TRANSCRIPTION OF THE RICKETTSIA FELIS OMPA GENE IN NATURALLY INFECTED FLEAS

JORGE E. ZAVALA-CASTRO, MELISSA SMALL, COLETTE KENG, DONALD H. BOUYER, JORGE ZAVALA-VELAZQUEZ, AND DAVID H. WALKER*

Centro de Investigaciones Regionales Dr. Huley Noguchi, e Facultad de Medicina, Departamento de Patología Tropical, Universidad Autonoma de Yucatan, Mérida, Yucatan, Mexico; Department of Pathology, University of Texas Medical Branch, Galveston, Texas

Abstract. Rickettsia felis is maintained transovarially in Ctenocephalides felis fleas in a widespread geographic distribution and is transmitted to humans and animals, including oppossums. This rickettsia is phylogenetically a member of the spotted fever group, most closely related to Rickettsia akari and R. australis. An unusual feature of this rickettsia is that the gene for the outer membrane protein A (OmpA) is interrupted by stop codons. To determine if this putatively dying gene is expressed, mRNA was extracted from laboratory-maintained, R. felis-infected cat fleas. Reverse transcriptase–polymerase chain reaction amplification of three segments of the ompA gene indicated that mRNA of ompA is actively transcribed in fleas. The cDNA sequences expressed represented mRNA of the first 1860-basepair segment of ompA, which includes domains I and II, part of domain III, the region from site 1836 to site 2180, despite the presence of several stop codons, and the open reading frame from site 2788 to site 3857. The detected sequences showed several differences in the amino acid composition when compared with the previously reported sequence.

INTRODUCTION

The family Rickettsiaceae contains many organisms that cause important diseases in humans around the world with a broad range of clinical manifestations. The genus Rickettsia can be divided in two groups according to their antigenic and molecular profiles: the spotted fever group (SFG) and the typhus group. Both groups share some antigenic proteins, including a 135-kD outer membrane protein (OmpB) and a 17-kD lipoprotein.

In addition, the SFG has a 190-kD transmembrane protein that contains tandemly repeated sequences flanked by conserved regions. The ompA gene has been described in almost all the members of the SFG. The gene length may differ between species depending on the number of repeated sequences, and this genetic variability becomes an excellent tool for molecular characterization among species of Rickettsia or for the identification of different isolates from the same species. Antigenic variation among Rickettsia species may be related, at least in part, to differences in the number, type, and order of the repeated units of ompA.

Recombinant constructs containing a fragment of OmpA stimulate protective immunity in mice, and functionally, monoclonal antibodies against OmpA inhibit the adhesion of Rickettsia rickettsii to L-929 cells. These results support a role of OmpA in cell invasion and reinforce the potential use of this protein for vaccine development.

We reported the DNA sequence of ompA of Rickettsia felis, confirming the taxonomic position of R. felis in the SFG, and described the presence of internal stop codons in the sequence. Its situation as a putatively dying gene makes this gene an excellent model for the study of the evolution of the genus and the pathogenic role of the outer membrane protein in rickettsial infection. This study was undertaken to investigate the expression of the sequences of ompA.

The factors involved in gene expression, the regulation of ompA, and the role of this protein in different environments where the rickettsia lives are unknown. Our findings not only support the active transcription of the ompA gene of R. felis in infected vectors, but also demonstrate the potential full transcription of apparently truncated genes in Rickettsia. The methodology strengthens our analysis of rickettsial gene expression for vaccine development and functional assays.

MATERIALS AND METHODS

Infected fleas. Ctenocephalides felis were generously provided by Dr. Lane Foil (Department of Entomology, Louisiana State University Agricultural Center, Baton Rouge, LA). The fleas were maintained frozen at −20°C.

The fleas were grouped in pools of 50 insects, disinfected with 70% ethanol for 30 minutes, and macerated under sterile conditions in 700 µL of lysis buffer (RNAqueous; Ambion, Austin, TX).

Isolation of RNA from C. felis. Total RNA was extracted from infected fleas using RNAqueous (Ambion), according to the manufacturer’s instructions. Total RNA was resuspended in 50 µL of elution buffer provided by Ambion, and DNase I (RNase free) (Ambion) was added to a final concentration of 2 units/µL and incubated at 37°C for 60 minutes. The DNase was inactivated, and RNA extraction was performed with the DNase-free kit (Ambion). One unit per microliter of SUPERase-In (Ambion) was added, and the RNA samples were stored at −80°C.

Reverse transcriptase–polymerase chain reaction (RT-PCR). Reverse transcription of the isolated RNA was performed using the Superscript III® one-step protocol (SuperScript III® One-Step RT-PCR System; Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions. The RT-PCRs were performed with primers designed for this work (Table 1 and Figure 1). The reactions were performed with 3 µg of total RNA, 0.5 µM of each primer, 1× Superscript III buffer, 1 unit/µL of SUPERase-In (Ambion), and RNase-free water to a final volume of 50 µL. Reverse transcription was performed at 58°C for 30 minutes with primer set 1, and at 60°C for 30 minutes with the primer sets 2 and 3. The enzyme was heat-inactivated for 2 minutes at 94°C, and PCR conditions were 15 seconds at 94°C, 30 seconds at 58°C, and 2 minutes at 68°C (primer set 1), and 15 seconds at 94°C, 30 seconds at 60°C, and 90 seconds at 68°C (primer sets 2 and 3).
for 40 cycles. As a control to ensure that there was no residual DNA or DNA contamination, a PCR was performed on the mRNA template without the inclusion of reverse transcriptase.

Sequence of the ompA RT/PCR products. The ompA RT/PCR products were cloned into the Topo TA pCR 2.1 vector (Invitrogen). Selected clones were sequenced three times using a Perkin Elmer (Wellesley, MA) ABI Prism 377 automated sequencer, and the sequences were compared with the ompA genes from other members of the genus Rickettsia in the GenBank database. The sequences were submitted to GenBank (accession numbers AY727036, AY729962, and AY817673).

Sequence data and phylogenetic analysis. The nucleotide and translated sequences from the cDNA products were compared with the sequence of R. felis ompA from colonized fleas reported previously (AF191026) and with the partial sequence of R. felis ompA from field-collected fleas (AJ563398). Sequence comparison was done using the ClustalW program at the European Bioinformatics Institute. The GenBank accession numbers of the ompA protein gene sequences are R. felis, AF191026;10 R. conorii Malish 7 U01028;5 R. australis AF149108;12 and R. rickettsi M31227.13

Phylogenetic analyses were performed using the maximum-parsimony and distance program of the PAUP 4.1 software.14 Distance matrix analyses were generated with the Kimura two-parameter model for multiple substitutions.15

RESULTS
Reverse transcriptase–polymerase chain reaction. To confirm that expression of R. felis ompA was truncated as suggested previously,16 primer sets were designed to detect the expression of the coding region, the region of domain III directly 3’ to the stop region that we designated the 3’ open reading frame (3’-ORF), and for the stop region present between these sequences. The RT-PCR product detected with primer set 1 consisted of the first 1,860-basepair coding region of the ompA sequence that contains domains I, II, and part of domain III according to the terminology of Anderson and others13 (Figure 2A). The nucleotide sequence of the cDNA fragment was submitted to the European Molecular Biology Laboratory (EMBL) (Heidelberg, Germany) nucleotide database with the accession number AY727036. The sequence showed 99% identity with R. felis and less than 60% homology when compared with the reported ompA sequences of other SFG rickettsiae. The translated cDNA sequence (R. felis 04) showed two differences when compared with the reported R. felis sequence from Mexican field-collected fleas and three differences in amino acid composition in domain II when compared with the published R. felis sequence (Table 2).

The RT-PCR product identified with the primer set 2 was a 1,050-basepair segment that corresponded to the region present from nucleotide site 2788 to site 3837 in the ompA sequence, directly 3’ downstream to the stop region. The nucleotide sequence of the 3’-ORF fragment was submitted to the EMBL nucleotide database with the accession number AY729962 (Figure 2B). The translated sequence showed two differences in amino acid composition when it was compared with the previously reported ompA sequence of R. felis (Table 2).

With primer set 3, we detected a 345-nucleotide transcript corresponding to the stop region located between the two sequences detected with the primer sets 1 and 2 (Figure 3). The five-nucleotide differences found in this fragment with respect to the previously reported sequence of R. felis modify the last 30 amino acids in the translated sequence.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Nucleotide position</th>
<th>Reference name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rf191.1fw</td>
<td>5’-ATGGCGAATATTTCTAAAAATTA-3’</td>
<td>1–1860</td>
<td>Primer set 1</td>
</tr>
<tr>
<td>Rf19.1800rev</td>
<td>5’-TAACTCACCCACCGTAAAGACGCG-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rf19.2788fw</td>
<td>5’-ACATTCCCTGCAAACACTTTACAG-3’</td>
<td>2788–3837</td>
<td>Primer set 2</td>
</tr>
<tr>
<td>Rf19.3837fw</td>
<td>5’-CCGGCGACCTACGGCA-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rf19.1836fw</td>
<td>5’-CTGCGCTAACGGTGGTGAGT-3’</td>
<td>1836–2180</td>
<td>Primer set 3</td>
</tr>
<tr>
<td>Rf19.2180fw</td>
<td>5’-CTGACCGTTGGCTAAAATTCCTAC-3’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 1**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Nucleotide position</th>
<th>Reference name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rf191.1fw</td>
<td>5’-ATGGCGAATATTTCTAAAAATTA-3’</td>
<td>1–1860</td>
<td>Primer set 1</td>
</tr>
<tr>
<td>Rf19.1800rev</td>
<td>5’-TAACTCACCCACCGTAAAGACGCG-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rf19.2788fw</td>
<td>5’-ACATTCCCTGCAAACACTTTACAG-3’</td>
<td>2788–3837</td>
<td>Primer set 2</td>
</tr>
<tr>
<td>Rf19.3837fw</td>
<td>5’-CCGGCGACCTACGGCA-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rf19.1836fw</td>
<td>5’-CTGCGCTAACGGTGGTGAGT-3’</td>
<td>1836–2180</td>
<td>Primer set 3</td>
</tr>
<tr>
<td>Rf19.2180fw</td>
<td>5’-CTGACCGTTGGCTAAAATTCCTAC-3’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**RESULTS**

Reverse transcriptase–polymerase chain reaction. To confirm that expression of R. felis ompA was truncated as suggested previously,16 primer sets were designed to detect the expression of the coding region, the region of domain III directly 3’ to the stop region that we designated the 3’ open reading frame (3’-ORF), and for the stop region present between these sequences. The RT-PCR product detected with primer set 1 consisted of the first 1,860-basepair coding region of the ompA sequence that contains domains I, II, and part of domain III according to the terminology of Anderson and others13 (Figure 2A). The nucleotide sequence of the cDNA fragment was submitted to the European Molecular Biology Laboratory (EMBL) (Heidelberg, Germany) nucleotide database with the accession number AY727036. The sequence showed 99% identity with R. felis and less than 60% homology when compared with the reported ompA sequences of other SFG rickettsiae. The translated cDNA sequence (R. felis 04) showed two differences when compared with the reported R. felis sequence from Mexican field-collected fleas and three differences in amino acid composition in domain II when compared with the published R. felis sequence (Table 2).

The RT-PCR product identified with the primer set 2 was a 1,050-basepair segment that corresponded to the region present from nucleotide site 2788 to site 3837 in the ompA sequence, directly 3’ downstream to the stop region. The nucleotide sequence of the 3’-ORF fragment was submitted to the EMBL nucleotide database with the accession number AY729962 (Figure 2B). The translated sequence showed two differences in amino acid composition when it was compared with the previously reported ompA sequence of R. felis (Table 2).

With primer set 3, we detected a 345-nucleotide transcript corresponding to the stop region located between the two sequences detected with the primer sets 1 and 2 (Figure 3). The five-nucleotide differences found in this fragment with respect to the previously reported sequence of R. felis modify the last 30 amino acids in the translated sequence.

**Figure 1.** Non-scale schematic representation of the Rickettsia felis outer membrane protein A (ompA) gene. The R. felis ompA sequence (GenBank accession no. AF191026) is represented by the open rectangle. The double head arrows at the top indicate the domains, and the filled circles indicate where the stop codons are present. The filled arrows below indicate the positions of the primers used for the reverse transcriptase–polymerase chain reaction (RT-PCR). ORF = open reading frame; nt = nucleotides.
Sequence analysis and taxonomy. As we proposed previously, we analyzed this transcript according to the protein domains by combining the data from the two sequences outside the repeats. Domain I of each OmpA protein started with the first methionine (residue 1) and ended at the beginning of the repeat region. The 3'-ORF region started with the first amino acid after the last stop codon (amino acid position 930 of the translated sequence) and ended at the last amino acid present in this ORF just before the next stop codon region (amino acid position 1288). The combined analysis of the domain I and the 3'-ORF in domain III of OmpA represented 968 amino acid residues.

Phylogenetic distances found for both R. felis sequences in this analysis were similar when compared with other SFG rickettsiae (Table 3), but were not identical. Both cDNA and genomic sequences showed slight differences that represent nearly 0.2% of the sequences analyzed.

DISCUSSION
Among the immunodominant surface proteins present in SFG rickettsiae, OmpA is considered as one of the most important. Because its presence discriminates between the two groups of the genus Rickettsia, it is used for diagnosis and strain and species characterization, and it is functionally involved in the pathogenesis of SFG rickettsial diseases.

Table 2

<table>
<thead>
<tr>
<th>Amino acid position</th>
<th>6</th>
<th>170</th>
<th>242</th>
<th>345</th>
<th>355</th>
<th>957</th>
<th>1154</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. felis†</td>
<td>L</td>
<td>G</td>
<td>P</td>
<td>N</td>
<td>I</td>
<td>G</td>
<td>V</td>
</tr>
<tr>
<td>R. felis 04‡</td>
<td>L</td>
<td>G</td>
<td>S</td>
<td>S</td>
<td>V</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>R. felis Mexican§</td>
<td>P</td>
<td>A</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

† R. felis GenBank accession number AF191026.
‡ R. felis 04 GenBank accession number AY727036.
§ R. felis Mexican GenBank accession number AF563398.

Table 3

<table>
<thead>
<tr>
<th>Amino acid position</th>
<th>R. australis</th>
<th>R. conorii</th>
<th>R. felis*</th>
<th>R. felis 04†</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. australis</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>R. conorii</td>
<td>0.13939</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>R. felis</td>
<td>0.21254</td>
<td>0.21170</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>R. felis 04</td>
<td>0.21435</td>
<td>0.21353</td>
<td>0.00181</td>
<td>–</td>
</tr>
<tr>
<td>R. rickettsii</td>
<td>0.14139</td>
<td>0.02256</td>
<td>0.21008</td>
<td>0.21191</td>
</tr>
</tbody>
</table>

* R. felis GenBank accession number AF191026.
† R. felis 04 GenBank accession number AY727036.
Although OmpA has been identified in almost all the members of the SFG and has been well characterized at the genomic level, active gene expression had only been demonstrated by indirect approaches such as immunofluorescence and immunoblotting. Previous studies of ompA and expression of other rickettsial proteins had been performed using plasmid vectors.1,3,16

The 1,860-basepair sequence of the ompA CDNA contains domains I, II and part of domain III, and the 1,050-basepair transcript contains the domain III ORF present 3' downstream of the stop codon region (Figure 1). Molecular heterogeneity is denoted by the difference of five amino acids and is similar to the intrastrain variation described in the amino acid composition of the ompB protein of different R. prowazekii strains.17 The differences of the ompA sequences of R. felis obtained from colonized fleas and the 170 amino acid partial sequence of R. felis ompA from field-collected fleas suggest that molecular heterogeneity also occurs in nature and likely represents the existence of different strains of R. felis. Although the occurrence of mutations has not been explored in the genome of any rickettsia maintained in colonized fleas, drastic changes were described in the ompA gene repeat domain of R. rickettsii that apparently developed during prolonged passages in cell culture.18 However, we cannot eliminate the possibility that the changes are the consequence of natural selective pressure during prolonged transovarial maintenance in a flea colony rather than an ecologic cycle also involving a mammalian host. This observation emphasizes the importance of further analysis of wild-type strains of R. felis obtained from fleas and animals infected in nature to determine the natural diversity of the ompA gene and whether ompA truncation is associated with long-term laboratory maintenance of fleas.

The amino acid changes in OmpA of R. felis were serine for proline, serine for asparagine, valine for isoleucine, aspartic acid for glycine, and alanine for valine with respect to the ompA sequence reported previously, and a proline for leucine. The slight differences in structure and biochemical properties among the others could also be important for antigenic recognition as reported in other organisms.19-21 Moreover, the sequences of R. felis show huge differences when compared with other rickettsial ompA sequences, and these variations could be used for diagnostic purposes using the protein fragments for detection of antibodies to R. felis in human patients and in animal reservoirs.

The stop region RT fragment contained five nucleotide differences when compared with the reported ompA gene of R. felis, including two deletions, one insertion, and two transversions. The indels and substitutions resulted in a change in the amino acid sequence of the last 30 amino acids in the sequence.

The genome of the genus Rickettsia contains several pseudogenes and ORFs with apparently noncoding function.22 The presence of these sequences suggests an active genome degradation process as a result of an adaptation of the different species to the diverse environments where they live, and could most likely be the result of speciation.23 As in other reported pseudogene sequences in many Rickettsia species, the ompA gene of R. felis contains several stop codons that interrupt their reading frame. The accumulation of these stop codons could be the result of the ongoing degradation process to become a pseudogene.

However, at this moment of the evolution of R. felis, the organization of the ompA gene resembles more a split gene than a pseudogene. It had been suggested that some of the truncated genes reported in the genus Rickettsia could retain some functional activity, such as the genes that encode the SpoT proteins of R. prowazekii, which represent a truncated version of the SpoT protein of Escherichia coli.22,24,25

Different transcription patterns with diverse complexity of several split genes had been described previously.24 These results suggest different transcription mechanisms for these genes, but also could reflect the presence of transcripts in different stages of processing. Our results support the concept of a full transcription event in the split genes despite the presence of stop codons or the noncoding region present between the ORFs.

The presence of the ompA transcript raises the possibility of its expression regardless of the condition of being a truncated gene, similar to that of R. peacockii, which has been found in nature only in ticks. A short sequence (600 basepairs) of the truncated ompA gene is detected.26 The putative OmpA protein of R. felis does not have the beta peptide for its transport to the outer membrane, but still has the signal peptide for its insertion into the inner membrane. The absence of the autotransporter segment makes this protein lack the properties needed to function for its own export to the outer membrane, and we are unsure if the remaining inner membrane segment is functional. Moreover, the changes do not affect the transcription terminator region of the gene, and it is also possible that the full sequence is transcribed. This observation is based on the RT-PCR amplification of the stop codon region and the domain III ORF adjacent to the 3’ end. Two predominant heat resistant proteins with molecular masses of 60 and 30 kD have been reported in R. felis.27 The molecular weights of these proteins are similar to the masses predicted for the transcript of 1,860 basepairs and the transcript of the 3’ ORF reported in this work (60.5 and 36 kD, respectively). Confirmation of this possibility would require further experimental evidence; the presence of both OmpA segments in R. felis would support the hypothesis of the functionality of some split genes by independent units or working together as subunits of one protein.

The gene degradation of ompA in R. felis could have several explanations. As discussed earlier in this report, the presence of this phenomenon could be associated with long-term laboratory maintenance in fleas where the protein might not require the function needed in the mammalian host. The degradation could also have been produced by evolutionary adaptation to their environment and/or could be the result of evolutionary natural genetic heterogeneity among strains.

The phylogenetic relationship of this R. felis variant to other SFG Rickettsia species is the same as reported for R. felis previously (Table 3)16; the differences between the R. felis sequences could be the result of strain variation. The sequence alignment of the R. felis ompA with the partial sequence of the R. akari ompA available in the GenBank
(AY727036) was not useful in determining potentially shared segments between the sequences of these two SFG rickettsiae not transmitted by ticks.

In this work, we report for the first time the full-length transcript of a putative dying gene detected by RT/PCR in infected fleas. This study was performed with an adapted methodology that could be considered a contribution to the study of gene expression in members of the genus Rickettsia and related intracellular bacteria, especially when axenic culture is not available.

The expression and isolation of OmpA in vitro is essential to determine its function in the pathogenesis of rickettsial diseases and to develop better methods for diagnosis and epidemiologic surveillance. It is also a potential target for new strategies for developing drugs and vaccines for the control of human illnesses caused by rickettsiae.

Received February 6, 2004. Accepted for publication May 24, 2005.

Acknowledgments: This research was supported by grants from the Fogarty International Center of the National Institutes of Health (D43 TW00903) and CONACyT (34436-M).

Authors’ addresses: Jorge E. Zavala-Castro, Laboratorio de Biología Celular, Centro de Investigaciones Regionales Dr. Hideyo Noguchi, Universidad Autónoma de Yucatán, Avenida Itzaes No. 490 x 59, Centro, CP 97000, Merida, Yucatan, Mexico. Telephone: 52-999-924-6412 extension 120, Fax: 52-999-923-6120, E-mail: zcastro@tunku.udy.mx. Melissa Small, Collect. Feng, Donald H. Bouyer, and David H. Walker, Department of Pathology, University of Texas Medical Branch, 301 University Boulevard, Keiller Building, Galveston, TX 77555-0609, Telephone: 409-772-2856, Fax: 409-772-1850, E-mails: dobouyer@utmb.edu and dwalker@utmb.edu. Jorge Zavala-Velazquez, Departamento de Patología Tropical, Universidad Autónoma de Yucatán, Avenida Itzaes No. 498 x 59 y 59a, Centro CP 97000, Merida, Yucatan, Mexico. Telephone: 52-999-930-0900, zavala@tunku.udy.mx.

Reprint requests: Jorge E. Zavala-Velázquez, Facultad de Medicina, Departamento de Patología Tropical, Universidad Autónoma de Yucatán, Avenida Itzaes No. 498 x 59 y 59a, Centro CP 97000 Merida, Yucatán, Mexico.

REFERENCES
