Molecular Typing of Isolates of *Rickettsia rickettsii* by Use of DNA Sequencing of Variable Intergenic Regions

Sandor E. Karpathy, Gregory A. Dasch, and Marina E. Eremeeva*

*Rickettsial Zoonoses Branch, Centers for Disease Control and Prevention, Atlanta, Georgia 30333*

Received 16 February 2007/Accepted 11 May 2007

*Rickettsia rickettsii*, the causative agent of Rocky Mountain spotted fever, is found throughout the Americas, where it is associated with different animal reservoirs and tick vectors. No molecular typing system currently exists to allow for the robust differentiation of isolates of *R. rickettsii*. Analysis of eight completed genome sequences of rickettsial species revealed a high degree of sequence conservation within the coding regions of chromosomes in the genus. Intergenic regions between coding sequences should be under less selective pressure to maintain this conservation and thus should exhibit greater nucleotide polymorphisms. Utilizing these polymorphisms, we developed a molecular typing system that allows for the genetic differentiation of isolates of *R. rickettsii*. This typing system was applied to a collection of 38 different isolates collected from humans, animals, and tick vectors from different geographic locations. Serotypes 364D, from *Dermacentor occidentalis* ticks, and Hlp, from *Haemaphysalis leporispalustris* ticks, appear to be distinct genotypes that may not belong to the species *R. rickettsii*. We were also able to differentiate 36 historical isolates of *R. rickettsii* into three different phylogenetic clades containing seven different genotypes. This differentiation correlated well, but not perfectly, with the geographic origin and likely tick vectors associated with the isolates. The few apparent typing discrepancies found suggest that the molecular ecology of *R. rickettsii* needs more investigation.

In the late 1800s, reports first described a fatal febrile illness affecting settlers in the Bitterroot Valley of western Montana. Howard Taylor Ricketts showed 100 years ago that this disease, Rocky Mountain spotted fever (RMSF), was caused by a bacterium later named in his honor (41, 47). Despite the localized geographic association of its name, RMSF is found throughout the continental United States as well as in Central and South America (12, 13, 39, 42, 43). RMSF is the most commonly fatal tick-borne bacterial disease reported in the world, with the fatality rate for untreated cases approaching 20% (12, 13, 30). Common clinical features include fever, chills, headache, malaise, myalgia, and rash. Early diagnosis and antibiotic therapy are critical to avoid severe disease and to decrease the chance of a fatal outcome, with doxycycline being the drug of choice for treatment of RMSF (11, 12, 25).

The etiological agent of RMSF is the obligately intracellular bacterium *Rickettsia rickettsii*. The life cycle of this bacterium involves both vertebrate and invertebrate hosts, with hard (ixodid) ticks serving as the vector and both the ticks and their mammalian hosts serving as reservoirs to maintain the bacterium in nature (7). In the United States, RMSF is spread primarily by two different tick vectors, *Dermacentor variabilis* in the eastern and midwestern part of the country (5, 44) and *Dermacentor andersoni* in the northwestern part of the country (34, 35). The brown dog tick, *Rhipicephalus sanguineus*, was recently confirmed to be a competent vector for *R. rickettsii* (9, 10, 14), while the rabbit tick, *Haemaphysalis leporispalustris*, has been implicated as a vector for low-virulence strains of *R. rickettsii* (2, 3, 32, 35) known as serotype Hlp. In California, a limited number of spotted fever isolates of serotype 364D that are closely related to *R. rickettsii* have also been recovered from *Dermacentor occidentalis* ticks (27, 35). However, formal identification of the Hlp and 364D serotypes as *R. rickettsii* has not been completed. *Dermacentor* ticks are not found in Central and South America (19); in these regions, *Amblyomma cajennense* and *Amblyomma aureolatum* have been implicated in the transmission of RMSF (23, 37, 42).

A number of phenotypic and genetic differences have been observed previously among isolates of *R. rickettsii*. During his early work on the virulence of *R. rickettsii*, Price classified isolates recovered from *D. andersoni* ticks collected in Montana into four categories (R, S, T, and U) based on the different degrees of pathology observed in a guinea pig model of infection (38). An in vitro endothelial cell system for differentiating virulence properties of *R. rickettsii* has also been described (17). Anacker and colleagues also reported differences in the virulence of *R. rickettsii* isolates in the guinea pig model, grouping their isolates into three categories: isolates with highest virulence, lesser virulence, and lowest virulence (2, 3). They also observed differences in the one-dimensional sodium do decyl sulfate-polyacrylamide gel electrophoresis mobilities of proteins recovered from whole-cell lysates of different isolates, but these differences allowed only for the differentiation of isolate Hlp (3). Philip et al. demonstrated antigenic differences between putative *R. rickettsii* isolates 364D and Hlp#2, but not other isolates, through the use of microimmunofluorescence with mouse immune sera (35). By AluI PCR-restriction fragment length polymorphism (PCR-RFLP) analysis of almost identical nucleotide regions of rompA, Regny et al. (40) (nucleotides 70 to 602) and Eremeeva et al. (18) (nucleotides 70 to 701) were able to differentiate Hlp#2 from a combined...
TABLE 1. *R. rickettsii* isolates used in this study

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Location, yr of isolation</th>
<th>Source</th>
<th>Passage history</th>
<th>Origin (collection)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheila Smith (VR149)</td>
<td>Missoula, MT, 1946</td>
<td>Patient suffering from RMSF</td>
<td>2GP + 8YS + 5V + 1GP + 3YS + 4V</td>
<td>ATCC (from M. Bozeman, WRAIR)</td>
<td>4</td>
</tr>
<tr>
<td>364D</td>
<td>Ventura County, CA, 1966</td>
<td>Cell culture isolate from <em>Dermacentor occidentalis</em> tick</td>
<td>6YS + 1V</td>
<td>RML</td>
<td>34</td>
</tr>
<tr>
<td>76RC</td>
<td>Georgia, 1976</td>
<td>Guinea pig isolate from blood clot from a patient who died of RMSF</td>
<td>1GP + 3YS + 2V</td>
<td>CDC</td>
<td>40</td>
</tr>
<tr>
<td>78RL</td>
<td>Georgia, 1978</td>
<td>Guinea pig isolate from a patient who suffered from RMSF</td>
<td>1GP + 4YS + 2V</td>
<td>CDC</td>
<td>40</td>
</tr>
<tr>
<td>80JC</td>
<td>Omaha, NE, 1980</td>
<td>Guinea pig isolate from blood clot from a patient who suffered from RMSF</td>
<td>1GP + 2YS + 4V</td>
<td>CDC</td>
<td>40</td>
</tr>
<tr>
<td>81WA</td>
<td>North Carolina, 1986</td>
<td>Cell culture isolate from a patient who suffered from RMSF</td>
<td>GP? + V? + 1YS + 1V</td>
<td>D. Walker (from C. Pretzman, ODPH)</td>
<td>1V D. Walker, personal communication</td>
</tr>
<tr>
<td>84JG</td>
<td>North Carolina, 1986</td>
<td>Guinea pig isolate from a patient who suffered from RMSF</td>
<td>GP? + 3V + 1YS + 5V</td>
<td>C. Pretzman, ODPH</td>
<td>1V D. Walker, personal communication</td>
</tr>
<tr>
<td>AZ-1</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from <em>R. sanguineus</em> tick</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>AZ-3</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from blood of a patient who died of RMSF</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>AZ-4</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from blood of a patient who died of RMSF</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>AZ-5</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from questing <em>R. sanguineus</em> tick</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>AZ-6</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from skin biopsy of a patient who suffered from RMSF</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>AZ-7</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from partially engorged <em>R. sanguineus</em> tick</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>AZ-8</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from <em>R. sanguineus</em> tick found in a house of a patient who died of RMSF</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>AZ-9</td>
<td>Arizona, 2004</td>
<td>Cell culture isolate from <em>R. sanguineus</em> tick found in house of a patient who died of RMSF</td>
<td>2-3V</td>
<td>CDC</td>
<td>15</td>
</tr>
<tr>
<td>Bitterroot Valley, MT, 1945</td>
<td>Bitterroot Valley, MT, 1945</td>
<td>Guinea pig isolate from <em>D. andersoni</em> tick</td>
<td>GP? + 55YS + 3V</td>
<td>ATCC (from M. Bozeman, WRAIR)</td>
<td>4</td>
</tr>
<tr>
<td>Brazil-A</td>
<td>Brazil, before 1943</td>
<td>Possible yolk sac isolate (exact passage history unknown); presumably from a tick (exact history unknown)</td>
<td>? + 28YS + 2L + 2V</td>
<td>D. Kelly, WRAIR</td>
<td>4</td>
</tr>
<tr>
<td>Brazil-B</td>
<td>Brazil, before 1943</td>
<td>Possible yolk sac isolate (exact passage history unknown); presumably from a tick (exact history unknown)</td>
<td>GP? + 4V</td>
<td>C. H. Melles, IAL</td>
<td>4</td>
</tr>
<tr>
<td>BSF Di6</td>
<td>Hanover County, VA, 1961</td>
<td>Yolk sac isolate from spleen and liver homogenate taken from <em>Dichelops marquesi virginiensis</em> (opossum)</td>
<td>5YS + 1V</td>
<td>M. Bozeman, WRAIR</td>
<td>5</td>
</tr>
<tr>
<td>BSF Mp40</td>
<td>Fairfax County, VA, 1961</td>
<td>Yolk sac isolate from spleen and liver homogenate taken from <em>Micronus pennsylvanicus</em> (vole)</td>
<td>4YS + 1V</td>
<td>M. Bozeman, WRAIR</td>
<td>5</td>
</tr>
<tr>
<td>BSF Rab1</td>
<td>Prince Edward County, VA, 1960</td>
<td>Yolk sac isolate from spleen and liver homogenate taken from <em>Sylvilagus floridanus</em> (rabbit)</td>
<td>5YS + 1V</td>
<td>M. Bozeman, WRAIR</td>
<td>5</td>
</tr>
<tr>
<td>Coleman</td>
<td>Charlottesville, VA, 1969</td>
<td>Yolk sac isolate from autopsy material (brain, liver, and spleen) from a patient who died of RMSF</td>
<td>4YS + 1V</td>
<td>McGhee, MCV</td>
<td>5</td>
</tr>
<tr>
<td>Colombia</td>
<td>Colombia, 1935?</td>
<td>Presumably yolk sac isolate (exact passage history unknown) from a patient suffering from RMSF</td>
<td>? + 19YS + 1L + 2V</td>
<td>C. Pretzman, ODPH</td>
<td>33</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Costa Rica</td>
<td>Yolk sac isolate from a patient who suffered from RMSF</td>
<td>? + 1YS + 3V</td>
<td>L. Fuentes, UCR</td>
<td>5</td>
</tr>
<tr>
<td>Hauke</td>
<td>Clermont, OH, 1986</td>
<td>Cell culture isolate from lung tissue of a patient who died of RMSF</td>
<td>5V + 1YS + 1V</td>
<td>C. Pretzman, ODPH</td>
<td>5</td>
</tr>
<tr>
<td>Hino</td>
<td>Oklahoma, 1964</td>
<td>Yolk sac isolate from a patient who died of RMSF</td>
<td>6YS + 1V</td>
<td>M. Bozeman, WRAIR</td>
<td>5</td>
</tr>
<tr>
<td>Hlp#2-A</td>
<td>Coyote Gulch, Bitterroot Valley, MT, 27 July 1948</td>
<td>Yolk sac isolate from <em>H. leporispalustris</em> tick</td>
<td>50YS + 9TC + 3-4YS + 8V</td>
<td>G. McDonald, RML</td>
<td>32</td>
</tr>
<tr>
<td>Hlp#2-B</td>
<td>Coyote Gulch, Bitterroot Valley, MT, 27 July 1948</td>
<td>Yolk sac isolate from <em>H. leporispalustris</em> tick</td>
<td>51YS + 9TC + 4YS + 6V</td>
<td>CDC</td>
<td>32</td>
</tr>
</tbody>
</table>

Continued on facing page
total of 17 \textit{R. rickettsii} isolates. Both groups also used Alu
PCR-RFLP analysis of the citrate synthase gene, \textit{gltA}, and
found that Montana isolates Bitterroot and Sheila Smith had
slightly different banding patterns from those of \textit{R. rickettsii}
isolates from either the Central or Eastern United States or
Central or South America (18, 40). Eremeeva et al. also re-
ported that isolates Lost Horse Canyon and Morgan shared
the same, less common \textit{gltA} Alu PCR-RFLP banding pattern
as isolates Bitterroot and Sheila Smith (18). In the same study,
Eremeeva et al. also showed that isolate Hlp\#2 contains a
unique nine-base-pair insertion in the sequence of a fossil gene
in comparison to other \textit{R. rickettsii} isolates. Finally, Gilmore
and Hackstadt (22) used HincII digestion of a 3.8-kb repeated
region fragment of \textit{rompA} to separate five \textit{R. rickettsii} isolates
into three groups.

In the public health setting, molecular typing of infectious
agents is important for tracing their origin and spread in out-
break investigations, the detection of disseminated antibiotic-
resistant strains in managed care facilities, the identification of
hypervirulent strains, and monitoring failures in live vaccina-
tion programs (29). Molecular typing also allows for the
study of bacterial population dynamics and may provide an
improved understanding of the ecological niches occupied
by specific pathogen types in the environment (24, 45). Mo-
molecular typing schemes based on the sequencing of inter-
genic regions (IGRs) have been developed for \textit{Rickettsia conorii}
and \textit{Rickettsia prowazekii} (21, 49), and genetic typing of
\textit{R. rickettsii} based on variable-number-tandem-repeat loci
was described recently (15, 18a, 48). In this work, we present
an IGR typing scheme for \textit{R. rickettsii} based on nucleotide
polymorphisms found within six sites. This typing method
was applied to a collection of 38 \textit{R. rickettsii} isolates from
human RMSF patients, animals, and ticks from different
geographic locations.

\textbf{MATERIALS AND METHODS}

\textit{R. rickettsii} isolates and DNA preparation. \textit{R. rickettsii} isolates (Table 1) were
cultivated in Vero cells (strain C1008; green monkey kidney cells) as described
elsewhere (17). DNAs for isolates Brazil-A, 84JG, Hlp\#2-A, Bitterroot, Colombi-
ia, Lost Horse Canyon, Morgan, PriceT, and Sheila Smith were prepared by
phenol-chloroform extraction from partially purified \textit{R. rickettsii} cells as de-
scribed previously (16). DNAs for the remaining isolates were isolated from
infected cell cultures by using a QAamp DNA Mini kit from QIAGEN (Va-
encia, CA). DNAs were eluted with AE buffer (QIAGEN) and stored at 4°C
prior to analysis. Hlp\#2-A and Hlp\#2-B refer to two samples of isolate Hlp\#2
with unique passage histories maintained by two different laboratories. Since the
nucleotide sequences of Hlp\#2-A and Hlp\#2-B are identical for all loci tested,
they are referred to jointly as isolate Hlp\#2 for the remaining of this commu-
nication. Likewise, Brazil-A and Brazil-B are samples of isolate Brazil with
differing passage histories. These samples are also identical and thus are referred
to singularly as isolate Brazil. All 38 isolates had \textit{rompA} 70p-602n gene frag-
ments (15, 18, 40) compatible with their typing as \textit{R. rickettsii}, although the
nucleotide sequences of the Hlp\#2 and 364D fragments each differ by three
bases from the Sheila Smith sequence.

PCR primer design. An automated preliminary annotation of the genomic
sequence of \textit{R. rickettsii} strain Sheila Smith (GenBank accession number
AADJ00000000) was used to identify suitable IGRs. IGRs of 200 to 500 nucle-
odites were chosen for further studies, and the nucleotide sequences were com-
pared to the homologous sequences of seven \textit{rickettsial} species, including
\textit{Rickettsia akari} strain CWPP-Hartford, \textit{Rickettsia canadensis} strain McKiel

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\textbf{Isolate} & \textbf{Location, yr of isolation} & \textbf{Source} & \textbf{Passage history} & \textbf{Origin (collection)} & \textbf{Reference} \\
\hline
Lost Horse Canyon & Lost Horse Canyon, Bitterroot Valley, MT, 1958 & Yolk sac isolate from \textit{Dermacentor andersonii} tick & 6YS + 2GP + 5YS + 3V & M. Peacock, RML & 8 \\
\hline
Morgan & North Carolina, 1975 & Yolk isolate from blood of a patient suffering from RMSF & 1M + 3YS + 5TC + 2-3YS + 8V & G. McDonald, RML & 35 \\
\hline
OSU 83-13-4 & Clermont County, OH, 1983 & Cell culture isolate from \textit{D. variabilis} tick & 6V + 1YS + 1V & C. Pretzman, ODPH & \\
\hline
OSU 84-21c & Franklin County, OH, 1985 & Cell culture isolate from \textit{D. variabilis} tick & 3V + 1YS + 1V & C. Pretzman, ODPH & \\
\hline
OSU 85-Lu1 & Lucan County, OH, 1985 & Cell culture isolate from \textit{D. variabilis} tick & 3V + 1YS + 1V & C. Pretzman, ODPH & \\
\hline
Panama2004 & Panama, 2004 & Cell culture isolate from brain tissue of a patient who died of RMSF & 2V & CDC & \\
\hline
Price T & Montana, before 1953 & Possible yolk sac isolate (exact passage history unknown) from \textit{D. andersonii} tick & ? + 3YS + 2L + 2V & J. Spielman, HSPH & 38 \\
\hline
Sawtooth & Sawtooth Canyon, Bitterroot Valley, MT, 1961 & Guinea pig isolate from \textit{D. andersonii} tick & >8T + 7V & C. Pretzman (from W. Burgdorfer, RML) & 6 \\
\hline
Stewart & Ft. Meade, VA, 1970 & Yolk sac isolate from blood clot from a patient who suffered from RMSF & 3YS + 1V & M. Bozeman, WRAIR & \\
\hline
Sukiewicz & Frederick, MD, 1973 & Yolk sac isolate from blood of a patient who suffered from RMSF & 3YS + 3V & M. Bozeman, WRAIR & 17 \\
\hline
von Schlemmer & Frederick, MD, 1970 & Isolate (isolation history unknown) from a patient who suffered from RMSF & ? + 6YS + 1V & M. Bozeman, WRAIR & \\
\hline
\end{tabular}
\end{table}
TABLE 2. Primers used in this study

<table>
<thead>
<tr>
<th>IGR Forward</th>
<th>Primer sequence (5’→3’)</th>
<th>Reverse</th>
<th>Amplion size (nt)*</th>
<th>Reference</th>
<th>NCBI accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>spo0l-abcT1</td>
<td>AAAAGTTTGGAAATAGTACTGAT</td>
<td>TTTCGTTAAACACATTATCCTCA</td>
<td>320</td>
<td>21</td>
<td>EF216182-EF216218</td>
</tr>
<tr>
<td>RR0155-pmb</td>
<td>TTATGCGACGCGGCTTATTACC</td>
<td>TTAGGCGAAAGGAGAGAACTTT</td>
<td>290</td>
<td>21</td>
<td>This study EF216032-EF216068</td>
</tr>
<tr>
<td>RR0345-locI</td>
<td>AGAGCCCTGCGAGTATTTA</td>
<td>AGGAGATGTGGGAGAAATCCTTT</td>
<td>238</td>
<td>21</td>
<td>EF215869-EF216105</td>
</tr>
<tr>
<td>cpA-ksgA</td>
<td>CTCATGACTCTGTTACTTAC</td>
<td>ATAGCCGTTTAATACGTTCGTT</td>
<td>405</td>
<td>21</td>
<td>This study EF215859-EF216105</td>
</tr>
<tr>
<td>RR1372-RR1373</td>
<td>TCGGCAGGATGACTTAC</td>
<td>CGGCAGAAATCGGTGTT</td>
<td>340</td>
<td>21</td>
<td>This study EF216114-EF216150</td>
</tr>
<tr>
<td>RR1240-dc58</td>
<td>CGGAGAAGCGGCGGATGA</td>
<td>ATTCGCGAAGATTTTGTGTT</td>
<td>357</td>
<td>21</td>
<td>EF215969-EF216005</td>
</tr>
<tr>
<td>IRNA3-adjR3</td>
<td>TTGAAAACACGAAACAGAAGGA</td>
<td>CCAGAACTTCTGAGCTTGAT</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>RCO102-RC0103</td>
<td>CGGAGAAGCGGCGGATGA</td>
<td>ATTCGCGAAGATTTTGTGTT</td>
<td>402</td>
<td>21</td>
<td>EF216025-EF216031</td>
</tr>
<tr>
<td>RCO280-23s rRNA</td>
<td>CAAAAGGCGAACAGGCTTA</td>
<td>CCTCTGCGCCTCATTGC</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>acrb-hupA</td>
<td>GGCACGTTAATACATAAGAGA</td>
<td>CTACCCCTTGTAGTAGTATG</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>pal-RC1201</td>
<td>TGAACGACACATATAGCACA</td>
<td>TCAAGTAGCTTATCCCTTCT</td>
<td>292</td>
<td>21</td>
<td>EF216014-EF216016</td>
</tr>
<tr>
<td>scb-czR</td>
<td>ATGGAGATGTGGGAGAAATCCTTT</td>
<td>GGGGGTTGGCCATTAATTACA</td>
<td>336</td>
<td>21</td>
<td>EF216167-EF216174</td>
</tr>
<tr>
<td>gran-ES-RC0970</td>
<td>CTTCAGCGTGTCTTCATTATA</td>
<td>AAAATCAGCTGAGAAAGAAGTAGAA</td>
<td>353</td>
<td>21</td>
<td>EF215926-EF215933</td>
</tr>
<tr>
<td>seck-pmA</td>
<td>GCAGCTGACGGGAGTAGG</td>
<td>AGTGCAGCGACACAAAGAA</td>
<td>370</td>
<td>21</td>
<td>EF216151-EF216158</td>
</tr>
<tr>
<td>RCO640-RC0605</td>
<td>AAAGGACGACTACGATGAAATC</td>
<td>AGTCGCGGATGTTACATTACTC</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>RCO409-rrmU</td>
<td>AACCTGGAAGCTGATGAAATC</td>
<td>GCTCTAGAAGCCGAAAT</td>
<td>400</td>
<td>21</td>
<td>EF216106-EF216113</td>
</tr>
<tr>
<td>yop-gotB</td>
<td>CTCTGCGCCTGATGTTATTAC</td>
<td>ATTCGCGACTGTTCCATAC</td>
<td>398</td>
<td>21</td>
<td>EF216107-EF216108</td>
</tr>
<tr>
<td>rnc-cxW</td>
<td>CGGAGAAGAAATGGGACGAGTCTTT</td>
<td>CCATTGTAATTTAATCCTTTCGCTC</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>dna-N-RC0585</td>
<td>TCTTGATGCTTCTGTAGG</td>
<td>TTGATAAATCCGCCGTTAG</td>
<td>457</td>
<td>21</td>
<td>EF215904-EF215910</td>
</tr>
<tr>
<td>lig-gg</td>
<td>TTTTGCTGTCTCCCTCATG</td>
<td>CCAAAATCAGCTGAGAAAGAAGTAGAA</td>
<td>395</td>
<td>21</td>
<td>EF215950-EF215957</td>
</tr>
<tr>
<td>rnc-RR076f</td>
<td>CCGATGCTTATGCTGTG</td>
<td>CCCCCCGGATGTTACATTACCTC</td>
<td>433</td>
<td>21</td>
<td>EF216197-EF216204</td>
</tr>
<tr>
<td>lco-bioY</td>
<td>AGGGGCGCCGGAAAGAT</td>
<td>TACGGCGGATGTTACATTACCTC</td>
<td>340</td>
<td>21</td>
<td>EF215911-EF215917</td>
</tr>
<tr>
<td>ntr-ypsi</td>
<td>AGCTCGTGTGTTACGATGAA</td>
<td>AAAAGGCGGAGTAGG</td>
<td>395</td>
<td>21</td>
<td>EF215950-EF215957</td>
</tr>
<tr>
<td>slxG-G-G-RG0568</td>
<td>TCCCAATAGGATGCTGTTCTT</td>
<td>TACATCGCGGATGTTACATTACCTC</td>
<td>225</td>
<td>21</td>
<td>EF215906-EF215903</td>
</tr>
<tr>
<td>fabB-fxpD</td>
<td>GAGAAGAAGAAATGGGACGAGTCTTT</td>
<td>CCATTGTAATTTAATCCTTTCGCTC</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>rr-psyB</td>
<td>GGGCGTTTAATACAAATTTTAGACA</td>
<td>CGAATTCTCCTTTGATAGGTTAATATGT</td>
<td>398</td>
<td>21</td>
<td>EF216107-EF216108</td>
</tr>
<tr>
<td>IRNA4-IRNA5</td>
<td>ATGCAGGATTCCAGCCTTTAGGGCTCGCCTTCAATTAACAA</td>
<td>AAAGGCCGTTAGGAAACACGGT</td>
<td>292</td>
<td>21</td>
<td>EF215964-EF215982</td>
</tr>
<tr>
<td>mgtU</td>
<td>GAGAAGAAGAAATGGGACGAGTCTTT</td>
<td>CCATTGTAATTTAATCCTTTCGCTC</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>dna-N-IRNA5</td>
<td>ATGCAGGATTCCAGCCTTTAGGGCTCGCCTTCAATTAACAA</td>
<td>AAAGGCCGTTAGGAAACACGGT</td>
<td>292</td>
<td>21</td>
<td>EF215964-EF215982</td>
</tr>
<tr>
<td>mgtU</td>
<td>GAGAAGAAGAAATGGGACGAGTCTTT</td>
<td>CCATTGTAATTTAATCCTTTCGCTC</td>
<td>NP</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>dna-N-IRNA5</td>
<td>ATGCAGGATTCCAGCCTTTAGGGCTCGCCTTCAATTAACAA</td>
<td>AAAGGCCGTTAGGAAACACGGT</td>
<td>292</td>
<td>21</td>
<td>EF215964-EF215982</td>
</tr>
</tbody>
</table>

a Size is based on the R. rickettsii Sheila Smith genome sequence from NCBI. NP, no product (a PCR product was not obtained from some or all isolates tested).
b The name of this primer was changed to more accurately define the IGR in R. rickettsii. In the cited text, it is named RC1137-dc5.
c This primer has been modified from the published R. conorii sequence to match the R. rickettsii Sheila Smith genome sequence. The underlined base is the modified base.
d Identification of these loci as IGRs was based on a preliminary annotation of the R. rickettsii Sheila Smith genome. The most current version of the annotation indicates that this region may contain as many as three hypothetical coding regions.

An initial subset of five isolates was used to screen seven candidate variable IGRs for nucleotide polymorphism. This subset consisted of isolates Bitterroot, Colombia, AZ-8, Hlp#, and 364D. These isolates were chosen to provide samples from different geographic locations and vector associations to maximize the potential for detecting genetic diversity. Four loci (RR0155-pmb, RR0345-locI, cpA-ksgA, and RR1372-RR1373) exhibited genetic diversity during this screen and were chosen for further study. In addition, primer pairs designed by Fournier et al. for the molecular typing of R. conorii isolates (21) were analyzed for their usefulness in the typing of R. rickettsii isolates. These primers were used to amplify and sequence loci from eight isolates (AZ-8, Hlp#, 364D, BSF Rab1, Hauke, Morgan, Panama2004, and PriceT), again chosen based on their geographic location, vector association, and the IGR sequence diversity detected using the
initial primers. Of the 27 Fournier primers tested, two IGRs (spolI-abcT1 and RR1240-tlc5) exhibited sufficient diversity during the sequencing screen to warrant further study, while amplicons were not obtained for all tested DNAs for six other IGR primer pairs (Table 2). Therefore, a total of 28 loci were screened (Table 2), of which 6 were chosen for sequencing from all 38 isolates and were included in the genotypic and phylogenetic analyses. Of the 22 IGR loci sequenced in the screen but not chosen for further study, 7 loci had no nucleotide differences among the tested isolates, 6 loci exhibited sequence differences only in both 364D and Hlp#2, 1 locus exhibited differences only in 364D, and 1 locus exhibited differences only in Hlp#2. The remaining seven loci tested showed various amounts of diversity, but the isolates grouped together as seen for the six chosen loci.

To ensure that no polymorphisms were due to mistakes in the reference genome sequence, all six selected loci were resequenced from isolate Sheila Smith; these sequences were identical to the Sheila Smith reference genome sequence in GenBank.

**Analysis of spolI-abcT1 IGR.** Two single nucleotide polymorphisms (SNPs) and two types of insertions were identified in the amplicons generated with the spolI-abcT1 primer pair (Fig. 1; Table 2). These genetic differences allowed for the separation of the isolates into four different genotypes (A1 to D1). Compared to the reference isolate Sheila Smith (genotype A1), all isolates except Sawtooth, Bitterroot, Lost Horse Canyon, and Morgan have a guanine-to-adenine transition at nucleotide 194 of the consensus amplicon, while isolate 364D (genotype C1) has a unique adenine-to-guanine transition at nucleotide 193. Isolate 364D (genotype B5) contains a guanine-to-adenine transition at nucleotide 219, and it shares a third SNP, an adenine-to-cytosine transversion of nucleotide 357. The 364D amplicon (genotype F4) shares two of these SNPs with Hlp#2: the transversion of nucleotide 179 and the transition of nucleotide 220, but not the others. Amplicons from all isolates except Sheila Smith, Sawtooth, Bitterroot, Lost Horse Canyon, and Morgan (genotypes A4 and B4) have a six-nucleotide insertion (AATTAT; consensus nucleotides 255 to 260). A polycytosine region consisting of 8, 9, or 10 cytosine residues is found for consensus nucleotides 270 to 279 of this locus. Amplicons from genotypes B4 (Bitterroot, Lost Horse Canyon, and Morgan), C4 (Colombia and all eight Arizona isolates), and F4 (364D) have 8 cytosines, while amplicoms from genotypes E4 (84JG, Hino, OSU 83-13-4, OSU 84-21c, 76RC, and Hauke) and G4 (Hlp#2) contain 10 consecutive cytosines. The remaining 19 amplicons (genotypes A4 and D4) all contain nine consecutive cytosines. Genotypes C4 and E4 were distinguished from genotype D4 (the largest group of isolates, which included all Central and South American strains and most Eastern and Midwestern U.S. strains) only by the number of cytosines in the polycytosine region.

**Analysis of the RR1372-RR1373 IGR.** Only two genotypes (A5 and B5) were present in the RR1372-RR1373 amplicons (Fig. 1; Table 2). Isolate 364D (genotype B5) contains a guanine-to-thymine transversion of nucleotide 36 and an adenine-to-guanine transition of nucleotide 193.

**Analysis of the RR1240-tlc5 IGR.** Four SNPs and three indels were detected in the RR1240-tlc5 amplicons, resulting in genotypes A6 to E6 (Fig. 1; Table 2). The Hlp#2 amplicon (genotype E6) had a thymine-to-guanine transversion at nucleotide 88, while the 364D amplicon (genotype D6) had two unique SNPs, namely, cytosine-to-adenosine transversions at nucleotides 147 and 255. All isolates except those with genotype A6 (Sheila Smith, Sawtooth, Bitterroot, Lost Horse Canyon, and Morgan) have a three-nucleotide insertion at consensus nucleotides 69 to 71 (ATA). The isolates with genotype B6 (all Central and South American isolates examined) can be distinguished from the remaining isolates (genotype C6) by a guanine-to-thymine transversion of nucleotide 131 in genotype C6.

**Phylogenetic analysis.** When the sequences of all six loci were concatenated and compared, the 38 isolates could be separated into nine genotypes. Maximum parsimony analysis of the concatenated sequences revealed the presence of five phylogenetic clades. Isolates 364D (unique at six of six loci analyzed) and Hlp#2 (unique at five of six loci analyzed) have unique genotypes both with respect to each other and compared to the other *R. rickettsii* isolates and were thus separated with great reliability (Fig. 2). The largest clade, clade I, contains 27 isolates in three closely related genotypes, comprising the eight Arizona isolates, a small group of 6 isolates (84JG, Hino, OSU 83-13-4, OSU 84-21c, 76RC, and Hauke), and the remaining 13 isolates. Clade II contains two genotypes, comprised of isolates Sawtooth and Sheila Smith and isolates Bit-
terroot, Lost Horse Canyon, and Morgan. A related clade, clade III, is comprised of two genotypes from Central and South America, one with three isolates (Brazil, Costa Rica, and Panama2004) and the other with only one isolate (Colombia). The closely related genotypes in each of these well-separated clades were distinguished only by the number of cytosines in the polycytosine region of the cspA-ksgA amplicon.

**DISCUSSION**

Fournier et al. examined the sequences of 52 IGRs and found only 4 that showed polymorphisms among 38 isolates of *R. conorii* subsp. *conorii* (21). One of these IGRs failed to provide additional discriminatory power to the analysis and was excluded from the study. Utilizing the remaining three IGRs (15, 5, and 2 genotypes), Fournier and colleagues were able to identify 27 unique genotypes. Phylogenetic analysis of these sequences grouped the isolates into three clusters, two of which correlated well with the geographic distribution of the isolates. In 2004, Zhu et al. used the same technique to type isolates of *R. prowazekii* (49). After analyzing 25 IGRs, only 2 contained nucleotide differences at the isolate level. These two IGRs contained only two SNPs, and one isolate exhibited a single 81-nucleotide repeat not found in the other isolates. Using these differences, Zhu and colleagues were able to identify four genotypes in a collection of 15 isolates. Two clusters were formed by phylogenetic analysis; however, no correlation between the phylogenetic groupings and the epidemiological characteristics of the *R. prowazekii* isolates was reported. However, another genetic method providing robust geographic groupings for *R. prowazekii* was recently developed (13a).

In the present study, the genotypic analysis and the phylogenetic tree resulting from maximum parsimony analysis of the concatenated sequences of six IGRs divided the 38 isolates of *R. rickettsii* into nine unique genotypes and five phylogenetic clades. There is a strong correlation between the phylogenetic grouping of most of the isolates and their geographic origin (Fig. 2). The three largest clades with multiple isolates contain all 36 isolates classically identified as *R. rickettsii* by serotyping and genotyping of the GltA and rOmpA genes. The largest of these clades contains three genotypes: genotype 1 contains only those isolates collected during an RMSF outbreak associated with *R. sanguineus* ticks in Arizona, while genotypes 2 and 3 (except isolate PriceT, which is a *D. andersoni* isolate from Montana [38]) encompass isolates collected from the Midwestern and Eastern United States that are associated with *D. variabilis*. Genotype 2 and 3 isolates are not separated geographically, and both types include patient isolates from western states (Hino from Oklahoma and 80JC from Nebraska). The second clade includes two genotypes that include the Sheila Smith isolate and *D. andersoni* ticks originating from the Bitterroot Valley of Montana, as well as a patient isolate from North Carolina (Morgan). Central and South American isolates are found in two genotypes in clade III.

With two exceptions, the genotyping system developed appears to associate specific genotypes with the different tick vectors of *R. rickettsii*. Among the eight isolates collected during a recent outbreak of RMSF in southeastern Arizona (14), five were isolated directly from *R. sanguineus* ticks, including one (AZ-7) that was removed from a human, and the other three isolates were all recovered from clinical specimens taken from humans who suffered from RMSF and were associated epidemiologically with *R. sanguineus* (15). Of the 20 isolates placed into genotypes 2 and 3, 13 were isolated from human RMSF patients from Ohio, Virginia, North Carolina, Maryland, Georgia, Nebraska, and Oklahoma. Twelve of these are thought to be associated with the tick *D. variabilis* because of where the patients resided when they became infected. The records for one of these patients mention that one *D. variabilis* tick was removed from the patient in the hospital. The three *R. rickettsii* isolates collected from animals in Virginia (vole, opossum, and rabbit) also have the same association based on tick habitat ranges and the collection of *D. variabilis* ticks from these sites (14, 44). Three additional isolates of *R. rickettsii* in this cluster were isolated directly from *D. variabilis* ticks collected from three different counties in Ohio. Three of the five
isolates found in the two genotypes of clade II were isolated from the primary vector of RMSF in the Northwestern United States, i.e., *D. andersoni*. However, *D. andersoni* is not found in North Carolina, where isolate Morgan was obtained (35). Similarly, the PriceT isolate recovered from *D. andersoni* in Montana (38) did not group with this cluster. A separate phylogenetic analysis of the same DNA stocks using analysis of variable-number-tandem-repeat loci also grouped PriceT and Morgan similarly to the IGR analysis (18a; Eremeeva, unpublished data). PCR-RFLP analysis of the GltA gene in 12 *R. rickettsii* isolates also grouped isolate Morgan with isolates Bitterroot, Sheila Smith, and Lost Horse Canyon (18). Thus, it appears that these clusters reflect the present fundamental genotypic characteristics of the isolate stocks available to us. At present, it cannot be excluded that isolates Morgan and PriceT may have been cross-contaminated with other isolates of *R. rickettsii* or mislabeled in the laboratory during their long passage and handling histories. It will be necessary to characterize additional new isolates of *R. rickettsii*, particularly from *D. andersoni*, to determine if other inconsistencies in vector-genotype associations exist. Clade III contains two samples with different passage histories of an isolate collected in Brazil prior to 1943, as well as isolates collected in Costa Rica, Colombia, and Panama. The Colombia and Panama2004 isolates are from patients, but the specific origins of the Costa Rica and Brazil isolates are less certain. *D. andersoni* and *D. variabilis* ticks are not found in Central and South America (19). It is thought that the most likely vector of *R. rickettsii* in these regions is a member of the genus *Amblyomma*, with most evidence pointing to *A. cajennense* or *A. aureolatum* (23, 37, 42).

Since the initial isolations of Hlp#2 and 364D, these iso-
lates, particularly Hlp#2, have frequently been associated with the species *R. rickettsii* despite several known differences between them and prototypical *R. rickettsii* isolates (2, 3, 32, 35). In their original paper describing rickettsial isolates from the rabbit tick, Parker and coworkers recovered seven Hlp isolates, of which they refer to only two by name (group 2 and group 3) (32). They reported that all seven Hlp isolates acted similarly in guinea pigs. In subsequent studies, some authors refer only to using isolate Hlp obtained from Parker, while others refer to isolate Hlp#. It is thought that most studies in fact used Hlp#, but this cannot be confirmed. In their paper, Parker et al. noted that although the Hlp isolates appeared serologically identical and provided protective immunity to other *R. rickettsii* isolates, there was a marked difference in the virulence of Hlp types observed in the guinea pig infection model (32). Whereas an unidentified “laboratory strain of Rocky Mountain spotted fever rickettsiae” showed fever and scrotal pathology typical of *R. rickettsii* infection, the Hlp isolates showed increased incubation periods, increased duration and decreased degree of fever, and decreased scrotal involvement. Anacker et al. (3) obtained similar results and noted that Hlp was even less virulent in the guinea pig than established *R. rickettsii* isolates (Morgan and Simpson) with decreased virulence. In 2001, Erremeeva et al. examined the cytotoxic effects of rickettsial infection on human endothelial cells and found that Hlp#2 actually caused more cellular injury than did some typical isolates of *R. rickettsii* (17).

Limited information exists concerning the relationship between isolate 364D and other isolates of *R. rickettsii*. 364D was originally isolated from *D. occidentalis* in Ventura County, CA, in 1966 (35). Early serological work by Philip et al. determined that while Hlp is closely related to *R. rickettsii* and Hlp#2, 364D is actually a different serotype from both (35). Later work showed that this isolate exhibits low virulence in a guinea pig model, has cytotoxicity in Vero cells, and kills chick embryos in 4 to 5 days after yolk sac inoculation (36). It has been suggested that 364D may also cause disease in humans, since analysis of convalescent-phase sera from patients in California thought to have RMSF showed specific antibodies to 364D serotype antigen (27).

Our data argue for a strong species or subspecies differentiation between Hlp#2 and 364D and between these two isolates and the other *R. rickettsii* isolates tested. Neither genotype has been isolated from patients manifesting symptoms of spotted fever. For five of the six loci analyzed for all isolates used in this study, Hlp#2 exhibits a genotype that is unique from all other isolates tested, while 364D exhibits a unique genotype for all six loci. There is high bootstrap support for an early divergence of both Hlp#2 and 364D from the other isolates (Fig. 2). When all 28 IGR loci were included, 364D had 15 unique sequences while Hlp#2 had 14 unique sequences. At three additional loci, 364D and Hlp#2 shared a sequence that was different from those of all classical *R. rickettsii* isolates tested. Additional genomic analysis of isolates of serotypes 364D and Hlp#2 is being done to determine if this strong differentiation of Hlp#2 and 364D isolates really indicates that they are novel subspecies lineages within the species *R. rickettsii* or if they should be reclassified as novel rickettsial species that are closely related to *R. rickettsii* (20). This distinction will be important because it will help to define whether these isolates are not isolates of *R. rickettsii* and do not need to be treated as select agents in the United States.

ACKNOWLEDGMENTS

This research was supported in part by an appointment of S. Karpathy to the Emerging Infectious Diseases (EID) Fellowship Program administered by the Association of Public Health Laboratories (APHL) and funded by the Centers for Disease Control and Prevention (CDC).

The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

We thank the many investigators who isolated and contributed the isolates used in this investigation. We also thank Christopher Paddock for his critical review of the manuscript and his insightful comments.

REFERENCES


17. Eremeeva, M. E., R. M. Klemt, L. A. Santucci-Donomot, D. J. Silverman, and...


Parker, R. R., E. G. Pickens, D. B. Lackman, E. J. Bell, and F. B. Thraikill.

Page, R. D. M.