Microbiology (2000), **146**, 2385–2394 Printed in Great Britain

Phylogeny of the genus *Pseudomonas*: intrageneric structure reconstructed from the nucleotide sequences of *gyrB* and *rpoD* genes

Satoshi Yamamoto,[1]† Hiroaki Kasai,[1] Dawn L. Arnold,[2] Robert W. Jackson,[2] Alan Vivian[2] and Shigeaki Harayama[1]

Author for correspondence: Satoshi Yamamoto. Tel: +81 43 203 0226. Fax: +81 43 203 0216. e-mail: yamamotost@nichirei.co.jp

[1] Marine Biotechnology Institute, Kamaishi Laboratories, Kamaishi City, Iwate 026, Japan

[2] Centre for Research in Plant Science, Faculty of Applied Sciences, University of the West of England, Coldharbour Lane, Bristol BS16 1QY, UK

Phylogenetic analysis of the genus *Pseudomonas* was conducted by using the combined *gyrB* and *rpoD* nucleotide sequences of 31 validly described species of *Pseudomonas* (a total of 125 strains). *Pseudomonas* strains diverged into two major clusters designated intrageneric cluster I (IGC I) and intrageneric cluster II (IGC II). IGC I was further split into two subclusters, the '*P. aeruginosa* complex', which included *P. aeruginosa*, *P. alcaligenes*, *P. citronellolis*, *P. mendocina*, *P. oleovorans* and *P. pseudoalcaligenes*, and the '*P. stutzeri* complex', which included *P. balearica* and *P. stutzeri*. IGC II was further split into three subclusters that were designated the '*P. putida* complex', the '*P. syringae* complex' and the '*P. fluorescens* complex'. The '*P. putida* complex' included *P. putida* and *P. fulva*. The '*P. syringae* complex' was the cluster of phytopathogens including *P. amygdali*, *P. caricapapayae*, *P. cichorii*, *P. ficuserectae*, *P. viridiflava* and the pathovars of *P. savastanoi* and *P. syringae*. The '*P. fluorescens* complex' was further divided into two subpopulations, the '*P. fluorescens* lineage' and the '*P. chlororaphis* lineage'. The '*P. fluorescens* lineage' contained *P. fluorescens* biotypes A, B and C, *P. azotoformans*, *P. marginalis* pathovars, *P. mucidolens*, *P. synxantha* and *P. tolaasii*, while the '*P. chlororaphis* lineage' included *P. chlororaphis*, *P. agarici*, *P. asplenii*, *P. corrugata*, *P. fluorescens* biotypes B and G and *P. putida* biovar B. The strains of *P. fluorescens* biotypes formed a polyphyletic group within the '*P. fluorescens* complex'.

Keywords: *gyrB*, *rpoD*, *Pseudomonas*, phylogeny, PCR

INTRODUCTION

Pseudomonads are ubiquitous bacteria in nature. They possess variable metabolic abilities that enable them to utilize a wide range of organic compounds, and occupy an important ecological position in the carbon cycle. They are also important as pathogens of animals and plants. Therefore, the ecology of pseudomonads in the biosphere has been a matter of interest. An essential prerequisite for a detailed investigation of the roles and evolution of pseudomonads is an accurate system of classification and identification. However, the classification of *Pseudomonas* strains is not fully established due to the lack of an accurate taxonomic system.

The classificatory criteria for the genus *Pseudomonas* have been revised along with progress in bacterial taxonomy. Today, DNA–DNA hybridization is the recommended standard for the delineation of a bacterial species (Wayne *et al*., 1987), but it has the drawback that it is not effective in the estimation of genetic distances between distantly related species. As a complement to DNA–DNA hybridization, sequence analysis of 16S rRNA or its gene (16S rDNA) is frequently used (Laguerre *et al*., 1994; Moore *et al*., 1996; Bennasar *et*

† **Present address:** Nichirei Corporation, Research & Development Center, 9 Shinminato, Mihama-Ku, Chiba-Shi, Chiba 261-8545, Japan.

Abbreviations: IGC, intrageneric cluster; NJ, neighbour-joining; UPGMA, unweighted pair group method with arithmetic averages.

The GenBank accession numbers for the sequences determined in this work are: *gyrB*, D37926, D37297, D86005–D86019 and AB039381–AB039492; *rpoD*, D86020–D86036 and AB039493–AB039624.

Table 1. List of *Pseudomonas* strains analysed

Species		Strains*
P. aeruginosa		IFO 3080, IFO 3446, IFO 12689^{T}, IFO 13275, MT350
P. agarici		NCPPB 1999, NCPPB 2289^{T}, NCPPB 2472
P. alcaligenes		IFO 14159^{T}
P. amygdali		NCPPB 2607^{T}, NCPPB 2609, NCPPB 2610
P. asplenii		NCPPB 959, NCPPB 1947^{T}
P. azotoformans		IAM 1603^{T}
P. balearica		DSM 6083^{T}
P. caricapapayae		NCPPB 1873^{T}, NCPPB 3080, NCPPB 3439
P. chlororaphis		IFO 3904^{T}, IMV B-107, NCIMB 9402, NCIMB 11881
	('*P. aureofaciens*')	NCIMB 9030$^{'T'}$, NCIMB 9265
P. cichorii		NCPPB 907, NCPPB 943^{T}, NCPPB 3109, NCPPB 3283
P. citronellolis		NCIMB 12783^{T}
P. corrugata		NCPPB 2445^{T}, NCPPB 3056, NCPPB 3316
P. ficuserectae		JCM 2400^{T}
P. fluorescens	biotype A	ATCC 17582, ATCC 49036, DSM 50108, IAM 12022^{T}
	biotype B	ATCC 17467, ATCC 17816
	biotype C	ATCC 17561, ATCC 17571, ATCC 17574
	biotype G	ATCC 17573
	biotype unknown	ATCC 12983, ATCC 31948, ATCC 31950, IFO 13334, NCIMB 9815
P. fulva		IAM 1529^{T}
P. iners		IAM 1419^{T}
P. marginalis	pv. *alfalfae*	NCPPB 2644^{T}, NCPPB 2645
	pv. *marginalis*	NCPPB 247, NCPPB 667^{T}, NCPPB 1689, NCPPB 2380, NCPPB 3183, NCPPB 3210
	pv. *pastinaceae*	NCPPB 806^{T}, NCPPB 807, NCPPB 949
'*P. marginata*'		HRI 17
P. mendocina		IMV B-171
P. mucidolens		IAM 12406^{T}
P. oleovorans		IFO 13583^{T}
P. pseudoalcaligenes	subsp. *pseudoalcaligenes*	IAM 12410, IFO 14167^{T}
P. putida	biovar A	ATCC 11172, ATCC 17485 (= PpG7)†, IFO 14164^{T}
	biovar B	ATCC 17484, ATCC 17522
	biotype unknown	A10L (= MBIC 1340)‡, ATCC 12633, ATCC 68832, BH§, IFO 14671, JCM 6156 (= mt-2), KF707‖, KF715¶, NCIMB 9816, PB4 (= MBIC 1487)#
	'*P. ovalis*'	IFO 3738
P. savastanoi	pv. *savastanoi*	NCPPB 639^{T}, NCPPB 3334
	pv. *glycinea*	HRI 1659A, NCPPB 1139, NCPPB 2411^{T}, NCPPB 3318
	pv. *phaseolicola*	HRI 882 (type of race 2), HRI 2709A (type of race 9), NCPPB 52^{T} (race 1)
P. stanieri		ATCC 27130^{T}
P. straminea		IAM 1598^{T}
P. stutzeri		ATCC 14405, IAM 12097, IFO 14165^{T}, IMV B-177
P. synxantha		IFO 3913^{T}
P. syringae	pv. *antirrhini*	ICMP 4303^{PT}
	pv. *coriandricola*	ICMP 12471^{PT}
	pv. *coronafaciens*	ICMP 3113^{PT}
	pv. *lachrymans*	HRI 789, ICMP 3988^{PT}
	pv. *maculicola*	HRI 5424, ICMP 3935^{PT}
	pv. *morsprunorum*	ICMP 5795^{PT}, N10
	pv. *pisi*	HRI 203^{neoPT}, HRI 299A (type of race 1)
	pv. *syringae*	HRI 1338, HRI 2242A, HRI 2692C, ICMP 3023^{T,PT}
P. taetrolens		IAM 1653^{T}
P. tolaasii		NCPPB 2192^{T}, NCPPB 2194, NCPPB 2412
P. viridiflava		HRI 2673C, HRI 2675C, HRI 2676C, ICMP 2848^{PT}

Table 2. PCR primers and sequencing primers used in this study

Primer UP-1E was designed from the conserved regions of the amino acid sequences of GyrB proteins of *E. coli* (SWISS-PROT code GYRB_ECOLI), *P. putida* (GYRB_PSEPU) and *Bacillus subtilis* (GYRB_BACSU); primer APrU was designed from the conserved GyrB sequences among 15 *Acinetobacter* strains including 12 genomospecies (Yamamoto & Harayama, 1996). The nucleotide sequences of universal sequence primers M13 reverse and M13(-21) were appended at the 5′ ends of the degenerated sequences of PCR primers UP-1E and APrU, respectively. Direct sequencing of the PCR fragments was performed by using the universal sequencing primers M13 reverse or M13(-21). PCR primers and sequencing primers for *rpoD* were also designed from the conserved regions of the amino acid sequences of RpoD proteins of *Buchnera aphidicola* (RP70_BUCAP), *E. coli* (RP70_ECOLI), *Salmonella typhimurium* (RP70_SALTY) and *P. aeruginosa* (RP70_PSEAE) (Yamamoto & Harayama, 1998). N, any; R, A or G; S, C or G; Y, C or T; M, A or C.

Primer	Target gene	Sequence (5′–3′)	Length (mer)
UP-1E	*gyrB*	CAGGAAACAGCTATGACCAYGSNGGNGGNAARTTYRA	37
M13R		CAGGAAACAGCTATGACC	18
APrU	*gyrB*	TGTAAAACGACGGCCAGTGCNGGRTCYTTYTCYTGRCA	38
M13(-21)		TGTAAAACGACGGCCAGT	18
70F	*rpoD*	ACGACTGACCCGGTACGCATGTAYATGMGNGARATGGGNACNGT	44
70Fs		ACGACTGACCCGGTACGCATGTA	23
70R	*rpoD*	ATAGAAATAACCAGACGTAAGTTNGCYTCNACCATYTCYTTYTT	44
70Rs		ATAGAAATAACCAGACGTAAGTT	23

al., 1998). However, the degree of resolution obtained with 16S rRNA sequence analysis is not sufficiently discriminatory to permit resolution of intrageneric relationships because of the extremely slow rate of evolution of 16S rRNA. Due to the gap between the valid genetic ranges of the two methods, a detailed intrageneric structure of the genus *Pseudomonas* remains to be resolved.

In this article, we discuss the intrageneric structure of the genus *Pseudomonas* on the basis of the nucleotide sequences of their genes for DNA gyrase B subunit (*gyrB*) and σ^{70} factor (*rpoD*). DNA gyrase is the enzyme responsible for introducing negative supercoils into bacterial chromosomes and plays a crucial role in the replication of chromosomes (Watt & Hickson, 1994). The σ^{70} factor, on the other hand, is one of the sigma factors that confer promoter-specific transcription initiation on RNA polymerase (Lonetto *et al*., 1992). Both proteins are ubiquitous in bacteria and essential for their cell growth. We previously reported that these protein-encoding genes evolved much faster than rDNAs and provided higher resolution than the use of 16S rRNA sequences (Yamamoto & Harayama, 1998). We further demonstrated that the phylogenetic clustering of *Acinetobacter* strains based on *gyrB* sequence analysis is almost equivalent to the genomic species delineated by DNA–DNA hybridization (Yamamoto *et al*., 1999). These results suggest that a phylogenetic analysis using the *gyrB* and *rpoD* sequences may fill the resolution gap between 16S rRNA sequence analysis and DNA–DNA hybridization studies.

Molecular phylogeny deduced from a single locus may be unreliable due to the stochastic nature of base substitutions or to rare horizontal gene transfer events. Consequently, we decided to use a combination of the *gyrB* and *rpoD* genes to establish the phylogenetic

* T, type strain; PT, pathotype strain; ATCC, American Type Culture Collection, Manassas, VA, USA; HRI, Horticulture Research International culture collection, Wellesbourne, UK; DSM, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany; IAM, Institute of Molecular and Cellular Biosciences (formerly Institute of Applied Microbiology), The University of Tokyo, Tokyo, Japan; ICMP, International Collection of Microorganisms from Plants (formerly PDDCC, Plant Diseases Division Culture Collection), Auckland, New Zealand; IFO, Institute for Fermentation, Osaka, Japan; IMV, Institute of Microbiology and Virology, Academy of Sciences of Ukraine, Kiev, Ukraine; JCM, Japan Collection of Microorganisms, Institute of Physical and Chemical Research, Wako, Japan; MBIC, Collection of the Marine Biotechnology Institute, Kamaishi, Japan; NCIMB, National Collection of Industrial Bacteria, Aberdeen, UK; NCPPB, National Collection of Plant-pathogenic Bacteria, Central Science Laboratory, York, UK.

† Barnsley (1976), Yen & Gunsalus (1985).

‡ Shimao *et al*. (1996).

§ Takeo *et al*. (1995).

‖ From Dr Furukawa, Faculty of Agriculture, Kyushu University, Fukuoka, Japan.

¶ Furukawa *et al*. (1989).

Yamamoto & Harayama (1995).

Fig. 1. For legend see facing page.

relationships of strains in the genus *Pseudomonas*. This strategy allowed, first, the comparison of the *gyrB*-based phylogenetic tree with the *rpoD*-based one to confirm the consistency of the analyses, and second, the improvement of the reliability of the phylogenetic tree.

METHODS

Bacterial strains and DNA preparation. Thirty-one validly described *Pseudomonas* species, comprising a total of 125 strains, were examined in this study (Table 1). Each bacterial sample was grown aerobically in nutrient broth at 30 °C. Chromosomal DNAs from *Pseudomonas* strains used as the PCR template were prepared by the methods described by Brenner *et al*. (1982) or by using the Puregene DNA Isolation kit (Gentra Systems) according to the supplier's instructions.

PCR amplification and sequencing of *gyrB* and *rpoD*. PCR amplification of *gyrB* and *rpoD* genes was done following the method described previously (Yamamoto & Harayama, 1995, 1998; Yamamoto *et al*., 1999). PCR amplification was performed with a Progene thermal cycler (Techne) by using PCR buffer (Perkin-Elmer) containing each of the deoxynucleoside triphosphates at a concentration of 200 μM, each of the primers at a concentration of 1 μM, 1 μg template DNA and 2·5 U AmpliTaq Gold DNA polymerase (Perkin-Elmer) in a total volume of 100 μl. A total of 35 cycles of amplification was performed, with the template DNA denaturation at 94 °C for 1 min, the primer annealing for 30 s and the primer extension at 72 °C for 2 min. The annealing temperature ranged from 58 to 63 °C depending on the G + C content of the template DNA. Higher annealing temperatures were used for higher G + C contents. Amplified products were electrophoresed on 0·8 % low-melting-temperature agarose gels (SeaPlaque GTG; FMC Bioproducts), and purified by using QIAquick (Qiagen) following the manufacturer's instructions. The nucleotide sequences of *gyrB* and *rpoD* genes were determined directly from the PCR fragments. PCR primers and sequencing primers used in this study are summarized in Table 2. Sequencing was carried out using an ABI PRISM BigDye Terminator Cycle Sequencing kit and a 377 DNA sequencer (Perkin-Elmer) according to the manufacturer's instructions.

Data analysis. GyrB and RpoD sequences translated from *gyrB* and *rpoD* sequences, respectively, were aligned using the CLUSTAL W computer program (Thompson *et al*., 1994). The *gyrB* and *rpoD* sequences were aligned manually according to the alignments of the GyrB and RpoD sequences, respectively. Phylogenetic trees were constructed with the PHYLIP computer program package (Felsenstein, 1989), using the neighbour-joining (NJ) method (Saitou & Nei, 1987) with genetic distances computed using Kimura's two-parameter model (Kimura, 1980). Phylogenetic trees were reconstructed not only from the data sets of *gyrB* (888–891 bp) and *rpoD* (798–816 bp), but also from the combined nucleotide sequences of these two genes (1686–1707 bp), assuming that the analysis using longer sequences would result in a better resolution and reliability. In the latter analysis, *gyrB* and *rpoD* sequences from the same strain were combined in series and treated as a single nucleotide sequence. The nucleotide sequences of the *gyrB* and *rpoD* genes from *Escherichia coli* K-12 (ECGYRB and ECORPSPRO) were used as the outgroup for phylogenetic tree reconstruction.

RESULTS

Reconstruction of phylogenetic tree

As the first step in the analysis, NJ trees from the *gyrB* and *rpoD* nucleotide sequences were reconstructed from the individual data sets. The basic topologies of these two NJ trees were similar to each other but slightly different in detail (data not shown). The most conspicuous difference between them was the branching order of the clusters that included the '*P. aeruginosa* lineage' strains and those that included the '*P. stutzeri* lineage' strains. In the *gyrB* NJ tree, these two clusters branched off after they had diverged from the remaining mass of *Pseudomonas*, whilst in the *rpoD* NJ tree, the '*P. stutzeri* lineage' cluster branched off first, followed by the '*P. aeruginosa* lineage', which diverged from the rest. Theoretically, the influence of stochastic drift on the rate of evolution cannot be eliminated from the molecular phylogeny. Hence, these minor discrepancies may have their origin in such drift. If this were the case, the use of longer sequences, for example the combined *gyrB* and *rpoD* sequences, in the analysis would give a more accurate estimate of the phylogeny.

We thus conducted the phylogenetic analysis by using the combined nucleotide sequences of the *gyrB* and *rpoD* genes. A NJ tree reconstructed from the combined *gyrB* and *rpoD* nucleotide sequences (*gyrB-rpoD* tree) is shown in Fig. 1. The topology of trees reconstructed from the combined sequences of *gyrB* and *rpoD* by the unweighted pair group method with arithmetic averages (UPGMA) (Sokal & Michener, 1958) and the maximum-parsimony method (Fitch, 1971) were almost identical to that of the NJ tree. One notable difference between them was the position of *P. straminea* IAM 1598^T. In the NJ and maximum-parsimony trees, *P. straminea* IAM 1598^T was excluded from both of the clusters of the '*P. aeruginosa* lineage' and the '*P. stutzeri* lineage', whereas in the UPGMA tree, *P. straminea* IAM 1598^T was included in the cluster of the '*P. aeruginosa* lineage'. The phenotypic characteristics of *P. straminea*, such as its considerably lower maximum growth temperature (no growth at 37 °C) in comparison with that of the '*P. aeruginosa* lineage' and the '*P. stutzeri* lineage' (can grow at 41 °C), may account for its divergence as in the NJ tree and maximum-parsimony tree. Consequently, we chose the combined *gyrB* and *rpoD* NJ tree as the basis of *Pseudomonas* phylogeny in this study.

Fig. 1. Phylogenetic tree of 125 *Pseudomonas* strains that have been assigned to 31 validly described species based on the nucleotide sequences of the *gyrB* and *rpoD* genes. The tree was reconstructed by using the NJ method, using the genetic distances computed by using the Kimura's two-parameter model (Kimura, 1980). The scale bar indicates a genetic distance of 0·1. The number shown next to each node indicates the percentage bootstrap values of 1000 replicates that exceeded 50 % (except for the important node). Sequences from *E. coli* K-12 were treated as the outgroup. The topological characteristics of the phylogenetic trees produced by the NJ method, the maximum-parsimony method and UPGMA were almost identical (data not shown).

Outline of the intrageneric structure of the genus *Pseudomonas*

The *Pseudomonas* strains examined, with the exception of two strains (*P. stanieri* ATCC 27130^T and *P. iners* IAM 1419^T), formed a monophyletic group that corresponds to the rRNA group I of Palleroni (1984). The intrageneric relationships of the authentic *Pseudomonas* species observed in the combined *gyrB* and *rpoD* NJ tree are summarized in Fig. 2. Within the genus *Pseudomonas*, two major intrageneric clusters were recognized. The first intrageneric cluster includes *P. aeruginosa*, *P. alcaligenes*, *P. balearica*, *P. citronellolis*, *P. mendocina*, *P. oleovorans*, *P. pseudoalcaligenes*, *P. straminea* and *P. stutzeri*. The second intrageneric cluster includes *P. agarici*, *P. amygdali*, *P. asplenii*, *P. azotoformans*, *P. caricapapayae*, *P. chlororaphis* (including ex-*P. aureofaciens* strains), *P. cichorii*, *P. corrugata*, *P. ficuserectae*, *P. fluorescens* (biotypes A, B, C and G), *P. fulva*, *P. marginalis* pathovars, *P. mucidolens*, *P. putida* (biovars A and B), *P. savastanoi* pathovars, *P. synxantha*, *P. syringae* pathovars, *P. taetrolens* and *P. tolaasii*. These two intrageneric divisions seem consistent with the '*P. aeruginosa* intrageneric cluster' and the '*P. fluorescens* intrageneric cluster' that have been designated using 16S rRNA gene sequence analysis (Moore *et al.*, 1996). However, phylogenetic relationships within each of the 'intrageneric clusters' in the combined *gyrB* and *rpoD* NJ tree were eminently different from that of the 16S rRNA gene sequence trees (Moore *et al.*, 1996; Verhille *et al.*, 1999). Consequently, in this article, we propose to designate the first and second major clusters as 'intrageneric cluster I' (IGC I) and 'intrageneric cluster II' (IGC II).

IGC I

IGC I includes the type species of the genus *Pseudomonas*, *P. aeruginosa*. Many of the constituents of this cluster were isolated from clinical specimens. The general characteristics of IGC I species are: a single polar flagellum; a higher range of G+C contents (60·6–66·3 mol%) than IGC II (59·0–63·6 mol%); growth at 41 °C, except for *P. straminea* (no growth at 37 °C); non-production of fluorescent pigments, except for *P. aeruginosa* and *P. straminea* (Palleroni, 1984). No consensus profile of carbon source utilization was recognized.

IGC I was further split into at least two subclusters, the '*P. aeruginosa* complex', which included *P. aeruginosa*, *P. alcaligenes*, *P. citronellolis*, *P. mendocina*, *P. oleovorans* and *P. pseudoalcaligenes*, and the '*P. stutzeri* complex', which included *P. balearica* and *P. stutzeri*. *P. aeruginosa* strains were monophyletic within the '*P. aeruginosa* complex'. The type strains of *P. oleovorans* and *P. pseudoalcaligenes*, IFO 13583^T and IFO 14167^T, respectively, were genetically almost identical. As we discussed previously, the phylogenetic position of *P. straminea* in IGC I was uncertain as its bootstrap probability was 46%. *P. straminea* possibly represents an independent third subcluster.

IGC II

The general characteristics of IGC II are: more than one polar flagellum; a lower range of G+C contents (59·0–63·6 mol%) than IGC I (60·6–66·3 mol%); inability to grow at 41 °C and ability of some strains to grow at 4 °C; production of fluorescent pigments by most strains (Palleroni, 1984). IGC II is more species-rich than IGC I despite the comparable extent of genetic diversity of the two clusters. Within IGC II, at least three distinct monophyletic groups were recognized. We designated these subclusters the '*P. putida* complex', the '*P. syringae* complex' and the '*P. fluorescens* complex'.

The '*P. putida* complex' included *P. putida* and *P. fulva*. All of the *P. putida* strains characterized as biovar A belonged in this group; however, all of the *P. putida* biovar B strains were included in the '*P. fluorescens* complex'. Strains of *P. putida* biovar B have been distinguished from biovar A mainly by their ability to grow on L-tryptophan and L-kynurenine. Collectively, the phenotypes of the *P. putida* biovar B strains resemble more closely those of *P. fluorescens* strains than those of biovar A. For example, *P. putida* biovar B strains can grow at 4 °C and can utilize a broader range of sugars than biovar A strains (Stanier *et al.*, 1966). These physiological differences support the phylogenetic separation of biovar B strains from the clade of *P. putida* including the type strain and biovar A strains.

The '*P. syringae* complex' was the clade of phytopathogenic *Pseudomonas*, representing a huge 'pathogenicity island'. The '*P. syringae* complex' included *P. amygdali*, *P. caricapapayae*, *P. cichorii*, *P. ficuserecta*e, *P. viridiflava*, *P. savastanoi* pathovars (pv. *savastanoi*, pv. *phaseolicola* and pv. *glycinea*) and *P. syringae* pathovars (pv. *antirrhini*, pv. *coriandricola*, pv. *coronafaciens*, pv. *lachrymans*, pv. *maculicola*, pv. *morsprunorum*, pv. *pisi* and pv. *syringae*).

The '*P. fluorescens* complex' was the most species-rich subcluster and at least 16 validly described species belonged in this clade. The majority of the constituents of the '*P. fluorescens* complex' were fluorescent, with the exception of *P. corrugata*. Many are saprophytic and/or psychrophilic (growth at 4 °C), and are frequently associated with food spoilage, especially at chilling temperatures (1–10 °C). Some members are also important as pathogens of plants and fungi. This cluster was further divided into two subpopulations, the '*P. fluorescens* lineage' and the '*P. chlororaphis* lineage'. The '*P. fluorescens* lineage' contained *P. fluorescens* biotypes A, B and C, *P. azotoformans*, *P. marginalis* pathovars, *P. mucidolens*, *P. synxantha* and *P. tolaasii*, whilst the '*P. chlororaphis* lineage' included *P. chlororaphis*, *P. agarici*, *P. asplenii*, *P. corrugata*, *P. fluorescens* biotypes B and G and *P. putida* biovar B. The strains of *P. fluorescens* biotypes, especially those of biotype B, were positioned either in the '*P. fluorescens* lineage' or in the '*P. chlororaphis* lineage'. It is clear that a reclassification of the *P. fluorescens* strains is required.

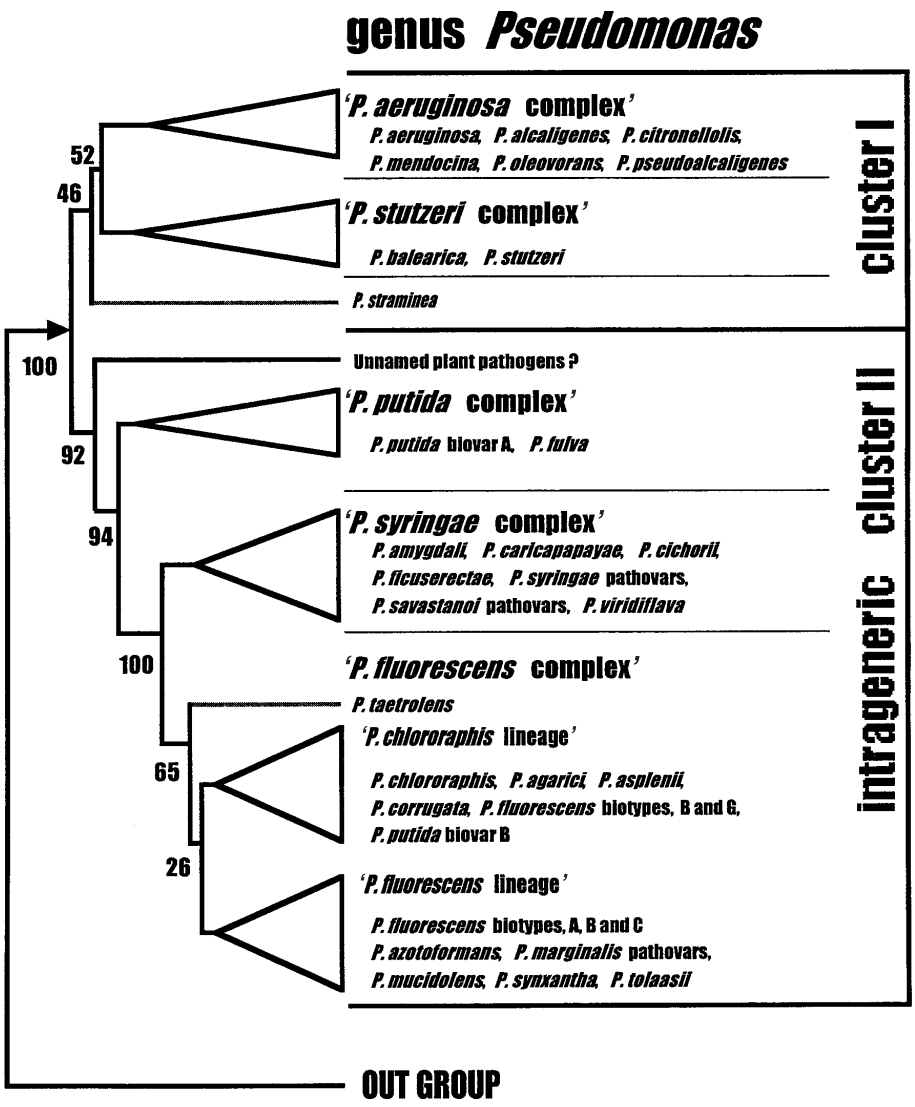

Fig. 2. Schematic dendrogram summarizing the intrageneric structure of the genus *Pseudomonas*. The number shown next to each node indicates the percentage bootstrap value of 1000 replicates (only those for the important nodes are indicated).

On the other hand, *P. agarici*, *P. asplenii*, *P. chlororaphis*, *P. corrugata* and *P. tolaasii* formed tight monophyletic branches and can be considered as solid species.

DISCUSSION

Phylogenetic relationships of pseudomonads resolved by using *gyrB* and *rpoD* sequences were eminently different from those resolved by using 16S rRNA sequences (Moore *et al.*, 1996; Verhille *et al.*, 1999). These discrepancies appear to have their origins in the eccentric evolutionary process of 16S rRNA genes. The secondary structures of 16S rRNA determined by the complementary sequences in the small helices are functionally important since 16S rRNA provides a scaffold for the assembly of ribosomal proteins into the small subunit and interacts with mRNA. The majority of base substitutions in 16S rRNAs between closely related organisms are located in these helices, called variable regions, and generally these substitutions are compensatory, i.e. they maintain the base pairing within the helices (Hancock *et al.*, 1988; Rousset *et al.*, 1991; Dixon & Hillis, 1993). We observed that the genetic distances in the variable regions of 16S rRNAs correlated poorly with the synonymous distances in the *gyrB* and *rpoD* genes (Yamamoto & Harayama, 1998). This observation suggests that the base substitutions in these helices might not be accumulated by successive point mutations, but might be caused by single-event mutations introducing multiple substitutions. Thus, the genetic distances calculated from the whole 16S rRNA sequences could be erroneous because the numbers of base substitutions outside the variable regions are much fewer. Frequently, the results of the 16S-rRNA-based

analysis did not correlate with the DNA–DNA reassociation values determined by DNA hybridization experiments that have been used as the criterion for bacterial species definition (Fox *et al.*, 1992; Stackebrandt & Goebel, 1994). In contrast, results of *gyrB* sequence analysis correlated very well with DNA reassociation values (Yamamoto *et al.*, 1999). Obviously, the use of appropriate protein genes like *gyrB* is more suitable for the delineation of intrageneric relationships because they evolved mainly by synonymous substitutions (the 'molecular clock').

The extent of genetic diversity among the *Pseudomonas* species was quite variable. The '*P. stutzeri* complex' and the '*P. putida* complex' consisted of single species, *P. stutzeri* and *P. putida*, respectively. In contrast, the '*P. fluorescens* complex', which is genetically less divergent than either the '*P. stutzeri* complex' or the '*P. putida* complex', contained no less than 16 species.

Three validly described, but as yet unconfirmed, *Pseudomonas* species (Kersters *et al.*, 1996), *P. azotoformans*, *P. fulva* and *P. straminea*, were newly confirmed as genuine members of the genus *Pseudomonas* in this study. Conversely, type strains of two valid *Pseudomonas* species, *P. stanieri* ATCC 27130^T and *P. iners* IAM 1419^T, were excluded from the genus.

Some of the phenotypic traits of the strains in the genus *Pseudomonas* did not reflect their phylogenetic relationships. The fluorescent strains were polyphyletic and belonged to both IGC I and IGC II. Within IGC II, non-fluorescent pseudomonads formed several independent clusters (e.g. *P. corrugata* in the '*P. fluorescens* complex' and *P. amygdali* in the '*P. syringae* complex'). We presume that the common ancestor of the genus *Pseudomonas* was originally fluorescent and that non-fluorescent species have appeared by losing their ability to produce a fluorescent pigment.

Phytopathogenic *Pseudomonas* species were also distributed throughout the genus *Pseudomonas*, not only in the '*P. syringae* complex'. The polyphyletic relationship of phytopathogenic pseudomonads suggests that pathogens in separate clades have independently acquired their phytopathogenicity, presumably by the intergeneric transmission of virulence factors. It is likely that the common ancestor of the '*P. syringae* complex' had acquired a set of factors that determine pathogenicity to plants, after the divergence of the '*P. syringae* complex' and the '*P. fluorescens* complex'. More recently, such genes would probably be horizontally acquired via plasmid transfer and involve virulence and/or avirulence genes such as *avrPpiB* described in *P. syringae* pv. *pisi* race 3 (Cournoyer *et al.*, 1995) or the pathogenicity island in *P. savastanoi* pv. *phaseolicola* (Jackson *et al.*, 1999). If so, virulence/avirulence genes would be expected to evolve along with the evolutionary course of the '*P. fluorescens* complex'. However, correlation was not necessarily observed between the phylogenetic relationships of phytopathogens and their host range. For example, strains of *P. viridiflava* (pathogenic on bean) were scattered at various phylogenetic positions in the '*P. syringae* complex', whilst those of *P. caricapapayae* (pathogenic on pawpaw) formed a tight cluster. Polyphyly of host range may be due to the horizontal transfer of virulence and/or avirulence genes.

Within the '*P. syringae* complex', *P. cichorii* strains formed an independent monophyletic cluster with two strains of *P. syringae* pv. *syringae*. *P. cichorii* strains are oxidase-positive, as are the majority of pseudomonads, whilst exceptionally, the *P. syringae* strains were oxidase-negative. This phenotypic difference corroborates the hypothesis that the '*P. cichorii* lineage' branched at an early stage after the acquisition of phytopathogenicity by the common ancestor, and subsequently the ancestor of the '*P. syringae* lineage' lost the ability to produce oxidase. The detached *P. syringae* cluster in the '*P. cichorii* lineage' might also have lost the ability to produce oxidase by an independent event.

As stated above, reclassification of *P. fluorescens* biotype strains, the major constituents of the '*P. fluorescens* complex', is required since they were polyphyletic. Thus, a DNA–DNA hybridization study accompanied by a phenotypic analysis should be carried out to propose new classificatory criteria for the '*P. fluorescens* complex'. Experiments with an optimal combination of strains for DNA–DNA hybridization should be designed, guided by data from the *gyrB* and *rpoD* sequence analysis rather than that from 16S rDNA analysis (Yamamoto & Harayama, 1998). The close relationship between *P. savastanoi* pathovars was clearly demonstrated in the *gyrB* and *rpoD* sequence analyses presented here. The new species *P. savastanoi* was created to include the *P. syringae* pathovars *glycinea*, *phaseolicola* and *savastanoi*, on the basis of their DNA relatedness obtained by DNA–DNA hybridization experiments (Gardan *et al.*, 1992).

Because of the high ratios of evolution in the *gyrB* and/or *rpoD* nucleotide sequences, it was very easy to design PCR primers or probes having specificity for clusters of species, subspecies or higher classes such as the 'lineages' and 'complexes' designated in this paper (data not shown). On the other hand, the design of such primers or probes based on the variable regions of 16S rRNA presents a problem because sequence similarity among the variable regions does not always guarantee a close phylogenetic relationship (Yamamoto & Harayama, 1998). In conclusion, classification, identification and detection systems for pseudomonads based on *gyrB* and/or *rpoD* sequences can be very useful in microbial ecology and other fields of bacteriology.

ACKNOWLEDGEMENTS

We are grateful to Ms Atsuko Katsuta, Ms Syouko Komukai, Ms Yuuka Takahashi and Ms Ikuko Hiramatsu for technical

assistance. We are also grateful to Dr Elena A. Kiprianova for providing us with bacterial strains. This work was performed as a part of The Industrial Science and Technology Frontier Program supported by New Energy and Industrial Technology Development.

REFERENCES

Barnsley, E. A. (1976). Role and regulation of the ortho and meta pathways of catechol metabolism in pseudomonads metabolizing naphthalene and salicylate. *J Bacteriol* **125**, 404–408.

Bennasar, A., Guasp, J. & Lalucat, J. (1998). Molecular methods for the detection and identification of *Pseudomonas stutzeri* in pure culture and environmental samples. *Microb Ecol* **35**, 22–53.

Brenner, D. J., Mcwhorter, A. C., Leete Kundson, J. K. & Steigerwalt, A. G. (1982). *Escherichia vulneris*: a new species of Enterobacteriaceae associated with human wounds. *J Clin Microbiol* **15**, 1133–1140.

Cournoyer, B., Sharp, J. D., Astuto, A., Gibbon, M. J., Taylor, J. D. & Vivian, A. (1995). Molecular characterization of the *Pseudomonas syringae* pv. *pisi* plasmid-borne avirulence gene *avrPpiB* which matches the *R*3 resistance locus in pea. *Mol Plant–Microbe Interact* **8**, 700–708.

Dixon, M. T. & Hillis, D. M. (1993). Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Mol Biol Evol* **10**, 256–267.

Felsenstein, J. (1989). PHYLIP – phylogeny inference package (version 3.2). *Cladistics* **5**, 164–166.

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* **20**, 406–416.

Fox, G. E., Wisotzkey, J. D. & Jurtshuk, P. J. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J Syst Bacteriol* **42**, 166–170.

Furukawa, K., Hayase, N., Taira, K. & Tomizuka, N. (1989). Molecular relationship of chromosomal genes encoding biphenyl/polychlorinated biphenyl catabolism: some soil bacteria possess a highly conserved *bph* operon. *J Bacteriol* **171**, 5467–5472.

Gardan, L., Bollet, C., Abu Ghorrah, M., Grimont, F. & Grimont, A. D. (1992). DNA relatedness among the pathovar strains of *Pseudomonas syringae* subsp. *savastanoi* Janse (1982) and proposal of *Pseudomonas savastanoi* sp. nov. *Int J Syst Bacteriol* **42**, 606–612.

Hancock, J. M., Tautz, D. & Dover, G. A. (1988). Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. *Mol Biol Evol* **5**, 393–414.

Jackson, R. W., Athanassopoulos, E., Tsiamis, G. & 7 other authors (1999). Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in *Pseudomonas syringae* pathovar phaseolicola. *Proc Natl Acad Sci U S A* **96**, 10875–10880.

Kersters, K., Ludwig, W., Vancanneyt, M., De Vos, P., Gillis, M. & Schleifer, K.-H. (1996). Recent changes in the classification of the pseudomonads; an overview. *Syst Appl Microbiol* **19**, 465–477.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

Laguerre, G., Rigottier-Gois, L. & Lemanceau, P. (1994). Fluorescent *Pseudomonas* species categorized by using polymerase chain reaction (PCR)/restriction fragment analysis of 16S rDNA. *Mol Ecol* **3**, 479–487.

Lonetto, M., Gribskov, M. & Gross, C. A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* **174**, 3843–3849.

Moore, E. R. B., Mau, M., Arnscheidt, A., Böttger, E. C., Hutson, R. A., Collins, M. D., Van de Peer, Y., De Wachter, R. & Timmis, K. N. (1996). The determination and comparison of the 16S rRNA gene sequence of species of the genus *Pseudomonas* (sensu stricto) and estimation of the natural intrageneric relationships. *Syst Appl Microbiol* **19**, 478–492.

Palleroni, N. J. (1984). Genus I. *Pseudomonas* Migula 1894, 237^{AL}. In *Bergey's Manual of Systematic Bacteriology*, vol. 1, pp. 141–199. Edited by N. J. Krieg & J. G. Holt. Baltimore: Wiliams & Wilkins.

Rousset, F., Pelandakis, M. & Solignac, M. (1991). Evolution of compensatory substitutions through G.U intermediate state in *Drosophila* rRNA. *Proc Natl Acad Sci U S A* **88**, 10032–10036.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

Shimao, M., Nakamura, T., Okuda, A., Abe, M. & Harayama, S. (1996). Characterization of transposon insertion mutants of mandelic acid-utilizing *Pseudomonas putida* strain A10L. *Biosci Biotechnol Biochem* **60**, 1051–1055.

Sokal, R. R. & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *Univ Kans Sci Bull* **28**, 1409–1438.

Stackebrandt, E. & Goebel, B. (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44**, 846–849.

Stanier, R. Y., Palleroni, N. J. & Doudoroff, M. (1966). The aerobic pseudomonads: a taxonomic study. *J Gen Microbiol* **43**, 159–271.

Takeo, M., Maeda, Y., Okada, H., Miyama, K., Mori, K., Ike, M. & Fujita, M. (1995). Molecular cloning and sequencing of the phenol hydroxylase gene from *Pseudomonas putida* BH. *J Ferment Bioeng* **79**, 485–488.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Verhille, S., Baida, N., Dabboussi, F., Izard, D. & Leclerc, H. (1999). Taxonomic study of bacteria isolated from natural mineral waters: proposal of *Pseudomonas jessenii* sp. nov. and *Pseudomonas mandelii* sp. nov. *Syst Appl Microbiol* **22**, 45–58.

Watt, P. M. & Hickson, I. D. (1994). Structure and function of type II DNA topoisomerases. *Biochem J* **303**, 681–695.

Wayne, L. G., Brenner, D. J., Colwell, R. R. & 9 other authors (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* **37**, 463–464.

Yamamoto, S. & Harayama, S. (1995). PCR amplification and direct sequencing of *gyrB* genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. *Appl Environ Microbiol* **61**, 1104–1109.

Yamamoto, S. & Harayama, S. (1996). Phylogenetic analysis of

Acinetobacter strains based on the nucleotide sequences of *gyrB* genes and on the amino acid sequences of their products. *Int J Syst Bacteriol* **46**, 506–511.

Yamamoto, S. & Harayama, S. (1998). Phylogenetic relationships of *Pseudomonas putida* strains deduced from the nucleotide sequences of *gyrB*, *rpoD* and 16S rRNA genes. *Int J Syst Bacteriol* **48**, 813–819.

Yamamoto, S., Bouvet, P. J. M. & Harayama, S. (1999). Phylogenetic structures of the genus *Acinetobacter* based on *gyrB* sequences: comparison with the grouping by DNA-DNA hybridization. *Int J Syst Bacteriol* **49**, 87–95.

Yen, K. M. & Gunsalus, I. C. (1985). Regulation of naphthalene catabolic genes of plasmid NAH7. *J Bacteriol* **162**, 1008–1013.

Received 4 April 2000; revised 17 July 2000; accepted 21 July 2000.