Phylogenetic analysis reveals the taxonomically diverse distribution of the *Pseudomonas putida* group

(Received June 6, 2016; Accepted June 28, 2016; J-STAGE Advance publication date: December 17, 2016)

Kenta Yonezuka,1,† Jun Shimodaira,2,† Michiro Tabata,1 Shoko Ohji,2 Akira Hosoyama,2 Daisuke Kasai,1 Atsushi Yamazoe,2 Nobuyuki Fujita,2 Takayuki Ezaki,3 and Masao Fukuda1,*

1 Department of Bioengineering, Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
2 Biological Resource Center, National Institute of Technology and Evaluation, 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
3 Department of Microbiology, Gifu University Graduate School of Medicine, 1-1 Yanagido, Gifu 501-1194, Japan

*Corresponding author: Masao Fukuda, Department of Bioengineering, Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan.
Tel: +81-258-47-9405 Fax: +81-258-47-9450 E-mail: masao@vos.nagaokaut.ac.jp
† Equal contributors.

None of the authors of this manuscript has any financial or personal relationship with other people or organizations that could inappropriately influence their work.

**Pseudomonas putida** is well-known for degradation activities for a variety of compounds and its infections have been reported. Thus, *P. putida* includes both clinical and nonclinical isolates. To date, no reports have examined the phylogenetic relationship between clinical and nonclinical isolates of the *P. putida* group. In this study, fifty-nine strains of *P. putida* group containing twenty-six clinical, and thirty-three nonclinical, isolates, were subjected to phylogenetic and taxonomic analyses based on 16S rRNA gene sequences and nine housekeeping gene sequences, including *argS, dnaN, dnaQ, era, gltA, gyrB, ppnK, rpoB, and rpoD*, to obtain insights into the diversity of species in this group. More than 97.6% similarity was observed among the 16S rRNA gene sequences of all the strains examined, indicating that the resolution of 16S RNA gene sequences is inadequate. Phylogenetic analysis based on the individual housekeeping genes listed above improved the resolution of the phylogenetic trees, which are different from each other. Multilocus sequence analysis (MLSA) based on the concatenated sequences of the nine genes significantly improved the resolution of the phylogenetic tree, which are different from each other. Multilocus sequence analysis (MLSA) based on the concatenated sequences of the nine genes significantly improved the resolution of the phylogenetic tree, and yielded approximately the same results as average nucleotide identity (ANI) analysis, suggesting its high reliability. ANI analysis classified the fifty-nine strains into twenty-six species containing seventeen singletons and nine strain clusters based on the 95% threshold. It also indicated the mixed distribution of clinical and nonclinical isolates in the six clusters, suggesting that the genomic difference between clinical and nonclinical isolates of the *P. putida* group is subtle. The *P. putida* type strain NBRC 14164T is a singleton that is independently located from the *P. putida* strains distributed among the six clusters, suggesting that the classification of these strains and the differentiation of species in the *P. putida* group should be re-examined. This study greatly expands insights into the phylogenetic diversity of the *P. putida* group.

Key Words: average nucleotide identity (ANI); clinical isolates; multilocus sequence analysis (MLSA); nonclinical isolates; *Pseudomonas putida* group

Introduction

The genus *Pseudomonas* has been isolated from a variety of habitats and is one of the most complicated genera containing a vast number of species (Silby et al., 2011). The genus *Pseudomonas* had been formerly classified into five rRNA homology groups based on DNA-DNA and rRNA-DNA hybridization experiments (De Vos and De Ley, 1983; Palleroni et al., 1973). De Vos et al. (1985, 1989) defined the rRNA homology group I as the authentic pseudomonads denoting the genus *Pseudomonas*, which were later defined as *Pseudomonas* (sensu stricto) by Moore et al. (1996) and Anzai et al. (2000). They used 16S rRNA gene sequences to determine the phylogenetic
affiliation of strains and indicated nine and seven lineages for the genus *Pseudomonas* (sensu stricto), respectively. Based on the current classification, group I contained *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, and *Pseudomonas putida* of Gammaproteobacteria, group II contained *Burkholderia mallei*, *Burkholderia cepacia*, and *Ralstonia picketti* of Betaproteobacteria, group III contained *Comamonas acidovorans* and *Comamonas testosteroni* of Betaproteobacteria, group IV contained *Brevundimonas diminuta* and *Brevundimonas vesicularis* of Alphaproteobacteria, and group V contained *Stenotrophomonas* (formerly *Xanthomonas*) *maltophilia* of Gammaproteobacteria. Because of the slow rate of 16S rRNA gene evolution, the resolution of phylogenies based on 16S rRNA gene sequences does not allow for intragenic differentiation. Yamamoto et al. (2000) employed the combined sequences of gyrB and rpoD genes to determine phylogenetic relationships of the genus *Pseudomonas* (sensu stricto), identifying five complexes: *P. aeruginosa*, *P. stutzeri*, *P. putida*, *P. syringae*, and *P. fluorescens*. The combined sequences of four housekeeping genes, *atpD*, *carA*, *recA*, and 16S rRNA, were used to verify the phylogenetic relationships of the genus *Pseudomonas* by Hilario et al. (2004). Mulet et al. (2010) employed the concatenated sequences of four housekeeping genes, 16S rRNA, gyrB, rpoD, and rpoB, to perform multilocus sequence analysis (MLSA) with as many as 107 type strains, revealing ten groups in the genus *Pseudomonas* (sensu stricto): *P. fluorescens*, *P. syringae*, *P. lutea*, *P. putida*, *P. anguilliseptica*, *P. straminea*, *P. aeruginosa*, *P. oleovorans*, *P. stutzeri*, and *P. oryzihabitans*. Gomila et al. (2015) used the concatenated sequences of the same four genes in MLSA of 253 strains, including 133 type strains, and identified the *P. putidagenosa* group in addition to the above ten groups indicated by Mulet et al. (2010).

*P. putida* is ubiquitous in the environment and is highly adaptable to various environments (Silby et al., 2011). *P. putida* is well-known for its ability to degrade a variety of compounds, such as benzoate (Nakazawa and Yokota, 1973), toluene (Ramos et al., 1995; Zylstra et al., 1988), naphthalene (Dunn and Gunsalus, 1973), phenol (Herrmann et al., 1988), and chloroethenes (Wackett and Gibson, 1988; Yonezuka et al., 2016). Conversely, infections by *P. putida* have been described in compromised patients, such as cancer patients (Anaissie et al., 1987; Martino et al., 1996) and newborns (Ladhani and Bhutta, 1998). *P. putida* infections have also been reported under a variety of conditions, such as wound infection and medical device contamination (Carpenter et al., 2008). Thus, *P. putida* strains include clinical isolates of opportunistic or moderate pathogens and nonclinical isolates of inhabitants in the natural environment, eliciting the question of what is the phylogenetic relationship between the clinical and nonclinical isolates. Mulet et al. (2013) performed MLSA using the concatenated sequences of three genes, 16S rRNA, gyrB, and rpoD, to assess the distribution of *P. putida* biovars A and B; a preferential distribution of biovar A within the *P. putida* group, and a broad distribution of biovar B among the *P. fluorescens* and *P. putida* groups, were observed. Additionally, the *P. putida* group included a variety of species, such as *P. cremoricolorata*, *P. entomophila*, *P. fulva*, *P. japonica*, *P. monteilii*, *P. mossellii*, *P. parafulva*, *P. plecoglossicida*, *P. putida*, *P. taiwanensis*, and *P. vranovensis*. Gomila et al. (2015) employed four concatenated gene sequences, 16S rRNA, gyrB, rpoD, and rpoB, revealing a similar *P. putida* group diversity. They confirmed their MLSA results using average nucleotide identity (ANI) analysis, which precisely separates strains among species based on the identities among the genome sequences (Goris et al., 2007; Richter and Rosselló-Móra, 2009). An increasing number of bacterial genome sequences is available with the benefit of next-generation DNA sequencing. ANI analysis technology can be crucial for separating strains among species instead of DNA-DNA hybridization, a technique used traditionally for this purpose. Performing both MLSA and ANI analysis with a variety of *P. putida* strains may offer greater insights into the diversity of species in this group. To date, no reports have examined the phylogenetic relationship between clinical and nonclinical isolates of the *P. putida* group, although *P. monteilii* (Elomari et al., 1997) and *P. mossellii* (Dabboussi et al., 2002) are established clinical isolates.

In this study, MLSA using nine housekeeping gene sequences and ANI analysis were performed on genome sequences of fifty-nine strains of *P. putida* group, which contained twenty-two novel genome sequences. We examined the accurate species diversity of the *P. putida* group and the phylogenetic relationship between the clinical and nonclinical isolates.

### Materials and Methods

**Pseudomonas strains.** Sixty genome sequences of *Pseudomonas* strains were subjected to analysis (Table S1). *P. oryzihabitans* NBRC 102199 was used as an outgroup strain. The fifty-nine genome sequences of the *P. putida* group contain those from nine type and twenty-four nontype strains of environmental isolates, and those from two type and twenty-four nontype strains of clinical isolates. Twenty-two nontype strains of clinical isolates were obtained from the culture collection of the Biological Resource Center of the National Institute of Technology and Evaluation (NBRC) and subjected to genome sequencing in this study. All these twenty-two strains were grown at 35°C overnight in a heart infusion broth (Becton Dickinson, Sparks, MD, USA). All of them were collected by Gifu University School of Medicine Pathogenic Bacterial Genetic Resource Stock Center (GMGC) and deposited in the NBRC. The remaining thirty-eight genome sequences, including that of the outgroup strain, were retrieved from the National Center for Biotechnology Information (NCBI) database.

**Genome sequencing.** Genomic DNA was extracted using the EZ1 Tissue Kit (Qiagen, Valencia, CA, USA) from the cells of each strain according to the manufacturer’s instructions. Libraries were prepared with the TruSeq DNA LT sample prep kit v2 (Illumina, Inc., San Diego, CA, USA). The genome sequences were determined by whole-genome shotgun sequencing with short paired-end
sequencing using the MiSeq platform with MiSeq reagent kit v2 (Illumina), and de novo assembly was performed using Newbler version 2.8 (Roche 454 Life Sciences, Branford, CT, USA). The assemblies were error-corrected manually using Consed software (Gordon and Green, 2013). Open reading frames (ORF) were predicted by Rapid Annotation using Subsystem Technology (RAST) server version 2.0 with the ClassicRAST annotation scheme (http://rast.nmpdr.org/rast.cgi).

**Phylogenetic analyses.** The sequences of 16S rRNA and nine housekeeping genes (argS, dnaN, dnaQ, era, gltA, gyrB, ppnK, rpoB, and rpoD) used in this study were extracted from predicted ORFs by BLAST searches using the corresponding sequence of P. putida KT2440 as a reference. The extracted gene sequences were manually examined for putative artifacts by multiple alignments using ClustalW version 2.1 at the DNA Data Bank of Japan database. The poorly-aligned sequences were trimmed using Gblocks version 0.91b (Castresana, 2000). The trimmed sequences were subjected to phylogenetic analysis using the software package MEGA version 6.06 with the neighbor-joining (NJ) or maximum likelihood (ML) algorithms (Saitou and Nei, 1987; Tamura et al., 2011, 2013). NJ analysis was performed using the pairwise distance (p-distance) model. ML analysis was performed employing the general time reversible model including estimation of invariant sites and assuming a discrete gamma distribution with five rate categories (GTR+G+I model). Phylogenetic trees were constructed based on individual and concatenated genes. To prepare a concatenated gene sequence, all nine gene sequences of each strain were combined in alphabetical order, argS-dnaN-dnaQ-era-gltA-gyrB-ppnK-rpoB-rpoD, totaling 13,831 bases (Table 1). The concatenated sequences were subjected to alignment and phylogenetic analysis using the software package MEGA version 6.06. Bootstrap values for individual nodes were calculated for 1,000 replicates (Felsenstein, 1985). P. oryzihabitans NBRC 102199T was used as an outgroup strain for the phylogenetic analyses.

**ANI analyses.** ANI analyses were performed using JSpecies version 1.2.1 and BLAST 2.2.22 as described by Goris et al. (2007). The distance matrix was generated based on the distance values (DV), which were calculated using the following formula: \( DV = 1 – \frac{ANI \text{ value}}{100} \), and was subjected to calculation using the NEIGHBOR program in the PHYLIP package version 3.695 (Felsenstein, 1989). A phylogenetic tree was generated from the resulting Newick format file using Njplot (Perrière and Gouy, 1996).

**Results**

**General genome characteristics**

The draft genome sequences of twenty-two clinical isolates obtained from GMGC were newly determined in this study. They were assembled into contigs ranging from 99 to 981 with an average of 252 and the estimated genome sizes ranging from 5.41 Mb (Pseudomonas sp. NBRC 111119) to 6.78 Mb (Pseudomonas sp. GTC 16482). The
Yonezuka et al.

Fig. 1. Phylogenetic tree based on the 16S rRNA gene sequences of strains of the *P. putida* group.
The NJ tree was constructed using the MEGA version 6.06. *P. oryzihabtans* NBRC 102199T was used as an outgroup. The bootstrap values were calculated with 1,000 replicates, and only those with 50% or more are shown. The type strains are indicated in boldface type and a superscript T. Clinical and nonclinical isolates including type strains are indicated by C and N in parentheses in front of the strain names, respectively. The scale bar represents 0.005 nucleotide substitutions per nucleotide position. The strains DOT-T1E, B001, and LS46 are excluded from this tree.

The phylogenetic analysis (Table S1). The genome sizes of the fifty-nine *P. putida* group strains range from 4.60 Mb (*P. putida* S610) to 6.87 Mb (*P. putida* H8234) with an average of 5.97 ± 0.47 Mb. The sizes of the clinical isolates range from 5.41 Mb (*P. putida* SB26) to 6.87 Mb (*P. putida* H8234) with an average of 6.09 ± 0.39 Mb. Those of the nonclinical isolates range from 4.60 Mb (*P. putida* S610) to 6.66 Mb (*P. japonica* NBRC 103040) with an average of 5.87 ± 0.50 Mb. The average genome size of the clinical isolates is a little superior to that of the nonclinical isolates. The G+C contents of the *P. putida* group strains range from 57.6 mol% (*P. putida* S610) to 64.4 mol% (*P. putida* H8234) with an average of 62.1 ± 0.8 mol%. The average G+C contents of the clinical and nonclinical isolates are 62.2 ± 0.8 and 62.1 ± 0.8 mol%, respectively and are mostly the same.

**Phylogenetic analysis based on individual housekeeping genes**

The phylogenetic tree based on 16S rRNA gene sequences of the fifty-six *P. putida* group strains, and the outgroup *P. oryzihabtans* NBRC 102199T, was constructed using the NJ method (Fig. 1). Because the 16S rRNA gene sequence of the strain DOT-T1E contained suspicious sequences, and those of the B001 and LS46 strains were too short to use for phylogenetic analysis, these three strains were excluded from the analysis. Forty-two (77.8%) nodes had a bootstrap value less than 50% and some terminal nodes contained multiple strains (Table 1). Focusing on the type strains of the *P. putida* group, their phylogenetic relationships were extremely close to each other, and the identities of the 16S rRNA gene sequences among them
Diverse distribution of *P. putida* group

ranged from 98.2% to 100%. The similarity of the sequence was more than 97.6% among all of the strains, including the clinical isolates and the type strains, indicating that all of the clinical isolates belong to the *P. putida* group. These results indicate that the resolution of the 16S rRNA gene sequence is too low to resolve the phylogenetic relationships among the strains of the *P. putida* group, consistent with previous reports (Anzai et al., 2000; Moore et al., 1996).

Hence, other housekeeping gene sequences than the 16S rRNA gene were employed in the phylogenetic analysis. Among the housekeeping genes used in previous studies (Adékambi et al., 2011; Mo et al., 2013; Mulet et al., 2010; Scortichini et al., 2013), we selected nine genes, *argS, dnaN, dnaQ, era, gltA, gyrB, ppnK, rpoB*, and *rpoD*, based on the following conditions: (i) all of the strains harbor the gene, (ii) the gene is present as a single copy in the genome, and (iii) the original gene sequence is longer than 700 bases. The G+C contents of the housekeeping gene sequences ranged from 56.3% to 66.6% (Table 1). The average, maximum, and minimum values of the p-distance obtained by phylogenetic analyses are indicated in Table 1. The average p-distance values of the housekeeping genes were 7.5- to 18.8-fold greater than that of the 16S rRNA gene, indicating that the phylogenetic analyses using these housekeeping genes are suitable for classifying the strains of the *P. putida* group at a high resolution.

The bootstrap values of the phylogenetic trees based on the housekeeping genes were dramatically elevated compared with the values in the tree based on the 16S rRNA gene. In the 16S rRNA tree, the branching points with bootstrap values of less than 50% accounted for 77.8% of the total branching points, whereas those in the housekeeping gene trees comprised 10.5% to 35.1% (Table 1). The branching points with bootstrap values of over 90% in the 16S rRNA tree accounted for 9.3% of the total branching points, but those in the housekeeping gene trees comprised 38.6% to 54.4%. These results indicate that phylogenetic analysis using housekeeping genes offers a greater resolution. However, the phylogenetic relationships...
of strains differed between the nine housekeeping genes (data not shown), revealing the uniqueness of each phylogenetic tree.

**Phylogenetic analysis based on concatenated housekeeping genes**

Because the phylogenetic analyses using individual housekeeping gene sequences yielded diverse results, the nine housekeeping gene sequences were concatenated in a 13.8-kb sequence in alphabetical order, *argS-dnaQ-era-gltA-gyrB-ppnK-rpoB-rpoD*, and used for further phylogenetic analysis. The 16S rRNA gene sequence was not included, as the phylogenetic analysis of the 16S rRNA gene did not include the 16S rRNA sequences of the DOT-T1E, B001, and LS46 strains.

The phylogenetic tree based on the MLSA is presented in Fig. 2. Phylogenetic relationships in the MLSA tree differ from those in the phylogenetic trees generated using individual housekeeping genes. As shown in Table 1, branching points with bootstrap values over 90% accounted for 80.7% of the MLSA tree, and those with bootstrap values of less than 50% accounted for 1.8%, indicating a significant improvement in the resolution.

In this study, the NJ method was used to generate the MLSA tree with fifty-nine strains of the *P. putida* group. To examine the reliability of the MLSA tree generated by the NJ method, we employed the ML method to generate the MLSA tree with these fifty-nine strains (Fig. S1). The resulting tree was quite similar to the tree generated with the NJ method, where only the position of the MTCC5279 strain slightly differed between the two trees. These results indicate that the MLSA tree presented in Fig. 2 can sufficiently dissect the phylogenetic relationships of the numerous strains of the *P. putida* group.

**ANI analysis based on BLAST**

Because the complete, or draft, genome sequence data are available for all of the strains used in this study, we
performed ANI analysis to elucidate the significance of the MLSA tree in the evolutionary relationship of these strains. The tree obtained with ANI analysis is presented in Fig. 3. The phylogenetic relationships among the strains in the ANI tree were quite similar to the relationships in the MLSA tree, indicating that the MLSA tree has as high a discerning potential as the ANI tree. Because ANI analysis uses a vast amount of nucleotide sequences, the branch lengths of the right-most terminal branches were longer than those of the MLSA tree. In the ANI analysis using BLAST, a 95% threshold for species delineation has been proposed (Goris et al., 2007; Richter and Rosselló-Móra, 2009). This 95% threshold is presented in Fig. 3 as horizontal dashed lines, which separates the fifty-nine P. putida strains into twenty-six species, including seventeen singletons and nine clusters from I to IX. The twenty-six species comprised eleven type-strain species and fifteen potential new species that included sixteen and forty-three strains, respectively.

**Discussion**

In this study, fifty-nine strains of the P. putida group were subjected to phylogenetic analyses after determining the genome sequences of twenty-two strains. Based on the 95% threshold, they were classified into twenty-six species by ANI analysis, which provided additional insights into the diversity of the P. putida group (Gomila et al., 2015; Mulet et al., 2013). Twenty-six strains were clinical isolates, which contained twenty-two newly sequenced strains, two previously reported strains, and two type strains (Table S1). Thirty-three strains were nonclinical isolates, including nine type strains. Nine housekeeping genes were used in the MLSA, and the results agreed with the ANI analysis, supporting the reliability of the MLSA tree. Thus, our results are expected to provide up-to-date detailed phylogenetic statistics on strains of the P. putida group, including both clinical and nonclinical isolates.

The twenty-six species of the P. putida group indicated by ANI analysis included seventeen singletons and nine strain clusters. These nine clusters were completely conserved among the phylogenetic trees of individual housekeeping genes and the MLSA tree. Focusing on the hierarchical tree structure of these nine clusters (Fig. 4), there is a difference between the individual housekeeping gene trees. Such gene-dependent differences in the phylogenetic trees suggest that the concatenated gene sequences that are used for MLSA yield better results. The MLSA tree was quite similar to the ANI tree, where only the position of cluster IV slightly differed between the two trees (Fig. 4). To articulate how the results of MLSA and ANI are similar, a dot plot of the pairwise comparisons of sixty strains between the identity percentages obtained in the MLSA and ANI values is presented in Fig. 5. All of the 1,770 dots are located close to the approximate line that was generated by a linear regression method with a correlation coefficient of 0.9841, suggesting a good correlation between the ANI and the MLSA trees. The dot plot indicates that the 95% threshold for species delineation in the ANI tree is equivalent to 98% in the MLSA tree. When a threshold of 98% was used for species delineation in the MLSA tree, the strain distribution among the species in the MLSA tree highly approximated the strains among the species in the ANI tree, except for the strains LF54, S610, and NBRC 16637T. The strain LF54, which was in cluster I with ANI values ranging from 95.08% to 95.52%, was excluded from the species represented by cluster I in the MLSA tree with MLSA values ranging from 97.30% to 97.58%. The strains S610 and NBRC 16637T that comprised cluster VIII with the ANI value of 95.65% were separated into different species in the MLSA tree with an MLSA value of 97.79%. These results suggest that the

**Fig. 4.** Schematic representation of the hierarchical relationships of the clusters generated by ANI analysis.

Panels (A), (B), (C), (D), (E), (F), (G), (H), (I), (J), and (K) represent the hierarchical relationships based on the phylogenetic trees of ANI, MLSA, argS, dnaN, dnaQ, era, gltA, gyrB, ppnK, rpoB, and rpoD, respectively. The results of MLSA using the NJ and ML methods are the same. The Roman numbers I through IX correspond to those in Fig. 3.

**Fig. 5.** Dot plot of the ANI value versus the MLSA sequence identity. Pairwise comparisons of sixty strains were performed, and each dot indicates an ANI value compared against the MLSA sequence identity. The dashed vertical line represents the 95% threshold value for species identification. The solid line represents the approximate line generated by linear regression. The relational equation and correlation coefficient (R) are indicated.
phyllogenetic relationship of the concatenated housekeeping gene sequences used in the MLSA reflects that of the genome sequences used in the ANI analysis. Based on the 95% threshold in the ANI tree (Fig. 3), twenty-one clinical isolates (out of twenty-six total clinical isolates) constituted eleven potential new species, which are separated from the type-strain and accounted for 79% of the total species of clinical isolates. Among the thirty-three nonclinical isolates, twenty-two strains constituted eight potential new species, which are separated from the type-strain and accounted for 44% of the total species of nonclinical isolates. These results suggest that clinical isolates are more diverse than nonclinical isolates, as they exhibited a greater species variety. Focusing on the separation between the clinical and nonclinical isolates in the ANI tree based on a 95% threshold for species delineation, clusters VI, VIII, and IX contained either clinical or nonclinical isolates, but not both (Fig. 3). The remaining six clusters contained both clinical and nonclinical isolates. A similar distribution of clinical and nonclinical isolates was observed in the MLSA tree. Equally, the nonclinical strain KT2440 isolated from the environment approximated the clinical isolate NBRC 111121 in cluster I. Molina et al. (2016) performed phylogenetic analysis with four clinical and eight nonclinical P. putida strains based on the amino-acid-sequence similarities between five proteins, and indicated that two different clades contained clinical and nonclinical isolates. They also indicated that one clade contained only nonclinical isolates such as the strains KT2440, F1, and DOT-T1E. In our phylogenetic analysis based on ANI, which appears to be more detailed than their analysis based on amino-acid-sequence similarity of proteins, the clade containing KT2440, F1, and DOT-T1E included clinical isolates. Our results with twenty-six clinical and thirty-three nonclinical isolates clearly indicate the mixed distribution of clinical and nonclinical isolates in the phylogenetic tree. These results suggest that the genomic difference between clinical and nonclinical isolates of the P. putida group is subtle and some limited genes are responsible for pathogenicity. Detailed genome sequence comparisons may achieve a separation between clinical and nonclinical isolates. Focusing on the aromatic- and alkane-degrading strains, the strains LS46, TRO1, B6-2, DOT-T1E, F1, YKD221, ND6, SJTE-1, BIRD-1, KT2440, and LF54 comprised cluster I, and the strains SJ9, SB3078, SB3101, DLL-E4, and S16 comprised cluster V (Fig. 3 and Table S1). Because these strains were independently isolated from different places by different researchers, they appeared to be specifically concentrated in the aromatic- and alkane-contaminated environment. Reports employing ANI analysis are becoming more prevalent, and ANI analysis has been proposed as a standard technique for classifying strains at the species level instead of DNA-DNA hybridization (Goris et al., 2007; Löffler et al., 2013; Olofsson et al., 2014; Richter and Rosselló-Móra, 2009; Stephan et al., 2014; Tong et al., 2015; Vanlaere et al., 2009). Based on the 95% threshold for species delineation, clusters I to IX in the ANI tree can constitute independent species. Although the type strain of P. monteilii NBRC 103158T is included in cluster IV, two P. monteilii strains reside in cluster V. Conversely, clusters I, II, III, IV, V, and VIII contain P. putida strains, such as LS46, DOT-T1E, F1, KT2440, GB-1, OUS82, B001, HB3267, DLL-E4, S16, and S610. The P. putida type strain NBRC 14164T is a singleton that is independently located close to cluster III (Fig. 3). These observations suggest that the classification of these strains and the differentiation of species in the P. putida group should be re-examined.

Acknowledgments

This work was supported by a grant from the Ministry of Economy, Trading, and Industry of Japan (Technology development for soil contamination countermeasures).

Supplementary Materials

Supplementary figure and table are available in our J-STAGE site (http://www.jstage.jst.go.jp/browse/jgam).

References


