Multilocus sequence analysis and 16S rRNA gene sequencing reveal that *Yersinia frederiksenii* genospecies 2 is *Yersinia massiliensis*

Roberto A. Souza, Priscilla F. M. Imori and Juliana P. Falcão

Brazilian Reference Center on *Yersinia* spp. other than *Y. pestis*, Departamento de Análises Clínicas, Toxicológicas e Bromatológicas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto – USP, Ribeirão Preto, SP, Brazil

Since *Yersinia frederiksenii* was first described in 1980, it has been recognized genotypically as a heterogeneous species, comprising three phenotypically indistinguishable genospecies. In this study, the sequence of the 16S rRNA gene and the concatenated sequences of six housekeeping genes (*glnA*, *gyrB*, *hsp60*, *recA*, *rpoB* and *sodA*) of all the currently known species of the genus *Yersinia* were used to determine the phylogenetic position of *Y. frederiksenii* genospecies 2 in the genus *Yersinia*. The phylogenetic analyses grouped the *Y. frederiksenii* genospecies 2 strains in a monophyletic group together with representative strains of *Yersinia massiliensis*. Moreover, the *Y. frederiksenii* genospecies 2 strains were also grouped apart from the other species of the genus *Yersinia* and far from the other two genospecies of *Y. frederiksenii*. All of the observations made in this study support the conclusion that *Y. frederiksenii* genospecies 2 should be reclassified as *Y. massiliensis*.

Since its description 33 years ago by Ursing *et al.* (1980), *Yersinia frederiksenii* has been recognized genotypically as a very heterogeneous species and composed of three phenotypically indistinguishable genospecies, as defined by DNA–DNA hybridization (DDH). Moreover, Dolina & Peduzzi (1993) concluded that strains classified as *Y. frederiksenii* belong to more than one species of the genus *Yersinia*, which could not be differentiated because of the inadequacy of the available methods used for identification at the time when this species was proposed.

Owing to the difficulty of separating the three genomic groups using biochemical tests and to the small number of strains studied by the time the classification was announced, it was decided that the three genospecies should be maintained as *Y. frederiksenii* (Demarta *et al.*, 2004; Ursing & Aleksić, 1995).

According to Ursing & Aleksić (1995), it is possible for a bacterial species to contain two or more genomic groups (genospecies) that are phenotypically similar if they are more related to each other than to another species. However, a genospecies can be defined as a separate species when it has phenotypic traits that distinguish it from its nearest genomic neighbour species or when more reliable tests become available (Wayne *et al.*, 1987).

In some cases, biochemical profiles are not sufficient to discriminate between closely related species, and atypical biochemical reactions may occur, rendering a bacterial identification based on biochemical characteristics difficult or leading to an erroneous identification (Delmas *et al.*, 2006; Iteman *et al.*, 1996).

For the reasons reported above, investigations of relatedness based on DNA can be considered more appropriate tools to achieve a complete and correct identification of bacterial strains.

DDH is considered the ‘gold standard’ method for species delimitation; however, this method is laborious and expensive (Stackebrandt & Ebers, 2006). In this way, several reports have demonstrated the usefulness of gene sequencing for the accurate identification of bacterial species, without the need to perform DDH experiments (Labeleda, 2011; Ramos *et al.*, 2011; Thompson *et al.*, 2005; Young *et al.*, 2008, 2010; Zeigler, 2003). Furthermore, unlike sequences, which should be deposited in public databases for inspection of quality, the DDH values supplied cannot be judged by any reviewer (Stackebrandt & Ebers, 2006).

The analysis of the 16S rRNA gene sequence has proven to be a powerful tool to trace the phylogenetic relationships within various bacterial genera and can be considered a valuable tool for assigning bacterial strains to a species (Dauga, 2002; Ibrahim *et al.*, 1993, 1997; Mignard & Flandrois, 2006; Souza *et al.*, 2010, 2011; Stackebrandt & Goebel, 1994). However, in some cases, the 16S rRNA gene resolution may be insufficient for discrimination at the

**Abbreviations:** DDH, DNA–DNA hybridization; MLSA, multilocus sequence analysis; VP, Voges–Proskauer.
intrageneric level (Demarta et al., 2004; Fox et al., 1992; Martens et al., 2008; Stackebrandt & Goebel, 1994; Woo et al., 2008). In these cases, the use of five or more housekeeping genes has been reported as more appropriate for phylogenetic analysis of closely related bacteria (Stackebrandt et al., 2002).

Specifically, multilocus sequence analysis (MLSA) appears to be an effective alternative methodology to DDH and offers a powerful tool to assign a novel strain as a member of a known species or to indicate its status as a member of novel species. Furthermore, MLSA makes possible the observation of phylogenetic traits among bacterial species, and its accuracy is considered equivalent to that of DDH. In this way, the use of MLSA as a tool for phylogenetic studies into bacterial systematics has been strongly supported and it can replace DDH in phylogenetic studies (Gevers et al., 2005; Zeigler, 2003).

In our study, we performed an *in silico* analysis using the full 16S rRNA gene sequence and the partial sequences of glnA, gyrB, hsp60, recA, rpoB and sodA genes available in the GenBank/EMBL/DBJ databases to elucidate the phylogenetic position of *Y. frederiksenii* genospecies 2 in the genus *Yersinia*, taking into account all the currently described species of the genus *Yersinia*. These six housekeeping genes were selected because they are known to resolve relationships among the yersiniae (Kotetishvili et al., 2005; Merhej et al., 2008; Souza et al., 2010, 2011).

The sequences of one representative strain of each of the 17 species of the genus *Yersinia* described to date, four *Y. massiliensis* strains described by Souza et al. (2011), the representative strains of *Y. frederiksenii* genospecies 2 and 3 plus nine strains of species that have in their 16S rRNA gene sequence the distinctive signature characteristic of *Y. frederiksenii* genospecies 2 found in the stem–loop of the V3 region, from position 455 through 477 (5’-CATTGGTTAATAACGCCAGTG-3’) of the *Escherichia coli* gene (IUB nomenclature) (Ibrahim et al., 1997) were used. This sequence of highest diagnostic value for the specific identification of *Y. frederiksenii* genospecies 2 was also found in the 16S rRNA gene sequence of all strains of *Y. massiliensis* studied here. The only difference found in the *Y. massiliensis* strains FCF 465, FCF 457 and FCF 216 of this study was an adenine (A) instead a guanine (G) at position 472.

Table 1 shows all of the strains of members of the genus *Yersinia* used in this study and the accession numbers of each locus in GenBank/EMBL/DDJB databases.

Phylogenetic trees were computed from the sequences obtained for each strain for the 16S rRNA gene (1450 bp) and also from a ‘superlocus’ consisting of concatenated glnA, gyrB, hsp60, recA, rpoB and sodA sequences (the ‘superlocus’ has a length of 2610 bp) obtained by MLSA assay. The clustering was performed with the software package BioNumerics 7.0 (AppliedMaths), using the UPGMA technique. The distance estimations were obtained by the Jukes–Cantor evolutionary distance method and the robustness of each cluster was checked using 1000 bootstrap replications.

UPGMA trees of members of the genus *Yersinia* based on 16S rRNA and MLSA demonstrated that the strains belonging to *Y. frederiksenii* genospecies 2 formed a monophyletic group together with representative strains of *Y. massiliensis* with 99.7% and 97.6% similarity in the 16S rRNA and MLSA assays, respectively, and far from the other two *Y. frederiksenii* genospecies (Figs 1 and 2). In agreement with other results, three independent *Y. frederiksenii* groups were recognized, each of them including a reference strain of genospecies 1 (ATCC 33641), 2 (ATCC 33644) and 3 (ATCC 29912) in both assays. The robustness of these clusters can be confirmed by the high bootstrapping values (Figs 1 and 2).

Stackebrandt & Ebers (2006) recommended that a 16S rRNA gene sequence similarity threshold ranging from 98.7 to 99% be the point at which DDH should be mandatory for testing the genomic uniqueness of novel strains. In our study, the cluster composed of strains of *Y. massiliensis* species and strains of *Y. frederiksenii* genospecies 2 in the 16S rRNA gene sequencing (Fig. 1) has 98.5% similarity with the cluster in which the representative strains of *Y. frederiksenii* genospecies 1 (ATCC 33641) and 3 (ATCC 29912) are located. Because the threshold value recommended by Stackebrandt & Ebers (2006) was very close to the similarity value found by us in the 16S rRNA gene sequencing, the use of DDH would be recommended to make clear the phylogenetic position of *Y. frederiksenii* genospecies 2 in the genus *Yersinia*. However, for the reasons reported above, we decided to perform the MLSA assay as a substitute for the DDH approach.

According to Christensen et al. (2007) two phenotypic characteristics must also be identified in order to recognize separated species. According to Souza et al. (2011), the species *Y. frederiksenii* differs phenotypically from *Y. massiliensis* in the Voges–Proskauer (VP), mucate and rhamnose tests. Interestingly, the nine strains of the genus *Yersinia* molecularly assigned to genospecies 2 of *Y. frederiksenii* studied were rhamnose-positive like *Y. frederiksenii*, but VP-negative and mucate-positive like *Y. massiliensis*. In this way, those nine strains are phenotypically similar to *Y. massiliensis*.

It is important to mention that at the time that *Y. frederiksenii* was described and phenotypically characterized (Ursing et al., 1980) it was not possible to perform a comparison with the type strain of *Y. massiliensis* since this species was described 28 years later (Merhej et al., 2008). For this reason together with the closer phenotypic similarity of *Y. frederiksenii* genospecies 2 to *Y. frederiksenii* genospecies 1 and 3 the representative strains of genospecies 2 were included in *Y. frederiksenii*.

Altogether, the phylogenetic analyses based on 16S rRNA and MLSA assays plus some phenotypic characteristics
Table 1. Strains of species of the genus Yersinia used in this study and the accession numbers of each locus in GenBank/EMBL/DDBJ databases

<table>
<thead>
<tr>
<th>Strain</th>
<th>16S rRNA</th>
<th>glnA</th>
<th>gyrB</th>
<th>hsp60</th>
<th>recA</th>
<th>sodA</th>
<th>rpoB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y. aldovae ATCC 35236T</td>
<td>AF366376</td>
<td>FJ518731</td>
<td>FJ518743</td>
<td>FJ518755</td>
<td>FJ518767</td>
<td>EF175607</td>
<td>EF175601</td>
</tr>
<tr>
<td>Y. aleksiciae CCUG 52872T</td>
<td>FJ717341</td>
<td>FJ717355</td>
<td>FJ717360</td>
<td>FJ717369</td>
<td>FJ717391</td>
<td>KC776738</td>
<td>KC776745</td>
</tr>
<tr>
<td>Y. bercovieri ATCC 43970T</td>
<td>FJ518720</td>
<td>FJ518732</td>
<td>FJ518744</td>
<td>FJ518756</td>
<td>FJ518768</td>
<td>EF175608</td>
<td>EF175606</td>
</tr>
<tr>
<td>Y. enterocolitica subsp. enterocolitica ATCC 9610T</td>
<td>KC788290</td>
<td>KC788312</td>
<td>KC788334</td>
<td>KC788356</td>
<td>KC788378</td>
<td>EF175609</td>
<td>EF175592</td>
</tr>
<tr>
<td>Y. enterocolitica subsp. enterocolitica ATCC 9610T</td>
<td>KC788290</td>
<td>KC788312</td>
<td>KC788334</td>
<td>KC788356</td>
<td>KC788378</td>
<td>EF175609</td>
<td>EF175592</td>
</tr>
<tr>
<td>Y. entomophaga CCUG 59867T</td>
<td>DQ400782</td>
<td>DQ400780</td>
<td>DQ400781</td>
<td>GU327740</td>
<td>GU327738</td>
<td>KC814163</td>
<td>GU327738</td>
</tr>
<tr>
<td>Y. frederiksenii ATCC 33641T</td>
<td>AF366379</td>
<td>AY333029</td>
<td>AY332923</td>
<td>AY332869</td>
<td>AY332975</td>
<td>EF175613</td>
<td>EF175605</td>
</tr>
<tr>
<td>Y. frederiksenii ATCC 33644</td>
<td>JX419426</td>
<td>JX419426</td>
<td>AJ639639</td>
<td>JX419427</td>
<td>FX419428</td>
<td>KC776741</td>
<td>KC776748</td>
</tr>
<tr>
<td>Y. intermedia ATCC 29909T</td>
<td>AF366380</td>
<td>AY333036</td>
<td>AY332930</td>
<td>AY332876</td>
<td>AY332981</td>
<td>EF175614</td>
<td>EF175603</td>
</tr>
<tr>
<td>Y. kristensenii ATCC 33638T</td>
<td>AF366381</td>
<td>AY333044</td>
<td>AY332937</td>
<td>AY332883</td>
<td>AY332989</td>
<td>EF175615</td>
<td>EF175596</td>
</tr>
<tr>
<td>Y. massiliensis CCUG 53443T</td>
<td>EF179119</td>
<td>FJ717356</td>
<td>EF175588</td>
<td>EF175174</td>
<td>FJ717390</td>
<td>EF175618</td>
<td>EF175599</td>
</tr>
<tr>
<td>Y. mollaretii ATCC 43969T</td>
<td>AF366381</td>
<td>AY333050</td>
<td>AY332943</td>
<td>AY332889</td>
<td>AY332995</td>
<td>EF175617</td>
<td>EF175604</td>
</tr>
<tr>
<td>Y. nurmii CCUG 60891T</td>
<td>AM182404</td>
<td>FJ717350</td>
<td>AY332930</td>
<td>AY332869</td>
<td>AY332975</td>
<td>EF175619</td>
<td>EF175602</td>
</tr>
<tr>
<td>Y. pekkanenii CCUG 60991T</td>
<td>JX419420</td>
<td>JX419421</td>
<td>JX419422</td>
<td>JX419423</td>
<td>JX419424</td>
<td>KC776740</td>
<td>KC776747</td>
</tr>
<tr>
<td>Y. pseudotuberculosis ATCC 29833T</td>
<td>JX419425</td>
<td>JX419426</td>
<td>AJ639639</td>
<td>JX419427</td>
<td>FX419428</td>
<td>KC776741</td>
<td>KC776748</td>
</tr>
<tr>
<td>Y. rohdei ATCC 43380T</td>
<td>AF366384</td>
<td>KC788329</td>
<td>KC788351</td>
<td>KC788373</td>
<td>KC788395</td>
<td>EF175619</td>
<td>EF175602</td>
</tr>
<tr>
<td>Y. ruckeri ATCC 29473T</td>
<td>AF366385</td>
<td>FJ518730</td>
<td>FJ518742</td>
<td>FJ518754</td>
<td>FJ518766</td>
<td>EF175620</td>
<td>EF175598</td>
</tr>
<tr>
<td>Y. similis CCUG 52882T</td>
<td>AM182404</td>
<td>FJ717350</td>
<td>FJ717362</td>
<td>FJ717371</td>
<td>FJ717390</td>
<td>KC776743</td>
<td>KC776751</td>
</tr>
<tr>
<td>Y. massiliensis FCF 216</td>
<td>JX419420</td>
<td>HM003824</td>
<td>GQ421620</td>
<td>GQ421624</td>
<td>GQ421626</td>
<td>HM003828</td>
<td>GQ421632</td>
</tr>
<tr>
<td>Y. massiliensis FCF 457</td>
<td>JX419421</td>
<td>HM003824</td>
<td>GQ421620</td>
<td>GQ421624</td>
<td>GQ421626</td>
<td>HM003828</td>
<td>GQ421632</td>
</tr>
<tr>
<td>Y. massiliensis FCF 465</td>
<td>JX419422</td>
<td>HM003825</td>
<td>GQ421621</td>
<td>GQ421625</td>
<td>GQ421627</td>
<td>HM003829</td>
<td>GQ421633</td>
</tr>
<tr>
<td>Y. massiliensis FCF 494</td>
<td>JX419423</td>
<td>HM003825</td>
<td>GQ421621</td>
<td>GQ421625</td>
<td>GQ421627</td>
<td>HM003831</td>
<td>GQ421635</td>
</tr>
<tr>
<td>Y. massiliensis FCF 1</td>
<td>KC776756</td>
<td>KC776757</td>
<td>KC776758</td>
<td>KC776759</td>
<td>KC776800</td>
<td>KC776822</td>
<td>KC776844</td>
</tr>
<tr>
<td>Y. massiliensis FCF 214</td>
<td>KC776757</td>
<td>KC776758</td>
<td>KC776759</td>
<td>KC776760</td>
<td>KC776761</td>
<td>KC776762</td>
<td>KC776763</td>
</tr>
<tr>
<td>Y. massiliensis FCF 276</td>
<td>KC776760</td>
<td>KC776761</td>
<td>KC776762</td>
<td>KC776763</td>
<td>KC776764</td>
<td>KC776765</td>
<td>KC776766</td>
</tr>
<tr>
<td>Y. massiliensis FCF 341</td>
<td>KC776753</td>
<td>KC776754</td>
<td>KC776755</td>
<td>KC776756</td>
<td>KC776757</td>
<td>KC776758</td>
<td>KC776759</td>
</tr>
<tr>
<td>Y. massiliensis FCF 400</td>
<td>KC776755</td>
<td>KC776756</td>
<td>KC776757</td>
<td>KC776758</td>
<td>KC776759</td>
<td>KC776760</td>
<td>KC776761</td>
</tr>
<tr>
<td>Y. massiliensis FCF 492</td>
<td>KC776757</td>
<td>KC776758</td>
<td>KC776759</td>
<td>KC776760</td>
<td>KC776761</td>
<td>KC776762</td>
<td>KC776763</td>
</tr>
<tr>
<td>Y. massiliensis FCF 512</td>
<td>KC776762</td>
<td>KC776763</td>
<td>KC776764</td>
<td>KC776765</td>
<td>KC776766</td>
<td>KC776767</td>
<td>KC776768</td>
</tr>
<tr>
<td>Y. massiliensis FCF 577</td>
<td>KC776759</td>
<td>KC776760</td>
<td>KC776761</td>
<td>KC776762</td>
<td>KC776763</td>
<td>KC776764</td>
<td>KC776765</td>
</tr>
</tbody>
</table>

demonstrate that Y. frederiksenii genospecies 2 represents a line distinct from Y. frederiksenii genospecies 1 (ATCC 33641T) and 3 (ATCC 29912) and more related to Y. massiliensis. In conclusion, our results provide substantial evidence that Y. frederiksenii genospecies 2 should be reclassified as Y. massiliensis.
Fig. 1. Phylogenetic tree calculated from the 16S rRNA gene sequence (1450 bp) using the Jukes–Cantor evolutionary distance method and UPGMA algorithm. Phylogenetic analysis was performed using the software package BioNumerics 7.0. Bootstrap values (based on 1000 resampled datasets) are shown behind the nodes and the similarity values are presented in front of the nodes. Figures in parentheses indicate genospecies of *Y. frederiksenii*. 
**Fig. 2.** Phylogenetic tree calculated from the alignment of concatenated partial sequences of the six housekeeping genes \( glnA \), \( gyrB \), \( hsp60 \), \( recA \), \( rpoB \) and \( sodA \) (‘superlocus’ of 2610 bp) using the Jukes–Cantor evolutionary distance method and UPGMA algorithm. Phylogenetic analysis was performed using the software package BioNumerics 7.0. Bootstrap values (based on 1000 resampled datasets) are shown behind the nodes and the similarity values are presented in front of the nodes. Figures in parentheses indicate genospecies of *Y. frederiksenii*. 

**MLSA**

- *Yersinia frederiksenii* ATCC 33641\(^T\) (1)
- *Yersinia frederiksenii* ATCC 29912 (3)
- *Yersinia enterocolitica* subsp. *enterocolitica* ATCC 9610\(^T\)
- *Yersinia enterocolitica* subsp. *palearctica* CCUG 52867\(^T\)
- *Yersinia pekkanenii* CCUG 60991\(^T\)
- *Yersinia kristensenii* ATCC 33638\(^T\)
- *Yersinia aldovae* ATCC 35236\(^T\)
- *Yersinia intermedia* ATCC 29909\(^T\)
- *Yersinia rohdei* ATCC 43380\(^T\)
- *Yersinia aleksiae* CCUG 52872\(^T\)
- *Yersinia bercovieri* ATCC 43970\(^T\)
- *Yersinia mollaretii* ATCC 43969\(^T\)
- *Yersinia frederiksenii* FCF 577
- *Yersinia frederiksenii* FCF 276
- *Yersinia frederiksenii* FCF 512
- *Yersinia frederiksenii* ATCC 33644 (2)
- *Yersinia massiliensis* FCF 494
- *Yersinia frederiksenii* FCF 341
- *Yersinia frederiksenii* FCF 342
- *Yersinia frederiksenii* FCF 400
- *Yersinia massiliensis* CCUG 53443\(^T\)
- *Yersinia massiliensis* FCF 216
- *Yersinia frederiksenii* FCF 214
- *Yersinia massiliensis* FCF 465
- *Yersinia massiliensis* FCF 457
- *Yersinia frederiksenii* FCF 1
- *Yersinia frederiksenii* FCF 492
- *Yersinia pestis* KIM
- *Yersinia pseudotuberculosis* ATCC 29833\(^T\)
- *Yersinia similis* CCUG 52882\(^T\)
- *Yersinia entomophaga* CCUG 59867\(^T\)
- *Yersinia ruckeri* CCUG 60891\(^T\)
- *Yersinia frederiksenii* FCF 1
- *Yersinia frederiksenii* FCF 492
- *Yersinia pseudotuberculosis* ATCC 29833\(^T\)
- *Yersinia similis* CCUG 52882\(^T\)
- *Yersinia entomophaga* CCUG 59867\(^T\)
- *Yersinia ruckeri* CCUG 29473\(^T\)
Acknowledgements

We thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for financial support (Proc. 2006/51434-7). During the course of this work, Roberto A. Souza and Priscilla Fernanda Martins Imori were supported by scholarships from FAPESP (Proc. 2009/08455-1 and Proc. 2009/10084-1, respectively).

References


