JOURNAL OF CLINICAL MICROBIOLOGY, Oct. 2003, p. 4647–4654 Vol. 41, No. 10
0095-1137/03/$08.00+0 DOI: 10.1128/JCM.41.10.4647–4654.2003

Use of 16S rRNA Gene Sequencing for Rapid Identification and Differentiation of *Burkholderia pseudomallei* and *B. mallei*

Jay E. Gee,[1]* Claudio T. Sacchi,[1,2] Mindy B. Glass,[1] Barun K. De,[1] Robbin S. Weyant,[3] Paul N. Levett,[1] Anne M. Whitney,[1] Alex R. Hoffmaster,[1] and Tanja Popovic[1]

Meningitis and Special Pathogens Branch, Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases,[1] and Office of Health and Safety,[3] Centers for Disease Control and Prevention, Atlanta, Georgia 30333, and Adolfo Lutz Institute, São Paulo, Brazil[2]

Received 19 May 2003/Returned for modification 7 July 2003/Accepted 14 July 2003

***Burkholderia pseudomallei* and *B. mallei*, the causative agents of melioidosis and glanders, respectively, are designated category B biothreat agents. Current methods for identifying these organisms rely on their phenotypic characteristics and an extensive set of biochemical reactions. We evaluated the use of 16S rRNA gene sequencing to rapidly identify these two species and differentiate them from each other as well as from closely related species and genera such as *Pandoraea* spp., *Ralstonia* spp., *Burkholderia gladioli*, *Burkholderia cepacia*, *Burkholderia thailandensis*, and *Pseudomonas aeruginosa*. We sequenced the 1.5-kb 16S rRNA gene of 56 *B. pseudomallei* and 23 *B. mallei* isolates selected to represent a wide range of temporal, geographic, and origin diversity. Among all 79 isolates, a total of 11 16S types were found based on eight positions of difference. Nine 16S types were identified in *B. pseudomallei* isolates based on six positions of difference, with differences ranging from 0.5 to 1.5 bp. Twenty-two of 23 *B. mallei* isolates showed 16S rRNA gene sequence identity and were designated 16S type 10, whereas the remaining isolate was designated type 11. This report provides a basis for rapidly identifying and differentiating *B. pseudomallei* and *B. mallei* by molecular methods.**

Burkholderia pseudomallei is the causative agent of melioidosis (37), whereas infection with *Burkholderia mallei* causes glanders (33). Both pathogens are listed as category B biothreat agents by the Centers for Disease Control and Prevention's Strategic Planning Workgroup because of their availability and potential to cause illnesses with high morbidity and mortality (27). There is documentation that *B. mallei* was used as a biowarfare agent in World War I (36).

Melioidosis is endemic in tropical regions, principally southeast Asia and northern Australia. Melioidosis is particularly problematic in Thailand, where it commonly affects rice farmers. Although melioidosis is generally considered a human disease, it is also manifested in a wide variety of animals, such as horses, rats, marine mammals, and even birds (8, 14, 21). In the mid-1970s, an outbreak in France caused widespread economic disruption and resulted in the destruction of a large number of animals, especially horses (21). The epidemiology of melioidosis is complicated by the environmental persistence of *B. pseudomallei*, which creates the reservoir of infection. Indeed, the presence of *B. pseudomallei* in soil samples years after the referenced outbreak in France demonstrates its ability to persist even in temperate climates (21). The growing recognition that melioidosis can pose a public health threat in temperate regions and reports that *B. pseudomallei* can be disseminated by animal carriage have increased interest in developing rapid diagnostic assays in areas where it is not endemic (3, 8, 21).

* Corresponding author. Mailing address: Meningitis and Special Pathogens Branch, Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, CDC, MS-D11, 1600 Clifton Rd., N.E., Atlanta, GA 30333. Phone: (404) 639-4936. Fax: (404) 639-4421. E-mail: JGee1@cdc.gov.

Glanders is predominantly an equine disease that can be transmitted to humans from infected equids. In the 19th century, glanders caused substantial economic damage during outbreaks among livestock. *B. mallei* is an obligate intracellular parasite that was eradicated from circulation in North America by 1938 through the destruction of large numbers of horses in both Canada and the United States. Glanders still appears sporadically in South America, Asia, and Middle Eastern countries (9).

The diagnosis of melioidosis and glanders relies on an extensive set of biochemical tests and observation of colony and cell morphology, which may take up to 7 days to complete (35). In endemic regions these tests are considered reliable; however, in regions where *B. pseudomallei* and *B. mallei* are seldom encountered, there is the possibility of misidentification of the organisms (15, 19). For both organisms, biosafety level 3 laboratory facilities are recommended when there is a potential risk of aerosolization of the pathogens because of the risk of laboratory-acquired infections. Therefore, molecular methods that reduce exposure of laboratory personnel to potentially infectious samples are needed (3, 32). Rapid identification by molecular methods may be useful in defining the source of infection, since melioidosis is primarily associated with environmental exposure to *B. pseudomallei* (7, 17), whereas glanders is associated with the handling of animals infected with *B. mallei* (9, 32).

Over the last few years, the increasing use of PCR, rapid template purification, and automated DNA sequencing has dramatically reduced the time necessary to yield a high-quality sequence. The use of 16S rRNA gene sequencing to study the relatedness of prokaryotic species is well established and has led to increased availability of 16S rRNA databases. The convergence of these technical and computational advances has

also enhanced the application of 16S rRNA gene sequence analysis to bacterial identification (2, 23, 25). It was recently reported that subtle sequence differences in the 16S rRNA gene could be used for species identification (28) and for subtyping and identifying hypervirulent bacterial clones (4, 20, 22). Consequently, the goals of our study were to determine the 16S rRNA gene sequences from *B. pseudomallei*, *B. mallei*, and closely related species to evaluate the sequences for diversity and to determine if this diversity could discriminate among study isolates sufficiently to provide a means for rapid identification of these two species.

MATERIALS AND METHODS

Bacterial strains. In this study, 56 *B. pseudomallei* and 23 *B. mallei* strains from a collection of over 300 were chosen to represent temporal (1949 to 2000), geographic (five continents), and origin (human, animal, and environmental) diversity (Table 1). A panel of 44 isolates closely related to *B. pseudomallei* and *B. mallei* were also sequenced for comparison and included *Pandoraea* spp., *Ralstonia* spp., *Burkholderia gladioli*, *Burkholderia cepacia*, *Burkholderia thailandensis*, and *Pseudomonas aeruginosa* (Table 2). All strains were stored at −70°C in defibrinated rabbit blood until they were tested. Identification of strains was carried out with a standard battery of biochemical tests (35).

Primer design. The previously published full-length 16S rRNA gene sequence (1,488 bp) of *B. pseudomallei* strain 1026b (GenBank accession no. U91839) (5) was queried against the draft version of the *B. pseudomallei* genome, strain K96243 (31), available on the Sanger Institute website. These sequence data were produced by the *B. pseudomallei* Sequencing Group at the Sanger Institute and can be obtained from http://www.sanger.ac.uk/Projects/B_pseudomallei/. The 16S rRNA gene sequence of strain 1026b was also used to query the *B. mallei* genome (strain ATCC 23344) on the preliminary sequence data on the Institute for Genomic Research (TIGR) website at http://www.tigr.org with the Blast algorithm (1).

With the sequences derived from the two websites, primers F229 (5′-CGC AAG CGA AAG TAT CAA GA-3′) and R1908 (5′-TTT ACA GCC GAT AAG CGT GAG-3′) were designed to amplify a 1.7-kb fragment that includes the full-length 16S rRNA gene for *B. pseudomallei* and *B. mallei* (Oligo Primer Analysis Software, version 6.57; Molecular Biology Insights, Inc., Cascade, Colo.).

Amplification of 16S rRNA genes. Whole-cell suspensions of bacteria were used for PCR. Bacteria were grown by plating one loop (1 μl) of stock cell suspension (heavy suspension of *Burkholderia* spp. in defibrinated rabbit blood, stored at −70°C until use) on Trypticase soy agar with 5% defibrinated sheep blood (BBL Microbiology Systems, Cockeysville, Md.) and incubating aerobically 1 to 2 days at 37°C. A single colony was suspended in 200 μl of 10 mM Tris (pH 8.0) in a 1.5-ml Millipore 0.22-μm filter unit (Millipore, Bedford, Mass.), heated at 95°C for 30 min, and centrifuged at 6,000 × *g* for 5 min. A DNA extract from *B. pseudomallei* strain K96243 was kindly provided by Mark Schell (University of Georgia) and resuspended in 50 μl of H_2O. Each final PCR (100 μl) contained 5 U of Expand DNA polymerase (Boehringer, Mannheim, Germany), 2 μl of DNA solution, undiluted or diluted 1:16 in H_2O, 10 mM Tris-HCl (pH 8.0), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each dATP, dCTP, dGTP, and dTTP, and 0.4 μM each primer. Reaction mixes were first incubated for 5 min at 95°C. Then, 35 cycles were performed as follows: 15 s at 94°C, 15 s at 60°C, and 1.5 min at 72°C. Reaction mixes were then incubated at 72°C for an additional 5 min. PCR products were purified with the Qiaquick PCR purification kit (Qiagen, Valencia, Calif.).

Strains used for comparison were also processed as stated for *B. pseudomallei* and *B. mallei* strains except that a set of universal primers, F8 (5′-AGT TTG ATC CTG GCT CAG-3′) and R1492 (5′-ACC TTG TTA CGA CTT-3′) (11), were used for amplification. Reaction mixes were first incubated for 5 min at 95°C. Then 35 cycles were performed as follows: 15 s at 94°C, 15 s at 50°C, and 1.5 min at 72°C. Reaction mixes were then incubated at 72°C for an additional 5 min.

16S rRNA gene sequencing. Sequencing primers were chosen from a panel of previously described oligonucleotides: F8 (described above), F357 (5′-TAC GGG AGG CAG CAG-3′), R357 (5′-CTG CTG CCT CCC GTA-3′), F530 (5′-CAG CAG CCG CGG TAA TAC-3′), R530 (5′-GTA TTA CCG CGG CTG CTG-3′), R790 (5′-CTA CCA GGG TAT CTA AT-3′), F790 (5′-ATT AGA TAC CCT GGT AG-3′) (34), F1068 (5′-GTC GTC AGC TCG TGT CGT GAG-3′), F1083 (5′-CGT GAC ATG TTG GGT TAA GTC-3′), F1390 (5′-GGG CCT TGT ACA CAC CG-3′), R1390 (5′-CGG TGT GTA CAA GGC CC-3′), R981 (5′-GGG TTG CGC TCG TTG CGG G-3′) (11, 28), and 16S5 (5′-AGT TTG ATC CTG GCT C-3′) (5). Amplification primers F229 and R1908 were also used as sequencing primers.

Sequencing was performed with an Applied Biosystems BigDye terminator cycle sequencing version 2.0 kit as per the manufacturer's instructions, except 6 μl of BigDye was used instead of 8 μl (Applied BioSystems, Foster City, Calif.). Sequencing products were purified with Centri-Sep spin columns (Princeton Separations, Adelphia, N.J.) and resolved with an Applied Biosystems model 3100 automated DNA sequencing system (Applied Biosystems).

Computer analysis of 16S rRNA gene sequences. (i) Determination of 16S types. The software used for all data analysis was the Genetics Computer Group Wisconsin Package version 10.2 (Accelrys, San Diego, Calif.) (10). The utilities used are shown in brackets. The raw trace files from the ABI 3100 sequencer were visually examined and edited [SEQMERGE]. For each *B. pseudomallei* and *B. mallei* strain, an inner segment of 1,488 bp was aligned with the previously published 16S rRNA sequence of *B. pseudomallei* strain 1026b (GenBank accession no. U91839) (5). A 16S type number was assigned to each different 16S rRNA sequence type, with 16S type 1 serving as the index sequence. Upon finding a novel 16S rRNA sequence, the amplification and sequencing of the gene were repeated to confirm the new 16S type. Numbering of base positions in this study is based from the beginning of the 1,488-bp segment. Position 1 corresponds to position 26 of the *Escherichia coli* system (5).

For *B. pseudomallei* and *B. mallei*, a 16S type was assigned for each unique sequence in the order it was discovered. Six of the 11 16S types are the result of single positions in each sequence that indicated a mixed base (Table 3). A mixed base occurs when there are multiple copies of a gene present with different bases at a given position (6). For example, if one copy of a gene has a C in a given position and another copy has a T in that position, then there will be overlapping peaks in the sequence trace file which may result in a pyrimidine (Y) base call. A purine (R) base call may result if G and A peaks overlap.

Comparisons of 16S rRNA gene sequences of *B. pseudomallei* and *B. mallei* to sequences of closely related strains. Initially, primers F229 and R1908 were used to attempt amplification of the 16S rRNA gene of the strains closely related to *B. pseudomallei* and *B. mallei*. *B. thailandensis* 2002721643 was amplified successfully; however, the 16S rRNA genes of the other strains often failed to amplify, so the universal primers 8F and 1492R were used as necessary. The 16S rRNA sequences from the amplicons that resulted from 8F and 1492R were nearly full length (1,429 bp to 1,466 bp). Primers 8F and 1492R worked well for both *B. pseudomallei* and *B. mallei*, but sequences were also shorter than full length.

The sequences were aligned by matching the bases of each sequence to the bases in the other sequences within each of the following subsets: *Pandoraea* spp., *Ralstonia* spp., *B. gladioli*, and *B. cepacia* complex isolates [PILEUP]. The sequences in each subset varied in length due to differences in sequence quality, so a core segment was selected for each subset that was common to all the sequences in a given subset. The level of divergence of the sequences in a subset was then determined with Jukes-Cantor correction [DISTANCES]. From each subset, the two most divergent sequences were compared to determine the level of similarity within a subset [BESTFIT] (Table 2). A representative sequence from each subset, as well as the 16S rRNA gene sequences from two *B. thailandensis* strains and a *P. aeurginosa* strain, were then compared to the 16S type 1 sequence to determine the level of similarity [BESTFIT] (Table 4).

The 16S rRNA gene sequence of 16S type 1 was compared to sequences previously submitted to GenBank by other laboratories to determine if 16S types could be assigned to those sequences. First the sequences were aligned so that base positions could be directly compared [PILEUP]. Then the aligned sequences were compared to see the number of bp differences among them [PRETTY] (10).

GenBank accession numbers. A total of 123 16S rRNA gene sequences were determined in this study (Tables 1 and 2). They were deposited in GenBank with accession nos. AY305738 to AY305760 (*B. mallei*), AY305763 to AY305818 (*B. pseudomallei*), AY268168 to AY268174 (*Pandoraea* spp.), AY268176 to AY268181 (*Ralstonia* spp.), AY268163 to AY268167 (*B. gladioli*), AY268140 to AY268162 (*B. cepacia*), AY268182 to AY268183 (*B. thailandensis*), and AY 268175 (*Pseudomonas aeruginosa*).

RESULTS

The 1,488-bp nucleotide sequences of the entire 16S rRNA gene from 56 *B. pseudomallei* and 23 *B. mallei* strains were generated, aligned, and compared. Differences were found at

TABLE 1. Designations of 56 *B. pseudomallei* and 23 *B. mallei* isolates analyzed in this study

Species (no. of strains)	Identifier	Other identifier	GenBank 16S rRNA gene accession no.	Geographic and/or temporal origin[a]	16S rRNA type
B. pseudomallei (56)	2002721184		AY305776	Human, Ecuador, 1962	1
	2002721638		AY305794	Human, Vietnam, 1963	1
	2002721161		AY305780	Human, US, 1970	1
	2002721162		AY305779	Human, Australia, 1970	1
	2002721171		AY305765	Human, Venezuela, 1976	1
	2002721630		AY305802	Environment, France, 1976	1
	2002721628		AY305804	Environment, Madagascar, 1977	1
	2002721181		AY305777	Human, US, 1979	1
	2002721096		AY305770	Human, US, 1981	1
	2002721102		AY305769	Human, US, 1983	1
	2002721635		AY305797	Human, Singapore, 1988	1
	2002721637		AY305795	Human, Pakistan, 1988	1
	2002721114		AY305787	Human, US, 1991	1
	2002721115		AY305786	Human, US, 1992	1
	2002721629		AY305803	Environment, Kenya, 1992	1
	2000032029		AY305818	Human, US, 1994	1
	2000032026		AY305815	India, 1995	1
	2002721123		AY305784	Human, Puerto Rico, 1998	1
	2002721124		AY305783	Human, US, 1999	1
	2000032028		AY305814	Human, US, 2000	1
	2003000540		AY305791	Human, US, 2002	1
	2002721617	NCTC 8016	AY305813	Sheep, Australia, 1949	2
	2002721186		AY305775	Human, US, 1966	2
	2002721146		AY305781	Human, US, 1969	2
	2002721209		AY305774	Monkey, US, 1969	2
	2002721177		AY305778	Human, US, 1977	2
	2002721639		AY305793	Human, Kenya, 1980	2
	2002721090		AY305771	Human, US, 1980	2
	2002721622		AY305809	Sheep, Australia, 1984	2
	2002721103		AY305768	Human, Netherlands, 1985	2
	2002721623		AY305808	Cow, Australia, 1985	2
	2002721633		AY305799	Human, Thailand, 1987	2
	2002721640		AY305792	Human, Papua New Guinea, 1989	2
	2002721625		AY305806	Environment, Singapore, 1991	2
	2002721634		AY305798	Human, Thailand, 1992	2
	2002721624		AY305807	Goat, Australia	2
	K96243[b]		AY305764	Human, Thailand, 1996	2
	2002721618		AY305812	Monkey, Philippines, 1990	3
	2002721619		AY305811	Monkey, Indonesia, 1990	3
	2002721620		AY305810	Horse, France, 1976	4
	2002721626		AY305805	Environment, Thailand, 1990	5
	2002721636	NCTC 10276	AY305796	Human, Bangladesh, 1960	6
	2002721145		AY305782	Human, Philippines, 1969	6
	2002721166		AY305766	Human, US, 1973	6
	2002721108		AY305767	Human, US, 1988	6
	2002721641		AY305790	Human, Fiji, 1992	6
	2002721646		AY305788	Human, Holland, 1999	6
	2002721647		AY305772	Human, UK, 1999	6
	2000032025		AY305816	Human, US, 2000	6
	2000032024		AY305817	Human, US, 2000	6
	2002721631		AY305801	Environment, Australia	6
	2002721632		AY305800	Environment, Australia	6
	2002721642		AY305789	Human, Malaysia	6
	2002721116		AY305785	Human, US, 1992	7
	2001029240		AY305773	Human, US, 2001	8
	2000032027		AY305763	US, 1968	9
B. mallei (23)	2002724306	NCTC 120	AY305739	UK, 1920	10
	2002724305	NCTC 3708	AY305740	Mule, 1932, India	10
	2002734301	NCTC 10260	AY305744	Human, Turkey, 1949	10
	2002724304	NCTC 10248	AY305741	Human, Turkey, 1950	10
	2002721273		AY305758	US, 1956	10
	2002721274		AY305759	US, 1956	10
	2002721275	ATCC 10399	AY305757	Horse, China, 1956	10
	2002721276		AY305756	US, 1956	10
	2002721277		AY305755	US, 1956	10
	2002734300	NCTC 10247	AY305745	Turkey, 1960	10

Continued on following page

TABLE 1—*Continued*

Species (no. of strains)	Identifier	Other identifier	GenBank 16S rRNA gene accession no.	Geographic and/or temporal origin[a]	16S rRNA type
	2002734299	NCTC 10229	AY305746	Hungary, 1961	10
	2000031063	ATCC 15310	AY305760	Horse, Hungary 1961	10
	2002721278		AY305754	Human, US, 1964	10
	2002721279		AY305753	Human, US, 1964	10
	2002721280		AY305752	France, 1972	10
	2000031304		AY305748	Human, US, 2000	10
	2002721648	ATCC 23344	AY305747	Human, China	10
	2000031064		AY305751	India	10
	2000031066		AY305749	India	10
	2000031065		AY305750	Turkey	10
	2002734302		AY305743	Turkey	10
	2002734307		AY305738		10
	2002734303	NCTC 3709	AY305742	Horse, India, 1932	11

[a] The source of the isolate is given when available. US, United States; UK, United Kingdom.
[b] Only DNA available.

eight single nucleotide positions, and no gaps were present. These positions were distributed throughout the gene. In seven of these positions (positions 157, 249, 651, 851, 968, 1232, and 1274), more than one nucleotide was detected, resulting in a mixed base (Table 3). Five 16S types (type 4, type 5, type 7, type 9, and type 11) contained a Y (C or T), and two (type 3 and type 6) contained an R (A or G). These results indicated that the strains contained multiple rRNA operons with slightly different 16S rRNA gene sequences.

There were nine different 16S types among the 56 *B. pseudomallei* strains. 16S type 1, type 2, type 6, and type 3 were identified in 21 (38%), 16 (29%), 12 (21%), and 2 strains (4%), respectively. The remaining five 16S types were each represented by a single strain, each differing from 16S type 1 by a single base (Table 3).

Twenty-two of 23 *B. mallei* 16S rRNA gene sequences were identical to each other and were designated 16S type 10. 16S type 10 differs from 16S type 1 by one base at position 75. In addition to the difference seen at position 75, one *B. mallei* sequence had a Y at position 249 instead of a C and was designated 16S type 11. Compared to 16S type 1, the other sequence types differed by 0.5 to 1.5 bases (Table 3).

Only one full-length 16S rRNA gene sequence of *B. pseudomallei* was available in the GenBank and European Molecular Biology Laboratory (EMBL) databases. The full-length sequence from *B. pseudomallei* strain 1026b (GenBank accession no. U91839) contained a T at position 1292 (5), whereas a C was in that position for all other sequences in both our *B. pseudomallei* set and previously submitted *B. pseudomallei* GenBank sequences.

All other *B. pseudomallei* 16S rRNA gene sequences in GenBank and EMBL that were examined were not full length; however, key regions could be compared. The partial sequences from strains H1 (GenBank accession no. AF093047), H2 (AF093053), L2 (AF093054), and V684 (AF093056) were consistent with 16S type 1. The partial sequences from strains K96243 (AF093055), V685 (AF093055), V688 (AF093057), V824 (AF093058), and V830 (AF093060) were consistent with 16S type 2.

We amplified and sequenced the 16S rRNA gene from *B. pseudomallei* strain K96243, and the sequence was a match to 16S type 2. The entire genome of this strain has been sequenced by The Sanger Institute and is presented in a draft version on the website (http://www.sanger.ac.uk/Projects/B_pseudomallei/). With our 16S type 1 sequence as a query in a Blast search, the results indicated that there were three copies of the 16S rRNA gene on chromosome 1 and one on chromosome 2. One of the copies on chromosome 1 corresponded to 16S type 1 whereas the remaining three copies matched 16S type 2.

Two full-length *B. mallei* 16S rRNA sequences were available in GenBank. The 16S rRNA gene sequence of *B. mallei* strain NCTC 10260 (GenBank accession no. AF110187) differed from 16S type 10 only by having a T instead of a C at position 525. The 16S rRNA gene sequence of *B. mallei* strain ATCC 23344 (AF110188) also differed from 16S type 10 at one position by having a C instead of a T at position 782. Both *B. mallei* strains NCTC 10260 and ATCC 23344 were in this study and both 16S rRNA gene sequences were perfect matches to 16S type 10.

ATCC 23344 is also the strain used for the *B. mallei* genome project. The full-length 16S rRNA gene sequence of strain ATCC 23344 (AF110188) in GenBank was queried against the rough draft online at the TIGR website with Blast (http://www.tigr.org/tdb/mdb/mdbinprogress.html) (11 March 2003). Only one sequence was returned and it differed by 1 bp from the query sequence. The sequence from the TIGR website had a T at position 782, which indicated that it was a perfect match to the sequence of 16S type 10.

A total of 44 strains representing species closely related to *B. pseudomallei* and *B. mallei* were analyzed. Among these strains, *B. thailandensis* had a 16S rRNA gene sequence most similar to those of *B. pseudomallei* and *B. mallei*. While the 1 bp difference between *B. pseudomallei* 16S type 1 and *B. mallei* 16S type 10 indicated a similarity of 99.933%, the 16S rRNA gene sequence for strain 2002721643 used in this study, a *B. thailandensis*, differed from 16S type 1 by 12 bp, which indicated a similarity of 99.059%.

One strain, 2002721627, was originally identified as *B. pseudomallei* prior to the description of *B. thailandensis*; however, the 16S rRNA gene sequence was a perfect match to the 16S rRNA gene sequence of *B. thailandensis* 2002721643. Con-

TABLE 2. Designations of 44 strains closely related to *B. pseudomallei* and *B. mallei*

Subset (no. of strains, % sequence similarity)[a]	Species	Identifier	Other identifier	GenBank 16S rRNA gene accession no.	Geographic and/or temporal origin[b]
B. cepacia complex (23, 98.880)	*B. cepacia*	1995030549		AY268146	Human, US, 1995
	B. cepacia	1995019362		AY268149	Human, US, 1995
	B. cepacia	1995035407		AY268148	Human, US, 1995
	B. cepacia	1996031071		AY268147	Human, US, 1996
	B. cepacia	1996031072		AY268145	Human, US, 1996
	B. cepacia	1996031073		AY268144	Human, US, 1996
	B. cepacia	1996033838		AY268154	Human, US, 1996
	B. cepacia	1996034502		AY268141	Human, US, 1996
	B. cepacia	1997019658		AY268140	Human, US, 1997
	B. cepacia	1997024024		AY268143	Human, US, 1997
	B. cepacia	1997030236		AY268150	Human, US, 1997
	B. cepacia	1997033973		AY268157	Human, US, 1997
	B. cepacia	1998000769		AY268155	Human, US, 1998
	B. cepacia	1998019618		AY268142	Human, US, 1998
	B. cepacia	1998019620		AY268158	Human, US, 1998
	B. cepacia	1998040346		AY268162	Human, US, 1998
	B. cepacia	1998040347		AY268160	Human, US, 1998
	Atypical *B. cepacia*	1998003405		AY268156	Human, US, 1998
	Atypical *B. cepacia*	1998003407		AY268153	Human, US, 1998
	B. cepacia	2001008925		AY268152	Human, US, 2001
	B. cepacia	2001038723		AY268161	Human, US, 2001
	B. cepacia	2002721599		AY268151	Human, US, 2001
	B. cepacia genomovar III	2001038720		AY268159	Human, US, 2001
B. gladioli (5, 99.724)	*B. gladioli*	1993027208		AY268167	Human, US, 1993
	B. gladioli	1995009253		AY268165	Human, US, 1995
	B. gladioli	1996012911		AY268166	Human, US, 1996
	B. gladioli	2002721590		AY268163	Human, US, 1999
	B. gladioli	2002721589		AY268164	Human, US, 2001
Pandoraea spp. (7, 99.110)	*P. apista*	2001005416	CCUG 38412	AY268172	
	P. norimbergensis	2001005417	CCUG 39188	AY268174	
	P. pnomenusa	2001008157		AY268168	Human, US, 2001
	P. pnomenusa, type strain	2001005418	CCUG 38742	AY268170	
	P. pulmonicola	2001005419	CCUG 38759	AY268173	
	Pandoraea sp.	2001005415	CCUG 39680	AY268171	
	Pandoraea sp.	2001032141		AY268169	Human, US, 2001
Ralstonia spp. (6, 99.863)	*Ralstonia* sp.	1999043680		AY268181	Human, US, 1999
	R. pickettii	2000032023		AY268176	Human, US, 2000
	R. pickettii	2000030635		AY268180	Human, US, 2000
	R. pickettii	2000030791		AY268179	Human, US, 2000
	R. pickettii	2002721591		AY268178	Human, US, 2000
	R. pickettii	2002721592		AY268177	Human, US, 2000
Others	*Burkholderia thailandensis*	2002721627		AY268182	Environment, Thailand
	Burkholderia thailandensis	2002721643		AY268183	
	Pseudomonas aeruginosa	2002721595	ATCC 27853	AY268175	

[a] The level of similarity is based on BESTFIT analysis of the two most divergent isolates based on DISTANCES results in GCG.
[b] US, United States.

sequently, standard biochemical testing was repeated. This testing included arabinose utilization which is the standard biochemical test to differentiate *B. pseudomallei* and *B. thailandensis* (5). Upon retesting, strain 2002721627 was found to be arabinose positive and was then reclassified as *B. thailandensis*.

For the other strains for which multiple strains were available for a given genus, subsets were selected and core sequences were determined and compared (Table 2). A 16S rRNA gene sequence of a representative strain was then compared to the 16S type 1 sequence with BESTFIT to determine the degree of similarity (Table 4). The comparisons of the sequences of closely related species and genera, *Pandoraea* spp., *Ralstonia* spp., *B. gladioli*, *B. cepacia*, *B. thailandensis*, and *Pseudomonas aeruginosa*, indicated that the level of similarity to the 16S type 1 sequence of *B. pseudomallei* was at most 99.059%. BESTFIT also indicated that with the exception of *B. thailandensis*, all these closely related strains had insertions and deletions (indels) in their sequences that made them easy to distinguish from the sequences of both *B. pseudomallei* and *B. mallei*. A search of the GenBank and EMBL sequence databases revealed that the *B. pseudomallei* and *B. mallei* 16S rRNA sequences that comprise the 16S types in this study differed substantially (at least 14 bp) from the previously submitted sequences of related species such as *B. thailandensis* (GenBank accession no. BSU91838).

TABLE 3. Position of base differences among 16S types of *B. pseudomallei* and *B. mallei*[a]

16S type	Difference at position:								No. of isolates	% of *B. pseudomallei* isolates[b]
	75	157	249	651	851	968	1232	1274		
1	C	G	C	C	C	T	G	C	21	38
2		A							16	29
3							R		2	4
4					Y				1	2
5		A				Y			1	2
6		R							12	21
7				Y					1	2
8							A		1	2
9								Y	1	2
10	T								22	NA
11	T		Y						1	NA

[a] A mixed base resulting from overlapping A and G peaks in the sequence trace file is noted as R for purine, whereas a mixed base due to overlapping C and T peaks is designated Y for pyrimidine.
[b] NA, sequence types 10 and 11 were found only in *B. mallei*.

DISCUSSION

The 16S rRNA sequences among the *B. pseudomallei* and *B. mallei* strains in this study varied by only 0.5 to 1.5 bp compared to 16S type 1. However, the two species could be consistently distinguished from each other because of a 1-bp difference at position 75. Furthermore, sequences of closely related species were also sufficiently divergent from those of both *B. pseudomallei* and *B. mallei* to allow easy discrimination.

The current diagnostic standard for the identification of *B. pseudomallei* and *B. mallei* is based on cell and colony morphology as well as on biochemical tests that may require up to 7 days to obtain results (35). Unfortunately, for patients with septicemia, death may occur in 2 days or less, so rapid methods of identification are needed (3, 32). Also, some laboratories may be using commercial test systems that have been shown to misidentify *B. pseudomallei* isolates as other bacteria, such as *Chromobacterium violaceum* (15). In a recent case of laboratory-acquired glanders, the infecting organism was initially identified as *Pseudomonas fluorescens* or *Pseudomonas putida*. 16S rRNA gene sequencing was used to confirm that the organism was indeed *B. mallei* (33).

Although PCR-based assays have been described for the identification of *Burkholderia* spp., none is currently in use as a standard diagnostic method (3, 13, 30). With the automation of sequencing and improvements in reagent efficiencies, both the time required for results and the cost have decreased substantially so that PCR-based assays and DNA sequencing are now increasingly used in clinical and public health laboratories for bacterial identification (23). Presently, a 16S rRNA gene sequence from bacterial DNA can be obtained in 9 h, which is significantly shorter than the minimum 2 days currently required for biochemical assays used to identify *B. pseudomallei* and *B. mallei* (3, 35).

TABLE 4. Similarities of 16S type 1 sequence to 16S rRNA gene sequences of six strains representing species closely related to *B. pseudomallei* and *B. mallei*

Isolate	% similarity of 16S rRNA gene sequence to 16S type 1[a]	Length of 16S rRNA core sequence used for comparison (bp)
B. thailandensis 2002721643	99.059	1,488
B. gladioli 1993027208	97.808	1,450
B. cepacia 1997019658	97.898	1,429
Pandoraea pnomenusa 2001008157	95.124	1,461
Ralstonia pickettii 2000030791	92.055	1,465
Pseudomonas aeruginosa 2002721595	83.528	1,466

[a] Sequence similarity is based on BESTFIT analysis in GCG.

We identified distinct 16S rRNA sequence groups among the study isolates in the *B. pseudomallei* panel. Nine 16S types that differed from 16S type 1 by 0.5 to 1.5 bp were found. A comparison of the derived 16S types with epidemiological data does not appear to correlate to any trend in geography, time, or origin of isolate. The 16S rRNA gene sequences of 22 out of 23 *B. mallei* isolates tested were identical in spite of the strains' diversity in terms of geography, date, and origin of isolation.

These results are consistent with recent results from a study done with multilocus sequence typing (MLST) by Godoy et al., who found that although they could derive 71 sequence types out of 128 *B. pseudomallei* isolates, there was no difference among their five *B. mallei* isolates. They also found that the *B. mallei* sequence type was grouped within the *B. pseudomallei* sequence types, supporting the idea that *B. mallei* is a clone of *B. pseudomallei* (12).

Each method offers different advantages. While multilocus sequence typing uses purified DNA and requires the sequencing of seven genetic loci per sample, our approach is accomplished with a simple DNA extraction and requires sequence data for only one gene (the 16S rRNA), making multilocus sequence typing more labor intensive and time-consuming (12). The speed of this method makes it highly preferable to MLST, especially for large-scale screens. In terms of the amount of data generated by the two methods, multilocus sequence typing clearly yields more detailed information on bacterial isolates because it involves the sequencing of seven genes. Consequently, the results are comparable to subtyping data obtained from pulsed-field gel electrophoresis, which may make MLST more amenable for epidemiological studies. Further studies are needed to fully assess the usefulness of 16S rRNA sequencing as a tool in epidemiologic investigations.

Although the full-length 16S rRNA gene sequences of *B.*

pseudomallei and *B. mallei* are available in GenBank and EMBL, it may not be possible to assign them 16S types because of variations in sequencing techniques and base calling. For example, if a cloned copy of the 16S rRNA gene is used for sequencing, only one of the alleles in the genome will be represented, and thus the ability to detect a mixed base will be lost (6, 26). The 16S types presented in this study are based on sequences amplified from whole-genome DNA preparations. In the case of the two *B. mallei* 16S rRNA gene sequences that are in GenBank (accession nos. AF110187 and AF110188), the annotations indicate that they are unpublished, so the sequencing protocol used is unknown.

It is worth noting that the full-length 16S rRNA gene sequence of *B. mallei* strain ATCC 23344 in GenBank (GenBank accession no. AF110188) does not match the one full-length 16S rRNA gene sequence found on the draft version of the genome for the same *B. mallei* strain on the TIGR website. The 16S rRNA gene sequence obtained in this study for strain ATCC 23344 does match 16S type 10, as does the one full-length sequence on the TIGR website for that strain. Since the *B. mallei* genome project is still in progress, it remains to be seen how many 16S rRNA alleles will be found.

It is also interesting that out of the four copies of the 16S rRNA gene found in *B. pseudomallei* strain K96243 on the draft version of the genome on the Sanger website, three are matched to 16S type 2 and one is matched to 16S type 1. Since the difference between 16S type 1 and 16S type 2 is a G versus an A at position 157, a mixed-base call of R was expected, which would have resulted in a 16S type 6. However, a study of the trace file indicated that K96243 is a 16S type 2. A peak for G is barely visible at position 157 and is at a similar level to background peaks. This indicates that the 3:1 ratio of A to G is sufficient to identify a 16S type 2. This observation also suggests that strains designated 16S type 6 may have a 1:1 ratio of A to G at position 157, since the overlapping A and G peaks are clearly above the background.

This work indicates that the 16S rRNA gene sequence of *B. pseudomallei* and *B. mallei* can be used to identify and distinguish the two species more quickly than can be done by currently used biochemical tests and by observation of colony and cell morphology. Further studies are needed to assess the potential of using the subtle variations in the 16S rRNA gene sequence as a subtyping method for these pathogens.

Currently used methods of subtyping, such as pulsed-field gel electrophoresis and ribotyping, have been useful in identifying strains of *B. pseudomallei* in outbreaks (16, 18) and have also been used to differentiate pathogenic *B. pseudomallei* strains from less virulent strains (24). Unfortunately, these methods tend to be time-consuming and labor intensive.

Recently, subtle differences in bacterial 16S rRNA sequences have also been used for subtyping. In the case of *Vibrio vulnificus*, two 16S rRNA types were determined. One was found in 17 of 18 clinical fatalities, whereas the other type was predominantly found in environmental isolates (22). We previously demonstrated that *Bacillus anthracis* can be differentiated from *Bacillus cereus* on the basis of a mixed base in the 16S rRNA gene (28). Another study from our laboratory showed substantial variability in the 16S rRNA sequence of *Neisseria meningitidis* (29), and so far 153 16S types have been identified based on 59 positions of difference (unpublished data).

We sequenced the 16S rRNA genes of 56 *B. pseudomallei* and 23 *B. mallei* strains representing temporal, geographic, and source diversity as well as the 16S rRNA genes of closely related isolates. This study indicates that the 16S rRNA gene sequences of *B. pseudomallei* and *B. mallei* can be used to discriminate reliably between the two species based on a 1-bp difference at position 75. Furthermore, the 16S rRNA gene sequences of closely related species were also sufficiently divergent to allow easy discrimination. While *B. mallei* appears to have less genetic variation, further studies will show whether 16S rRNA gene sequencing may assist in defining the molecular epidemiology of melioidosis.

ACKNOWLEDGMENTS

We thank Rich Meyer, Mark Schell, David DeShazer, and Steve Harvey for providing some of the strains and DNA used in this study.

Preliminary sequence data were obtained from the Institute for Genomic Research website at http://www.tigr.org and the Sanger Institute website at http://www.sanger.ac.uk/Projects/B_pseudomallei/. Sequencing of *B. mallei* at TIGR was accomplished with support from USAMRIID, NIH, and NIAID. Sequencing of *B. pseudomallei* at the Sanger Institute was accomplished with support from Beowulf Genomics.

REFERENCES

1. **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.
2. **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.
3. **Bauernfeind, A., C. Roller, D. Meyer, R. Jungwirth, and I. Schneider.** 1998. Molecular procedure for rapid detection of *Burkholderia mallei* and *Burkholderia pseudomallei.* J. Clin. Microbiol. **36:**2737–2741.
4. **Bootsma, H. J., H. G. van der Heide, S. van de Pas, L. M. Schouls, and F. R. Mooi.** 2000. Analysis of *Moraxella catarrhalis* by DNA typing: evidence for a distinct subpopulation associated with virulence traits. J. Infect. Dis. **181:** 1376–1387.
5. **Brett, P. J., D. DeShazer, and D. E. Woods.** 1998. *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. Int. J. Syst. Bacteriol. **48**:317–320.
6. **Cilia, V., B. Lafay, and R. Christen.** 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. Mol. Biol. Evol. **13:**451–461.
7. **Dance, D. A.** 2002. Melioidosis. Curr. Opin. Infect. Dis. **15:**127–132.
8. **Dance, D. A.** 2000. Melioidosis as an emerging global problem. Acta Trop. **74:**115–119.
9. **Derbyshire, J. B.** 2002. The eradication of glanders in Canada. Can. Vet. J. **43:**722–726.
10. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
11. **Eden, P. A., T. M. Schmidt, R. P. Blakemore, and N. R. Pace.** 1991. Phylogenetic analysis of *Aquaspirillum magnetotacticum* with polymerase chain reaction-amplified 16S rRNA-specific DNA. Int. J. Syst. Bacteriol. **41:**324–325.
12. **Godoy, D., G. Randle, A. J. Simpson, T. I. Pitt, R. Kinoshita, and B. G. Spratt.** 2003. Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *B. mallei.* J. Clin. Microbiol. **41:**2068–2079.
13. **Hagen, R. M., Y. P. Gauthier, L. D. Sprague, D. R. Vidal, G. Zysk, E. J. Finke, and H. Neubauer.** 2002. Strategies for PCR based detection of *Burkholderia pseudomallei* DNA in paraffin wax embedded tissues. Mol. Pathol. **55:**398–400.
14. **Hicks, C. L., R. Kinoshita, and P. W. Ladds.** 2000. Pathology of melioidosis in captive marine mammals. Aust. Vet. J. **78:**193–195.
15. **Inglis, T. J., D. Chiang, G. S. Lee, and L. Chor-Kiang.** 1998. Potential misidentification of *Burkholderia pseudomallei* by API 20NE. Pathology **30:** 62–64.
16. **Inglis, T. J., S. C. Garrow, C. Adams, M. Henderson, M. Mayo, and B. J. Currie.** 1999. Acute melioidosis outbreak in Western Australia. Epidemiol. Infect. **123:**437–443.
17. **Inglis, T. J., S. C. Garrow, M. Henderson, A. Clair, J. Sampson, L. O'Reilly, and B. Cameron.** 2000. *Burkholderia pseudomallei* traced to water treatment plant in Australia. Emerg. Infect. Dis. **6**:56–59.

18. **Lew, A. E., and P. M. Desmarchelier.** 1993. Molecular typing of *Pseudomonas pseudomallei*: restriction fragment length polymorphisms of rRNA genes. J. Clin. Microbiol. **31:**533–539.
19. **Lowe, P., C. Engler, and R. Norton.** 2002. Comparison of automated and nonautomated systems for identification of *Burkholderia pseudomallei.* J. Clin. Microbiol. **40:**4625–4627.
20. **Mayer, L. W., M. W. Reeves, N. Al-Hamdan, C. T. Sacchi, M. K. Taha, G. W. Ajello, S. E. Schmink, C. A. Noble, M. L. Tondella, A. M. Whitney, Y. Al-Mazrou, M. Al-Jefri, A. Mishkhis, S. Sabban, D. A. Caugant, J. Lingappa, N. E. Rosenstein, and T. Popovic.** 2002. Outbreak of W135 meningococcal disease in 2000: not emergence of a new W135 strain but clonal expansion within the electrophoretic type-37 complex. J. Infect. Dis. **185:**1596–1605.
21. **Mollaret, H. H.** 1988. L'affaire du jardin des plantes ou comment la melioidoise fit son apparition en France. Med. Mal. Infect. **18:**643–654.
22. **Nilsson, W. B., R. N. Paranjype, A. DePaola, and M. S. Strom.** 2003. Sequence polymorphism of the 16S rRNA gene of *Vibrio vulnificus* is a possible indicator of strain virulence. J. Clin. Microbiol. **41**:442–446.
23. **Patel, J. B.** 2001. 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. Mol. Diagn. **6:**313–321.
24. **Pitt, T. L., S. Trakulsomboon, and D. A. Dance.** 2000. Molecular phylogeny of *Burkholderia pseudomallei.* Acta Trop. **74:**181–185.
25. **Rantakokko-Jalava, K., S. Nikkari, J. Jalava, E. Eerola, M. Skurnik, O. Meurman, O. Ruuskanen, A. Alanen, E. Kotilainen, P. Toivanen, and P. Kotilainen.** 2000. Direct amplification of rRNA genes in diagnosis of bacterial infections. J. Clin. Microbiol. **38:**32–39.
26. **Reischl, U., K. Feldmann, L. Naumann, B. J. Gaugler, B. Ninet, B. Hirschel, and S. Emler.** 1998. 16S rRNA sequence diversity in *Mycobacterium celatum* strains caused by presence of two different copies of 16S rRNA gene. J. Clin. Microbiol. **36:**1761–1764.
27. **Rotz, L. D., A. S. Khan, S. R. Lillibridge, S. M. Ostroff, and J. M. Hughes.** 2002. Public health assessment of potential biological terrorism agents. Emerg. Infect. Dis. **8:**225–230.
28. **Sacchi, C. T., A. M. Whitney, L. W. Mayer, R. Morey, A. Steigerwalt, A. Boras, R. S. Weyant, and T. Popovic.** 2002. Sequencing of 16S rRNA gene: a rapid tool for identification of *Bacillus anthracis.* Emerg. Infect. Dis. **8:**1117–1123.
29. **Sacchi, C. T., A. M. Whitney, M. W. Reeves, L. W. Mayer, and T. Popovic.** 2002. 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.
30. **Sirisinha, S., N. Anuntagool, T. Dharakul, P. Ekpo, S. Wongratanacheewin, P. Naigowit, B. Petchclai, V. Thamlikitkul, and Y. Suputtamongkol.** 2000. Recent developments in laboratory diagnosis of melioidosis. Acta Trop. **74:**235–245.
31. **Songsivilai, S., and T. Dharakul.** 2000. Multiple replicons constitute the 6.5-megabase genome of *Burkholderia pseudomallei.* Acta Trop. **74:**169–179.
32. **Sprague, L. D., G. Zysk, R. M. Hagen, H. Meyer, J. Ellis, N. Anuntagool, Y. Gauthier, and H. Neubauer.** 2002. A possible pitfall in the identification of *Burkholderia mallei* with molecular identification systems based on the sequence of the flagellin fliC gene. FEMS Immunol. Med. Microbiol. **34:**231–236.
33. **Srinivasan, A., C. N. Kraus, D. DeShazer, P. M. Becker, J. D. Dick, L. Spacek, J. G. Bartlett, W. R. Byrne, and D. L. Thomas.** 2001. Glanders in a military research microbiologist. N. Engl. J. Med. **345:**256–268.
34. **Stackebrandt, E., and O. Charfreitag.** 1990. Partial 16S rRNA primary structure of five *Actinomyces* species: phylogenetic implications and development of an *Actinomyces israelii*-specific oligonucleotide probe. J. Gen. Microbiol. **136:**37–43.
35. **Weyant, R. S., C. W. Moss, R. E. Weaver, D. G. Hollis, J. G. Jordan, E. C. Cook, and M. I. Daneshvar.** 1995. *Pseudomonas pseudomallei*, p. 486–487. *In* W. R. Hensyl (ed.), Identification of unusual pathogenic gram-negative aerobic and facultatively anaerobic bacteria, 2nd ed. Williams & Wilkins, Baltimore, Md.
36. **Wheelis, M.** 1998. First shots fired in biological warfare. Nature **395:**213.
37. **Zysk, G., W. D. Splettstosser, and H. Neubauer.** 2000. A review on melioidosis with special respect on molecular and immunological diagnostic techniques. Clin. Lab. **46:**119–130.