Sequence analysis*:** The assembled sequences in contig form were then compared in Gen Bank database of National Centre for Biotechnology Information NCBI BLAST

(Altschul *et al.,* 1990) at http://www.ncbi.n1m.nih.gov/blast/Blast.cgi to find out the maximum identity.

RESULTS

The corresponding band (1.5 kb), after 1% agarose gel electrophoresis, confirmed the amplification of 16S rDNA fragments as shown in Figure 2A, while Figure 2B represent the corresponding bands after gel elusion and purification process. The products after removing unincorporated dye terminators were analyzed on DNA sequencer (ABI PRISM 3130 genetic analyzer with 16 capillaries) and the length & sequence of each of four sequences was determined using capillary electrophoresis, resulting in raw sequence data. The two strands of the DNA were sequenced separately, generating four sequences, one with each primer. Four isolates were sequenced using the same technique (Table 1). Some missing sequences and some wrong readings (Fig. 3) present in the data were edited accordingly by matching the electropherogram with the raw sequence data (Fig. 3) hence getting more precise gene sequence data. All three or four primer sequences needed to be assembled in a single sequence by eliminating the overlapping sequences for identification. Cap3 (Fig. 4), an online sequence assembling program (http://doua.prabi.fr/software/cap3) fulfilled the

Table 1. Identification of rhizobial isolates based on 16S rRNA gene sequence analysis.

Isolate code	Length of gene sequenced	Closest Genbank similarity	% Similarity	Strain Identification	Genbank Accession No.
SR-3	1304	*B. yuanmingense* MN-S	99%	*B. yuanmingense*	JQ315248
SR-5	1377	*B. yuanmingense* MN-S	100%	*B. yuanmingense*	JQ315249
SR-10	1360	*B. japonicum* gene	99%	*B. japonicum*	JQ315251
SR-11	1331	*B. japonicum* strain 311b6	100%	*B. japonicum*	JQ315252

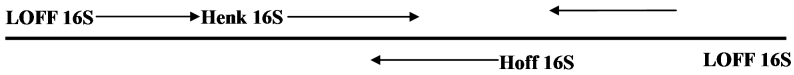

Figure 1. Proposed model for sequencing full length of 16S rRNA gene with double set of primers.

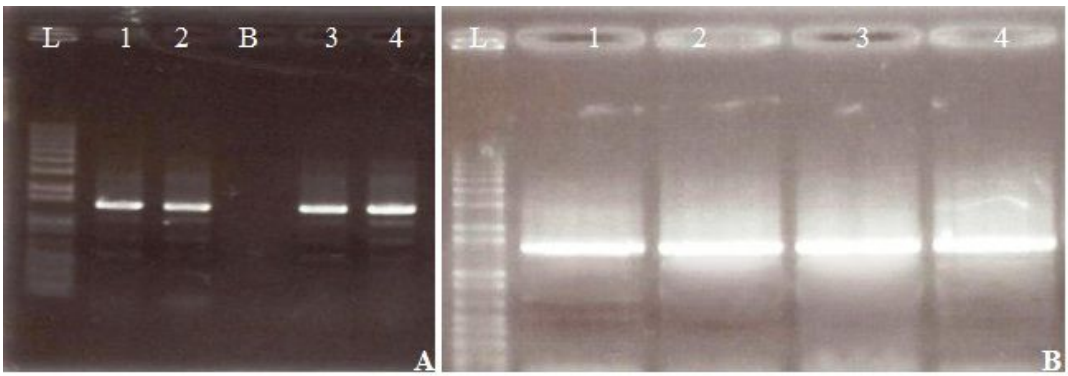

Figure 2. PCR amplification (A) and gel purification (B) of 16S rRNA gene of the putative *Bradyrhizobium* isolates isolated from soybean nodules.

L= 1Kb Ladder, 1= SR-3, 2= SR-5, 3= SR-10, 4= SR-11, B= blank

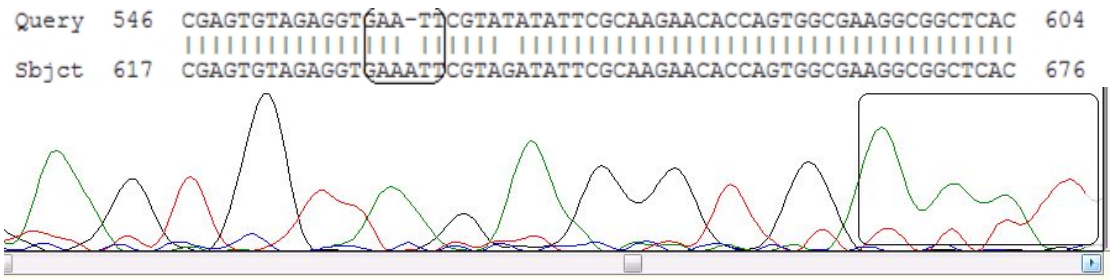

Figure 3. Editing of raw sequence data manually by matching the sequence data with electropherogram peaks.

purpose and overlapping of forward and reverse sequences confirmed the efficient assembling of the whole length of 16S rRNA gene (Fig. 5). The assembly details (Fig. 5), added up in the efficiency of the program. Alignment of the sequences from unknown bacterial isolates against already submitted sequences in the GenBank database of National Centre for Biotechnology Information (BLASTn) (Fig. 6) resulted in the identification of bacteria up to species level (Table 1).

In this study four isolates were identified through improved 16S rRNA gene amplification technique. The isolates were identified as *Bradyrhizobium japonicum* and *Bradyrhizobium yuanmingense*. The 16S rRNA gene sequences were submitted to Gen Bank through the accession numbers of JQ315248, JQ315249, JQ315251 and JQ315252. The sequence length of 16S rRNA gene for all of

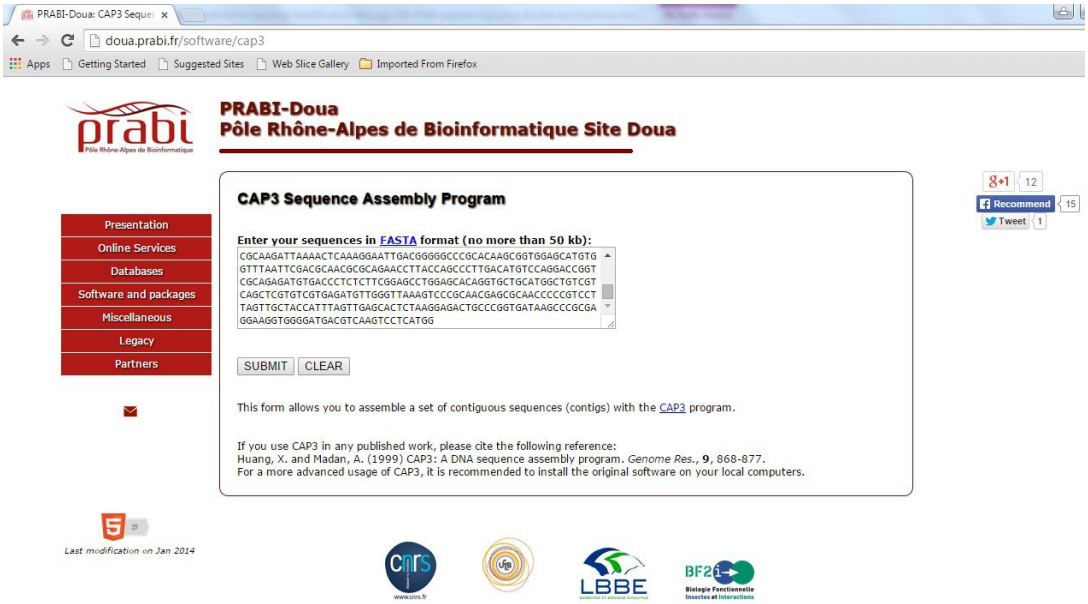

Figure 4. Cap3 an online sequence assembly program used for assembling forward sequences with the reverse sequences

Figure 5. An outlook of assembling Loff 16SF, Henk 16S and Loff 16SR sequence, the overlapping of sequence confirming the accuracy of sequence assembly.

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value
EU010398.1	Bradyrhizobium japonicum strain CCBAU 15618 16S ribosomal RNA ge	1448	1448	100%	0.0
EF394153.1	Bradyrhizobium japonicum strain CCBAU 33143 16S ribosomal RNA ge	1448	1448	100%	0.0
NR_036865.1	Bradyrhizobium japonicum strain 3I1b6 16S ribosomal RNA, partial se	1447	1447	99%	0.0
GU552894.1	Bradyrhizobium japonicum strain CCBAU 15216 16S ribosomal RNA ge	1447	1447	99%	0.0

Figure 6. Alignment of the sequence of rhizobial isolates in NCBI through BLAST.

the strains were above 1300 bp (Table.1).

DISCUSSION

Precise identification of the microorganisms is inevitable in clinical as well as microbiological research. However, phenotypic methods used to identify a bacterial strain are commonly associated with a number of drawbacks as the phenotypic properties are often unstable and their expression may depend on the changing environmental conditions of growth substrate, temperature, pH levels (Rosselló-Mora and Amann, 2001) and improper phenotypic tool kits (Janda and Abbott, 2002) with old tests of 1975. 16S rDNA sequencing technique has the ability to provide unambiguous data for rare isolates (Drancourt *et al*., 2000) and it can provide accurate identification at the species level (Clarridge *et al.*, 2001).

In the current study double set of primers was used to overcome the problem of shorter sequence length that may lead to incorrect identification of microbes. The universal primers generally used in 16S gene sequencing studies do not target the total diversity of bacteria (Rajendhran and Gunasekaran, 2011). In this context using two or more sets of primers have been recommended to overcome the PCR and sequencing issues in perfect bacterial identification (Klindworth *et al*., 2012; Starke *et al.,* 2014). Double sets of primers helps in sequencing the entire 16S rRNA gene length (1500 bp) which is necessary to distinguish between particular taxa or strains (Sacchi *et al*., 2002a). In all of our microbes identified the sequence length obtained by using double set of primers was more than 1300 bp (Table.1) as compared to previously reported Sanger method reads of 800 and 900 bp (Herman and De Ridder, 1992; Nossa *et al.,* 2010). The proposed model mechanism is illustrated in Fig.1 revealing that the second set of primers helps the middle portion of the gene to get sequenced that may remain unsequenced in using single set of primers.

Ambiguities in raw sequence data obtained from sequencer can be resolved by comparing the electropherogram and raw sequence data visually but it is not the common practice in identification. This case study revealed its importance as we got 99-100% similarity when compared in gene database library.

The next challenge in identification procedure is the assembling of the forward and reverse sequence data to get a single sequence in contig form. Generally, Justbio, an online program (http://www.justbio.com) is used for the same purpose in which complementation of the reverse sequence is carried out before assembling of both forward and reverse sequence data. In the current study, Cap3 an online assembly program (http://doua.prabi.fr/software/cap3) was used in which no complementation of the reverse sequence was required, just the sequence data in fasta format was submitted and single sequence in contig form was obtained. The question arises here whether the DNA sequence must have overlapping to be accurate or just the forward or reverse sequence may become sufficient to have an adequate sequence for the identification of a bacterial strain. In some of the studies only the forward sequence (Bosshard *et al.,* 2003) is used for identification, however, multiple overlaps are required in considering microheterogeneity (Sacchi *et al*., 2002b). By using CAP3 assembly program all 4 primer sequences were assembled and overlapping confirmed the whole length sequence of 16S rRNA gene hence getting over 1300 bp sequence length with 99-100% similarity with *B. japonicum* (JQ315251, JQ315252) and *B. yuanmingense* (JQ315248, JQ315249) (Table.1) a strain that is reported first time from Pakistan to have symbiosis with soybean (Ali, 2014). Hayat *et al.* (2012) identified 14 of compost bacteria using 16S rRNA gene sequencing, however, the amplified PCR products were sequenced through the services of MACROGEN, Korea.

***Conclusion*:** 16S rDNA sequencing provides the better identification of unknown bacterial isolates as compared to old phenotypic techniques and identification kits. The use of double sets of primers instead of single set, has the potential to make the technique more precise and accurate. More experimentation is required however for validity of this improved method in the identification of unknown bacteria from other fields of studies.

***Acknowledgements*:** The research work was a part of Ph.D research carried out at University of Manitoba, Winnipeg, Manitoba, Canada. The Higher Education Commission (HEC) Islamabad, Pakistan is highly acknowledged for providing funds for the said research work. We are highly thankful to Mr. Nasir Rahim, Assistant Professor, University

of Poonch, Rawalakot, AJK for fully co-operating in isolation of *Rhizobium* from soybean root nodules grown in Rawalakot, AJK.

REFERENCES

Ali, A. 2014. Evaluating *Bradyrhizobium*-soybean symbiosis from semi-arid to humid regions of Pakistan for enhanced crop production, Ph. D thesis. Dept. Biotech. QAU, Islamabad, Pakistan.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

Amann, R.I., W. Ludwig and K.-H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59: 143-169.

Bosshard, P., S. Abels, R. Zbinden, E. Böttger and M. Altwegg. 2003. Ribosomal DNA sequencing for identification of aerobic gram-positive rods in the clinical laboratory (an 18-month evaluation). J. Clin. Microbiol. 41: 4134-4140.

Clarridge, J.E., S.M. Attorri, Q. Zhang and J. Bartell. 2001. 16S ribosomal DNA sequence analysis distinguishes biotypes of Streptococcus bovis: Streptococcus bovis biotype II/2 is a separate genospecies and the predominant clinical isolate in adult males. J. Clin. Microbiol. 39: 1549-1552.

Coombs, J.T. and C.M. Franco. 2003. Isolation and identification of actinobacteria from surface-sterilized wheat roots. Appl. Environ. Microbiol. 69: 5603-5608.

Drancourt, M., C. Bollet, A. Carlioz, R. Martelin, J.P. Gayral and D. Raoult. 2000. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J. Clin. Microbiol. 38: 3623-3630.

Hayat, R., R.A. Sheirdil, M. Iftikhar-ul-Hassan and I. Ahmed. 2012. Characterization and identification of compost bacteria based on 16S rRNA gene sequencing. Ann. Microbiol. 63: 905-912.

Henckel, T., M. Friedrich and R. Conrad. 1999. Molecular analyses of the methane-oxidizing microbial community in rice field soil by targeting the genes of the 16S rRNA, particulate methane monooxygenase, and methanol dehydrogenase. Appl. Environ. Microbiol. 65: 1980-1990.

Hoffmann, T., H.P. Horz, D. Kemnitz and R. Conrad. 2002. Diversity of the particulate methane monooxygenase gene in methanotrophic samples from different rice field soils in China and the Philippines. Syst. Appl. Microbiol. 25: 267-274.

Janda, J.M. and S.L. Abbott. 2002. Bacterial identification for publication: when is enough enough? J. Clin. Microbiol. 40: 1887-1891.

Klindworth, A., E. Pruesse, T. Schweer, J. Peplies, C. Quast, M. Horn and F.O. Glöckner. 2012. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. doi: 10.1093/nar/gks808.

Mignard, S. and J. Flandrois. 2006. 16S rRNA sequencing in routine bacterial identification: a 30-month experiment. J. Microbiol. Methods 67: 574-581.

Nossa, C.W., W.E. Oberdorf, L. Yang, J.A. Aas, B.J. Paster, T.Z. DeSantis, E.L. Brodie, D. Malamud, M.A. Poles and Z. Pei. 2010. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. World J. Gastroenterol. 16: 4135.

Rajendhran, J. and P. Gunasekaran. 2011. Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. Microbiol. Res. 166: 99-110.

Ramadhar, T.R., C. Beemelmanns, C.R. Currie and J. Clardy. 2013. Bacterial symbionts in agricultural systems provide a strategic source for antibiotic discovery. J. Antibiot. 67: 53-58.

Herman, L. and H. De Ridder. 1992. Identification of *Brucella* spp. by using the polymerase chain reaction. Appl. Environ. Microbiol. 58: 2099-2101.

Rosselló-Mora, R. and R. Amann. 2001. The species concept for prokaryotes. FEMS Microbiol. Rev. 25: 39-67.

Sacchi, C.T., A.M. Whitney, L.W. Mayer, R. Morey, A. Steigerwalt, A. Boras, R.S. Weyant and T. Popovic. 2002a. Sequencing of 16S rRNA gene: a rapid tool for identification of *Bacillus anthracis*. Emerg. Infect. Dis. 8: 1117-1123.

Sacchi, C.T., A.M. Whitney, M.W. Reeves, L.W. Mayer and T. Popovic. 2002b. Sequence diversity of Neisseria meningitidis 16S rRNA genes and use of 16S rRNA gene sequencing as a molecular subtyping tool. J. Clin. Microbiol. 40: 4520-4527.

Simmon, K.E., S. Mirrett, L.B. Reller and C.A. Petti. 2008. Genotypic diversity of anaerobic isolates from bloodstream infections. J. Clin. Microbiol. 46: 1596-1601.

Somasegaran, P. and H.J. Hoben, 1994. Handbook for rhizobia: methods in legume-Rhizobium technology. 1st ed. Springer-Verlag New York.

Starke, I.C., W. Vahjen, R. Pieper and J. Zentek. 2014. The influence of DNA extraction procedure and primer set on the bacterial community analysis by pyrosequencing of barcoded 16S rRNA gene amplicons. Mol. Biol. Int. 2014: 1-10.

Woese, C.R., O. Kandler and M.L. Wheelis. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Nat. Acad. Sci. 87: 4576-4579.

Woo, P.C., K.H. Ng, S.K. Lau, K.-T. Yip, A.M. Fung, K.-W. Leung, D.M. Tam, T.-L. Que and K.-Y. Yuen. 2003. Usefulness of the MicroSeq 500 16S ribosomal DNA-based bacterial identification system for identification of clinically significant bacterial isolates with ambiguous biochemical profiles. J. Clin. Microbiol. 41: 1996-2001.

Yarza, P., M. Richter, J. Peplies, J. Euzeby, R. Amann, K.-H. Schleifer, W. Ludwig, F.O. Glockner and R. Rossello-Mora. 2008. The All-Species Living Tree project: A 16S r*RNA*-based phylogenetic tree of all sequenced type strains. Syst. Appl. Microbiol. 31: 241-250.

Zhuo-Hua, W., L. Hong-Hao and M. Hui-Wen. 2000. Recovery of DNA from agarose gel with home-made silica milk. Wuhan Univ. J. Nat. Sci. 5: 373-376.