Phylogenetic diversity of Klebsiella pneumoniae and Klebsiella oxytoca clinical isolates revealed by randomly amplified polymorphic DNA, gyrA and parC genes sequencing and automated ribotyping

Sylvain Brisse and Jan Verhoef

The infra-specific phylogenetic diversity and genetic structure of both Klebsiella pneumoniae and Klebsiella oxytoca was investigated using a combination of randomly amplified polymorphic DNA (RAPD) analysis, sequencing of gyrA and parC genes, and automated ribotyping. After RAPD analysis with four independent primers of 120 clinical isolates collected from 22 European hospitals in 13 countries, K. pneumoniae isolates fell into three clusters and K. oxytoca isolates fell into two clusters, while Klebsiella planticola isolates formed a sixth cluster. Each cluster was geographically widespread. K. pneumoniae cluster I (KpI) accounted for 80% of the isolates of this species and included reference strains of the three subspecies K. pneumoniae subsp. pneumoniae, K. pneumoniae subsp. ozaenae and K. pneumoniae subsp. rhinoscleromatis. Clusters KpII and KpIII were equally represented, as were the two K. oxytoca clusters. Individualization of each cluster was fully confirmed by phylogenetic analysis of gyrA and parC gene sequences. In addition, sequence data supported the evolutionary separation of K. pneumoniae from a phylogenetic group including K. oxytoca, Klebsiella terrigena, K. planticola and Klebsiella ornithinolytica. Automated ribotyping using MluI appeared suitable for identification of each Klebsiella cluster. The adonitol fermentation test was found to be useful for cluster identification in K. pneumoniae, since it was negative in all strains of clusters KpIII and in some KpII strains, but always positive in cluster KpI. The usefulness of gyrA and parC sequence data for population genetics and cluster identification in bacteria was demonstrated, even for the phylogenetic positioning of quinolone-resistant isolates.

Keywords: Klebsiella taxonomy, population genetics, Klebsiella planticola, quinolone resistance-determining region, molecular epidemiology

INTRODUCTION

Bacteria of the genus Klebsiella are increasingly important opportunistic pathogens associated with severe hospital-acquired infections such as septicemia, pneumonia and urinary tract infections. An estimated 8% of all nosocomial bacterial infections in the United States and in Europe are caused by klebsiellae, mainly by Klebsiella pneumoniae and Klebsiella oxytoca (Podschun & Ullman, 1998). The ecological habitats of Klebsiella include surface water, sewage, soils and plants, as well as mucosal surfaces of mammals. In humans, K. pneumoniae can be present in the nasopharynx and in the intestinal tract.

K. pneumoniae and K. oxytoca exhibit a high degree of genetic heterogeneity, as was demonstrated by capsular typing (Ørskov & Ørskov, 1984), O-antigen
Klebsiella planticola associated with the carriage of plasmids encoding within- and between-hospital outbreaks, often aslular antigens (Mizuta between virulence have been demonstrated in the mouse model International Journal of Systematic and Evolutionary Microbiology the genes encoding subunit A of DNA gyrase (gyrA) and subunit C of topoisomerase IV (parC), two proteins that correspond to the main targets of fluoroquinolones in Klebsiella (Deguchi et al., 1997; Brisse et al., 1999). Correspondence of ribotyping data with the identified phylogenetic clusters was investigated.

METHODS

Selection of a test collection of 120 strains for RAPD analysis. The isolates included in the study were sent from 23 European university hospitals located in 13 different countries. Between April 1997 and January 1999, 445 K. pneumoniae isolates, 238 K. oxytoca isolates and 50 K. planticola isolates were collected. The isolates studied were collected from sepsicaemia, nosocomial pneumonia infections, respiratory tract infections and skin and soft tissue infections. Only one isolate per patient was submitted.

For this study, a geographically representative set of ciprofloxacin-susceptible (MIC < 2 mg l⁻¹) strains, including 64 K. pneumoniae isolates, 26 K. oxytoca isolates and 4 K. planticola isolates, was randomly selected from various university hospitals out of the strain bank described above. The selected isolates were derived from (82), respiratory tract infections (9) and wound and soft tissue infections (3). Additionally, 22 K. pneumoniae ciprofloxacin-resistant (MIC > 2 mg l⁻¹) isolates and four K. oxytoca ciprofloxacin-resistant isolates were included in the study. This set included isolates from blood infections (15), respiratory tract infections (2), urinary tract infections (7) and wound and soft tissue infections (2). The K. pneumoniae isolates were collected in Austria (7), Belgium (1), France (15), Germany (6), Greece (11), Italy (8), The Netherlands (3), Poland (11), Portugal (3), Spain (13), Switzerland (2), United Kingdom (3) and Turkey (3). The K. oxytoca isolates were collected in Belgium (2), France (5), Germany (4), Greece (2), Italy (6), The Netherlands (1), Poland (3), Spain (5) and Switzerland (2). Finally, the K. planticola strains were isolated from Austria (2), Belgium (1) and Spain (1). In addition to these clinical isolates, three type strains and five reference strains were examined: ATCC 13883T (K. pneumoniae subsp. pneumoniae), ATCC 33531T (K. planticola), ATCC 33257T (K. terrigena), NCTC 11228 (K. pneumoniae), ATCC 11297 [K. pneumoniae subsp. ozaenae type 5 (E)], ATCC 6908 (K. pneumoniae subsp. rhinoscleromatis), NCTC 49131 (K. oxytoca) and ATCC 31898 (K. ornithinolytica).

Bacterial identification and biochemical tests. Isolates were considered as belonging to the genus Klebsiella based on a confident identification as ‘K. pneumoniae or K. oxytoca’ after testing the isolates in the VITEK apparatus (BioMérieux). K. planticola and K. terrigena are not included in the VITEK database and can be misidentified as K. oxytoca or K. pneumoniae. Determination of the species was performed manually using a combination of the indole test (positive only for K. oxytoca and some K. planticola isolates), the histamine test (positive only in K. planticola and K. terrigena) and the melezitose test (positive in K. terrigena and some K. oxytoca isolates), according to Monnet & Freney (1994). Data on adonitol fermentation, lactose fermentation, and ornithine and lysine decarboxylase activities were obtained with the VITEK apparatus.

Susceptibility testing. Ciprofloxacin (Bayer) susceptibility testing was performed using the reference broth micro-
Phylogenetic clusters in *K. pneumoniae* and *K. oxytoca*

**DNA preparation.** Bacterial DNA was prepared by a quick alkaline lysis according to Saulnier et al. (1997). Briefly, one colony (approx. 30 mg bacterial cells) was suspended in 25 µl 0.5 M NaOH. After incubation at room temperature for 30 min, 25 µl 1 M HCl and 450 µl sterile water were added and the supernatant was collected after centrifugation (5 min, 13000 r.p.m.).

**RAPD analysis.** RAPD characterizations were performed with four different primers: A10, 5′-GGGATCCGAT-3′; F4, 5′-GGGATCGGC-3′; R16, 5′-CTCTGCGCCG-3′; N9, 5′-GTGCTGCTGGT-3′. These primers had been selected, after testing 20 decameric primers (listed in Brisse et al., 2000), on a set of five *K. pneumoniae*, one *K. planticola* and three *K. oxytoca* strains, because they gave easily readable patterns (containing a limited number of well-separated bands) in all strains tested. Amplification reactions were performed in a final volume of 25 µl containing 10 mM Tris/HCl, pH 9.0, 1.5 mM MgCl₂, 50 mM KCl, 0.01% gelatin, 0.1% Triton X-100, 100 µM each dNTP, 200 nM of the chosen primer, 0.5 units *Taq* DNA polymerase (Super *Taq*, HT Biotechnology) and 2.5 µl bacterial DNA. The samples were subjected to four low-temperature hybridization cycles (94 °C, 1 min; 26 °C, 1 min; 72 °C, 1 min) followed by 40 more stringent cycles (94 °C, 45 s; 40 °C, 45 s; 72 °C, 2 min) and a final elongation step of 5 min at 72 °C. Amplification was performed in a Perkin-Elmer 9600 thermal cycler. RAPD products were analysed by electrophoresis in 1.6% agarose gels in 0.5× TBE buffer, stained with ethidium bromide and visualized by ultraviolet light. Patterns were interpreted by visual inspection for the presence or absence of bands. Reproducibility was controlled by analysing at regular time intervals, for each primer, a set of seven control strains and clustering analysis (see results section) of these controls always showed consistent results.

**PCR and sequence determination.** For PCR amplification of the gyrA and parC portions homologous to the quinolone resistance-determining region (QRDR) of *E. coli* (Yoshida et al., 1990), we designed new primers based on the alignment of the gyrA and parC *E. coli* sequences with those of other bacteria available in the public sequence databases. These new primers were designed to amplify a portion of the genes broader than those generally used to determine solely mutations in the QRDR. The following primers were used. gyrA gene: primers gyrA-A (5′-CGCGTACTATACGGCATGGATACGTA-3′) and gyrA-C (5′-ACCCCTGATCCT-GTCAGG-3′) were used for all Klebsiella species. parC gene: primers parC-A (5′-CTGCTGTGATATCAGGCTTTTA-3′) and parC-C (5′-GCCTATGGATATCAGGCTTTTA-3′) were used for K. pneumoniae groups KpI and KpII (see text) and K. terrigena; primers parC-F1 (5′-ATGGACGCTGATGGATATCAGGCTTTTA-3′) and parC-R1 (5′-CGGCGGATTACGGATATCAGGCTTTTA-3′) were used for *K. oxytoca*, *K. planticola* and *K. ornithinolytica*. Finally, a combination of primers parC-A and parC-R1 was used for isolates belonging to *K. pneumoniae* group KpIII. PCR amplifications were performed in the same reaction mix as for the RAPD analysis (see above), with 200 nM each primer. Annealing temperatures of 50 °C were used for gyrA and for all combinations of parC primers except the parC-F1 + parC-R1 combination (52 °C). The samples were subjected to 35 cycles of 1 min at 94 °C, 30 s at the annealing temperature and 30 s at 72 °C, followed by a final elongation step of 5 min at 72 °C. PCR-amplified DNA was sequenced according to the dye terminator method in both directions using the Ready Reaction Dye Terminator Cycle Sequencing Kit (Perkin-Elmer), according to the manufacturer’s recommendations. Products were resolved and analysed using an ABI 377 automated sequencer (Perkin-Elmer) following the manufacturer’s recommendations.

**Automated ribotyping.** Strains were grown overnight on Columbia agar with sheep blood (Oxoid). Automated ribotyping was performed under the conditions described by the manufacturer of the RiboPrinter microbial characterization system (Qualicon) (Sethi, 1997). Restriction enzymes that were validated with respect to their usefulness were EcoRI and *MuI*. A large DNA probe harbouring the genes for both the small and large subunit rRNAs of *E. coli* was employed. Each set of sample data was normalized, using an adjacent standard marker set, by the RiboPrinter integrated software. The normalized banding patterns were exported using GelConvert software (Qualicon). Parsimony analysis of the exported banding patterns was performed with the software GelCompar version 4.1 (Applied Maths), using the unweighted pair-group method with arithmetic averages (UPGMA) based on the Pearson correlation coefficient and using an optimization coefficient (inter-lane band position shift tolerance) of 1.2%, identical to the optimization coefficient used by the RiboPrinter system.

**Data analysis.** The Jaccard distance was used as described in Tibayrenc et al. (1993) to quantify the RAPD pattern differences among the stocks and UPGMA was used for cluster analysis using the NEIGHBOUR program of the PHYLP package, version 3.5c (Felsenstein, 1993). Molecular evolutionary relationships among gene sequences were examined by the neighbour-joining method (Saitou & Nei, 1987), based on pairwise genetic distances estimated on the basis of all substitutions using the Jukes–Cantor parameter. The significance of the branching order was evaluated by bootstrap analysis with 100 replicates. All distance-based analyses were performed using the PHYLP package version 3.5c. Parsimony analysis was performed using PAUP version 4 for Windows (Swofford, 1998). Trees were drawn with the TREEVIEW program, version 1.4 (Page, 1996).

**RESULTS**

**RAPD analysis**

The selected set of 120 test strains and the reference strains *K. pneumoniae* ATCC 13883T, *K. pneumoniae* NCTC 11228 and *K. oxytoca* NCTC 49131, were subjected to RAPD analysis with the four primers A10, F4, R16 and N9. When the patterns obtained separately with the four primers were considered together, a total of 77 bands were scored, all of them being variable on the sample (i.e. absent in at least one strain). A high degree of variability was found in the three species analysed. In *K. pneumoniae*, 70 genotypes were distinguished and all *K. oxytoca* and *K. planticola* isolates were distinct. The data were then used to compute a Jaccard dissimilarity index between the strains and the matrix of pairwise distances was used for clustering analysis using the UPGMA procedure (not shown). The *K. pneumoniae* isolates fell into three major clusters. The biggest one, named group KpI, comprised 69 clinical isolates and also included *K. pneumoniae* subsp. *pneumoniae* reference strains.
Table 1. Strains analysed and their characteristics

<table>
<thead>
<tr>
<th>Species</th>
<th>RAPD cluster*</th>
<th>Strain codes</th>
<th>Amino acid alterations in:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Gyra</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>Kpl</td>
<td>sb10, sb12, sb17, sb20, sb25, sb35, sb44, sb52, sb68, sb93, sb108, sb112, sb113, sb114</td>
<td>None</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>Kpl</td>
<td>sb115, sb28, sb319, sb412, sb413, sb4149</td>
<td>Ser-83 = Tyr</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>Kpl</td>
<td>sb118, sb26, sb29, sb138, sb148, sb151</td>
<td>Ser-83 = Phe</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>Kpl</td>
<td>sb152</td>
<td>Ser-83 = Ile</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>Kpl</td>
<td>sb153, sb162</td>
<td>Ser-83 = Phe, Asp-87 = Asn</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>Kpl</td>
<td>sb154</td>
<td>Ser-83 = Tyr, Asp-87 = Tyr</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>KplII</td>
<td>sb18, sb59</td>
<td>None</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>KplII</td>
<td>sb39, sb129</td>
<td>Ser-83 = Phe, Asp-87 = Asn</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>KplIII</td>
<td>sb1, sb31, sb55, sb69</td>
<td>None</td>
</tr>
<tr>
<td>K. pneumoniae subsp. pneumoniae</td>
<td>spI</td>
<td>ATCC 138837, NCTC 11228</td>
<td>None</td>
</tr>
<tr>
<td>K. pneumoniae subsp. ozaenae type 5 (E)</td>
<td>spII</td>
<td>ATCC 11297</td>
<td>None</td>
</tr>
<tr>
<td>K. pneumoniae subsp. rhinoscleromatis</td>
<td>spII</td>
<td>ATCC 6908</td>
<td>None</td>
</tr>
<tr>
<td>K. oxytoca</td>
<td>Kpl</td>
<td>sb9, sb18</td>
<td>None</td>
</tr>
<tr>
<td>K. oxytoca</td>
<td>Kol</td>
<td>sb157</td>
<td>Thr-83 = Ile</td>
</tr>
<tr>
<td>K. oxytoca</td>
<td>KolI</td>
<td>NCTC 49131</td>
<td>None</td>
</tr>
<tr>
<td>K. oxytoca</td>
<td>KolII</td>
<td>sb159</td>
<td>Thr-83 = Ile</td>
</tr>
<tr>
<td>K. oxytoca</td>
<td>KolII</td>
<td>sb131</td>
<td>Thr-83 = Ile</td>
</tr>
<tr>
<td>K. oxytoca</td>
<td>Kol</td>
<td>sb158</td>
<td>Thr-83 = Ile, Asp-87 = Gly</td>
</tr>
<tr>
<td>K. planticola</td>
<td>KpI</td>
<td>sb7, sb161</td>
<td>None</td>
</tr>
<tr>
<td>K. planticola</td>
<td>KpI</td>
<td>ATCC 33531†</td>
<td>None</td>
</tr>
<tr>
<td>K. amnithelvetica</td>
<td>spI</td>
<td>ATCC 31998</td>
<td>None</td>
</tr>
<tr>
<td>K. terrigena</td>
<td>spII</td>
<td>ATCC 33257†</td>
<td>None</td>
</tr>
</tbody>
</table>

* Kplan: K. planticola RAPD cluster; ND, not determined.

ATCC 138837 and NCTC 11228. Clusters KpI and KpII each comprised eight strains. The K. oxytoca isolates fell into two major clusters, group KoI with 17 isolates and group KoII with 11 clinical isolates and strain NCTC 49131. Finally, the four K. planticola isolates fell into a single branch. Only one K. pneumoniae strain and two K. oxytoca strains did not fall in any of these clusters, but appeared loosely associated in a single, separate branch.

None of the six clusters was restricted to a particular geographic area and no correlation was found between the cluster and the clinical origin of the sample (not shown).

Molecular phylogeny obtained with the gyrA gene

For molecular phylogeny analyses, a genetically diverse collection representative of all RAPD groups was selected (Table 1). These strains included 23 ciprofloxacin-resistant strains belonging to KpI, KpII, KoI and KoII clusters. In addition, the eight reference strains were included in the study collection (Table 1).

The sequence of a portion of 383 nt, encompassing the QRDR of the gyrA gene, was determined in the 56 selected strains. A total of 73 nucleotide sites were polymorphic (defined as different in at least one strain) among the Klebsiella strains.

The neighbouring-join tree constructed on the basis of the gyrA sequences is depicted in Fig. 1 (right). The tree revealed a strong clustering that was in agreement with the RAPD grouping: the three K. pneumoniae groups, the two K. oxytoca groups and K. planticola were each clustered separately. Additionally, the bootstrap analysis revealed that each of these six clusters was strongly supported, with the exception of cluster KpI (Fig. 1). Similar results were obtained using maximum-parsimony (not shown).

The amounts of nucleotide differences within and between the different clusters are given in Table 2. The mean sequence divergence among the three Klebsiella pneumoniae groups was at least 5-60 ± 0-89 nt (KpI–KpIII comparison), whereas the mean sequence divergence within the groups was never more than 2-0 ± 1 nt (for KpIII). The quantitative measure of the distinctness of the clusters’ gene sequences (parameter k in Palys et al., 1997) was above 3-0 ± 0-92 (KpI–KpII comparison) for all three comparisons. Similarly, the two K. oxytoca clusters were well distinguished by gyrA sequence variability (Table 2).

The correspondence observed between RAPD grouping and the gyrA sequence was also true for the quinolone-resistant strains, since these isolates fell into their corresponding RAPD group (Table 1 and Fig. 1). In the ciprofloxacin-resistant isolates, convergent non-synonymous mutations in the gyrA gene were observed at codon 83 and, additionally, deduced alterations of the amino acid at position 87 of GyrA were observed in some isolates (Table 1). However, when the phylogenetic analysis was performed after removal of the nucleotide positions corresponding to these two codons, no difference was observed in terms of clustering into the six clusters. Thus, the selection of quinolone resistance-conferring alterations in GyrA does not affect phylogenetic positioning in the clusters.
**Table 2.** Sequence divergence (given as the number of nucleotide differences) within and between the sequence clusters identified in *K. pneumoniae* and *K. oxytoca*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>gyrA</th>
<th>parC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean divergence within group†</td>
<td>Mean divergence between groups†</td>
</tr>
<tr>
<td>KpI vs KpII</td>
<td>1.70 ± 0.74</td>
<td>2.0 ± 1.0</td>
</tr>
<tr>
<td>KpI vs KpIII</td>
<td>2.0 ± 0.10</td>
<td>5.96 ± 0.89</td>
</tr>
<tr>
<td>KpII vs KpIII</td>
<td>2.10 ± 1.10</td>
<td>9.50 ± 3.55</td>
</tr>
<tr>
<td>KpI vs KpIII</td>
<td>2.0 ± 0.10</td>
<td>5.62 ± 0.92</td>
</tr>
</tbody>
</table>

*Mean ± SE for pairwise divergence within each of the groups shown in order of appearance in the comparison column.
†Mean ± SE for pairwise divergence between the groups, based on all pairwise comparisons of strains from different groups.
‡K parameter: ratio of the between-group divergence to the mean of the within-group divergence levels. A ratio above 2 indicates that the groups can be considered as separate sequence similarity clusters (Palys et al., 1997).

**Molecular phylogeny obtained with the parC gene**

The sequence of a portion of 319 nt encompassing the QRDR of the *parC* gene was determined in the same set of 56 strains. A total of 73 polymorphic sites were observed. The neighbour-joining tree built on the basis of the *parC* sequences is depicted in Fig. 1 (left). The *parC* tree was in strong agreement with the RAPD and...
**Fig. 2.** Overview of the *Mlu*I riboprints obtained for the 41 *Klebsiella* strains analysed, showing correspondence with the RAPD and sequence clusters. *K. planticola* ATCC 33531² was included twice and illustrates the reproducibility of the automated characterization. * indicates that strain ATCC 11297 is not included in group KpIII. For cluster analysis, the UPGMA method was used based on the matrix of Pearson coefficient correlation. In the dendrogram scale, correlation levels were converted to percentage similarity levels. All sequence clusters were clearly distinguished and individualized, with one exception represented by strain ATCC 6908 (*K. pneumoniae* subsp. *rhinoscleromatis*), which belongs to cluster KpI based on *gyrA* and *parC* sequence analyses (Fig. 1; see text for discussion). The arrow indicates the DNA fragment of approximately 3.8 kb that was specific for all strains of *K. pneumoniae* cluster KpI. Additionally, a DNA fragment of approximately 1.0 kb was specific for *K. pneumoniae*.

*gyrA* clustering. In addition, a sharper delineation of the three *K. pneumoniae* groups and the two *K. oxytoca* groups was observed, as compared with *gyrA* tree. This resulted from the fact that the sequence divergence between the three *K. pneumoniae* clusters and between the two *K. oxytoca* clusters, was higher than that observed with the *gyrA* gene (Table 2), indicating a higher rate of nucleotide substitution in *parC*. In contrast to the *gyrA* results, even group Kpl had strong bootstrap support (99%) in the *parC* analysis. The parameter of distinctness also indicated a clear separation, on the basis of *parC* sequences, of the three *K. pneumoniae* groups and of the two *K. oxytoca* groups (Table 2). As for *gyrA*, removal of the *parC* codons associated with quinolone resistance had no impact on the phylogenetic clustering.

**Phylogenetic relationships among Klebsiella species and subspecies**

With respect to the relatedness among the groups, subspecies and species, the *gyrA* and *parC* trees yielded a number of observations. Both genes of the reference strains of *K. pneumoniae* subsp. *ozaenae* and *K. pneumoniae* subsp. *rhinoscleromatis* showed sequences typical of group Kpl (Fig. 1). When the branching
order of the *K. pneumoniae* groups was considered. KpI and KpII were branched together and KpIII was in an external position in both trees. However, the bootstrap support of the KpI–KpII association was not strong (Fig. 1).

*K. pneumonia* and *K. oxytoca* each appeared monophyletic with both genes. In contrast, *K. planticola* and *K. ornithinolytica* were not separated.

*K. planticola* appeared more closely related to *K. oxytoca* than to *K. pneumoniae*. This was mainly supported by the *gyrA* sequences: *K. planticola* ATCC 33531<sup>T</sup> *gyrA* sequence showed 35 nt differences compared to *K. oxytoca* NCTC 49131 and 44 nt to *K. pneumoniae* ATCC 13883<sup>T</sup>. *K. terrigena* appeared closely related to *K. planticola* on the basis of *gyrA* sequence comparisons (Fig. 1, right), but this association was not supported by the *parC* sequence comparison (Fig. 1, left).

**Identification of *K. pneumoniae* and *K. oxytoca* groups by automated ribotyping and biochemical tests**

To initiate attempts to identify the discrete sequence clusters found within *K. pneumoniae* and *K. oxytoca*, we explored the correspondence of *gyrA* and *parC* clustering with ribotyping. We analysed by automated ribotyping 89 strains using EcoRI. In each group, almost every strain was distinguished as a distinct ribotype. However, clustering analysis of these patterns did not group the strains according to their sequence cluster (data not shown). Subsequently, we analysed 41 strains representative of each cluster using MluI. Interestingly, there was good correspondence between ribotyping results obtained and the clusters defined on the basis of RAPD, *gyrA* and *parC* sequencing (Fig. 2). All strains belonging to KpI fell into a single branch, with the two following important exceptions: strain ATCC 6908, the reference strain of *K. pneumoniae* subsp. *rhinoscleromatis*, was grouped with strains of group KpII; and strain ATCC 11297, the reference strain of *K. pneumoniae* subsp. *ozaenae*, did not appear closer to strains of KpI than to strains of group KpIII. The profile of strain ATCC 11297 lacked the band of approximately 9 kb that is shared among all strains of KpI and KpIII, which may explain its position out of cluster KpI in the dendrogram. The ribotype patterns in groups KpI and KpIII were very similar, the main difference between them being a fragment of 3·8 kb, which was conserved among all strains of KpI, including strain ATCC 11297, and was not observed in any strain of group KpIII (Fig. 2).

Each group of *K. oxytoca* formed a well delineated branch and so did the strains belonging to the *K. planticola*-*K. ornithinolytica* sequence cluster (Fig. 2). However, the two latter taxa could not be distinguished by their ribotype. Indeed, the *K. planticola* clinical isolates sb67 and sb161, which were ornithine-decarboxylase-negative, exhibited patterns that were very similar to that of *K. ornithinolytica* ATCC 31898.

We also investigated whether the ability to ferment adonitol, which is variable within *K. pneumoniae* (Farmer et al., 1985), showed distinctive distributions among the sequence clusters. The adonitol test was positive in all tested strains of KpI (65), and in all strains of *K. oxytoca* (28) and *K. planticola* (4). In contrast, it was always negative in strains of KpIII (8) and it was negative in three out of eight KpII strains.

**DISCUSSION**

To investigate the genetic structure of *K. pneumoniae* and *K. oxytoca*, the two *Klebsiella* species of major clinical importance, we combined RAPD analysis with sequencing of protein-coding genes. We found complete agreement between the grouping as defined by RAPD analysis and the sequence clustering based on both *gyrA* and *parC* sequencing. Agreement of RAPD clustering with multilocus enzyme electrophoresis data has been reported (Tibayrenc et al., 1993; Desjardins et al., 1995). Therefore, although RAPD analysis has limitations in terms of reproducibility and lack of portability (van Belkum et al., 1995), it represents a suitable first screening approach for identification of distinct genetic lineages. Another approach to classify bacterial diversity has stemmed from the observation that bacterial strains, within given species or DNA groups, fall into clusters of closely related organisms based on sequence similarity of protein-coding genes (Ambler, 1996; Palys et al., 1997). Sequencing is more time-consuming and expensive than RAPD, but yields data that are easily portable and provides information on the evolution of well defined genes (Maiden et al., 1998).

As a consequence of its low evolutionary rate, 16S rRNA sequencing has failed to distinguish many closely related but ecologically distinct groups of bacteria (Palys et al., 1997). Due to their higher rate of neutral mutation, protein-coding genes are often more suitable for relatedness determinations at or below the species level. Our interest in the chromosomal genes *gyrA* and *parC* as phylogenetic tools stemmed from the fact that they are ubiquitous among bacterial strains and species, and that a considerable amount of sequence data are currently being gathered, in many bacterial species, to determine mutations conferring resistance to quinolones (Hooper, 1995). Our results show that sequencing data of *gyrA* and *parC* genes are also useful for phylogenetic delineation of genetic groups in *K. pneumoniae* and *K. oxytoca*, even when quinolone-resistant strains are considered. The congruence between the clustering based on both genes and the agreement with RAPD clustering indicate that *gyrA* and *parC* genes are not prone to frequent horizontal transfer among the clusters and therefore represent reliable phylogenetic markers. Whether these two genes are also useful for phylogenetic purposes in other bacterial species deserves to be explored.
This study indicates the existence of discrete sequence clusters in the species *K. pneumoniae* and *K. oxytoca*. This was supported by agreement between RAPD, gyrA and parC phylogenetic clustering analyses and by their correspondence with ribotyping. Palys et al. (1997) considered a distinctness parameter k value greater than or equal to 2 as a criterion for the delineation of sequence clusters. The *K. pneumoniae* and *K. oxytoca* clusters identified herein clearly fulfill this criterion (Table 2). The reference strain for *K. pneumoniae* subsp. *ozaenae* included in this study belongs to the *K. pneumoniae* subsp. *ozaenae* subsp. *pneumoniae* group. Biochemical characterization revealed a profile typical of *K. pneumoniae* subsp. *pneumoniae* strains, but not in the two other taxa (*Farmer et al., 1985*). The adonitol fermentation test could provide a means to identify strains of groups KpII and KpIII, but this needs confirmation on a larger number of strains.

Whether other strains of *K. pneumoniae* subsp. *ozaenae* and *K. pneumoniae* subsp. *rhinoscleromatis* also fall within group KpI remains to be investigated. Ferragut et al. (1989) also observed, based on protein profile analysis, the clustering of strains belonging to *K. pneumoniae* subsp. *pneumoniae*, *K. pneumoniae* subsp. *ozaenae* and *K. pneumoniae* subsp. *rhinoscleromatis* into several electrophoretic groups and the three subspecies appeared scattered across these groups.

In *K. oxytoca*, investigation of the chromosomal gene encoding β-lactamase revealed the existence of two variants of this gene among *K. oxytoca* strains (Fournier et al., 1996). Preliminary experiments suggest that the dimorphism of β-lactamase genes corresponds to the two *K. oxytoca* sequence clusters (S. Brisse & T. van Himbergen, unpublished).

Automated ribotyping using *Mlu I* appears to be a promising tool for identification of the seven *Klebsiella* sequence clusters, as strains of different clusters were always distinguished and as ribotype patterns were grouped in agreement with the sequence clusters. The only result that appeared at odds with the sequence clustering was the association of strain ATCC 6908 (*K. pneumoniae* subsp. *rhinoscleromatis*) with group KpII, although this strain exhibited a unique pattern. A probable explanation for this discrepancy is that ribotype clustering is not fully reliable, since clustering relies on a limited number of DNA fragments and since the lack or change of position of a single band could dramatically affect the position of a strain. Nevertheless, it is interesting that *K. pneumoniae* subsp. *rhinoscleromatis* and *K. pneumoniae* subsp. *ozaenae* reference strains could be distinguished from all other KpI strains by ribotyping with *Mlu I*, as this was not the case with sequencing data (Fig. 1). Ribotyping with *EcoRI* did not appear suitable for cluster identification, but its higher discriminatory power is useful to subdivide *Mlu I* ribogroups.

Knowledge of the phylogenetic relationships between species and subspecies within the genus *Klebsiella* provides a framework for studies on the distribution of phenotypic properties implicated in virulence or epidemiological differences between clones. In *E. coli*, it has been shown that factors associated with pathogenicity are unequally distributed among the genetic clusters (Boyd & Hartl, 1998). It will be interesting to investigate whether the groups identified in *K. pneumoniae* and *K. oxytoca* are characterized by ecological or clinically relevant properties, as predicted...
by theoretical models of bacterial evolution (Palys et al., 1997). In addition, DNA–DNA hybridization studies will be useful to determine whether the *K. pneumoniae* and *K. oxytoca* sequence clusters correspond to distinct genomic species.

**ACKNOWLEDGEMENTS**

We are grateful to B. Aliyary Ghraboghy for technical help in sequencing. D. Monnet and O. Bellon are thanked for advice on assimilation tests for species identification. We are grateful to Professor P. A. D. Grimont and Professor F. R. Mooi for critical reading of the manuscript. Sylvain Brisse was supported by a European Union Training and Mobility of Researchers grant.

**REFERENCES**


