Genotyping of the genus *Proteus* by *rpoB* sequence analysis

Giuseppe Giammanco[1], Sarina Pignato[1], Patrick A.D. Grimont[2], Francine Grimont[2], Giovanni M. Giammanco[3]

[1]Dipartimento "G.F. Ingrassia" - Igiene e Sanità Pubblica, Università di Catania, Catania, Italy; [2]Unité des Entérobactéries, Unité INSERM 389, Institut Pasteur, Paris, France; [3]Dipartimento di Igiene e Microbiologia, Università di Palermo, Palermo, Italy

Correspondence to: Giuseppe Giammanco, Dipartimento "G.F. Ingrassia" - Igiene e Sanità Pubblica, Università di Catania, via Santa Sofia, Catania, Italy. E-mail: giugiam@unict.it

Abstract

Background. The genetic relationship among different species within the genus *Proteus* has still not been clarified and previous studies by rRNA gene restriction patterns (ribotypes) suggested a high genetic variability in the presently recognised nomenspecies.
Methods. The usefulness of *rpoB* sequencing for inter- and intraspecies discrimination in the genus *Proteus* was evaluated on 11 type and clinical strains belonging to the four described species: *P. penneri*, *P. vulgaris*, *P. mirabilis*, and *P. myxofaciens*.
Results. The *rpoB* sequencing proved able to characterise the different species, showing six well defined *rpoB* sequence groups in the 1153 bp region analysed. *P. myxofaciens* and *P. mirabilis* could be clearly differentiated from the types and clinical strains of the other species showing sequence divergences of 19.5-23% and 6.6-7.6%, respectively. Two groups of *P. vulgaris* sequences could be described, one of them including the new strain type, differing from each other by 3.3-3.6% of their nucleotides (nt) and for both of them 6.5-6.8% nt differences from the stand-alone former *P. vulgaris* strain type were found. *P. penneri* differed by only 2.3-3% from all *P. vulgaris* strains, however, differed from the former *P. vulgaris* strain type by 5.5-5.8%.
Conclusions. Our results confirm previous ribotyping data regarding the validity of the presently recognised nomenspecies within the genus *Proteus*, but provides further evidence for the existence of genetic differences within the *P. vulgaris* species.

Key words: Proteus, genetic analysis, rpoB sequencing, ribotyping

Introduction

Bacteria of the genus *Proteus* are part of the normal flora of the intestinal tract of humans and animals and are widespread in the environment. Some of the recognised species act as opportunistic pathogens causing primary and secondary infections. In particular, *P. mirabilis* accounts for approximately 3% of nosocomial infections in the United States, mainly causing the urinary tract diseases: cystitis, pyelonephritis, and prostatitis.[1] Furthermore, *P. mirabilis* and *P. penneri* have been implicated in cases of bacteremia, neonatal meningoencephalitis, empyema, osteomyelitis and subcutaneous abscesses.[2] These two species were also implicated as the causes of hospital-acquired infections.[3,4] As known for other opportunistic pathogens, predisposing factors for urinary tract infections in hospital patients are catheterisation and surgery of the urinary tract, and in outpatients the factors include diabetes and urinary tract abnormalities.

Several phenotypic typing methods have been used in the past for both taxonomic and epidemiological studies of the genus *Proteus*, including serology, phage typing, bacteriocyn typing, and antibiotic resistance patterns.[5-8] In more recent years, molecular typing methods such as ribosomal DNA (rDNA) fingerprints (ribotyping) and arbitrarily primed PCR have been applied in the epidemiological investigation of outbreaks of *P. mirabilis* colonisation [9] and 16S rRNA sequencing has been used to assess the prevalence of the different genomospecies of *P. vulgaris* among clinical isolates and their predilection for certain anatomic sites.[10] Sequence analysis of the RNA polymerase ß-subunit encoding gene (*rpoB*) has been proposed as a novel tool for bacterial identification.[11] Sequences of *rpoB* from several strain types of enterobacterial species have been reported and comparison of *rpoB* and 16S rRNA sequencing data has confirmed the ancestral branching of *Providencia stuartii* and *P. mirabilis*.[11]

In a previous paper we described the rRNA gene restriction patterns (ribotypes) obtained for the strain types and clinical isolates belonging to the different species within the genus *Proteus*.[12] In this study we evaluated the usefulness of *rpoB* sequencing for inter- and intraspecies discrimination in the genus *Proteus* and compared the results with previously obtained ribotyping data.

Methods

Bacterial strains

A total of 11 *Proteus* strains belonging to the species *P. penneri* (3 strains), *P. vulgaris* (5 strains), *P. mirabilis* (2 strains), and *P. myxofaciens* (1 strain) were studied. Reference strains were obtained from the National Collection of Type Cultures (NCTC), London, United Kingdom; from the National Collection of Industrial and Marine Bacteria (NCIMB), Aberdeen, Scotland; and from the Collection of the Institut Pasteur (CIP), Paris, France. Clinical isolates were identified by standard biochemical methods from urine specimens at the Department of Hygiene and Microbiology, University of Palermo, Palermo, Italy.

DNA extraction and rpoB amplification

Genomic DNA was extracted by a commercial nucleic acid extraction kit (IsoQuick, InterBiotech ORCA Research, Othell, WA) according to the manufacturer's instructions. A portion of the coding region of the *rpoB* was amplified in a 40-cycle PCR reaction with denaturation at 94°C for 10 seconds after initial denaturation of 94°C for 90 seconds, annealing at 50°C for 20 seconds, elongation at 72°C for 10 seconds and final extension at 72°C for 5 minutes. PCR amplification was performed in a 50 µl reaction volume containing 1 µg of DNA in 10 mM Tris-HCl, pH 8.3, 8 µl of deoxynucleotide triphosphate solution (200 µM each dATP, dCTP, dGTP, and dUTP), 3 mM $MgCl_2$, 5 µl of 10x buffer, 1 µl of each primer (10 µM), and 0,1 µl of Taq DNA polymerase (Gibson BRL Life Technologies, Gaithersburg, MD, US). The upstream and downstream primer sequences for *rpoB* were CM7 (5'-AACCAGTTCCGCGTTGGCCTGG-3') and CM31b (5'-CCTGAACAACACGCTCGGA-3'), respectively.[11] The PCR products were assessed by electrophoresis in 1% agarose gel with 0.5% ethidium bromide.

Sequencing and phylogenetic analysis

Partial *rpoB* gene nucleotide sequences (1153 nt) were obtained from the PCR amplicons generated with primers CM7 and CM31b. The PCR products were sequenced in duplicates by Genome Express (Montreuil, France) with primers CM81 (5'-CAGTTCCGCGTTGGCCTG-3') and CM31b(seq) (5'-TGAACAACACGCTCGG-3').[11] Sequence alignment was performed using CLUSTAL W.[13] Phylogenetic analysis was carried out using the software MEGA version 3.0 [14] using Kimura 2-parameter model as a method of substitution and the neighbour joining method to reconstruct the phylogenetic tree. The statistical significance of the phylogenies inferred was estimated by bootstrap analysis with 1,000 pseudoreplicate data sets.

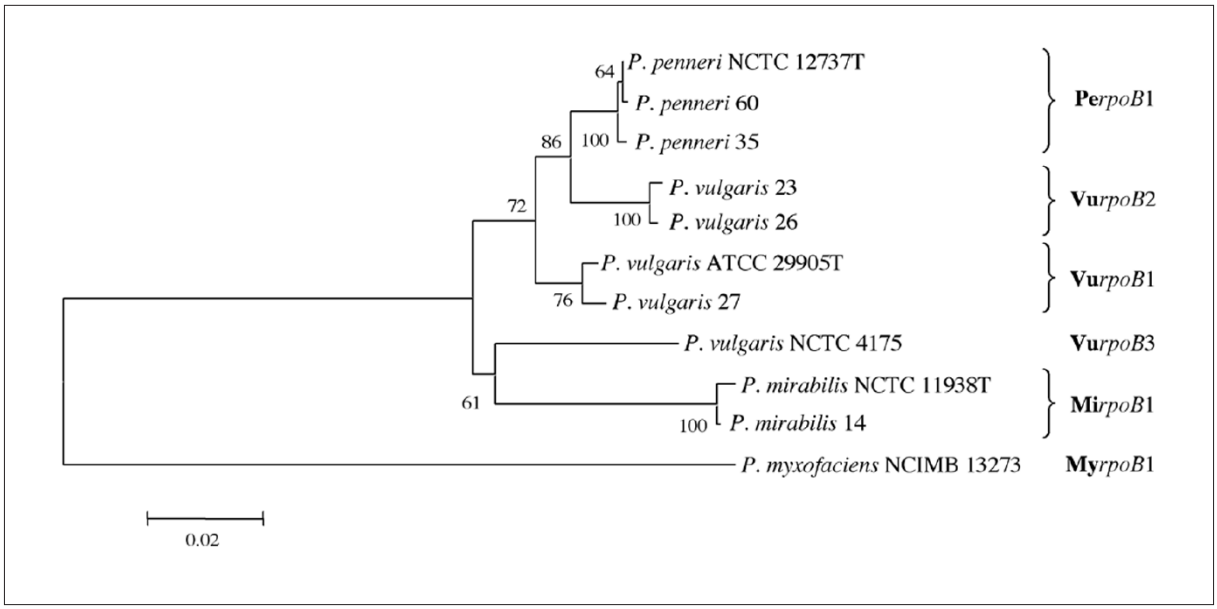
Results

We examined the *Proteus* strains by *rpoB* sequencing showed between 0.1 and 23% nt differences in the *rpoB* 1153 bp region analysed. On the basis of *rpoB* sequence divergence, the type strain of *P. myxofaciens* was clearly discriminated from the type and clinical strains of the other species. In fact, its *rpoB position* sequence diverged by 23% from those of *P. mirabilis*, by 22.2% from the former strain type of *P. vulgaris* NCTC 4175, and by 19.5 to 22% from the type and clinical strains of *P. vulgaris* and *P. penneri*. *P. mirabilis* strains showed 6.6-7.6% nucleotide differences compared to *P. vulgaris* and *P. penneri* strains. *P. penneri* differed by only 2.3-3% from all *P. vulgaris* strains but differed from the former *P. vulgaris* strain type by 5.5-5.8%. Two groups of *P. vulgaris* sequences could be described, one of them including the new strain type, differing from each other by 3.3-3.6% of their nucleotides and for both of these there was a nucleotide difference of 6.5-6.8% from the stand-alone former *P. vulgaris* strain type.

To study the phylogenetic relationships among the four species of the genus *Proteus*, a phylogenetic tree was derived from *rpoB* sequencing results. As shown in Figure 1, six *rpoB* groups supported by >60% bootstrap values could be described. *P. penneri* strains, the clinical strains and the new type strain of *P. vulgaris* clustered in three separate groups within the same branch of the tree, while the former *P. vulgaris* type strain stood apart along with the *P. mirabilis* strains. Finally, *P. myxofaciens* constituted a well defined separate branch.

The phylogenetic tree obtained from the *rpoB* sequencing data was compared with that obtained by ribotyping for the same *Proteus* strains after genomic DNA digestion by *Eco*RV or *Hinc*II, in order to evaluate the discriminative power of these two tools for the identification and subspecies typing within the genus *Proteus*.

Figure 1. Neighbor-joining phylogenetic tree based on the comparison of the *rpoB* sequences of the type strains and clinical isolates of the four species in the genus *Proteus*. Bootstrap values are given at branch nodes. Scale bar indicates percent nucleotide substitutions.



As shown in Table 1, the strain groups corresponded exactly with both molecular tools, but ribotyping was also able to differentiate ribotypes within ribogroups.[12]

Discussion

In a previous study, strain types of the four species of the genus *Proteus* showed remarkably different rRNA gene restriction patterns after digestion by *Eco*RV and *Hinc*II.[12] Furthermore, clinical isolates of *P. mirabilis* and *P. penneri* had patterns identical or very similar to those of the respective strain types, so that they could be grouped in two distinct ribogroups (groups of related patterns) corresponding to the two species, while minor differences observed among patterns of the strains of the same ribogroups were assumed to define ribotypes within each ribogroup. On the contrary, two distinct ribogroups were detected within clinical strains of *P. vulgaris*, with the ancient *P. vulgaris* type strain NCTC 4175 differing considerably from both ribogroups as well as from the newly proposed strain type ATCC 29905 which was included in the Vu*E*1-Vu*H*1 ribogroup. These results raised concerns relating to the correspondence of the current species nomenclature to the genetic evolution of the genus *Proteus*.

Sequencing of *rpo*B gene has been proposed as a powerful tool for universal bacterial genetic

Table 1. Comparison of *rpoB* groups established according to the percentage of nucleotide substitutions and phylogenetic analysis with ribogroups and ribotypes obtained in a previous study [12].

Species	Strain no.	Origin	Ribogroup and ribotype		*rpoB* group
			*Eco*RV	*Hinc*II	
P. mirabilis	NCTC 11938T	Type strain	**Mi***E*1a	**Mi***H*1b	**Mi***rpoB*1
	14	Italy 1990	**Mi***E*1a	**Mi***H*1b	**Mi***rpoB*1
P. vulgaris	ATCC 29905T	Type strain	**Vu***E*1a	**Vu***H*1a	**Vu***rpoB*1
	27	Italy 1992	**Vu***E*1a	**Vu***H*1c	**Vu***rpoB*1
	23	Italy 1991	**Vu***E*2a	**Vu***H*1d	**Vu***rpoB*2
	26	Italy 1991	**Vu**E2e	**Vu***H*1a	**Vu***rpoB*2
	NCTC 4175	Former type strain	**Vu***E*3a	**Vu***H*3a	**Vu***rpoB*3
P. penneri	NCTC 12737T	Type strain	**Pe***E*1a	**Pe***H*1a	**Pe***rpoB*1
	60 (CIP 6.88)	France	**Pe***E*1a	ND	**Pe***rpoB*1
	35	Italy	**Pe***E*1a	**Pe**H1b	**Pe***rpoB*1
P. myxofaciens	NCIMB 13273	Type strain	**My***E*1a	**My***H*1a	**My***rpoB*1

identification and has also proven to be useful for the phylogenetic analysis of named species.[11] In our study, partial *rpo*B sequencing of the strains belonging to the four species of the genus *Proteus* clearly confirmed previous ribotyping data. In fact, while the strain types and clinical isolates of *P. mirabilis*, *P. penneri* and *P. myxofaciens* grouped into three distinct *rpo*B designed groups, Mi*rpo*B, Pe*rpo*B and My*rpo*B, respectively, the new type strain and clinical isolates of *P. vulgaris* grouped into two distinct *rpo*B groups, Vu*rpo*B1 (new type strain ATCC 29905 and clinical strain 27) and Vu*rpo*B2 (clinical strains 23 and 26) respectively. The ancient *P. vulgaris* type strain NCTC 4175 was in a different branch of the phylogenetic tree named Vu*rpo*B3. Moreover, comparable genetic distances were observed between the two different clades of *P. vulgaris* strains and *P. penneri* sequences. Alternatively from what we observed with rRNA gene restriction patterns, rpoB sequencing revealed no subgroup clustering (types).

In 1982, three biogroups within *P. vulgaris* were identified on the basis of some biochemical characters and biogroup 1 was found to represent a new species named *P. penneri*.[15] The results of the present study confirm the ribogrouping of *P. penneri* and allow for it to be consider genetically close to but distinct from two clades of *P. vulgaris*. In consideration of the homogeneity of the rDNA-digested patterns of different clinical strains, it has been suggested that *P. penneri* could be the most recently differentiated species within the genus *Proteus*.[12] The *rpo*B sequencing results provide evidence for genetic relationship between *P. penneri* and two clades of *P. vulgaris*. On the contrary, the ancient *P. vulgaris* type strain (NCTC 4175) ranged in a separate branch, confirming the aberrant position already revealed by ribotyping [12] and further supporting its substitution with strain ATCC 29905 as the strain type for the species.[16,17] In fact, strain NCTC 4175 resided in DNA group 3 composed of only two strains not biochemically representative of the majority of the isolates identified as *P. vulgaris*.

In conclusion, *rpo*B sequencing proved able to characterise the different species of the genus *Proteus* on a molecular basis. The correspondence between *rpo*B groups and ribogroups confirms the validity of the presently recognised nomenspecies within the genus, but our results provide further evidence for the existence of genetic differences within *P. vulgaris*, as previously shown by DNA-DNA hybridisation.[18] When compared with ribotyping, *rpo*B sequencing methods did not perform as well for fine typing and epidemiological purposes.

References

1) Centers for Disease Control and Prevention. National nosocomial infections surveillance (NNIS) report, data summary from October 1986-April 1996, issued May 1996. A report from the National Nosocomial Infections Surveillance (NNIS) System. Am J Infect Control 1996;24(5):380-8.

2) O'Hara MC, Brenner FW, Miller JM. Classification, identification, and clinical significance of *Proteus, Providencia, and Morganella,* Clin Microbiol Rev 2000;13(4):534-46.

3) Burke JP, Ingall D, Klein JO, Gezon HM, Finland M. *Proteus mirabilis* infections in a hospital nursery traced to a human carrier. N Engl J Med 1971;284(3):115-21.

4) Krajden S, Fuksa M, Petrea C, Crisp LJ, Penner JL. Expanded clinical spectrum of infections caused by *Proteus penneri*. J Clin Microbiol 1987;25(3):578-9.

5) Bergan T. Phage typing of *Proteus.* Methods Microbiol 1978;11:243-58.

6) Hawkey PM, Penner JL, Linton AH, Hawkey CA, Crisp LJ, Hinton M. Speciation, serotyping, antimicrobial sensitivity and plasmid content of *Proteeae* from the environment of calf-rearing units in south west England. J Hyg 1986;97(3):405-17.

7) Larsson P. Serology of *Proteus mirabilis* and *Proteus vulgaris.* Methods Microbiol 1984;14:187-214.

8) Senior BW. Typing of *Proteus* strains by proticine production and sensitivity. J Med Microbiol 1977;10(1):7-17.

9) Bingen EC, Boissinot C, Desjardins P, et al. Arbitrarily primed polymerase chain reaction provides rapid differentiation of *Proteus mirabilis* isolates from a pediatric hospital. J Clin Microbiol 1993;31(5):1055-9.

10) Janda JM, Abbott SL, Khashe S, Probert W. Biochemical identification and characterization of DNA groups within *Proteus vulgaris* complex. J Clin Microbiol 2001;39(4):1231-4.

11) Mollet C, Drancourt M, Raoult D. *rpoB* sequence analysis as a novel basis for bacterial identification. Mol Microbiol 1997;26(5):1005-11.

12) Pignato S, Giammanco GM, Grimont F, Grimont PAD, Giammanco G. Molecular characterization of the genera *Proteus, Morganella,* and *Providencia* by ribotyping. J Clin Microbiol 1999;37(9):2840-7.

13) Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22(2):4673-80.

14) Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment, Brief Bioinform. 2004;5(2):150-63.

15) Hickman FW, Steigerwalt AG, Farmer JJ III, Brenner DJ. Identification of *Proteus penneri* sp. nov., formerly known as *Proteus vulgaris* indole negative or *Proteus vulgaris* biogroup I. J Clin Microbiol, 1982;15(6):1097-102.

16) Brenner DJ, Hickman-Brenner FW, Holmes B. Replacement of NCTC 4175, the current type strain of *Proteus vulgaris*, with ATCC 29905. Request for an opinion. Int J Syst Bacteriol 1995;45(4):870-1.

17) Judicial Commission of the International Committee on Systematic Bacteriology. Replacement of strain NCTC 4175, since 1963 the neotype strain of *Proteus vulgaris*, with strain ATCC 29905-opinion 70. Int J Syst Bacteriol 1999;49:1949.

18) O'Hara CM, Brenner FW, Steigerwalt AG. Classification of *Proteus vulgaris* biogroup 3 with recognition of *Proteus hauseri* sp. nov., nom. rev, and unnamed *Proteus* genomospecies 4, 5 and 6. Int J Syst Evol Microbiol 2000;50:1869-75.