An Analysis of the Genomic Variability of the Phytopathogenic Mollicute Spiroplasma kunkelii

Pablo Carpane, Ulrich Melcher, Astri Wayadande, María de la Paz Gimenez Pecci, Graciela Laguna, William Dolezal, and Jacqueline Fletcher

ABSTRACT

Corn stunt disease has become a factor limiting maize production in some areas of the Americas in recent years. Although resistant maize genotypes have been developed in the past, this resistance has been unstable over time or in some geographical locations. To better understand disease components that could affect the stability of host resistance, we assessed the genome variability of the etiologic agent, Spiroplasma kunkelii. Isolates were obtained from a number of areas, and characterized molecularly by amplification of several regions of the spiroplasma chromosome and sequencing of specific gene fragments. The degree of polymorphism between isolates of different geographic origins was low, and the level of genomic variability was similar within isolates of different countries. Polymorphism among isolates was found in viral insertions and in the sequence of Skarp, a gene that encodes a membrane protein implicated in attachment to insect cells. The results suggest that the genome composition of this species is highly conserved among isolates. Hence, it is unlikely that the instability of maize resistance is due to generation of new pathotypes of S. kunkelii. Instead, other components of this complex pathosystem could account for the breakdown of resistance.

The mollicute Spiroplasma kunkelii is one of three phytopathogenic spiroplasmas known, together with S. citri and S. phoeniceum (20). S. kunkelii is transmitted in a propagative manner by Dalbulus maidis and, less frequently, by several other leafhopper species (38), and is one of the causal agents of corn stunt, an important maize disease in the Americas (3,8,39). The prevalence of this disease has been increasing in recent years due to the extension of almost year-long maize cultivation (26,51), and to the lack of resistant maize germplasm (50).

Although maize genotypes resistant to corn stunt have been obtained by selecting disease-free plants in the field (47,48), such genotypes have usually become susceptible after a short period of time or in geographical locations other than those for which they were developed (50). At present, neither the target(s) nor the mechanism(s) of resistance have been characterized in maize, and it is not known whether the insect vector or the pathogen itself can overcome such resistance. For this reason, as part of a strategy to understand the components of corn stunt disease, we sought to estimate genomic variability of S. kunkelii, and to infer its potential ability to generate new strains that could overcome host resistance.

Several mechanisms that cause variability in the genome of spiroplasmas and that could ultimately lead to appearance of new pathotypes have been identified. Changes in the DNA sequence can occur because of nucleotide substitutions during DNA replication or repair, deletion, rearrangement or duplication of DNA fragments (33), integration of extrachromosomal DNA into the spiroplasma chromosome (42,43,49), or DNA exchange between spiroplasma cells (5). In particular, spiroplasma viruses can be an important source of genomic variability (33,34,36,57), because viral DNA sequences usually integrate into the chromosome in one or more copies, either intact or fragmented (1,34,36,49,57).

Additionally, the genes of several membrane-associated proteins identified in the genomes of Spiroplasma species (7,14,58) are similar to those of highly variable proteins in Mycoplasma species (10,12,44) and are potentially variable in Spiroplasma spp. as well.

Although a high level of genomic variability is characteristic of other members of the class mollicutes, such as Mycoplasma species (10,17,44), information available on the genus Spiroplasma is limited and unlikely to reflect a clear picture of its variability. For example, strains of S. citri collected in different countries or different years were not consistently distinguished using a REP-polymerase chain reaction (PCR) approach (37); but were discriminated by random amplified polymorphic DNA (RAPD)-PCR (35) and by analysis of viral insertions located along the spiroplasma chromosome (6). Moreover, this species can undergo rearrangements of its DNA sequence over brief time intervals when subjected to different propagation histories (21).

In the only characterization of S. kunkelii isolates performed so far, the sequence of the gene encoding the membrane protein spiralin was almost identical across isolates collected in different areas of Brazil. However, the recent availability of the almost-complete genome sequence of S. kunkelii CR2-3X (http://www.genome.ou.edu/spiro.html) and the availability of a set of isolates covering a broad geographical range should provide a better understanding of genome variability of this species.

The hypothesis tested in this work was that the genome compositions of S. kunkelii isolates of different geographical origins...
vary substantially. The objectives were to estimate the DNA sequence variability of *S. kunkelii* isolates from diverse geographical and ecological habitats in order to assess the potential of this species to generate new genotypes.

**MATERIALS AND METHODS**

*Spiroplasma isolation, culture, and DNA extraction.* Strains of *S. kunkelii* were obtained, through cooperation with researchers in Latin America and the United States, from a variety of different geographical locations and collection dates (Table 1). Spiroplasmas were isolated from symptomatic plants, which had been tested previously by PCR using primers F2-R6 (4) to confirm the presence of *S. kunkelii*.

The spiroplasmas were isolated from leaf midribs of maize and cultured in LD8A3 broth as described previously (30). Spiroplasma cultures were incubated at 31°C until the titer reached about 10^8 cells/ml, when they were aliquoted and stored (–80°C). Total DNA was extracted using the CTAB method (16).

**Detection of DNA polymorphisms.** Several primer pairs were used to amplify different regions of the spiroplasma genome (Table 2). BOX, REP, and ERIC primers (31) recognize repetitive elements present in Eubacteria. Primers recognizing the conserved ends of viral insertions (20) were designed to detect polymorphism in the distance between viral insertions: R8A2BR (5′-agatgtggtgacatttactaca-3′) and R8A2BL (5′-atttgtgtgactttc-3′) for virus SpV1 R8A2B; and C74R (5′-tgtacatttggagtmgttg-3′), C74L1 (5′-aaaaaaaataatgctttgctt-3′), and C74L2 (5′-tattctcctagaataaac-3′) for virus SpV1 C74. Reverse complementary primers (with the same name plus the letter “i” at the end) were designed to detect polymorphism in the length of viral insertions. In addition to primer pairs, 10-mer primers (Operon Technologies, Alameda, CA) were chosen randomly (35) for RAPD analysis.

PCR tests were performed using 20 ng of DNA template and 2 units of GoTaq DNA polymerase (Promega, Madison, WI) in a reaction volume of 25 µl, with 1× buffer, provided by the supplier. The components of the final reaction included: 3 mM MgCl2, 1.25 mM of each dNTP, and 4 µM of each primer. Reactions were initially incubated for 10 min at 95°C followed by 35 cycles of 1 min at 94°C, 1 min at each corresponding temperature (for primers annealing, Table 2), and 5 min at 72°C. The reaction ended with an extension step for 10 min at 72°C. PCR products were electrophoresed in 3% agarose gels at 10 V/cm for 90 min, stained with ethidium bromide, and visualized under ultraviolet light (46).

**Sequencing of specific DNA segments.** Selected isolates were submitted to the OSU Recombinant DNA/Protein Core Facility for sequencing of selected DNA segments by the Sanger dideoxy method (see results). Amplification of several membrane-associated protein genes was performed using primers F2-R6 (60°C) for the gene encoding spiralin (4), 32F1-32R (66°C) for the p32 gene (28), SkB-SkABC (47°C) for the gene ScA, coding for P58 (12), and P89L-P89R (60°C) for Skarp1, the gene encoding P89 (P89L: 5′-gatctgaaaaatggagga-3′ and P89R: 5′-ctcagaaaataggacagag-3′). Viral insertions were amplified using the primers R8A2BRI-R8A2BLi (62°C) and C74RI-C74LIi (60°C). PCs were performed as described above, using for each primer combination the annealing temperature indicated between parentheses, and an extension time of 2 min. Single bands of the expected sizes were obtained in all cases. Specific PCR products were purified from amplification components with a GENECLEAN SPIN Kit (BIO 101, Vista, CA), to remove salts and enzymes required for the amplifications, and were used as templates for sequencing using the amplification primers to prime synthesis.

**Data analysis.** Paired primer PCR and RAPD tests were performed three times each, and only those bands present in all three replications were considered in the analysis. The presence-absence of bands in these tests was assessed visually and converted to binary-coded matrices. For the calculation of pairwise genetic distances, the RAPD scores and the frequencies of bands were combined. The data were analyzed using a simple matching algorithm, and unweighted pair group method using arithmetic averages (UPGMA). The band presence/absence data were used to construct a dendrogram by a MEGA (version 4) program (8). 

**TABLE 1. Isolates of *Spiroplasma kunkelii* used in this study**

<table>
<thead>
<tr>
<th>Isolate</th>
<th>State</th>
<th>Country</th>
<th>Date collected</th>
<th>Number of passage</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR01</td>
<td>Tucumán</td>
<td>Argentina</td>
<td>7/28/2002</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>AR02-AR11</td>
<td>Chaco</td>
<td>Argentina</td>
<td>1/5/2006</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>AR12-AR14</td>
<td>Santiago del Estero</td>
<td>Argentina</td>
<td>1/5/2006</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>AR15</td>
<td>Santa Fe</td>
<td>Argentina</td>
<td>5/20/2006</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>AR16</td>
<td>Córdoba</td>
<td>Argentina</td>
<td>5/20/2006</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>AR17</td>
<td>Tucumán</td>
<td>Argentina</td>
<td>5/20/2006</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>AR18-AR23</td>
<td>Córdoba</td>
<td>Argentina</td>
<td>1/3/2007</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>BR01-BR03</td>
<td>Minas Gerais</td>
<td>Brazil</td>
<td>3/5/2007</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>CA01-CA31</td>
<td>California</td>
<td>United States</td>
<td>10/2/2006</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>FL01</td>
<td>Florida</td>
<td>United States</td>
<td>5/25/1987</td>
<td>Unknown</td>
<td>(35)</td>
</tr>
<tr>
<td>ME01</td>
<td>Unknown</td>
<td>Mexico</td>
<td>Unknown</td>
<td>Unknown</td>
<td>This study</td>
</tr>
<tr>
<td>ME02-ME06</td>
<td>Jalisco</td>
<td>Mexico</td>
<td>11/28/2005</td>
<td>1</td>
<td>This study</td>
</tr>
<tr>
<td>R8A2B (S. citri)</td>
<td>Unknown</td>
<td>Morocco</td>
<td>1985</td>
<td>Unknown</td>
<td>(35)</td>
</tr>
<tr>
<td>23-6 (S. floricola)</td>
<td>United States</td>
<td>Maryland</td>
<td>1978</td>
<td>Unknown</td>
<td>(13)</td>
</tr>
<tr>
<td>TS2 (S. melliferum)</td>
<td>Maryland</td>
<td>United States</td>
<td>1/6/1982</td>
<td>9</td>
<td>(11)</td>
</tr>
<tr>
<td><em>S. phoenicum</em></td>
<td>Unknown</td>
<td>Unknown</td>
<td>1982</td>
<td>42</td>
<td>(35)</td>
</tr>
</tbody>
</table>

* a Number of passage after isolation of tissues at which DNA was extracted for polymorphism detection.
into a binary form, (1) for presence and (0) for absence. Molecular Genetic Evolutionary Analysis (MEGA4, 53) was used to generate a phylogenetic tree using the neighbor-joining (45) and maximum parsimony (18) algorithms. To determine the robustness of the phylogenetic tree, the data were bootstrapped 1,000 times (19). *S. melliferum*, the spiroplasma species closest to *S. kunkelii* and *S. citri*, was used as the outgroup. Genetic diversity (Δ) was estimated by determining diversity for all the pairs of isolates for each geographical location, using a Jukes-Cantor model in Mega4. For nucleotide sequences, alignments, bootstrapping and construction of consensus phylogenetic trees by maximum parsimony (MP) were done with Mega4. The maximum parsimony tree was obtained using the close-neighbor-interchange algorithm, with prior exclusion of uninformative characters.

**RESULTS**

Among the 20 primer pairs or single primers tested, a total of 80 bands were generated for evaluation for polymorphisms. Of these, eight exhibited polymorphisms, one primer pair (C74R-C74L1) being responsible for two polymorphisms (Table 2). Four polymorphic loci were amplified by RAPD primers, while four polymorphic loci were amplified by targeting conserved ends of viral insertions, three of them being polymorphic for distance between two consecutive insertions. Consistent with the failure of BOX, ERIC, and REP primers to reveal polymorphisms in *S. citri* DNA (35), no *S. kunkelii* polymorphisms were detected using these primer pairs.

Figure 1 shows the phylogenetic tree generated from the molecular analyses of polymorphisms in the different amplification reactions. Although spiroplasma species were outgrouped somewhat more reliably (bootstrap values about 40%), the distance among isolates was so small that bootstrap separation was weak, which indicates a close relationship among all *S. kunkelii* isolates. The low bootstrap values at most of the branches indicate that even the largest distinctions among isolates were not reliable. No positive correlation could be drawn between geographical origin of isolates and structure in different clades of the tree, although most of the isolates from Argentina tended to be grouped together, as were a significant proportion of the California isolates. The six Mexican isolates seemed to be more dissimilar from one another.

To quantify the genetic diversity, average pair wise diversity (Δ) was estimated for each geographic location from which more than one isolate was collected. The fact that values for California (0.310 ± 0.303), Argentina (0.244 ± 0.225), Mexico (0.230 ± 0.159), and Brazil (0.084 ± 0.073) were similar to each other.

![Fig. 1. Phylogenetic tree of *Spiroplasma kunkelii* isolates constructed by molecular data obtained by analysis of random amplified polymorphic DNA and virus insertions and inferred using the maximum parsimony method. The percentage of replicated trees in which the associated isolates clustered together in the bootstrap test (1,000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. *S. melliferum* was used as the outgroup.](image-url)
indicates that the genome structure in all populations is similarly diverse.

To further characterize genomic variability of S. kunkelii, DNA regions expected to be variable among specific isolates were sequenced and compared. Isolates (Table 1) ME01 through ME06; BR01 through BR03; AR01 through AR03; CA01, CA02, and CA24; CR2 and FL01 were selected for this analysis because they represent different geographical origins (from the United States to Argentina) and different collection dates and were collected in areas expected to differ in pathogen variability. Mexico is hypothesized to have the largest variability since it is the geographic origin of maize.

Using primers designed to detect the length of viral insertions, and PCR conditions of high stringency, single amplification products were obtained. The ampiclon of the viral insertion R8A2B, which was 100% identical across all isolates sequenced, is part of the putative transposase gene (ORF 3) of a virus related to SpV1 R8A2B (accession no. NC_001365.1), sharing 680 of 720 residues with it. Results of the sequence of the SpV1 C74 viral insertion were unreliable, because more than one nucleotide residue per position was detected, possibly due to the presence of multiple fragments of similar size that differed in their internal sequences. Amplification products from the Sp1 gene (spiralin) and the ScA gene (P58) were also identical across isolates sequenced, matching that on the S. kunkelii database (www.genome.ou.edu/spiro.html). No amplification was obtained from the p32 gene in the isolates of S. kunkelii included in this work or from S. citri isolates R8A2 or BR3-3X.

In relation to the Skarp1 gene (P89), as the sequences of the initial set of isolates were polymorphic, the whole set of isolates (Table 1) was submitted to sequence analysis. A reliable sequence was obtained from a few isolates (Fig. 2). By using less stringent conditions during PCR (60 and 55°C of annealing temperature), amplification products were obtained from isolates AR02, AR08, AR22, ME01, ME03, and FL01, while no amplification products were obtained from the remaining isolates. However, because these sequences had more than one nucleotide residue per position they were not useful for comparison. Figure 2 shows that the isolates ME06, BR01, and CR2 have exactly the same sequence as that of the Skarp1 gene (NC_006400.1). Variability was found among the other isolates, although it was supported by only seven parsimony informative sites, providing only weak support for the inference of relatedness among them. Additionally, the isolates AR01, AR03, AR04, AR06, AR07, AR09, AR10, AR12, AR17, CA15, and CA21 gave an amplification product whose sequences were identical to Skarp2, a similar gene found in S. kunkelii (DQ333197.1).

**DISCUSSION**

In this work, we used several approaches to assess genome variability among isolates of S. kunkelii, with the ultimate goal of identifying molecular signatures or profiles that could be correlated with geographic origin or date of isolation. In addition to RAPDs using 17 primers that were effective to identify variation among S. citri isolates (35), we also tested the primers ERIC, REP, and BOX, commonly used to characterize bacterial species and isolates (31). The lack of strain discrimination of this method, in our hands, could be due to the high G+C content of these primers (over 65%), which could result in their recognition of conserved parts of the S. kunkelii genome which is rich in A+T. Additionally, we expected that viral insertions could be useful for discriminating among S. kunkelii isolates, because the genome sequence of strain CR2 reveals that these insertions are abundant and of different lengths. However, only four polymorphic fragments were obtained from virus-based primers in this study, unlike abundant polymorphisms detected in other plant-pathogenic bacteria when repetitive elements were used as probes (23,40).

After integration of extrachromosomal DNA into bacterial chromosomes, gradual changes in the nucleotide sequence may occur over time if the integrated DNA provides no selective advantage to the host organism (25). In our work, the finding that sequence of part of the gene coding for a putative transposase of viral origin is 100% identical across isolates could indicate that either this sequence plays an important role in the metabolism of S. kunkelii, or that mechanisms of DNA replication and repair are highly efficient. The fact that the integrated sequence is different from that of the virus (40 out of 720 nucleotide residues changed) may indicate that this sequence was first modified after its insertion, and was stably maintained thereafter. Alternatively, this gene could be variable in the SpV1 R8A2B virus, with the one inserted in the S. kunkelii genome differing from that sequenced (NC_001365.1).

Sequences of S. kunkelii genes encoding for the membrane proteins spiralin and P58 are identical across isolates. Similarly, low variability (98% identical) was found among isolates of S. kunkelii from different regions of Brazil (24). In contrast, the gene for spiralin in S. citri (22) showed positive selection for amino acid changes. Unfortunately, those sequences are not available to determine if those nucleotide substitutions lead to positive selection for the gene encoding spiralin in S. kunkelii.

Using the primers for the plasmid-located Skarp1 gene (NC_006400.1), the sequence obtained corresponded to the target gene in most of the isolates. However, in some other isolates the product amplified was from Skarp2, a closely related gene located in the spiroplasma chromosome (54). In the latter isolates the ampiclon was identical (100% identity) to that of isolate CR2, obtained from Costa Rica in 1992. In contrast, there was some variability in the Skarp1 sequence among Mexican isolates. Because maize is thought to have originated in Mexico (15), and since isolates obtained from other geographical locations were identical to each other, this variability could reflect longer-term evolution in association with plant and insect hosts. Unfortunately, the number of informative sites for this gene was too low to draw definitive conclusions about S. kunkelii phylogeny. Because of its location in a plasmid-located operon also containing genes similar to those of a type IV secretion system (1,14), it also may function in conjugation. The failure to detect this gene in some S. kunkelii isolates, despite using a variety of PCR conditions, could be due to mutation in the sites complementary to the primers, or to loss of the plasmid during isolation and culture. Considering that the Skarp1 gene product has been associated with attachment to

![Fig. 2. Phylogenetic tree of Spiroplasma kunkelii isolates inferred using the maximum parsimony (MP) for the sequence of part of the Skarp1 gene. The MP tree was obtained using the close-neighbor interchange method. The tree is drawn to scale, with branch lengths in the units of the number of changes over the whole sequence. All positions containing gaps and missing data were eliminated from the data set. There were a total of 775 positions in the final data set, out of which seven were parsimony informative. Skarp2 was considered the outgroup.](image-url)
insect cells (58), and that spiroplasmas adhere to insect epithelial cells lining the gut before moving into the hemocoele, it would be interesting to test the insect transmissibility of these isolates.

Regarding the gene encoding the membrane-associated protein P32 (29), no amplification product was obtained from the S. citri isolate used, nor S. kunkelii isolates, including CR2, which retains insect transmissibility (9). Furthermore, the S. kunkelii database contains no significant match to the p32 gene as annotated in S. citri (Gene ID: 3678468). Hence, it is likely that this gene is present either in the part of the chromosome remaining to be sequenced (5% as of October 2009), or in the extrachromosomal DNA, as in S. citri (28), and hence potentially lost during isolation and culture of these isolates. For this reason, it could be useful to confirm these results with Southern hybridization, to determine if S. kunkelii indeed has a copy of this or a similar gene, specifically in the insect-transmissible strain CR2-3X.

The high conservation of genes encoding membrane proteins of S. kunkelii was unexpected. Based on the similarity of some of these proteins to highly variable mycoplasma adhesins (10,32,44), we anticipated finding similar variability among membrane-located spiroplasma proteins. However, that proteins spiralin (29), P58 (56), and P89 (58) are associated with insect transmission could indicate a selection pressure to avoid changes in sequence. In this sense, mycoplasmas can avoid being recognized by the immune system of mammals by changing the DNA sequence coding for their membrane proteins (44). Alternatively, although there is no direct evidence for S. kunkelii and the leafhopper D. maidis, it has been found that S. paulsonii is not recognized by the immune system of its insect host, Drosophila melanogaster (27). If this were the case for D. maidis, then the resulting selection pressure would be directed to maintain the sequences of these proteins, as seen in this study.

Genome composition analyses by REP-PCR, RAPD, and sequencing of several genes indicate that all of the S. kunkelii isolates included in this study are genetically highly similar, despite their origin from different geographical locations over a time span of 24 years. These results are in contrast to previous findings in S. citri, a closely related phytopathogenic spiroplasma species, in which abundant variability was characterized (33,41,43,55,57). Many of the isolates included in the S. citri studies had been maintained for extended periods of time outside the insect and/or plant hosts, either by grafting (21,57), or by culturing (6,41,43,57). In the latter case, viral DNA can become progressively integrated into the spiroplasma chromosome (43), accounting for some of the variability observed. Furthermore, as a consequence of culture or grafting maintenance, isolates can lose insect transmissibility (53), and hence would not play an important role in natural variability observed in the field. In our study of S. kunkelii strains, we used low-passage cultures, passage 1 in most of the cases; the maximum passage number (nine) was strain CR2, which retained highly transmissible in recent biological tests (9). Hence, these strains should constitute a relatively accurate representation of existing S. kunkelii genome variability in nature. Our results, obtained using low-passage S. kunkelii isolates, resemble those reported by Mello et al. (35), in which the low-passage isolates of S. citri yielded 17 out of 159 (10.7%) polymorphic amplification products, compared to 8 out of 80 (10%) in S. kunkelii. Furthermore, Mello and co-workers found no correlation between molecular phenotype and geographical location or collection date, similar to our finding for S. kunkelii.

The high level of similarity observed between S. kunkelii isolates limits the potential to make inferences about the spread of this pathogen in both space and time. It is possible that the analysis of other genes or genome parts not considered in our study (that will be facilitated by the eventual publication of the whole genome sequence), could reveal variability not seen by us. However, these analyses suggest that S. kunkelii avoids high rates of mutation, resulting in a relatively low rate of genomic change. Hence, it is unlikely that the observed reemergence of corn stunt in formerly resistant maize cultivars was caused by spiroplasma genomic variation leading to the overcoming of host resistance, but rather to changes in the other elements of the corn stunt pathosystem such as insect vector dynamics or environmental conditions. Furthermore, these results suggest that maize resistance targeted to S. kunkelii could be stable in both time and geographical locations, as implied by the low variability detected in the genome of this pathogen.

ACKNOWLEDGMENTS

This project, funded by a Pioneer Hi-Bred International Scholarship, was part of the senior author’s Ph.D. dissertation research at Oklahoma State University.

LITERATURE CITED

Spiralin polymorphism in strains of *Spiroplasma citri* is not due to differences in posttranslational palmitoylation. J. Bacteriol. 178:2934-2940.


