Use of tuf as a target for sequence-based identification of Gram-positive cocci of the genus Enterococcus, Streptococcus, coagulase-negative Staphylococcus, and Lactococcus

Xuerui Li1, Juan Xing1, Baoyu Li1, Pu Wang2 and Jixing Liu1*

Abstract

Background: Accurate identification of isolates belonging to genus Enterococcus, Streptococcus, coagulase-negative Staphylococcus, and Lactococcus at the species level is necessary to provide a better understanding of their pathogenic potential, to aid in making clinical decisions, and to conduct epidemiologic investigations, especially when large blind samples must be analyzed. It is useful to simultaneously identify species in different genera using a single primer pair.

Methods: We developed a primer pair based on the tuf gene (encoding elongation factor) sequence to identify 56 Gram-positive cocci isolates.

Results: The target sequences were amplified from all 56 samples. The sequencing results and the phylogenetic tree derived from the partial tuf gene sequences identified the isolates as three enterococcal species, two lactococcal species, two staphylococcal species, and six streptococcal species, as well as eight isolates that were novel species of the genus Streptococcus. Partial gene sequence analysis of the sodA, dnaK, and 16S RNA genes confirmed the results obtained by tuf gene sequencing.

Conclusion: Based on the uniform amplification of the tuf gene from all samples and the ability to identify all isolates at both the genus and species levels, we conclude that the primer pair developed in this research provides a powerful tool for identifying these organisms in clinical laboratories where large blind samples are used.

Keywords: Enterococcus, Streptococcus, Staphylococcus, Lactococcus, tuf, species identification

Background

Bacterial species of the genera Enterococcus, Streptococcus, coagulase-negative Staphylococcus, and Lactococcus are Gram positive cocci in the class Bacillii. Enterococcus, streptococcus, and staphylococcus inhabit a wide range of environments, including the gastrointestinal tracts of humans and animals. Lactococcus is a genus of lactic acid bacteria that were formerly included in the genus Streptococcus, and have been widely used in the production of cheese and milk products.

Accurate identification of isolates belonging to these genera at the species level is necessary to provide a better understanding of their pathogenic potential, to aid in making clinical decisions, and to conduct epidemiologic investigations. Because species identification based on phenotypic characterization is time-consuming and can produce ambiguous results [1-6], molecular identification methods have taken precedence. Of the molecular methods used, polymerase chain reaction (PCR) sequencing-based methods are powerful tools for identifying species both within [7-10] and between genera [11-14]. When large blind samples must be analyzed, it is useful to simultaneously identify species in different...
genera using a single primer pair. Other than the 16S rRNA gene-based methods [15], several PCR-based sequencing tools have been developed to identify species in the class Bacilli. rpoB gene-based PCR sequencing has been used for accurate detection and identification of species in the genera Streptococcus, Enterococcus, Gemella, Abiotrophia, Granulicatella [11], and Staphylococcus [16]. The groES and groEL genes have also been used as a target for identification of Abiotrophia, Granulicatella, and Gemella species [12].

The tuf gene, which encodes the elongation factor EF-Tu, is involved in peptide chain formation. It is a ubiquitous and highly evolutionarily conserved part of the core genome, and is more discriminative than the 16S rRNA gene for identifying strains belonging to the genera Enterococcus [17], Staphylococcus [18], and Streptococcus [19]. In this study, we developed a simple tuf gene based PCR and sequencing assay to identify isolates belonging to the genera Enterococcus, Staphylococcus, Streptococcus, and Lactococcus. Analysis based on the partial tuf gene sequence revealed that the target could be amplified from all isolates used in this study, and that it is superior to previous techniques for differentiating the strains at the species level.

**Methods**

**Bacterial strains**

The strains examined in this study included *Streptococcus suis* (S.suis)serotype 2 strain 9801, S. suis 05ZYH33, one Lactococcus lactis subsp. lactis isolate, 48 cocci isolates that was isolated by our lab between 2006 and 2011, and four cocci strains purchased from the Chinese General Microbiology Culture Collection Center (CGMCC), which included the following: *Streptococcus peryi* (S.peryi) CGMCC1.1010, *Streptococcus salivarius* CGMCC 1.2498, *Streptococcus mutans* CGMCC 1.2499, and *Streptococcus bovis* CGMCC1.2502. *Streptococcus cremoris* (L. lactis subsp. cremoris) CICC 20175 was also purchased from the Chinese Center of Industrial Culture Collection (CICC).

**Isolation and sequencing of genomic DNA**

Genomic DNA was isolated and purified using a GenElute Bacterial Genomic DNA Kit (Sigma) according to the manufacturer’s instructions. PCR primers complementary to highly conserved regions of the tuf gene were designed (The primers were designed on the basis of *Streptococcus suis* P1/7 tuf gene sequence. However, before designing these, the tuf gene sequences of all the published sequences in genus *Streptococcus* were aligned to identify sequences that were suitable for primer design. Table 1) and used to amplify partial fragments of the tuf gene from all isolates. The amplification was performed in the personal thermal thermocycler (BioRad MJ-Mini), and PCR conditions were optimized as follows: a total reaction mix of 50 μl contained 0.5 μl (2.5U/ul) of PrimerSTAR HS DNA polymerase (Takara), 5× PrimerSTAR buffer, 4 μl dNTP Mixture (250 μM each), 1 μl of each primer and 1 μl of genomic DNA template. PCR was performed in a DNA thermal cycler with 30 cycles of 94°C for 30 s, 50–55°C for 1 min, 72°C for 1 min 30 s, followed by a final extension step of 72°C for 10 min. The PCR products were collected using Takara MiniBEST DNA Fragment Purification Kit and ligated into PMD18 vector (Takara) and three positive clones were sent to the Shanghai Sangon Biotechnology Company for sequencing. The PCR products were sequenced using a BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems) using a PRISM3730 Genetic Analyzer (Applied Biosystems).

For the isolates that were determined to be either Staphylococcus or Enterococcus species, partial sodA gene regions were amplified and sequenced to confirm the result, while *Streptococcus* species isolates were confirmed by amplification and sequencing of the partial dnaK gene. Isolates that could not be identified at the species level following comparison with available tuf gene sequence data published in Genbank and EMBL were further analyzed by 16S rRNA gene amplification and sequence analysis. PCR primers are listed in Table 1.

**Table 1 Primers used in this study**

| Primers       | Sequence (5_ to 3_)          | Gene    | Primers from          |
|---------------|------------------------------|---------|-----------------------|
| Tuf-F         | S'-CCAATGCGCCACAAACTCGT -3'  | tuf     | This study            |
| Tuf-R         | S'-CCTGAACCAACAGTACGT -3'    |         |                       |
| D1            | S'-CCITAYICTAYGAYGCIYTIGARCC-3' | sodA    | reference [35,36]     |
| D2            | S'-ARRTARTAIGCRTGYTCCCAIACRTC-3' | dnaK    | This study            |
| dnaK-F        | S'-CTTGGTGGTGACGACTTTGAC -3' |         |                       |
| dnaK-R        | S'-CCACCCATTTGCTTGCATACCA-3' |         |                       |
| 16s F         | S'-AGAGTTTGATCTCGGCTACG -3'  | 16S RNA | Reference [37]        |
| 16s R         | S'-AAAGGAGGTTGATCCAGCC-3'    |         |                       |
Phylogenetic relationships
The phylogenetic relationships among species were analyzed using the neighbor-joining method in MEGA 5.0 [20]. For this analysis, distances between the sequences were calculated using Kimura’s two-parameter model [20-23]. Levels of similarity were determined among species. Bootstrap values were obtained for 500 randomly generated trees.

Nucleotide sequence accession numbers
Nucleotide sequences determined in this study were submitted to GenBank (http://www.ncbi.nlm.nih.gov/genbank/). tuf gene nucleotide sequences were submitted under GenBank accession numbers JX436506–JX436520; sodA nucleotide sequences were submitted under GenBank accession numbers JX436496–JX436497; L. lactis subsp lactis strain LVR1001 dnaK sequence was submitted under GenBank accession number 436495; and 16S rRNA nucleotide sequences were submitted under GenBank accession numbers JQ255459–JQ255462.

Results
Identification of isolates
Partial tuf gene sequences were amplified from all 56 strains and then sequenced and compared. In the identified isolates, tuf was 827 bp long in all Staphylococcus warneri strains and 830 bp in all other strains.

NCBI BLASTN analysis revealed that the 48 isolates collected by our lab included eight Streptococcus uberis strains that have 99.8%–100% sequence similarity to each other. The nucleotide sequences of these tuf genes shared approximately 99.8% nucleotide similarity with that of Streptococcus uberis 0140J. Six Streptococcus thermophilus (S. thermophilus) isolates were identified, all of which showed 100% sequence similarity to the tuf gene of S. thermophilus ND03, while a further eight isolates were identified as Staphylococcus warneri (S. warneri), which showed 100% nucleotide sequence similarity to each other, and 99.5% (777/781 bp) similarity to the tuf gene of S. warneri ATCC 27836. Six strains were identified as Staphylococcus hyicus (S. hyicus) and shared 100% identity with each other and 99.9% (705/706 bp) identity with S. hyicus strain CIP 81.58. Seven Enterococcus faecalis (E. faecalis) strains were also identified and showed 99% identity to the tuf sequence of E. faecalis V583. Five Enterococcus avium (E. avium) isolates shared 99.6%–100% sequence similarity with each other and 98.6% (706/717 bp and 707/717 bp) sequence similarity to E. avium ATCC 14025. Finally, we identified eight isolates as belonging to the genus Streptococcus, but could not make a species identification based on comparison with published tuf sequence data in GenBank or EMBL. These sequences shared 99.0%–100% sequence similarity with each other, and the highest tuf gene homology was with S. suis A7 (94.8%, CP002570.1 SSUA7_0486). The levels of partial tuf gene sequence divergence between the isolates and representatives of other Staphylococcus species were in the range 5.1–14.2%.

The tuf gene diversity was calculated using Kimura’s two-parameter mode using Mega 5.0. At the genus level, the tuf sequence diversity among different species within the genus Streptococcus ranged from 1.1–12.9%. Sequence diversity ranged from 9–12.3% among different species in the genus Enterococcus. For genus Staphylococcus, diversity ranged from 4.7–10.1%.

Confirmation of results using sodA or dnaK gene sequences
Because the isolates might be misidentified when using a single gene sequence for identification, we used a second partial gene sequence to confirm the result obtained by tuf gene analysis. The sodA gene was used for confirmation of the isolates identified as belonging to the genera Staphylococcus and Enterococcus. For the isolates in the genus Streptococcus, a partial dnaK sequence was amplified and sequenced. The sequencing results confirmed the results obtained from the partial tuf gene analysis.

When S. warneri strains used in this research showed 99.5% (777/781 bp) nucleotide sequence (nt) similarity to S. warneri ATCC 27836, it also showed 99.5% (767/771) nt sequence similarity to Staphylococcus pasteuri (S. pasteuri) FI64. For this reason, the partial sodA sequences were then used for analysis. The sodA gene of the S. warneri strains isolated in our lab were found to show 99.8% (428/429 bp) sequence similarity to the sodA gene of S. warneri CIT S00-147. However, they were clearly distinguished from S. pasteuri strain CIP 103540 T, with a nucleotide sequence similarity of only 94.3% (396/420 bp).

Identification of S. pernyi CGMCC1.1010 as Enterococcus mundtii
The strain S. pernyi CGMCC1.1010 was obtained from the CGMCC. In a recent work by Cui-Fang et al., this strain was suggested to be an enterococcal species [24].

Based on partial tuf gene analysis, we determined that the tuf gene of S. pernyi CGMCC1.1010 has 98.9% nucleotide identity (792/801 bp) to that of E. mundtii strain ATCC43186. Furthermore, analysis of the partial sodA sequence revealed that sodA of S. pernyi CGMCC1.1010 has 99.8% nucleotide identity (437/438 bp) to that of E. mundtii strain ATCC 43186. Thus, we confirmed that CGMCC1.1010 is actually an E. mundtii strain.

Phylogenetic analysis of eight Streptococcus isolates
Phylogenetic analysis was carried out using the neighbor-joining method (reference strains and accession numbers are listed in Additional file 1. See Figure 1).
Phylogenetic analysis identified 40 of the 48 isolates at both the genus and species levels and agreed with results of the BLAST analysis. The eight isolates that shared 90–100% sequence similarity with each other, but could not be identified at the species level by BLAST analyses were clustered into the same group. They were further analyzed by 16S rRNA gene amplification and sequencing. Preliminary 16S rRNA gene sequence analysis (1550 bp) found that the eight isolates have more than 99.7% similarity to each other. BLAST analyses of these 16S rRNA sequences determined that the isolates showed the greatest sequence similarity to *Streptococcus minor* (*S. minor*) strain LMG 21735 (96.8%, AY232833.1) and *S. minor* strain ON59 (96.8%, AB559734.1). Because isolates that show ≤ 98.7% 16S rRNA sequence similarity are always considered members of a different species [25], we designated these isolates as a novel species of *Streptococcus* and named this species *Streptococcus parasuis*. Further biochemical research is being conducted to classify and confirm this suspected novel species.

**Discussion**

The application of tuf gene analysis to molecular identification has been evaluated for many bacterial species, including *Staphylococcus*, *Streptococcus*, *Enterococcus*, *Campylobacter*, and *Aeromonas* species [26,27]. In the current study, we assessed the feasibility of sequencing partial tuf genes for identifying Enterococcus, Staphylococcus, Lactococcus, and Streptococcus species. Sequencing results, a phylogenetic tree derived from the partial tuf gene sequences, and confirmatory sodA, dnaK, and 16S RNA gene sequence analyses revealed that the partial tuf gene sequence used here could identify the majority of isolates at both the genus and species levels. In addition to identifying isolates of known species, the partial tuf gene sequences could also be used to identify suspected novel species, with a total of eight isolates considered to belong to a novel species, based on tuf and 16S rRNA gene sequence analysis. The data presented here demonstrate that PCR amplification of the tuf gene fragment using our specifically designed primer pair, followed by sequence analysis, is a suitable molecular approach for the identification of Enterococcus, Staphylococcus, Lactococcus, and Streptococcus isolates at the species level.

As it is often difficult to identify isolates at the species level relying only on single gene based primers [28-31], partial gene sequences of sodA, dnaK, or the 16S rRNA gene were also used to confirm our results.

Reliable discrimination between closely related species depends on the variability of the target gene. A high level of variability might be helpful for good discrimination between species, but it can also be a disadvantage because of instability of species-specific signatures and
difficulties in developing reliable primers or probes. Because of the high level of variability, the higher the discriminating power, the higher the proportion of strains that are not amplified using a single primer pair [32–34]. In this research, the primer pair was based on S. suis sequences. Because the primers were not degenerate, we expected to have to adjust the nucleotide sequences to get good amplification from all strains. However, the primer pair worked well and all the appropriate target sequences were amplified from 14 species in four genera (56/56).

The tuf gene-based primer pair designed in the present study may be helpful for the accurate detection and identification of Enterococcus, Staphylococcus, Lactococcus, and Streptococcus species, as well as related genera of medical interest.

Conclusions
In conclusion, this study confirms that tuf is a good alternative molecular marker for both phylogenetic analysis and species identification of clinical isolates when large blind samples are used. It should be applied to phylogeny as a first-line genomic technique.

Additional file

Additional file 1: Sources and gene accession numbers of the bacterial reference strains used in this study.

Abbreviations
Tuf: Elongation factor Tu; S.suis: Streptococcus suis; S.pernyi: Streptococcus peryi; S. thermophilus: Staphylococcus thermophilus; S. warneri: Staphylococcus warneri; S. hyicus: Staphylococcus hyicus; E. faecalis: Enterococcus faecalis; E. avium: Enterococcus avium; S. pasteurii: Staphylococcus pasteurii; E. mundtii: Enterococcus mundtii; S. minor: Streptococcus minor.

Competing interests
Authors have no competing interests.

Authors’ contributions
Xuelei Li, Jiaxin Li and Baoyu Li co-worked on data collection and organisation, performed statistical analysis of the data and contributed to writing and interpretation of the manuscript. Pu Wang wrote the manuscript. Jixing Liu contributed to the design and writing of the manuscript. All authors have read and approved the manuscript.

Acknowledgements
This study was supported by the Gansu Natural Science Foundation 325061-A25-071, Gansu Science and Technology Project 8801NKDA035, and by the National Basic Research Priorities Program of China: 2008FY210200.

Author details
1State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Grazing Animal Diseases of the Ministry of Agriculture, Key Laboratory of Veterinary Public Health of the Ministry of Agriculture; Animal Infectious Diseases Research Laboratory, Lanzhou Veterinary Research Institute, CAAS, Lanzhou 730046, China. 2The Wistar Institute, Philadelphia, PA 19104, USA.

Received: 24 August 2012 Accepted: 15 November 2012 Published: 27 November 2012

References
1. Calvo JJ, et al: Osteomyelitis caused by Staphylococcus schleiferi and evidence for misidentification of this Staphylococcus species by an automated identification system. J Clin Microbiol 2000, 38:3887–3889.
2. Flynn CE, Ruoff KL: Identification of Streptococcus milleri group isolates to the species level with a commercially available rapid test. J. Clin. Microbiol 1995, 33:2704–2706.
3. Leven MJ, et al: Rapid and economical method for species identification of clinically significant coagulase-negative staphylococci. J Clin Microbiol 1995, 33:1060–1063.
4. Moore DF, et al: Comparison of 16S rRNA sequencing with conventional and commercial phenotypic techniques for identification of enterococci from the marine environment. J. Appl. Microbiol 2006, 100:1272–1281.
5. Teng Li, et al: GroEL sequence determination, phylogenetic analysis, and species differentiation for viridans group streptococci. J. Clin. Microbiol 2002, 40:3172–3178.
6. Utbach EB, et al: The ldh phylogeny for environmental isolates of Lactococcus lactis is consistent with rDNA genotypes but not with phenotypes. Applied Environ. Microbiol 1997, 63:694–702.
7. Fihman V, et al: Lactococcus garvieae endocarditis: identification by 16S rRNA and sodA sequence analysis. J Infect 2006, 52(1):e3–e6.
8. Ghebremedhin B, Lauer F, König W, König B: Genetic classification and distinguishing of Staphylococcus species based on different partial gap, 16S rRNA, hsp60, rpoB, sodA, and tuf gene sequences. J Clin Microbiol 2008, 46(3):1019–1025.
9. Glazunova OO, Raout D, Roux V: Partial recA gene sequencing: a new tool for identification and phylogeny within the genus Streptococcus. Int J Syst Evol Microbiol 2010, 60(Pt 9):2140–2148.
10. Rahkila R, et al: Identification of enterococci from broiler products and a broiler processing plant and description of Enterococcus vilkiensissp. nov. Appl Environ Microbiol 2011, 77(4):1196–1203.
11. Drancourt MN, et al: rpoB gene sequence-based identification of aerobic Gram-positive cocci of the genera Streptococcus, Enterococcus, Gemella, Abiotrophia, and Granulicatella. J. Clin Microbiol 2004, 42:497–504.
12. Hung WC, et al: Use of groESL as a target for identification of Abiotrophia, Granulicatella, and Gemella species. J Clin Microbiol 2010, 48:3532–3538.
13. Sakamoto M, Ohkuma M: Usefulness of the hsp60 gene for the identification and classification of Gram-negative anaerobic rods. J Med Microbiol 2010, 59(Pt 11):1293–1302.
14. Wang T, et al: Improved identification of Gordonia, Rhodococcus and Tsukamurella species by 16F-end 16S rRNA gene sequencing. J Med Microbiol 2011, 43(1):58–63.
15. Woo PC, et al: 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 2003, 41:1996–2001.
16. Drancourt M, Raout D: rpoB gene sequence-based identification of Staphylococci species. J Clin Microbiol 2002, 40(4):1333–1338.
17. Ke D, et al: Development of a PCR assay for rapid detection of enterococci. J Clin Microbiol 1999, 37(11):3497–3503.
18. Martineau F, et al: Development of a PCR assay for identification of staphylococci at genus and species levels. J. Clin Microbiol 2001, 39(7):2541–2547.
19. Picard FJ, et al: Use of tuf sequences for genus-specific PCR detection and phylogenetic analysis of 28 streptococcal species. J Clin Microbiol 2004, 42(8):3686–3695.
20. Tanaka K, et al: MEGAS: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011, 28(10):2731–2739.
21. Felsenstein J: Confidence limits on phylogenies: An approach using the bootstrap. Evolution 1985, 39:783–791.
22. Kimura M: A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 1980, 16:111–120.
23. Saitou N, Nei M: The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 1987, 4:406–425.
24. Cui-Fang S, et al: Identification of the Pathogen for Antheraea pernyi Empty-gut Disease by Using 16S rRNA Gene Sequence. Science of Silkworm Biology 2011, 37(5):0931–0936.
25. Achtman M, Wagner M: Microbial diversity and the genetic nature of microbial species. Nat Rev Microbiol 2008, 6(6):431–440.

26. Carpaij N, et al: Comparison of the identification of coagulase-negative staphylococci by matrix-assisted laser desorption ionization time-of-flight mass spectrometry and tuf sequencing. Eur J Clin Microbiol Infect Dis 2011, 30:1169–1172.

27. Loonen AJ, et al: Comparative study using phenotypic, genotypic, and proteomics methods for identification of coagulase-negative staphylococci. J Clin Microbiol 2004, 42:4988–4995.

28. Becker K, et al: Development and evaluation of a quality-controlled ribosomal sequence database for 16S ribosomal DNA-based identification of Staphylococcus species. J Clin Microbiol 2004, 142:4988–4995.

29. Bialkowska-Hobrzanska H, et al: Typing of coagulase-negative staphylococci by Southern hybridization of chromosomal DNA fingerprints using a ribosomal RNA probe. Eur J Microbiol Infect Dis 1990, 9:588–594.

30. Buyser D, et al: Evaluation of ribosomal RNA gene probe for the identification of species and sub-species within the genus Staphylococcus. J Gen Microbiol 1992, 138:889–899.

31. Sakamoto M, Kitahara M, Benno Y: Parabacteroides johnsonii sp. nov., isolated from human faeces. Int J Syst Evol Microbiol 2007, 57:293–296.

32. Kim BJ, et al: Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (rpoB). J Clin Microbiol 1999, 37:714–71720.

33. Mignard S, Flandrois JP: Identification of Mycobacterium using the EF-Tu encoding (tuf) gene and the tmRNA encoding (ssrA) gene. J Med Microbiol 2007, 56(Pt 8):1033–1041.

34. Roth A, et al: Differentiation of phylogenetically related slowly growing mycobacteria based on 16S–23S rRNA gene internal transcribed spaces sequences. J Clin Microbiol 1998, 36:139–147.

35. Poyart C, Quesnes G, Trieu-Cuot P: Sequencing the gene encoding manganese-dependent superoxide dismutase for rapid species identification of enterococci. J Clin Microbiol 2000, 38(1):415–418.

36. Poyart C, et al: Rapid and accurate species-level identification of coagulase-negative staphylococci by using the sodA gene as a target. J Clin Microbiol 2001, 39:4296–4301.

37. Lane DJ: 16S/23S rRNA sequencing. In Nucleic Acid Techniques in Bacterial Systematics. Edited by E. Stackebrandt & M. Goodfellow. Chichester, UK: Wiley; 1991:115–175.

doi:10.1186/1476-0711-11-31
Cite this article as: Li et al.: Use of tuf as a target for sequence-based identification of Gram-positive cocci of the genus Enterococcus, Streptococcus, coagulase-negative Staphylococcus, and Lactococcus. Annals of Clinical Microbiology and Antimicrobials 2012 11:31.

Submit your next manuscript to BioMed Central and take full advantage of:
• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit