Use of the \textit{gyrB} gene for the identification of \textit{Pandoraea} species

Tom Coenye *, John J. LiPuma

Department of Pediatrics and Communicable Diseases, University of Michigan Medical School, 8301 MSRB III, Box 0646, 1150 W. Med. Ctr. Dr.,
Ann Arbor, MI 48109-0646, USA

Received 9 November 2001; accepted 27 November 2001
First published online 29 January 2002

Abstract

The recently described genus \textit{Pandoraea} consists of five named species and four unnamed genomospecies, several of which have been identified in clinical specimens including respiratory secretions from persons with cystic fibrosis. We investigated whether it is possible to distinguish species of the genus \textit{Pandoraea} by means of restriction fragment length polymorphism (RFLP) analysis and direct sequencing of the \textit{gyrB} gene. Sixty-seven \textit{Pandoraea} isolates were included. Species-specific RFLP patterns were obtained following digestion of the PCR-amplified \textit{gyrB} gene with \textit{MspI}. Specificity of RFLP groupings was confirmed by direct sequencing of several representative isolates. Our results indicate that RFLP analysis and sequencing of the \textit{gyrB} gene are useful for the identification of \textit{Pandoraea} species. We also found that further taxonomic studies within the \textit{β-Proteobacteria} using the \textit{gyrB} gene would benefit from the development of additional primers allowing more efficient amplification of the \textit{gyrB} gene. Our data also indicate that the taxonomic status of \textit{Pandoraea} genomospecies 2 should be reinvestigated. © 2002 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: \textit{gyrB}; Identification; Cystic fibrosis; \textit{Pandoraea}; \textit{Burkholderia cepacia} complex

1. Introduction

It is generally accepted that the phylogenetic relationships between microorganisms can be deduced from sequence comparisons of conserved macromolecules. Ribosomal RNA genes are one of the best targets for phylogenetic studies because they are universally present and functionally constant and have a mosaic structure of highly conserved and more variable domains [1]. The direct sequencing of 16S and or 23S rDNA molecules by PCR technology provides a phylogenetic framework which serves as the backbone for modern microbial taxonomy [2]. However there is no threshold value for 16S rDNA homology for species recognition and due to the slow evolution of the 16S rDNA genes, recently diverged species may not be recognizable [3–6]. To circumvent these limitations it has been suggested that phylogenetic information derived from other genes could be used to complement the information obtained from 16S rDNA sequence analysis [5,6]. Other genes that have been used for bacteri-
species is also not straightforward and reliable identification often requires a polyphasic approach [18,19]. The aim of this study was to determine whether restriction fragment length polymorphism (RFLP) analysis and direct sequencing of the gyrB gene could be used for the identification of Pandoraea species.

2. Materials and methods

2.1. Bacterial strains and growth conditions

All 67 Pandoraea isolates included in this study have previously been described [18–20]. We included five P. norimbergensis isolates, three P. pulmonicola isolates, 26 P. apista isolates, 13 P. sputorum isolates and 15 P. pnomenusa isolates. In addition we also included strains belonging to the unnamed Pandoraea genomospecies 1, 2, 3 (one isolate each) and 4 (two isolates). Strains were grown aerobically on Mueller Hinton broth (Becton Dickinson) supplemented with 1.8% (w/v) agar and incubated overnight at 32°C. All strains were previously identified using a polyphasic–taxonomic approach and/or genus- and species-specific 16S rDNA-based PCR assays [18–20].

2.2. DNA preparation

DNA was prepared by heating one or two colonies (picked from an overnight grown culture) at 95°C for 15 min in 20 µl lysis buffer, containing 0.25% (v/v) SDS and 0.05 M NaOH. Following lysis, 180 µl sterile distilled water was added to the lysis buffer and the DNA solutions were stored at −20°C.

2.3. RFLP analysis of the gyrB gene

PCR reactions were performed in 25-µl reaction mixtures, containing 2 µl DNA solution, 1 U Taq polymerase (Qiagen), 250 mM (each) deoxynucleotide triphosphate (Gibco), 1 × PCR buffer (Qiagen) and 20 pmol of each oligonucleotide primer. The primers used were the universal primers described by Yamamoto and Harayama [10]: UP-1 (5’-GAAGTCATCATGACCGTTCTGCAY-3’ and UP-2r (5’-AGCA-GGTTACGGATGTGCAGCCRTCNACRTCGNCR-T-CNGTAT-3’). Amplification was carried out using a PTC-100 programmable thermal cycler (MJ Research). After initial denaturation for 2 min at 94°C, 30 amplification cycles were completed, each consisting of 1 min at 94°C, 1 min at 65°C and 1 min at 72°C. A final extension of 10 min at 72°C was applied. Restriction digests were performed in 20-µl reaction mixtures, containing 5 U restriction enzyme, 1 × appropriate buffer, 0.2 µl bovine serum albumin (10 mg ml⁻¹) and 8 µl PCR product. Restriction digests were carried out at the temperature recommended by the manufacturer for 3 h. Restriction enzymes used included MspI (Promega), RsaI, TaqI, HaeIII and DdeI (Gibco). Restriction fragments were separated in 10 × 15-cm 2% agarose gels (Genepure) in 0.5 × TBE buffer at 100 V for 75 min. A 100-bp DNA ladder (Gibco) was included on all gels to allow standardization and sizing. Gels were stained with ethidium bromide and visualized using the GelDoc System (Bio-Rad). Densitometric analysis, normalization and interpolation of the patterns and numerical analysis using the Dice coefficient were performed using the Quantity One 4.1 (Bio-Rad) and Molecular Analyst (Bio-Rad) software packages.

2.4. Partial sequence determination of the gyrB gene

The gyrB gene was amplified using PCR, as described above for the RFLP analysis. The PCR product was purified using the Promega Wizard PCR Prep kit (Promega) according to the manufacturer’s instructions. Sequence analysis was performed with an Applied Biosystems 3700 DNA sequencer and the protocols of the manufacturer (PE Applied Biosystems) using the BigDye Terminator Cycle Sequencing Ready Reaction kit. The sequencing primer used was the universal sequencing primer UP-1S (5’-GAAGTCATCATGACCGTTCTGCAY-3’). All the sequence reads were trimmed to approximately 550 bp. Phylogenetic trees based on the neighbor-joining method [21] were constructed using the MegAlign (DNASTar Inc.) software package.
2.5. Nucleotide accession numbers

All partial gyrB sequences have been deposited in GenBank with the following accession numbers: AF439524 (LMG 18819\textsuperscript{T}), AF439525 (LMG 18106\textsuperscript{T}), AF439526 (LMG 13019), AF439527 (LMG 16407\textsuperscript{T}), AF439528 (LMG 18087\textsuperscript{T}), AF439529 (LMG 18379\textsuperscript{T}), AF439530 (HI2800), AF439531 (HI2801), AF439532 (HI2802), AF439533 (LMG 20601), AF439534 (HI2804), AF439535 (LMG 20602) and AF439536 (LMG 20603).

3. Results

3.1. RFLP analysis of the gyrB gene

Using universal primer pair UP-1 and UP-2\textsubscript{r}, the complete gyrB gene (approximately 1200 bp) was successfully amplified from 66 of the 67 Pandoraea strains included in this study. Despite repeated attempts we were not able to amplify the gyrB gene from Pandoraea genomospecies 1 isolate R-5199. The efficiency of amplification of the gyrB

Fig. 2. Dendrogram derived from the UPGMA linkage of Dice coefficients between the gyrB RFLP patterns generated by digestion with MspI.
gene from all three *P. pulmonicola* isolates was low, but the PCR yielded sufficient product for our analyses (data not shown). We investigated the ability of several restriction enzymes to separate the different *Pandoraea* species. RFLP analysis with *Msp*I, *Rsa*I, *Taq*I, *Hae*III and *Dde*I revealed that RFLP types generated by digestion with *Msp*I were the most discriminatory among the enzymes tested (data not shown). RFLP analysis with *Msp*I was performed on all strains (except R-5199); resultant RFLP patterns consisted of three or four bands in the size range of 134–749 bp (Fig. 1). Following numerical analysis RFLP patterns could be grouped into six clusters, while one strain (*Pandoraea* genomospecies 3 LMG 20602) occupied a separate position (Fig. 2). *P. norimbergensis*, *P. pulmonicola* and *P. apista* isolates formed clusters I, II and III respectively. The two representatives of *Pandoraea* genomospecies 4 formed cluster IV. Cluster V is composed of isolates belonging to *P. sputorum* and *Pandoraea* genomospecies 2. Two subclusters can be seen in cluster VI, both comprising *P. pnomenusa* isolates.

### 3.2. Sequence analysis of the gyrB gene of *Pandoraea* species

Using sequencing primer UP-1S we obtained high-quality sequence information on approximately 550 nucleotides of the *gyrB* gene. Using the neighbor-joining method [21], the phylogenetic tree shown in Fig. 3 was constructed. The three *P. pnomenusa* isolates included formed one cluster and shared between 81.1 and 88.1% sequence similarity. Similarity values between the *P. apista* isolates ranged from 92.9 to 95.4%. Both *P. norimbergensis* sequences shared 93.7% sequence similarity, while *P. sputorum* LMG 18819T and *Pandoraea* genomospecies 2 LMG 20601 shared 90.0% sequence identity. Similarities between those different groups were between 88.2 and 68.1%. Isolates representing *Pandoraea* genomospecies 3 and 4 and *P. pulmonicola* occupied distinct positions in the phylogenetic tree.

### 4. Discussion

Comparison of 16S rDNA sequences has been used as gold standard in bacterial phylogeny [1,2], but its resolution may be too low to distinguish closely related species [3–6]. In this study we examined whether it was possible to use the *gyrB* gene as a complement to 16S rDNA sequences for the identification of *Pandoraea* species. *Pandoraea* species share high levels of 16S rDNA sequence similarity and correct identification using conventional methods is difficult [18,19].

As can be seen from Figs. 1 and 2, *gyrB* RFLP patterns were identical within a given species and different patterns were seen in different *Pandoraea* species. There were two exceptions: two slightly different RFLP profiles were present in *P. pnomenusa* isolates (cluster VI), while *Pandoraea* genomospecies 2 could not be differentiated from *P. sputorum*. To confirm the data obtained by RFLP analysis of the *gyrB* gene we determined partial *gyrB* sequences for isolates from each cluster and, as could be expected, isolates from the same RFLP cluster also shared high *gyrB* sequence similarities and could easily be differentiated from members of other RFLP clusters. There seems to be more variation in the *gyrB* sequences of *P. pnomenusa* isolates than in the sequences of other *Pandoraea* isolates, confirming the variation seen in the RFLP patterns of isolates of this species. The high similarity of the RFLP patterns of *Pandoraea* genomospecies 2 and *P. sputorum* was confirmed by sequence analysis (*P. sputorum* LMG 18819T and *Pandoraea* genomospecies 2 LMG 20602 share 90.0% sequence similarity). These isolates also have almost identical 16S rDNA sequences and biochemical characteristics [19]. *Pandoraea* genomospecies 2 also cross-reacts with the PCR primers developed for the identification of *P. sputorum* [20]. An uncharacteristically low DNA–DNA hybridization value has been reported between LMG 18819T and LMG 20602 (4%) while interspecies DNA–DNA hybridization values within the genus *Pandoraea* are generally between 25 and 45% [18,19]. Based on these findings the taxonomic status of *Pandoraea* genomospecies 2 should be reinvestigated.

The universal primers UP-1 and UP-2r were originally
developed based on three protein sequences [10]. So far, the main focus of research on the use of gyrB sequences in identification and phylogeny has been on members of the α and γ subclasses of the Proteobacteria and on members of the Cytophaga-Flavobacterium-Bacteroides phylum; very few members of the β subclass of the Proteobacteria have been investigated so far. Although we were able to successfully amplify the gyrB gene from most Pandoraea isolates, we were unable to amplify gyrB from Pandoraea genomospecies 1 R-5199 and amplification of gyrB from all P. pulmonicola isolates was problematic. Several additional primers have been developed that allow amplification of gyrB fragments of specific bacterial taxa (for an overview see http://cod.mbio.co.jp/icb/index.php). Our data suggest that the development of primers more suitable for amplification and sequencing of the gyrB gene from organisms belonging to the β-Proteobacteria might be necessary in order to extend the application of gyrB analysis to these taxa.

In summary, our results indicate that RFLP analysis and direct sequencing of the gyrB gene are a useful tool for the identification of Pandoraea species as all species could easily be identified using both approaches. Our results also indicate that the taxonomic status of Pandoraea genomospecies 2 should be reinvestigated. Further application of the gyrB gene in studies concerning β-Proteobacteria might benefit form the development of novel primers more suitable for the amplification of the gyrB gene from organisms belonging to this group.

Acknowledgements

This work was supported by a grant from the Cystic Fibrosis Foundation (USA) (to J.J.L.). T.C. is supported by the Caroll Haas Research Fund in Cystic Fibrosis.

References