CHARACTERIZATION OF THE *SPIROPLASMA* P58 MULTIGENE FAMILY

By

JANA LYNN COMER

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1999

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in Partial fulfillment of
the Degree of
MASTER OF SCIENCE
May, 2008

CHARACTERIZATION OF THE *SPIROPLASMA* P58 MULTIGENE FAMILY

Thesis Approved:

Dr. Ulrich Melcher

Thesis Advisor

Dr. Robert Matts

Dr. Steven Hartson

Dr. A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGEMENTS

This work could not have been accomplished without the support, encouragement, and guidance of my family, colleagues, and friends. The efforts of many individuals have contributed to my professional development, and I would particularly like to thank the following: My adviser, Dr. Ulrich Melcher, who in addition to his devotion to teaching and keen insight has fostered in me a lasting interest in learning. Dr. Jacqueline Fletcher for her crucial guidance on this project as well as being my mentor and friend. My former and present committee members, Dr. Robert Matts, Dr. Steve Hartson, Dr. Carol Bender, and Drs. Richard and Margaret Essenberg for their direction and advice.

I would like to thank Dr. Bharat Joshi, Dr. Anamika Ray, and all of my labmates and friends who have provided me with lively discussions, encouragement, and frank advice. In addition, I would like to gratefully acknowledge Dr. FengchunYe, Dr. Bharat Joshi, and Ananda Chandrasekaran for sharing their unpublished data.

Most of all, I thank my husband, Gene Comer, and my family for their time, love, and patience.

TABLE OF CONTENTS

CHAPTER I 1

INTRODUCTION 1
Background 1
Significant Chromosome Rearrangement in *S. citri* BR3 Lines 1
Immunological Detection of P58 in BR3 Lines 4
P58 Multigene Families 4

LITERATURE REVIEW 10
Taxonomy 10
Multigene Families 11
Spiroplasma and *Mycoplasma* Viruses 12
Host Range 14
Adhesins 15
Symptoms *in Planta* 17
Spiroplasma Sequencing Projects 18

CHAPTER III 20

EVOLUTION OF THE SPIROPLASMA P58 MULTIGENE FAMILY 20
Introduction 20
Materials and Methods 21
Results 25
Discussion 41
Future Work 45

LITERATURE CITED 48

APPENDIX A 54

EVALUATING TRANSPOSON MUTAGENESIS FOR GENETIC ANALYSIS OF MOLLICUTE (S. citri) GENOMIC DNA 54
Introduction 54
Materials and Methods 55
Results 56
Discussion 58

LITERATURE CITED 59

LIST OF TABLES

TABLE 1. PRIMERS USED TO VALIDATE RECOMBINATION JUNCTIONS. 24

TABLE 2. SHINE DALGARNO SEQUENCES. 28

TABLE 3. SIMILARITY BETWEEN *S. CITRI* AND *S. KUNKELII* GENES...................................... 36

TABLE 4. PORTAL PROTEIN AND SMALL SUBUNIT TERMINASE POSITIONS IN *S. KUNKELII* GENOME. 38

LIST OF FIGURES

FIGURE 1. *S. CITRI* BR3-G DELETION AREA I. 3

FIGURE 2. *S. CITRI* AND *S. KUNKELII* P58 AND FLANKING REGION. 25

FIGURE 3. PREDICTED OPEN READING FRAMES (ORFS) OF P58 AND P123. 27

FIGURE 4. P58 NUCLEOTIDE PERCENT IDENTITY OF 5' AND 3' SECTIONS. 30

FIGURE 5. P58 AMINO ACID PERCENT IDENTITY OF N-TERMINAL AND C-TERMINAL SECTIONS. 32

FIGURE 6. P58 REGIONS AND RECOMBINATION JUNCTIONS. 34

NOMENCLATURE

arp1 gene	Adherence Related Protein 1 gene
arp2 gene	Adherence Related Protein 2 gene
C-terminal	carboxyl terminus
Contig	contiguous nucleotide segment
kb	kilobase
kDa	kilodalton
MALDI-TOF	Matrix-Assisted LASER Desorption/Ionization Time of Flight
MAViT	Multiple Alignment Visualization Tool
MSRO	*Melanogaster* Sex Ratio Organism
mol%	mole percent
nm	nanometers
nt	nucleotides
NSRO	*Nebulosa* Sex Ratio Organism
N-terminal	amine terminus
ORF	Open Reading Frame
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
2D-PFGE	Two Dimensional Pulsed Field Gel Electrophoresis
PSI-BLAST	Position-Specific Iterated Basic Local Alignment Search Tool
rDNA	ribosomal deoxyribonucleic acid
rRNA	ribosomal ribonucleic acid
SARP1	Spiroplasma Adherence Related Protein 1
ScA	*S. citri* P58 sequence A
ScA-P58	*S. citri* P58 sequence A protein product
ScARP	*S. citri* Adhesin Related Protein
scarp gene	*S. citri* adherence related protein gene
ScB	*S. citri* P58 sequence B
ScB-P58	*S. citri* P58 sequence B protein product
ScC	*S. citri* P58 sequence C
ScC-P58	*S. citri* P58 sequence C protein product
SD	Shine Dalgarno
SkA	*S. kunkelii* P58 sequence A
SkA-P58	*S. kunkelii* P58 sequence A protein product
SkB	*S. kunkelii* P58 sequence B

SkB-P58	*S. kunkelii* P58 sequence B protein product
SkC	*S. kunkelii* P58 sequence C
SkC-P58	*S. kunkelii* P58 sequence C protein product
SkD	*S. kunkelii* P58 sequence D
SkD-P58	*S. kunkelii* P58 sequence D protein product
SkE	*S. kunkelii* P58 sequence E
SkE-P58	*S. kunkelii* P58 sequence E protein product
SkF	*S. kunkelii* P58 sequence F
SkF-P58	*S. kunkelii* P58 sequence F protein product
SkG	*S. kunkelii* P58 sequence G
SkG-P58	*S. kunkelii* P58 sequence G protein product
tBLASTn	protein-nucleotide 6-frame translation Basic Local Alignment Search Tool
tRNA	transfer ribonucleic acid
Vaa	variable adherence associated

CHAPTER I

INTRODUCTION

Background

Spiroplasma kunkelii and *S. citri* are closely related wall-less prokaryotic phytopathogens. *S. kunkelii*, the causative agent of corn stunt disease of maize, and *S. citri*, a citrus and crucifer pathogen, were the first plant pathogenic "mycoplasma-like organisms" to be discovered and cultured *in vitro*, respectively (Davis *et al.* 1972; Saglio *et al.* 1971; Saglio *et al.* 1973). *S. citri* BR3-3X, a phytopathogen, was isolated from horseradish with brittle root disease and is transmitted by a phloem-feeding leafhopper, *Circulifer tenellus* (De Almeida *et al.* 1997). Spiroplasmas lack a cell wall and are bound by a single trilaminar cell membrane (Saglio *et al.* 1971). Like other members of the class Mollicutes, *S. citri* has a small genome, 1,600 kb in *S. citri* BR3-3X (Ye *et al.* 1996). Spiroplasmas and their close relatives, the mycoplasmas, use UGA to encode tryptophan rather than as a stop codon.

Significant Chromosome Rearrangement in *S. citri* BR3 Lines

S. citri has an unstable genome characterized by frequent chromosomal deletions, insertions, and inversions (Melcher *et al.* 1999). Viral sequences, such as those of SpV1-R8A2 B and SpV1-C74, as well as homologous and site specific recombination have changed the organization of the *S. citri* chromosome. Integrated insertion elements

of SpV1-R8A2 B and SpV1-C74 viruses, belonging to IS3 and IS30 families, respectively, and putatively encoding transposases, are present in *S. citri*. One sequence, which arose during attempts to infect *S. citri* with a SpV1 viral DNA carrying an insert, resulted in two homologous recombination events (Marais *et al.* 1996). In addition, symmetric inverted repeats, a feature of site specific recombination, were present in some sequences. In this case, recombination occurred between circular free viral replicative form and SpV1 viral sequences present in the host, *S. citri*, chromosome.

Extensive chromosomal rearrangement has been found in one derivative of *S. citri* BR3-3X during lengthy maintenance (Ye *et al.* 1996). *S. citri* BR3 was originally isolated from horseradish plants with brittle root disease (Fletcher *et al.* 1981). One derivative, BR3-G, which was maintained by plant to plant transmission through grafting, has lost insect transmissibility. Another derivative, BR3-T, was maintained by transmission through the natural insect vector *C. tenellus*, and it retained insect transmissibility.

When *S. citri* BR3-3X and the two derivatives, BR3-G and BR3-T were analyzed by pulsed field gel electrophoresis (PFGE), an apparent increase in genome size was detected: 1600 kb in BR3-3X, 1870 kb in BR3-G, and 1750 in BR3-T {Ye, 1996 #81}. Hybridization with the virus SpV1-R8A2 DNA revealed multiple bands in all lines, and very different patterns among the three lines. Construction of physical maps revealed markers in opposite order in a region between BR3-3X and BR3-G, suggesting a genomic inversion, and missing markers at the borders of the inversion suggested deletions at the inversion borders. A 9.6 kb segment of BR3-3X corresponding to one of the areas deleted from BR3-G was sequenced (Figure 1).

Figure 1. *S. citri* BR3-G deletion area I.
The 9.6 kb contiguous sequence of *S. citri* BR3-3X that includes one of the areas deleted in *S. citri* BR3-G. A transposase (Tn'ase) is located 5' of the deleted area. Five ORFs were in the deleted area, and predicted to encode proteins of 58-, 12-, 54-, and 123 kDa and an incomplete, interrupted SpV1-related C4 ORF (C4-like). An ORF predicted to encode a protein of 18 kDa is 3' of the deleted area.

One of the open reading frames (ORFs) in the deleted region was predicted to encode a protein of 58 kDa that has limited amino acid sequence similarity with *Mycoplasma hominis* P50 and *Mycoplasma genitalium* MG191, both adhesins. Other ORFs in the deleted region are predicted to encode proteins of 12, 54, and 123 kDa. An SpV1-R8A2 B (Renaudin *et al.* 1990) related bacteriophage transposase gene is located 5' of the deletion area, and an incomplete, interrupted SpV1-related C4 ORF (Melcher *et al.* 1999) is located 3' of the deletion area. The large genomic inversion and two deletion areas alter the genome of BR3-G significantly from that of BR3-3X. The characterized deletion area contains several potentially functional genes, and the other deletion area may as well. Genetic rearrangements such as these suggest an extreme instability in the spiroplasma genome.

Sequences surrounding P58 contain several features needed for gene expression (Ye *et al.* 1997). A putative Shine-Dalgarno sequence, GAGG, is present nine nt upstream of the start codon ATG, 65 nt upstream is a putative Pribnow box, TATAAAT, preceded by a putative -35 region, TGTTGAGAAATA. Dot matrix analysis revealed a hairpin immediately downstream of the coding region, which could be a factor-independent transcription terminator. That P58, P12, P54, and P123 are transcribed

independently was suggested by northern hybridizations of total RNAs using the 9.6 kb segment as a probe.

Immunological Detection of P58 in BR3 Lines

Two anti-P58 antibodies were raised (Ye *et al.* 1997). One was against a synthetic peptide corresponding to residues 198-210 of the P58 protein. The other antibody was made by inserting a segment of the P58 gene, corresponding to amino acid residues 145-263, into the reading frame of a maltose binding protein gene, and using the purified fusion protein as an antigen. Both antibodies detected a protein of approximately 60 kDa from *S. citri* BR3-3X, BR3-T, and BR3-G. Using the anti-MBP-P58 serum in western blotting, a 60 kDa band was detected in both membrane and cytoplasmic fractions of *S. citri* BR3-3X, BR3-T, and BR3-G. When spiroplasma protein preparations were subjected to Triton X-114 phase partitioning, P58 was detected only in the detergent phase, and the 60 kDa band was missing from the membrane preparation after treatment with trypsin. Thus, P58 was detected in the membrane and at least partially (residues 145-263) on the cell surface. The role of P58 in insect interactions remained unclear, however, because P58 was detected in the insect non-transmissible BR3-G.

P58 Multigene Families

P58 multigene families exist in both *S. citri* and *S. kunkelii*. In addition to the characterized *S. citri* P58 gene (ScA), at least two other P58-like sequences (ScB, ScC) are present in *S. citri* BR3-3X (Ye *et al.* 1997). Southern blot analysis using a probe corresponding to the 5' region of P58 revealed three reactive bands in the insect

transmissible lines *S. citri* BR3-3X and BR3-T, but only two in the insect non-transmissible line BR3-G (Ye, unpublished data). At least one reactive band was present in all but one of the *S. citri* strains tested, and multiple reactive bands were present in *S. kunkelii* CSS-Mex (Joshi 2002, unpublished data). Partial P58 sequences have been reported from *S. insolitum, S. phoeniceum, S. poulsonii,* and the sex ratio organisms of *Drosophila nebulosa* and *D. melanogaster*, *Spiroplasma* sp NSRO and *Spiroplasma* sp MSRO (Montenegro *et al.* 2005). Multiple P58-like sequences may be the result of a mechanism of antigenic variation, or they may be remnants of viral sequences that integrated into the *Spiroplasma* chromosome.

That the P58 multigene family may play a role in insect interaction was suggested by the observation that the characterized *S. citri* P58 gene is not present in the insect non-transmissible line *S. citri* BR3-G. However, because the characterized P58 gene was deleted in BR3-G, but a P58 protein was detected in BR3-G, we know that in this line at least one other P58-like gene is transcribed and its mRNA is translated (Ye *et al.* 1997). Thus, the role of P58 in insect interactions remains unclear.

Research Objectives

The central question of this work is: Does P58 function as an adhesin? In many instances, this question would be answered with a knockout mutation then complementation of the gene in question. However, genetic manipulation by homologous recombination is inefficient in *Spiroplasma* as compared to many bacteria (Razin *et al.* 1998). This is likely due at least in part to lack of a functional *recA* gene (Carle *et al.* 2006; Marais *et al.* 1996; Roe *et al.* 2002). However, homologous

recombination does occur (Melcher *et al.* 1999), and has been used to produce knockout mutants (Andre *et al.* 2005; Duret *et al.* 2003; Gaurivaud *et al.* 2000; Lartigue *et al.* 2002). Among the approaches attempted were the use of *oriC* plasmids as disruption vectors (Duret *et al.* 1999; Gaurivaud *et al.* 2000). However, the *oriC* plasmids frequently integrated at the *oriC* origin of replication instead of at the target gene site (Duret *et al.* 1999). To minimize this, the length of the *oriC* sequence was reduced and tetracycline resistance was conferred after a single cross over event (Lartigue *et al.* 2002). Because P58 is a multigene family, in addition to the difficulties of making one knockout mutant, one would have to address the issues of knocking out all three genes. Because of these limitations, options other than knockout mutants were explored.

That two SpV1-related elements flank the region including one of the P58 genes (ScA) (Ye *et al.* 1996) suggests that multiple genomic sequences could also be remnants of inserted viral elements. Multiple integrated sequences of spiroplasma viruses have, in part, led to the frequent rearrangement of genomes of several strains of *S. citri* (Melcher *et al.* 1999). Insertion of IS elements, and site-specific and homologous recombination, involving SpV1-related sequences, may contribute to genomic variation in several *Spiroplasma* species.

To answer whether P58 functions as an adhesin, the following four questions were posed:

1. What are the gene structures of ScB and ScC?

If ScB and ScC function as adhesins, one would expect them to be full length and free of any frameshift or premature stop mutations. In addition, elements such as promoters, terminators, and other elements likely needed

for expression would be expected to be present, and adhesins would be expected to be transcribed. In contrast, inserted viral elements may have mutations that render them non-functional.

2. How do the P58 genes of *S. citri* compare to those of *S. kunkelii*?

Both *S. citri* and *S. kunkelii* infect leafhopper vectors in a persistent and circulative manner, and thus, they must pass similar physical barriers within the insect body. The adhesins employed by *S. citri* and *S. kunkelii* may be similar. However, because spiroplasma-insect interactions are complex and specific, it would not be surprising if the adhesins involved in *S. citri-C. tenellus* interactions differ from those in *S. kunkelii-Dalbulus maidis* interactions (Fletcher *et al.* 2006).

3. Are the P58 genes viral in origin?

Spiroplasma viruses SpV1 and SpV3 have been documented to insert into the genome of spiroplasmas, including *S. citri* and *S. kunkelii* (Renaudin and Bove 1994). In addition, some SpV1 virus genomes have likely evolved by recombination (Sha *et al.* 2000). One of the segments deleted in the *S. citri* BR3-G chromosomal rearrangement was flanked by SpV1 viral elements (Ye *et al.* 1996). If P58 is an inserted viral element, one would expect to find, in addition to random mutations, similarity to other viral genes.

4. Has recombination occurred within the P58 genes?

Intragenic recombination is a common mechanism of antigenic variation in mycoplasmas (Razin 1999), but is less common among

bacteriophages, in which intergenic recombination often occurs, but the gene is conserved (Desiere *et al.* 2001). Stretches of DNA sequence between polymorphisms that are significantly longer than would be expected to occur by chance is one way to indicate probable recombination.

These research objectives will provide circumstantial evidence that the P58 multigene family either plays a role in an adhesion or is an inserted viral element. Specifically, research objective 1 will rule out mutations that would render the P58 genes non-functional. Research objective 2 will determine whether a P58 multigene family is present in *S. kunkelii*, and whether the genes are full length and free of any mutations that may render them non-functional. Objective 2 will also determine evolutionary relationships between members of the *S. kunkelii* and *S. citri* P58 multigene families. Objective 3 will detect relationships to viral elements, of which both tailed viruses and filamentous viruses have been detected in *S. citri* and *S. kunkelii*. If found, intragenic variation, objective 4, might provide clues that the P58 multigene family was capable of recombination consistent with antigenic variation. This study will not provide a direct correlation between the presence of P58 and attachment to insect cells or transmission to plants. However, if the results of all objectives suggest a role as an adhesin for one or more members of the P58 multigene family, the basis for future work would be enhanced. Evidence that P58 is an inserted viral element would create uncertainty about P58's function as an adhesin. Regardless of its role as an adhesin, study of the P58 multigene family will provide insight to the mechanisms of chromosomal rearrangement,

recombination, and genetic flux of spiroplasmas, which could influence the understanding of bacterial evolution.

CHAPTER II

LITERATURE REVIEW

Taxonomy

S. citri and *S. kunkelii* belong to the class *Mollicutes*, which arose from the walled, gram-positive Clostridial lineage of Eubacteria (*Clostridium ramosum* and *C. innocuum*) approximately 600 million years ago, about the same time as the proliferation of arthropods (Gasparich 2002). Mollicutes evolved to their present genome size of 530 kb (Bermuda grass white leaf phytoplasma) (Lee *et al.* 2000) to 2,220 kb (*S. ixodetis*) (Gasparich 2002) genomes from an ancestor of about 2,000 kb (Razin *et al.* 1998). Mollicutes have characteristically low G + C content, ranging from 24 to 33 mol%. All members of the genus *Spiroplasma* can be filtered through 220 nm pores, ferment glucose, are resistant to penicillin, require cholesterol, are unable to hydrolyze urea, and are helical and motile (Gasparich 2002). In addition to the characters required for genus classification, 16S rDNA determination and serological determination are required for species classification.

Mollicutes lack several enzyme pathways typical of most prokaryotes, including those for cell wall biosynthesis, the tricarboxylic acid cycle, and electron transport (Razin *et al.* 1998). Genome sequences of seven mycoplasmas and *Ureaplasma parvum*, formerly *U. urealyticum*, reveal that these mollicutes, except *M. mycoides*, a mycoplasma closely related to spiroplasmas, lack most of the genes for synthesis of purines,

pyrimidines, and nucleotides (Chambaud *et al.* 2001; Fraser *et al.* 1995; Glass *et al.* 2000; Himmelreich *et al.* 1996; Jaffe *et al.* 2004; Minion *et al.* 2004; Papazisi *et al.* 2003; Sasaki *et al.* 2002). However, *S. kunkelii* strain CR2-3X (Zhao *et al.* 2003) has some genes in these pathways including those for cytidine deaminase, cytidylate kinase, GMP reductase, purine-nucleoside phosphorylase, thymidine kinase, thymidine phosphorylase, thymidylate synthase and uridine kinase (Melcher 1999, unpublished data), and strain CSS-M has those for adenylosuccinate synthase, deoxyguanosine kinase, and folylpolyglutamate synthase/dihydrofolate synthetase (Bai and Hogenhout 2002).

Multigene Families

Mollicutes lack some multigene families that are present in other bacteria. For example, most spiroplasmas have a single rRNA operon (Amikam *et al.* 1984). In contrast, *Escherichia coli* and *Bacillus subtilis* have seven and ten rRNA operons, respectively (Bove 1993). Duplication of mycoplasma tRNA is less prevalent than tRNA duplication in other bacteria. However, multigene families do exist in mycoplasmas, and many contribute to antigenic variation (Razin *et al.* 1998). Several mechanisms, including genomic rearrangement, lead to the variable expression of mycoplasma surface proteins (Citti and Rosengarten 1997).

The P50 adhesin of *M. hominis* is one example of the mechanism of variable expression through genomic rearrangement. Antigenic variation of *M. hominis* P50 and its homolog, Vaa (variable adherence associated), in strain 1620 is accomplished through both insertion/deletion of interchangeable cassettes in the domain-encoding portion of the gene, and mutation in poly A tracts of the gene (Boesen *et al.* 2001). P50 is a surface

exposed membrane protein from *Mycoplasma hominis* strain FBG that, when blocked by monoclonal antibodies, resulted in *M. hominis* metabolism inhibition and reduced viability as measured by reduction in arginine catabolism and reduction in *M. hominis* CFU per mL as measured by plating serial dilutions (Feldman *et al.* 1992). That P50 is important in adhesion was demonstrated by the blocking of adherence of *M. hominis* to HeLa cells after addition of anti-P50 monoclonal antibody (Henrich *et al.* 1993). Immunoelectron microscopy of negatively stained *M. hominis* revealed that P50 is evenly distributed over the cell surface.

Spiroplasma and *Mycoplasma* Viruses

The four types of viruses that have been observed in spiroplasmas (SpV1 to SpV4) have the following morphologies: naked rod, long contractile tails, short contractile tails, and naked isometric, respectively. (Renaudin and Bove 1994). SpV4 viruses, which are about the same size as ΦX174, infect *S. melliferum*, and has a circular and single stranded DNA genome. SpV4 viruses apparently, like spiroplasmas and other mollicutes, use UGA as a codon for tryptophan instead of as a stop codon.

SpV1 viruses, like SpV4 viruses, have a circular, single-stranded DNA genome (Renaudin and Bove 1994). Although first observed in *S. citri*, SpV1 viruses have been found in several other spiroplasma species, including *S. melliferum*, *S. mirum*, and *Drosophila* spiroplasmas. Plectroviruses, of which SpV1 is one type, infect mollicutes (Buchen-Osmond 2002).

In an effort to identify conserved features of plectroviruses, sequences of two *S. citri* infecting plectroviruses, SpV1-R8A2 B, and SpV1-C74, were compared to the

sequence of a virus that infects *S. melliferum*, SVTS2 (Sha *et al.* 2000). Ten of the 14 SVTS2 ORFs were distinguishable as SpV1-R8A2 B/SpV1-C74 homologues by sequence alignment and genome order. However, while SpV1-R8A2 B and SpV1-C74, have transposase genes of the IS30 and IS3 families, no transposase homologue was found in SVTS2, and yet, the presence of an inverted imperfect repeat suggests that SVTS2 may be a circularized insertion element. Thus, comparisons of these three viruses indicate insertion is common among SpV1 viruses, and insertion of multiple viruses could create avenues for recombination and chromosome rearrangement.

Indeed, several copies of SpV1 virus sequences have been detected in the genomes of *S. citri* BR3 lines (Ye *et al.* 1996). When restricted DNA fragments of BR3-3X and two derivatives (BR3-G and BR3-T), were probed with SpV1-R8A2 virus DNA, the viral DNA hybridized to several fragments. SpV1 viruses have contributed to genome rearrangements (Melcher *et al.* 1999). In addition, multiple dispersed SpV1-like sequences throughout the *S. citri* chromosome have provided sites for homologous and site-specific recombination events.

SpV2 and SpV3 are two types of tailed bacteriophages, that have been detected in *S. citri* (Abedon and Ackermann 2000). In fact, tailed viruses were the first viruses to be discovered in spiroplasmas (Cole *et al.* 1973). SpV2 and SpV3 have long, non-contractile tails and short, non-contractile tails, respectively. SV-C2, the only representative of SpV2 observed in *S. citri*, has been observed in *S. citri* strain R8A2 preparations but has not been experimentally propagated; SV-C2 presumably has a dsDNA genome.

SpV3 viruses have dsDNA genomes of about 16 to 21 kb that may be circularly permuted (Cole *et al.* 1977; Dickinson and Townsend 1984; Dickinson *et al.* 1984). One

SpV3 virus, *ai*, is lysogenic and has a non-covalently closed genome with cohesive ends (Dickinson and Townsend 1984). *S. citri* strain BR6, which was isolated at the same time and geographic location as *S. citri* BR3, is resistant to infection by SpV3 viruses *ai, ag,* and AV9/3 (Dickinson *et al.* 1984). An SpV3 DNA probe hybridized with extrachromosomal elements from 11 of 14 *S. citri* strains tested (Gasparich *et al.* 1993).

Host Range

S. citri and *S. kunkelii* are transmitted by leafhopper vectors. One other spiroplasma, *S. phoeniceum*, which infects periwinkle, is phytopathogenic. In addition to the three phytopathogenic spiroplasmas, nonphytopathogenic spiroplasmas have been documented in true bugs, beetles, flies, bees, wasps, butterflies, and ticks, and crustaceans (Fletcher *et al.* 1998; Nunan *et al.* 2005; Wang *et al.* 2005). While most spiroplasmas maintain a benign relationship with their arthropod host, some cause disease. For example, *S. melliferum* causes spiroplasmosis in honeybees, *S. poulsonii* causes a sex ratio disorder by killing male progeny, and *S. mirum* causes suckling mouse cataract disease (Gasparich 2002). Although spiroplasmas are found in a wide variety of arthropods, plant mollicute vectors identified to date are phloem-feeding members of the class Homoptera and are either leafhoppers or psyllids (Fletcher *et al.* 1998). Spiroplasmas of some arthropods remain in the intestine (orders Trichoptera, Diptera, Lepidoptera, and Coleoptera), but some (such as the order Mecoptera) are able to enter the hemolymph. *S. citri* is transmitted in the United States primarily by *C. tenellus* (Melcher and Fletcher 1999), and *S. kunkelii* is transmitted primarily by *D. maidis*. The phytopathogenic spiroplasmas' transmission is both propagative and persistent (Fletcher

et al. 1998). Thus, the spiroplasmas enter the gut, must cross the gut barriers, such as the plasmalemma and basal lamina, multiply in the hemolymph, and traverse salivary gland barriers to be deposited in a new host plant.

Adhesins

In mycoplasmas, tip structures are important for attachment to and orientation of the host cell (Rottem 2003). Mycoplasma tip structures are made of a complex array of adhesins, proteins involved in receptor binding, and accessory proteins, proteins necessary for cytadherence but not directly involved in receptor binding. The tip organelle of *Mycoplasma pneumoniae* is a tapered cell extension containing an electron dense core (Biberfeld and Biberfeld 1970). Tip structures have been observed in *S. kunkelii* CSS-M*, S. melliferum, S. citri* Asp1*, and S. floricola* (Ammar *et al.* 2004). In thin sections of *D. elimatus*, which had acquired *S. kunkelii* CSS-M by feeding on diseased plants, the spiroplasma helices oriented perpendicular to the apical plasma membrane of the insect midgut epithelial cell, their tips were aligned between the cell microvilli. An electron dense core was noted in this tip structure. In addition, pilus-like structures were described on the surface of *S. melliferum* and *S. kunkelii* CSS-M. An electron micrograph showed an anti-58 kDa antibody-tagged proteins localized on what may be tip structures of *S. citri* BR3 (Fletcher 1981, unpublished data).

Mycoplasmas are able to move by gliding and glide in the direction of their tip structures (McBride 2001). *M. mobile*, the fastest moving mycoplasma, glides by a five step process: initial binding, tight binding, stroke, release, and return (Seto *et al.* 2005). To accomplish this, a “spike” or “leg” protein, Gli349, binds to the surface (Uenoyama *et*

al. 2004) while force is generated by or through another protein, Gli521 (Seto *et al.* 2005). A third protein, Gli123, acts to position the proteins at the base of the tip structure (Uenoyama and Miyata 2005), and the system is driven by a novel nucleoside triphosphatase, P42 (Ohtani and Miyata 2006).

One putative adhesin, SARP1, or "*Spiroplasma* adherence related protein 1," was discovered when a reduction in the amount of an 89 kDa protein in a total cell protein preparation, afrter treatment with proteinase K or pronase coincided with a reduction of *S. citri* binding to a monolayer of *C. tenellus* cells (Yu *et al.* 2000). Upon providing the *S. citri* with fresh medium, the 89 kDa protein regenerated, and binding returned to normal levels as well. The gene encoding SARP1, *arp*1, was cloned and sequenced (Berg *et al.* 2001), and SARP1 was found to contain an N-terminal signal peptide, 6 tandem repeats, and a transmembrane domain. The complete plasmid containing *arp*1 was sequenced and found to contain several open reading frames (ORFs), as well as a chromosomal homologue named *arp*2 (Joshi *et al.* 2005). Two *arp* genes have also been found in an *S. kunkelii,* one on a plasmid (pSKU146) (Davis *et al.* 2005), and another with similarity in the C-terminal region (Wei *et al.* 2006).

In studies on another *S. citri* strain, GII3, eight genes designated *scarp* were found on plasmids (pSci1-5) (Berho *et al.* 2006). Western blotting with anti-ScARP serum revealed reactivity in *S. citri* transmissible lines (GII3 and GII39a2) but not in non-transmissible lines (44, R8A2, ASP-1) (Berho *et al.* 2006). Another plasmid-encoded putative adhesin, the cytoplasmic P32 (Killiny *et al.* 2006), which is encoded by pSci6, was detected in *S. citri* transmissible lines (GII3, Corsica, Cyprus, Palmyra) but not in non-transmissible lines (R8A2, 44, 26, ASP-1) using 2D electrophoresis and MALDI-

TOF spectrometry. When the non-transmissible strain *S. citri* 44 was complemented with P32, spiroplasmas were found between the basal lamina and the plasmalemma, of cells containing secretion granules, whereas no degradation of the membrane was seen in the uncomplemented strain. However, P32 alone does not confer transmissibility. Complementation of the non-transmissible *S. citri* 44 with pSci6, which contains a gene encoding P32 and truncated versions of *scarp*, does confer transmissibility (Berho *et al.* 2006).

Symptoms *in Planta*

Genes important for spiroplasma transmission and proliferation can be detected by the study of mutants which do not produce symptoms in plants. *S. citri* mutants GMT 553 and GMT 470, obtained by random Tn4001 mutagenesis, do not produce symptoms in periwinkle plants exposed to microinjected *C. haematoceps* (Foissac *et al.* 1997). GMT 553 grew to high titers in the microinjected leafhoppers and was transmitted by the leafhoppers to plants or to Parafilm sachets; however, the time required to reach peak titer in the plant was twice as long as for wild type. The gene in which Tn4001 is inserted in GMT 553 was identified as belonging to the 5' end of the first gene the fructose operon, *fruR*, the activator of the fructose operon. In this mutant, none of the mRNA of the fructose operon was transcribed (Gaurivaud *et al.* 2001). GMT 470 multiplied poorly in the leafhopper, and was not transmitted to periwinkle plants. In addition, it grew more slowly than did the wild type GII3 or GMT 553 in broth medium. A calcium transporting ATPase is disrupted in GMT 470.

A xylitol-resistant *S. citri* mutant (xyl 3), also deficient in fructose utilization, showed no symptoms in periwinkle until week 10, whereas wildtype started showing symptoms in week 1, but reached high titers in planta (10^6 CFU/mL at week 5) (Gaurivaud *et al.* 2000). Xyl 3 was altered in the *fruA* gene, the gene encoding fructose permease of the phosphoenolpyruvate:fructose phosphotransferase system (fructose PTS). Delayed symptoms possibly due to lack of fructose utilization may be due to either excess lactic acid production or competition for fructose in the phloem, or both (Bove *et al.* 2003).

Spiroplasma Sequencing Projects

S. kunkelii CR2-3X is currently being sequenced through a collaborative project between the University of Oklahoma, Advanced Center for Genome Technology and the USDA-ARS in Beltsville, MD using a strain of *S. kunkelii* isolated by Jacqueline Fletcher, Oklahoma State University (Roe *et al.* 2002). The *S. kunkelii* genome has been sequenced with 9x coverage, is currently in 50 contiguous nucleotide segments (contigs), and is publicly available at http://www.genome.ou.edu/spiro.html. One contig, 85 kb long, contains several features, including gene information and storage, cellular functions, metabolism, and a possible phase variable region (Zhao *et al.* 2003). In addition, 16 putative ABC systems have been identified (Zhao *et al.* 2004), suggesting that membrane trafficking may be an important mechanism of acquiring nutrients for parasitic spiroplasmas. A physical and genetic map has been constructed using 1D and 2D pulsed field gel electrophoresis (PFGE) (Dally *et al.* 2006). *S. kunkelii* CR2-3X DNA was analyzed after digestion with six different enzymes in all possible pair combinations

and analyzed with 2D PFGE. In addition, 27 single copy genes were mapped by Southern hybridization. A single restriction site for I-*Ceu*I, which recognizes a highly conserved 26 kb sequence in the 23S rRNA genes of many bacteria (Liu *et al.* 1993), indicated the presence of a single rRNA operon. Ordering the contigs of the physical and genetic map will aid the completion of the *S. kunkelii* genome sequencing project.

The partial chromosome sequence of *S. citri* strain GII3-3X has recently been made publicly available (Carle *et al.* 2006). Seventy-seven chromosomal contigs were assembled, representing about 92 % of the chromosome. In addition, 8 circular contigs, one SVTS2 viral DNA sequence and 7 plasmids, were sequenced. Thirty-eight of the contigs were annotated, and the largest portion, besides those of unknown function (47.2%), were found to be of viral origin (20.5%), followed by macromolecule metabolism (10.8%), small molecule metabolism (8.2%), cellular processes (6.2%), transposon-related (3.6%), and structural elements (3.5%).

CHAPTER III

EVOLUTION OF THE *SPIROPLASMA* P58 MULTIGENE FAMILY[1]

Introduction

To investigate the gene structures of ScB and ScC and compare the P58 genes of *S. citri* to those of *S. kunkelii*, the P58 multigene family and flanking sequences of both *Spiroplasma* species were analyzed. *S. kunkelii* CR2-3X sequences were retrieved from the *S. kunkelii* genome sequencing project database, publicly available at www.genome.ou.edu/spiro.html. In addition to 9x sequence coverage, a physical and genetic map is available (Dally *et al.* 2006). Partial sequence of *S. citri* GII3-X is also available (Carle *et al.* 2006). Ninety-two percent of the chromosomal and 8 circular contigs are available in the NCBI Core Nucleotide Database (www.ncbi.nlm.gov). In the *S. citri* and *S. kunkelii* P58 families, regions of low percentage identity between some of the P58 family members were discovered in the 5' portion of the P58 sequences. To investigate a possible intragenic recombination event, these regions of low identity and the surrounding regions of high identity among some of the P58 family members were analyzed. Categories for the 5' (A or B) and 3' (X or Y) portions of the P58-like sequences are proposed. BLASTsearch revealed similarity to a terminase gene in portions of the P58 genes, and to a P50 adhesin from *Mycoplasma hominis* in other

---

[1] Portions of the following chapter have been published in *Biochemical Genetics* and appear in this dissertation with the journal's permission.

portions. A region of possible higher rate of recombination in which one or more recombination events occurred, including that of an adhesin and a bacteriophage gene, is discussed. Recombination events, such as those described here, could provide a mechanism for variation in spiroplasma evolution and in niche adaptation.

Materials and Methods

S. kunkelii genome data

S. kunkelii CR2-3X genome data was obtained from www.genome.ou.edu/spiro.html. The *S. kunkelii* Genome Sequencing Project is a collaborative project between the University of Oklahoma, Advanced Center for Genome Technology and the USDA-ARS in Beltsville, MD using a strain of *S. kunkelii* isolated by Jacqueline Fletcher at Oklahoma State University. Collection of genome sequence data was described by Zhao and colleagues (2003).

Using the *S. citri* predicted amino acid P58 sequence (Acc. No. AAB00186.1) as a query sequence, *S. kunkelii* sequences selected for study had expect values less than $1x10^{-50}$ from a tBLASTn search performed on 7-15-04. Contigs retrieved from the *S. kunkelii* database, designated herein by letters of the alphabet, were of the following sizes (in bp): SkA, 9,311; SkB, 63,951; SkC, 201,453; SkD, 90,800; SkE, 31,893; SkF, 4,221; SkG, 19,359. For ease of analysis, large *S. kunkelii* contigs (SkB, SkC, SkD, SkE, and SkG) were trimmed to no closer than 4kb to the P58 region.

Other P58-like sequences from *Spiroplasma* spp. used for comparison were as follows: *S. phoeniceum* (Acc. No. AJ628438), *S. insolitum* (Acc. No. AJ628437), *S.*

poulsonii (Acc. No. AJ628436), *Spiroplasma* sp. NSRO (Acc. No. AJ628435), and *Spiroplasma* sp. MSRO (Acc. No. AJ628434).

ScB and ScC PCR amplification

S. citri BR3-G was grown to exponential phase in LD8 broth medium (Chen and Davis 1979) at 31°C, and DNA was isolated using the CTAB method, as described by Berg and colleagues (2001). Regions on either side or in the P58 region were used to design primers for ScB and ScC. The forward primer 5'-CCATTTTTCTTTTTACAA-3', designed from a region in which ScA, SkA, SkB, SkC, SkF, and SkG were identical, was used in the primer pair for both ScB and ScC. The reverse primer 5'-TTTACCCATTAAACTTGC-3', designed from a region in which ScA, SkA, SkB, and SkC were identical, was used to amplify the ScB P58 sequence. The primer 5'-ACAGCACCTTGTACTTGT-3', designed from a region in which SkE, SkF, and SkG were identical, was used as the reverse primer to amplify ScC.

PCR of the ScB region was performed in a 50 µl reaction volume with final concentrations as follows: 1x reaction buffer (Promega), 200 µ*M* dNTPs, 3.5 m*M* $MgCl_2$, 0.485 µg *S. citri* BR3-G DNA for ScB, 0.2 µ*M* primers, and 2.5 u Taq (Promega). The reaction had an initial denaturation step of 1 min. at 95 °C, then 35 cycles as follows: 94 °C 30 sec., 42 °C 1 min., 72 °C 3 min. A final elongation step of 5 min. was performed at 72 °C. PCR of the ScC region was identical with the following exceptions: 1.12 µg *S. citri* BR3-3X DNA was used as template, elongation at 72 °C was 1 min. 30 sec., and the final elongation step was 10 min. The PCR products were cloned using a TOPO TA cloning kit (Invitrogen) and sequenced by primer walking.

Sequence comparisons

Only regions sharing greater than 70% with at least one other P58-like sequence were included in the comparisons. Global alignments were performed using ClustalX as implemented by AlignX (Invitrogen). A multiple sequence alignment of the *S. citri* and *S. kunkelii* CR2-3X P58 sequences was imported into MAViT (Multiple Alignment Visualization Tool, available at www.biochic.com. MAViT uses a sliding window of 100 nucleotides to calculate a percent identity between each pair of sequences of a multiple sequence alignment. Any percentage identity at or below 50% is shown as black. Percentage identity greater than 50% ranges from dark red to bright red.

Junctions between regions of high (>70%) and low (<70%) identity were determined. The pairwise percentage identities on either side of these boundaries were calculated. Pairwise predicted amino acid alignments were also assessed for percentage identity in N-terminal and C-terminal regions using the region boundaries.

Validation of recombination junctions

PCR primers chosen for validation of the recombination junctions are shown in Table 1.

Table 1. Primers used to validate recombination junctions.		
SkA[1]	(r)	GATTTCAAAAGTCATATC
SkB[1]	(r)	TTTAAGTTTGGTAAAGCG
SkC[1]	(r)	GAAACTGCCATCTTCTAA
SkE[1]	(f)	TAAAGATATAAATTTTTC
SkF[1]	(f)	AAAATAAATGAGGTAATA
SkG[1]	(f)	ATTTGTATCATTGCTAGT
SkABC[2]	(f)	GAAGAGAAGAATGTGACC

1. Primers specific to a recombination junction. 2. Primers shared among P58 sequences. r: reverse; f: forward

PCR reactions were carried out in a reaction volume of 50 µl with final concentrations as follows: 1x reaction buffer (Promega), 200 µ*M* dNTPs, 3.0 m*M* $MgCl_2$, 0.1 µg *S. kunkelii* CR2-3X DNA, 0.2 µ*M* primers, and 2.5 u Taq (Promega). Thermocycler conditions were as previously described, except for the use of 24 cycles with annealing temperatures as follows: SkA and SkG, 42°C; SkB and SkC, 45°C; SkE and SkF, 38°C. The final elongation step was 10 min. at 72°C.

Local PSI-BLASTs

BLASTpgp and BLASTall executables and the nonredundant database were downloaded from www.ncbi.nlm.nih.gov/BLAST. “Phage portal protein” or “phage terminase small subunit” was used as a query in a keyword search at www.ncbi.nlm.nih.gov. Recovered sequences were used as queries in local BLASTpgp searches against the nonredundant database, and the resulting matrices were used to search the *S. kunkelii* database using tBLASTn.

Comparison of *S. citri* and *S. kunkelii* genes

Predicted amino acid sequences of *S. citri* genes were used as a query in tBLASTn search of the *S. kunkelii* CR2-3X genome database. *S. kunkelii* sequences corresponding to the *S. citri* gene were retrieved and conceptually translated. Amino acid alignments were performed using ClustalX as implemented by AlignX.

Results

For discussion purposes, those sequences having similarity to P58 have been designated as regions, and the contigs containing P58-like regions have been assigned letters of the alphabet (ScA-ScC, and SkA-SkG). Flanking regions are sequences immediately surrounding the P58 region. The P58 region has been divided into section 1 and section 2 (Figure 2).

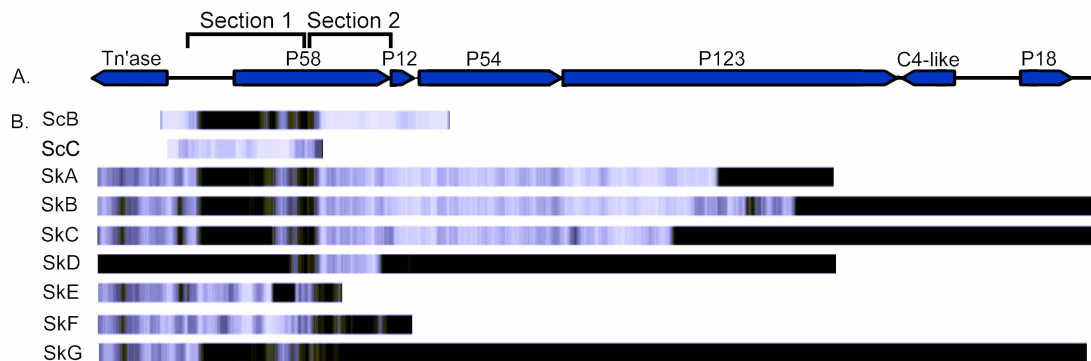

Figure 2. *S. citri* and *S. kunkelii* P58 and flanking region.
A. *S. citri* BR3-3X genome section containing P58 and adjacent ORFs.
B. P58-like genes and surrounding sections showing relative degree of similarity to *S. citri* sequence. Fifty percent identity and lower is shown in black. From 50% to 100% the shading ranges from dark blue to light blue.

Two new *S. citri* P58-like contigs, ScB and ScC, were cloned and sequenced (Acc. Nos. DQ344811 and DQ344812, respectively) and seven *S. kunkelii* P58 contigs (SkA-SkG) were detected in the *S. kunkelii* genome sequence (Figure 2B). Expect values of a tBLASTn search using the ScA P58 predicted amino acid sequence as a query were as follows: SkA, $1x10^{-178}$; SkB, $1x10^{-180}$; SkC, $1x10^{-149}$; SkD, $1x10^{-125}$; SkE, $3x10^{-55}$; SkF, $1x10^{-166}$; SkG, $1x10^{-71}$.

Frameshifts/deletions/premature stops

To assess the possibility that one or more of the P58 family members is a pseudogene, the P58 region and flanking regions of each contig were scanned for apparent frameshifts, premature stop codons, and deletions (Figure 3).

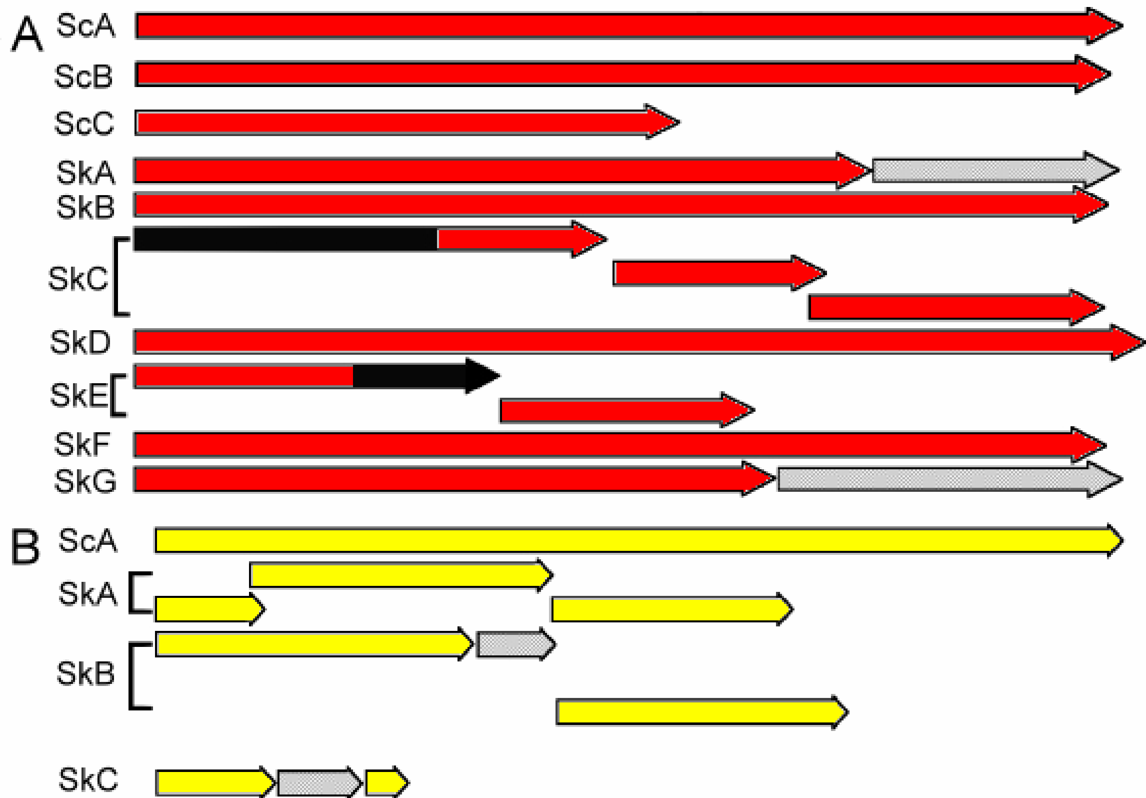

Figure 3. Predicted open reading frames (ORFs) of P58 and P123. *S. citri* and *S. kunkelii* ORFs of P58 (A) or P123 (B). Artemis Release 7 was used to mark ORFs of 100 amino acids or more. Shading represents sequence present after a stop codon. Black indicates a deletion as compared to ScA.

SkA contains a premature stop codon in the P58 gene at position 372, relative to the initial nucleotide of the first codon of the ORF. SkC contains one deletion beginning 5' of the P58 gene and extending into the 5' portion of the gene. Two frameshift mutations are also present in the P58 region of SkC, at positions 229 and 354. The P58 region of SkE contains a 163 nucleotide deletion after position 110 and a frameshift mutation at position 165. The SkF P58 ORF was approximately the same length as the ScA P58 ORF, 486 predicted amino acid residues, but its nucleic acid sequence was similar to that of ScA only in the 5' region of the P58 gene. A stop codon in SkG was present in the

P58 region at position 327. One *S. kunkelii* P58 gene, that on contig SkB, the *S. citri* P58 genes on ScB, and at least the sequenced portion of ScC were free of any premature stop, frameshift or deletion mutations and contained high sequence similarity (>70%) to that of another P58 was present throughout the entire reading frame.

As shown in Table 2, Shine Dalgarno (SD) sequences with at least 4 nucleotides of the sequence AGGAGG were present in all the P58 sequences except SkC.

Table 2.[1] Shine Dalgarno Sequences.

S. citri 16S		AGAAAG-GAGGTGATCC	
ScA	-19	ATAAAT-GAGGTAATTA	-4
ScC	-19	ATAAAT-GAGGTAATTA	-4
SkE	-19	ATAAAT-GAGGTAATAA	-4
SkF	-19	ATAAAT-GAGGTAATA	-4
SkG	-18	AAAAAGAGAG-TGATAAT	-2
SkA	-18	AAAAAGAGAG-TGATAAT	-2
SkB	-18	AAAAAGAGAG-TGATAAT	-2
ScB	-22	AGTAAG-GA--TGATAAGA	-7
SkD	-17	AAAAAG-GAG-TAAAAA	-1

1. Positions relative to P58 start codon.

Purine-rich sequences necessary for translation are present upstream of the P58-like genes. The SD sequences can be placed into three different groups. Those of ScA, ScC, SkE, and SkF are 100% identical from position -28 to -1, and include the sequence GAGG at -10. Those of SkA, SkB, and SkG are identical to each other from -30 to -1, and have four of the six nucleotides corresponding to the SD sequence. ScB has the sequence AGGA at position -15, and SkD has the sequence AGGAG at -7. Sequences representing portions of the SD sequence (AGGAGG) were present upstream of each of

the P58 ORFs except SkC, and may indicate that proteins could be translated from the P58 ORFs.

Mutations were present in regions flanking P58. Beginning at the P123 ORF, the SkA contig contains two frameshift mutations at positions 116 and 366. The frameshift mutation at position 366 is also present in the P123 ORF of *S. citri* R8A2 (Melcher et al. 1999). SkB contains one frameshift mutation codon at position 400. One premature stop mutation is present in the P123 region of contig SkC at position 127. Thus, of the ORFs detected downstream of the P58 region of ScA (P12, P54, and P123), similar sequences were detected, but mutations were present in potential ORFs found in the regions surrounding P58-like sequences in SkA, SkB, and SkC.

Neighborhoods

To corroborate the computer sequence assembly, PCR and then end sequencing were performed across the putative recombination junctions of the P58 regions of contigs SkA, SkB, SkC, SkE, SkF, and SkG. Due to the length and the number of polymorphisms in the P58 region of SkD, specific primers could not be designed for SkD. All PCRs generated products of expected size and end sequencing confirmed computer assembly.

P58 regions and flanking regions were examined for degrees of identity with each other. Sequences similar to an IS30 family transposase gene were located 5' of the P58-like regions on all contigs except SkD; and except in contig SkD, the high degree of similarity (77.0%-99.1% identity, by pairwise comparison at the nucleotide level) included part of the transposase-P58 intergenic region (Figure 2). The presence of

sequence similar to a viral related transposon in regions near P58, except SkD, and the additional viral sequence near the P58 sequence in ScA (the incomplete, interrupted SpV1-related C4 ORF) indicates the possibility that P58 or an ancestral sequence may be an inserted viral element.

The P58-like gene sequences were not all uniformly similar to each other, however. To analyze their similarity, subsections of the sequence were examined on both the nucleic acid and amino acid levels (Figure 2). As shown in Figure 4, section 1 (-357-673) of contigs ScA and ScC and two of the *S. kunkelii* contigs, SkE and SkF, shared 89.6%-96.7% identity.

		S. citri			*S. kunkelii*						
		A	B	C	A	B	C	D	E	F	G
S. citri	A		59.6	95.6	58.1	61.4	58.7	62.4	89.8	89.6	57.3
	B	99.2		56.8	82.7	82.6	83.1	79.8	56.0	58.3	82.7
	C	68.5	69.9		61.0	61.1	58.6	59.7	90.0	90.0	61.0
S. kunkelii	A	89.9	90.4	65.8		99.7	93.7	91.9	57.6	60.7	99.2
	B	90.0	90.1	65.8	99.2		93.8	92.3	57.8	61.9	99.0
	C	88.3	88.9	64.4	95.4	95.4		88.1	41.1	65.9	93.8
	D	90.1	90.9	66.0	99.6	99.5	97.5		63.7	62.3	91.5
	E	66.7	62.2	83.6	61.7	61.7	36.9	61.0		96.7	58.4
	F	60.2	60.0	83.6	58.7	59.9	58.3	59.7	95.0		61.6
	G	51.7	52.5	83.6	51.5	56.5	50.1	52.5	96.3	97.6	

Figure 4. P58 nucleotide percent identity of 5' and 3' sections. Pairwise nucleotide sequence alignment of *S. citri* and *S. kunkelii*. 5' and 3' sections of P58 were analyzed separately. Section 1, corresponding to part of the intergenic region 5' of P58 and the 5' half of P58 (see Figure 2A) is in the upper right and section 2, corresponding to the 3' half of P58 is in the lower left. Sequences with percent identities higher than 70% are shown in yellow; those less than 70% are shown in gray.

Conversely, contigs ScB, SkA, SkB, SkC, and SkG shared considerable identity (79.8%-99.7%) with one another, but were much less similar (41.1%-65.9%) to ScA, ScC, SkE, and SkF.

Section 2 (746-1484) (Figure 2) shows a different pattern of conservation among several members of the multigene family. In this region, the P58 gene ScA is more similar to the corresponding parts of contigs ScB, SkA, SkB, SkC, and SkD (88.3%-99.2%) than to those of ScC, SkE, SkF, and SkG (51.7%-68.5%). ScC and SkE are truncated within section 2, and their similarity with the other contigs can be assessed only to the point of the truncation. Within section 2, SkG shared 83.6%-97.6% identity with ScC, SkE, and SkF in a segment of 261 nucleotides. The percent identity between SkF and SkG was 49.6% in the remaining portion of section 2. SkF and SkG have little similarity to each other or to the other contigs 3' of section 2. SkA, SkB, and SkC shared sequence similar to ScA 3' of the P58-like region. The region of similarity extended into the P12, P54, and P123 ORFs of the SkA, SkB, and SkC contigs. Thus, ScA, ScC, SkE, and SkF have one type of 5' end and ScB, SkA, SkB, SkC, SkD, and SkG have another type of 5' end. Of the 3' ends, one type is present in ScB, SkA, SkB, SkC, SkD, SkE, and SkF, and another type is present in ScA, ScC, and SkG.

Because the junction of regions of different identity occurred in the middle of the P58 ORF, the N-terminal (corresponding to the coding sequence of section 1) and C-terminal (corresponding to section 2) halves of the P58 predicted amino acid sequences were analyzed separately (Figure 5).

		S. citri			*S. kunkelii*						
		A	B	C	A	B	C	D	E	F	G
S. citri	A		50.0	91.4	47.0	45.6	53.5	43.0	88.9	89.4	45.2
	B	99.2		48.4	73.4	73.9	71.5	78.6	45.9	49.3	73.9
	C	61.1	63.2		48.9	49.0	45.9	40.1	85.4	85.8	46.6
S. kunkelii	A	88.2	89.0	61.1		99.1	96.0	98.6	41.1	47.1	99.1
	B	88.7	89.4	61.1	98.4		96.2	100	41.1	47.1	99.1
	C	83.4	85.7	16.7	91.3	92.1		93.1	ND	55.0	96.9
	D	89.5	91.6	61.1	98.9	100	95.7		43.3	44.9	98.6
	E	40.4	39.4	88.9	36.2	37.2	27.9	42.7		99.4	27.2
	F	36.8	36.4	88.9	35.1	35.9	34.3	36.5	92.0		45.8
	G	37.5	36.4	88.9	34.1	35.2	33.0	35.2	93.8	93.8	

Figure 5. P58 amino acid percent identity of N-terminal and C-terminal sections.
Pairwise alignment percent identity of *S. citri* and *S. kunkelii* P58 predicted amino acid sequences. N-terminal and C-terminal sections of P58 predicted amino acid sequence were analyzed separately. N-terminal section of amino acid sequence in upper right. C-terminal section of amino acid sequence in lower left. Sequences with percent identities higher than 70% are shown in yellow; less than 70% are shown in gray. ND: not determined. An alignment of SkC-P58 and SkE-P58 C-terminal section was not possible due to the low percent identity and truncation of the SkE contig.

Here, conceptually translated amino acid sequences are named for their contigs. For example, the P58 ORF located on SkA was named SkA-P58. As shown in the upper right half of Figure 5, N-terminal (section 1) identities among ScA-P58, ScC-P58, SkE-P58, and SkF-P58 were high, ranging from 88.9% to 99.4%. The identities among ScB-P58, SkA-P58, SkB-P58, SkC-P58, SkD-P58, and SkG-P58 were also high, ranging from 71.5% to 100%. Thus, ScA-P58, ScC-P58, SkE-P58, and SkF-P58 have one type of N-

terminal end and ScB-P58, SkA-P58, SkB-P58, SkC-P58, and SkG-P58 have another type of N-terminal end.

In the C-terminal half of the molecule (section 2), as shown in the lower left of Figure 5, ScA-P58 was more similar to ScB-P58, SkA-P58, SkB-P58, SkC-P58, and SkD-P58 (83.4%-99.2%) than it was to ScC-P58, SkE-P58 (before the point at which the sequence is truncated), SkF-P58, or SkG-P58 (36.4%-61.1%). ScA-P58, ScB-P58, SkA-P58, SkB-P58, SkC-P58, and SkD-P58 share a high degree of predicted amino acid percent identity in the C-terminal region (83.4%-100%), and a low degree of identity with ScC-P58, SkE-P58, SkF-P58, and SkG-P58 (27.9%-63.2%). ScC-P58, SkE-P58, SkF-P58, and SkG-P58 share 88.9%-93.8% identity to each other in the C-terminal region. Considering just the C-terminal ends, ScA-P58, ScB-P58, SkA-P58, SkB-P58, SkC-P58, and SkD-P58 have one type of end, while ScC-P58, SkE-P58 and SkF-P58, and SkG have a different type of C-terminal end.

Among the various P58-like sequences, the putative recombination junctions within the P58 gene did not occur in precisely the same place (relative to ScA, as shown in Figure 6).

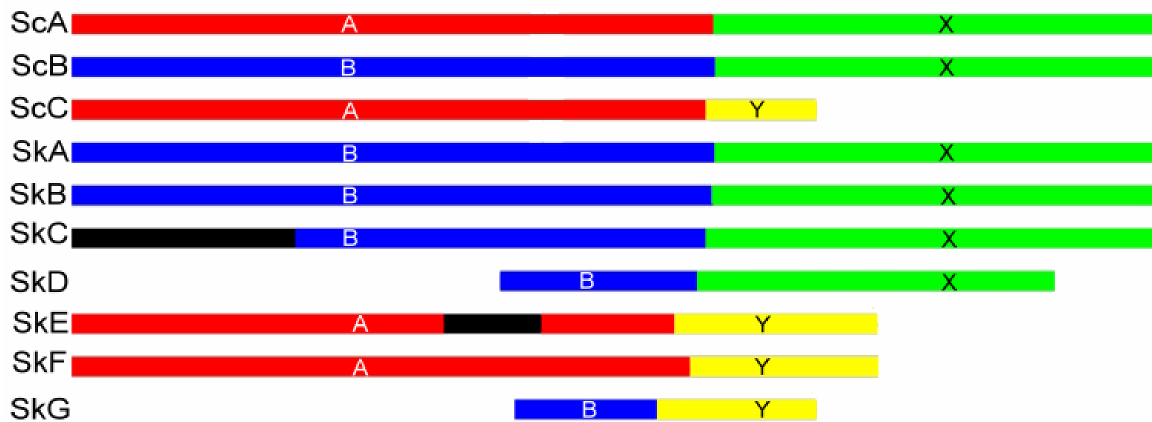

Figure 6. P58 regions and recombination junctions.
Regions colored similarly had greater than 70% identity at the nucleotide level. 5' P58 regions were of two types: "A" (red) or "B" (blue). Black boxes indicate a deletion relative to ScA. 3' regions were of two types: "X" (green) or "Y" (yellow).

While the ScB, SkA, and SkB junctions were present in close proximity to one another (positions 746, 745, and 740, respectively), the SkC and SkD junctions occurred at positions 729 and 715, respectively. ScC, SkE, SkF, and SkG were at more distant positions: 674, 678, 704, and 649, respectively. Junctions of differences in similarity likely represent recombination junctions, and the relative positions of these junctions may have implications for the mechanism of recombination.

GeneConv (Sawyer 1999) was used to detect probable recombination among members of the P58 multigene family. GeneConv uses a nucleotide substitution model to detect recombination, or gene conversion, events, by finding fragments (stretches of DNA sequence between polymorphisms) that are significantly longer than would be expected to occur by chance. A recombination fragment between ScA and ScB was detected (P=<0.05) in the P58 coding region from position 880 past the end of the P58

coding region to the end of the sequenced region. A single polymorphism is present in this fragment at position 1243. While not detected as a fragment by GeneConv, in the remaining portion of section 2 only two polymorphisms were present at positions 791 and 832. A recombination fragment was also detected in ScC from 674 to the end of the region sequenced. Thus, in addition to identification by inspection, recombination fragments were also detected by statistical methods.

Whether the range of high and low similarity levels among the P58 genes of *S. citri* and *S. kunkelii* was unusually large was unknown. As shown in Table 3, sequence alignment of each of 33 *S. citri* genes with their respective homologues in *S. kunkelii* CR2-3X revealed that most (79%) *S. citri* genes tested share a high degree of identity with the corresponding *S. kunkelii* genes (>80%), with the notable exception of the major membrane protein, spiralin (60.0%). That most of the genes shared between *S. citri* and *S. kunkelii* are greater than 80% identical indicates that low similarity, as found in the different types of P58 ends, is unusual; however, lower similarity in the major membrane protein, spiralin, indicates that variation may be more common among surface proteins.

Table 3. Similarity between *S. citri* and *S. kunkelii* genes.

Acc. No.	Gene	%ID	Length
AAC35871	ribosomal protein L16	99.3	137
BAA03630	DnaA	98.5	265
BAA03630	DnaA	98.2	265
CAA79524	DNA gyrase subunit A	97.8	227
CAA79524	DNA gyrase subunit A	97.8	227
AAC45548	(p)ppGpp3'-pyrophosphohydrolase	97.5	749
P47767	DNA-dependent RNA polymerase subunit beta	96.9	1302
AAC35870	ribosomal protein S3	96.8	252
AAC35875	ribosomal protein L24	96.6	58
AAC35869	ribosomal protein L22	96.4	112
AAC45550	alanyl-tRNA synthetase	96.3	136
AAL51057	mreB2	96.1	349
AAC35874	ribosomal protein L14	95.9	122
A43668	fibril protein	95.7	515
CAA79523	DNA gyrase subunit B	94.4	640
AAB69999	pyruvate kinase	94.3	157
AAC35873	ribosomal protein S17	94.1	85
AAF08320	Fructose-specific IIABC component	92.6	687
P20275	6-phosphofructokinase	92.4	327
AAF08321	1-phosphofructokinase	91	310
AAC45549	scm1	90.5	409
P19216	elongation factor Ts	89.9	296
CAA79522	DnaN protein	88.7	363
AAB69994	ribosomal protein S2	87.9	359
AAC35872	ribosomal protein L29-like protein	81.4	339
CAC10363	P89	65.0	798
AAB69997	Spiralin	60.0	241
CAC10363	P89	53.9	798
CAA79523	DNA gyrase subunit B	47.1	640
T43229	restriction modification system S subunit	40.5	294
AAF08319	Fructose operon transcriptional regulator	37.1	233
T43229	restriction modification system S subunit	36.7	294

1. Multiple listings of certain accession numbers reflect the multiple occurrences of these genes in *S. kunkelii*. 2. Number of amino acids. 3. Shaded area reflects percent identity less than 70%

BLAST searches

PSI-BLAST searches were performed using one representative of each type of N-terminal and C-terminal end as query (section 1 of ScA-P58 and SkA-P58, and section 2 of ScA-P58 and SkF-P58). PSI-BLAST revealed that section 1 of both ScA-P58 and SkA-P58 are closely related to the large subunit of bacteriophage terminases (expect: ScA-P58, Acc. No. NP_958178, $1x10^{-45}$, iteration 4; SkA-P58, Acc. No. YP_059360, $9x10^{-50}$, iteration 4). Bacteriophage terminases are composed of large and small subunits, and are responsible for packaging DNA into the heads of tailed viruses. Terminases function in strand separation and have activity as an ATPase, a site specific endonuclease, and a putative DNA translocase (Catalano 2000). Searches of the Prosite database revealed ATP/GTP binding domains in ScB-P58, SkA-P58, SkB-P58, and SkG-P58 but not in ScA-P58, SkC-P58, SkD-P58, SkE-P58, or SkF-P58. This indicates that section 1 of P58 may be the remnant of a portion of an inserted viral element of a tailed phage.

If functional bacteriophage terminases are present in *S. kunkelii*, then it is likely that tailed phage sequences occur in the spiroplasma genome. A tBLASTn search of the *S. kunkelii* database using *Streptococcus pyogenes* SSI-1 phage-associated single stranded binding protein (Acc. No. BAC63381) as a query revealed a phage gene sequence (expect $4x10^{-7}$) 5.2 kb downstream of the P58 region of SkC, one of four results having expect values less than 0.005. Another phage associated ssb protein hit (expect $1x10^{-9}$) was found 672 bp downstream of a phage associated recombinase (expect $2x10^{-30}$).

To further assess the presence of tailed phage sequences in the *S. kunkelii* genome, PSI-BLAST searches of the *S. kunkelii* database were performed using phage predicted amino acid sequences to form PSI-BLAST matrices (Table 4).

Table 4. Portal protein and small subunit terminase positions in *S. kunkelii* genome.

Portal Protein				Terminase Small Subunit			
Query	Contig	Expect	Length	Query	Contig	Expect	Length
EAL11265	SkC	$2x10^{-03}$	177	CAD43931	SkC	$6x10^{-4}$	177
NP_050632	SkC	$1x10^{-121}$	1329	NP_680543	SkC	$6x10^{-8}$	555
YP_151587	SkH	$4x10^{-04}$	300	NP815175	SkC	$9x10^{-9}$	609
AAV78275	SkI	$4x10^{-04}$	435	AAR83223	SkI	$1x10^{-111}$	1326
CAG43695	SkJ	$3x10^{-38}$	1088	NP_815175	SkJ	$5x10^{-4}$	279
NP_775254	SkJ	3x10-05	789	NP_680543	SkL	$4x10^{-4}$	174
EAL11265	SkM	$2x10^{-03}$	198	NP_680543	SkL	$3x10^{-6}$	516
YP_006983	SkN	$3x10^{-05}$	963	AAR83223	SkO	$4x10^{-5}$	966
AAL49568	SkO	$3x10^{-06}$	576	NP_680543	SkQ	$3x10^{-4}$	75
AAV89004	SkO	$2x10^{-05}$	255	NP_815175	SkU	$2x10^{-05}$	339
AAV78275	SkP	$3x10^{-03}$	189	NP_680543	SkW	$9x10^{-05}$	468
AAQ12193	SkP	$2x10^{-09}$	1005	AAR83223	SkX	$4x10^{-08}$	600
AAS02087	SkR	$6x10^{-04}$	552	P39786	SkZ	$1x10^{-06}$	1389
NP_050632	SkT	$1x10^{-08}$	1017	NP_815175	SkAA	$8x10^{-05}$	369
NP_775254	SkX	$1x10^{-03}$	450	NP_680543	SKbb	$3x10^{-03}$	369
CAH21051	SkX	$2x10^{-03}$	375	NP_680543	SkZ	$3x10^{-3}$	369
AAV89004	SkY	$6x10^{-09}$	771	NP_815175	SkZ	$8x10^{-5}$	369
CAH21051	SkZ	$2x10^{-03}$	522	AAR83223	SkAA	$9x10^{-77}$	1344
AAQ12193	SkC	$6x10^{-16}$	672	NP_680543	SkBB	$4x10^{-03}$	219
BAB07251	SkCC	$8x10^{-07}$	1353				
CAE35605	SkAA	$2x10^{-03}$	231				
YP_024701	SkBB	$6x10^{-87}$	1287				
AAF03962	SkBB	$1x10^{-04}$	963				
AAV78275		$3x10^{-06}$	858				

Because they interact with the large subunit terminase, predicted amino acid sequences of the small subunit terminase and portal protein were selected for the searches. One thousand and six searches were performed with small subunit terminase and 739 with

portal protein queries. Similarity to portal protein genes was detected in at least 23 sites in the *S. kunkelli* genome, and similarity to small subunit terminase genes was detected in at least 16 sites. Similarity to both portal proteins and small subunit terminase were detected on the contig containing one of the P58-like regions, SkC. One sequence on SkC, 35 kb from the P58 region, is similar to both an *Enterobacteria* phage Sf6 portal protein (Acc. No. AAQ12193) and an *Enterococcus faecalis* V583 terminase small subunit (Acc. No. NP_815175). An IS911 transposase-like sequence, which could have been associated with tailed phage, was detected by PSI-BLAST in at least 25 sites in the *S. kunkelii* genome, including 2 sites on contig SkB 17 and 20 kb from the P58 region, 1 site on SkC at a position 20 kb from the P58 region, and 3 sites on contig SkD 18, 16, and 44kb from the P58 region. Helicases, which could also have been derived from tailed phage, were detected in at least 19 sites in the *S. kunkelii* genome, including 2 sites on SkB 17 and 20 kb from the P58 region, 1 site on SkC 20 kb from the P58 region, and 3 sites on SkD 27, 45 and 5 kb from the P58 region. Thus, tailed phage sequences are present in the *S. kunkelii* genome and are located near P58. This supports the hypothesis that P58, at least section 1, originated from an inserted viral element of a tailed phage.

PSI-BLAST revealed that section 2 of ScA-P58 is related to the C-terminus (residues 331-464) of *Mycoplasma hominis* P50/Vaa lipoprotein adhesin (expect: $3x10^{-35}$, Acc. no. CAA04901). P50/Vaa in *M. hominis* is composed of a conserved N-terminus and a variable number of C-terminal repeats, which undergo recombination, duplication, and truncation at a high frequency (Henrich *et al.* 1998). If an adhesin and bacteriophage sequence were ancestors of a recombination event, one would expect the N-terminus of the P50/Vaa adhesin to be in close proximity to bacteriophage sequence.

A PSI-BLAST search of the *S. kunkelii* database using *M. hominis* P50 (Acc. No. CAA04901) detected P50 N-terminus (residues 1-106) as part of a 483 amino acid ORF (expect $1x10^{-5}$, iteration 5). PSI-BLAST searches using the C-terminal part of the 483 amino acid ORF as a query detected similarity ($3x10^{-46}$, iteration 3) to a bacteriophage protein (Acc. No. CAD19144). Thus, no similarity to phage genes was detected in section 2, and instead, similarity to a mycoplasma adhesin was found.

Some sequence information is available for another strain of *S. kunkelii*, CSS-M. Searches of the Genome Survey Sequences database, using ScA-P58 as a query sequence, detected three P58-like sequences, BH235075, a 599 nt sequence with an expect value of $2x10^{-40}$, BH234891, a 370 nt sequence with an expect value of $8x10^{-22}$, and BH234931, a 200 nt sequence with an expect value of $1x10^{-10}$. BH235075 and BH234931 were similar to ScA-P58 in the C-terminal half of P58. BH235075 was similar to residues 339-457 of the predicted amino acid sequence, and BH235931 was similar from position 396 to 596, the latter position being 3' of the P58 stop codon. Of the various P58 versions, BH235075 most closely resembled SkD-P58, as determined by BLAST search, with an expect value of 10^{-172} and matching 308 of 310 aligned nucleotide residues. BH235931 most closely resembled SkC, matching 199 of 200 aligned residues. BH234891 was similar to ScA-P58 in residues 133-218 in the N-terminal half of the P58 predicted amino acid sequence. BH234891 most closely resembled SkC-P58, as determined by BLAST search, with all 370 residues matching. Available sequence information for the N-terminal halves of P58 sequences in *S. insolitum, S. phoeniceum, S. poulsonii, Spiroplasma* sp. MSRO, and *Spiroplasma* sp. NSRO revealed that they are similar to the N-terminal halves of SkA-P58 and SkB-P58

(95.1%-97.3% identical). In contrast, pairwise alignments with ScA-P58 yielded identities ranging from 51.1% to 51.8%. Thus, the presence of P58 is not limited to *S. citri* BR3-3X and *S. kunkelii* CR2-3X, but P58-sequences are also distributed among other strains of *S. citri* and *S. kunkelii* as well as different *Spiroplasma* species.

PSI-BLAST searches revealed the presence of a P58-like gene ($1x10^{-93}$, iteration 2) and a P123-like gene ($1x10^{-145}$, iteration 2) in another mollicute, *Ureaplasma parvum*, but not in *Mycoplasma* spp., the closest relatives of *Spiroplasma*. Alignment of the predicted amino acid sequence of Up-P58 with ScA-P58 revealed that the similarity is distributed throughout the sequence. BLAST searches did not detect homologues of P12 or P54, ORFs that were present in between ScA P58 and P123 ORFs in *S. citri* BR3-3X. However, ORFs of appropriate length are present in homologous positions. P58, as well as another ORF (P123) that was present in some P58 regions, was identified in *U. parvum* by limited sequence similarity. This indicates that recombinations among P58 sequences may include adjacent sequences, and P58 is present in mycoplasmas other than those in the genus *Spiroplasma*.

Discussion

In *S. kunkelii*, P58-like sequences and some flanking regions are duplicated. Because *Spiroplasma* genomes have evolved to be reduced in size, retention of a multigene family may indicate a significant function in these genes. P58 is present in several strains (Chandrasekaran 2000) and species (Montenegro *et al.* 2005) of *Spiroplasma*, as well as in another mollicute, *Ureaplasma parvum* (Glass *et al.* 2000). Most of the copies of P58 contain mutations that probably

prevent function, and it would be interesting to know, in future work, if the truncated products of SkA and SkG are actually produced, and if the remainder of the ScC putative coding sequence was full length and free of mutations. However, one *S. kunkelii* P58 predicted amino acid sequence, SkB-P58, and at least one other *S. citri* P58, ScB-P58, are full length and free of any frameshift mutations or deletions and could be functional.

For discussion purposes here, the P58 region of ScA is represented as AX, having an "A" type 5' end (Figure 6) and an "X" type 3' end. ScB, SkA, SkB, SkC, and SkD are represented as BX, with a "B" type 5' end, and an X type 3' end. ScC, SkE, and SkF are termed AY, with a 5' end similar to that of ScA but a different 3' end, designated "Y". SkG has a "B" type 5' end, and part of the "Y" type 3' end, and thus is designated BY. Thus, the P58 family members characterized in this study fall into four groups: AX, BX, AY, and BY. If it is true that two ancestral combinations of A or B, and X or Y recombined, then four sequences, AX, BX, AY, and BY would be expected, and these sequences were observed.

Detection of a gene conversion event between ScA and ScB and ScC by GeneConv within the P58 coding region provides support for the hypothesis of a recombination event between groups AX and BX. Recombination events between ScA and *S. kunkelii* contigs SkA, SkB, and SkC, suggested by other data reported here (Figure 2B) to have occurred, were not detected by GeneConv, possibly because the *S. kunkelii* contigs are more diverged from ScA in section 2 than from ScB and ScC (Figure 2B). *S. citri* and *S. kunkelii* contigs ScA, ScB, SkA, SkB, and SkC may have been formed via one recombination event that occurred prior to speciation of *S. citri* and *S. kunkelii*, and

selective pressure may have prevented divergence between ScA and ScB and, ScA and ScC. Another possibility is that each of the contigs represents a separate recombination event. In that case, the P58 region may represent a place in the genome that has a higher rate of recombination. While SkA and SkB may have been duplicated after recombination, recombination junctions occurring at slightly different sites (Figure 6) in the other P58-like sequences supports the explanation of separate recombination events.

The similarity of the N-terminal portion of P58 amino acid sequences (A and B) to those of bacteriophage terminases suggests that the multiple P58-like sequences may be remnants of viral sequences that integrated into the *S. citri* chromosome. One would expect to find mutations in non-functional viral sequences, and indeed, many P58-like sequences had mutations that probably rendered them non-functional. The P58-like sequences could have been duplicated after insertion, or may be the result of multiple viral insertion events.

The similarity of the C-terminal portion of P58 (X) to *M. hominis* P50/Vaa is likely due to a reciprocal exchange between P50/Vaa and bacteriophage genes. As described above, ORFs containing an "X" type 3' end have sequences similar to phage terminase sequences 5' of the regions similar to *M. hominis* P50/Vaa, and, elsewhere in the genome, the 5' portion of *M. hominis* P50/Vaa gene is adjacent to bacteriophage sequence.

At least one of the *S. citri* P58 genes is a surface exposed membrane protein (Ye *et al.* 1997), a context unusual for a bacteriophage terminase, and the predicted transmembrane regions of both P58 and *M. hominis* P50/Vaa are in the N-terminal half (Boesen *et al.* 2001; Ye *et al.* 1997). If P58 is viral in origin, P58

may have adopted a new function. In this case, one would expect more divergence among P58 genes than between other respective genes in *S. citri* and *S. kunkelii*. Indeed, most *S. citri* and *S. kunkelii* genes tested have a high degree of sequence similarity (Table 3).

The effect of inserted viral sequences on the stability of the *Spiroplasma* genome has been documented (Ye *et al.* 1992). SpV1 viral sequences have been implicated as a major factor in the variation of the *S. citri* genome (Melcher *et al.* 1999). Although not yet sequenced, tailed phages, which usually encode terminases, were reported to infect spiroplasmas and insert into the chromosome (Dickinson and Townsend 1984).

If P58 does play a role in spiroplasma-insect interactions, the role could be direct, such as adhesion to the insect cell, or indirect, such as being involved in the production of a metabolite or other component necessary for transmission. One P58 gene may be sufficient for transmission through the insect; thus, mutations in the other P58-like genes would not affect insect transmission. *S. citri* and *S. kunkelii* are transmitted by different insect vectors. Because spiroplasma-insect interactions are complex and specific, it would not be surprising if the adhesins differed between *S. citri-Circulifer tenellus* interactions and *S. kunkelii-D. maidis* interactions (Fletcher *et al.* 2006).

Spiroplasmas must adapt to very different environments in the plant host and insect vector. Thus, expressing different genes at different times is likely necessary for survival. Because of the limited *Spiroplasma* genome size, spiroplasmas probably do not have a full complement of all the genes they could

possibly need in the different environments, and thus, likely rely heavily on mechanisms such as flux of genetic information (Melcher and Fletcher 1999). If so, chromosomal recombinations such as those described here may affect the evolution of *Spiroplasma* more than that of other bacteria. Thus, events such as the P58 recombination could be mechanisms of variation for *Spiroplasma* and for adaptation to the different niches they encounter.

Future Work

Continuation of the study of the P58 multigene family is important because while many advances in understanding the path of the spiroplasma through the insect gut and salivary gland barriers have been made (Fletcher *et al.* 1998; Kwon *et al.* 1999), some of the exact proteins that are involved in adhesion to the gut barrier remain unidentified. The lack of ScA-P58 is a likely reason for the non-pathogenicity of BR3-G, and if so, may play a role in *S. citri*-*C. tenellus* interactions. Thus, continuation of the study of the P58 multigene family may provide insight into movement of spiroplasma through the insect. Regardless of its status as an adhesin, understanding of the P58 multigene family would provide information as to the role of recombination on genetic flux and evolution of spiroplasmas.

That one copy of P58 is deleted in a non-pathogenic strain of *S. citri* BR3-3X, at least one copy of P58 is surface exposed (Ye *et al.* 1997), and a 58 kDa protein, which may be on a proposed tip structure, is on the surface of *S. citri* BR3-3X, (Fletcher, unplubished data) makes P58 an interesting candidate as at least one of the adhesins

involved in the *S. citri* BR3-*C. tenellus* interaction. Addressing the role of P58 would raise several interesting questions, such as the following:

Is P58 expressed, and if so, where is it localized?

This could be addressed by a PAGE of crude *S. citri* lysate, western blot using the anti-P58 anti-sera that was previously made to a P58 kDa protein, and peptide mass fingerprinting. If P58 is the only protein present, an electron micrograph would localize it to the surface. If necessary, P58 could be separated from other 58 kDa proteins present by an immunoadsorbent column or 2D-PAGE.

What is the full sequence of ScC, and is it expressed?

Obtaining the full sequence of ScC, could be accomplished by DNA sequencing. By doing so, several pieces of information would be obtained. For example, identification of other sites of recombination could be identified. In addition, likely mutations such as possible additional sites of recombination, premature stop codons, and point mutations could be identified. One of these mutations would likely mean that this gene is not expressed. This could be accomplished by cloning restriction fragments and using the 3' portion of ScA as a probe for Southern blot to select the correct clone.

Alternatively, peptides of 58 kDa proteins could be sequenced by mass spectrometry. In addition to sequence information, peptide

sequencing of ScC-P58 would confirm expression of this protein, and since the remainder of the sequence is available, it could provide insights as to function through database searches. In addition, other 58 kDa proteins, including ScA-P58 and ScB-P58, may be present, confirming their expression. Using mass spectrometry would also be useful to determine whether the premature stop codons of SkA and SkG prevent their expression.

Does *S. citri* form tip structures, and if so, does P58 localize to them?

An electron micrograph showed an anti-58kDa protein on what may be tip structures (Fletcher, unpublished data). In mycoplasmas, tip structures are important for attachment to and orientation of the host cell (Rottem 2003). Tip structures have also been observed in several species of *Spiroplasma* (Ammar *et al.* 2004). Immunogold labeling and electron microscopy may implicate the function of one or more P58s and localize the protein or proteins.

The prospect of further research on the P58 multigene family creates exciting possibilities of understanding the mechanisms of genome structure and differential gene expression in organisms with small genomes, such as mollicutes in addition to a better understanding of spiroplasma-insect interactions.

LITERATURE CITED

Abedon, S. T. and H. W. Ackermann. (2000). "Bacteriophage Names 2000." Retrieved 2/25/02, 2002, from www.phage.org.

Amikam, D., G. Glaser and S. Razin (1984). "Mycoplasmas (Mollicutes) have a low number of rRNA genes." *J. Bacteriol.* **158**(1): 376-378.

Ammar, E. D., D. Fulton, X. Bai, T. Meulia and S. A. Hogenhout (2004). "An attachment tip and pili-like structures in insect- and plant-pathogenic spiroplasmas of the class Mollicutes." *Arch. Microbiol.* **181**(2): 97-105.

Andre, A., M. Maucourt, A. Moing, D. Rolin and J. Renaudin (2005). "Sugar import and phytopathogenicity of *Spiroplasma citri*: glucose and fructose play distinct roles." *Mol. Plant Microbe Interact.* **18**(1): 33-42.

Bai, X. and S. A. Hogenhout (2002). "A genome sequence survey of the mollicute corn stunt spiroplasma *Spiroplasma kunkelii*." *FEMS Microbiol. Lett.* **210**: 7-17.

Berg, M., U. Melcher and J. Fletcher (2001). "Characterization of *Spiroplasma citri* adhesion related protein SARP1, which contains a domain of a novel family designated sarpin." *Gene* **275**: 57-64.

Berho, N., S. Duret, J. L. Danet and J. Renaudin (2006). "Plasmid pSci6 from *Spiroplasma citri* GII-3 confers insect transmissibility to the non-transmissible strain *S. citri* 44." *Microbiol.* **152**(9): 2703-2716.

Berho, N., S. Duret and J. Renaudin (2006). "Absence of plasmids encoding adhesion-related proteins in non-insect-transmissible strains of *Spiroplasma citri*." *Microbiol.* **152**(3): 873-886.

Biberfeld, G. and P. Biberfeld (1970). "Ultrastructural features of *Mycoplasma pneumoniae*." *J. Bacteriol.* **102**(3): 855-861.

Boesen, T., N. U. Fedosova, M. Kjeldgaard, S. Birkelund and G. Christiansen (2001). "Molecular design of *Mycoplasma hominis* Vaa adhesin." *Protein Science* **10**(12): 2577-2586.

Bove, J. M. (1993). "Molecular features of mollicutes." *Clin. Infect. Dis.* **17 Suppl 1**: S10-31.

Bove, J. M., J. Renaudin, C. Saillard, X. Foissac and M. Garnier (2003). "*Spiroplasma citri,* a plant pathogenic mollicute: relationships with its two hosts, the plant and the leafhopper vector." *Annu. Rev. Phytopathol.* **41**: 483-500.

Buchen-Osmond, C. ICTVdb: The Universal Virus Database of the International Committee on Taxonomy of Viruses [Internet]. (2002: modified 7/12/07) [cited 2/14/08] International Committee on Taxonomy of Viruses. www.ncbi.nlm.nih.gov/ICTVdb/.

Carle, P., C. Saillard, A. Blanchard, N. Carrere, S. Carrere, S. Duret, S. Eveillard, P. Gaurivaud, G. Gourgues, J. Gouzy, A. Henry, P. Salar, F. Laigret, J. M. Bove, J.

Renaudin and X. Foissac. Partial chromosome of *Spiroplasma citri* GII3-3X. Direct sequence submission. 6/27/06 [cited 7/20/06]. Institut National de la Recherche Agronomique. NCBI Core Nucleotide Database: www.ncbi.nlm.nih.gov.

Catalano, C. E. (2000). "The terminase enzyme from bacteriophage lambda: a DNA-packaging machine." *Cell Mol. Life Sci.* **57**(1): 128-148.

Chambaud, I., R. Heilig, S. Ferris, V. Barbe, D. Samson, F. Galisson, I. Moszer, K. Dybvig, H. Wroblewski, A. Viari, E. P. Rocha and A. Blanchard (2001). "The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*." *Nucleic Acids Res.* **29**(10): 2145-2153.

Chandrasekaran, A. (2000). Biochemistry and Molecular Biology, Oklahoma State University.

Chen, T. A. and R. E. Davis (1979). "Cultivation of spiroplasmas". *The Mycoplasmas*. R. F. Whitcomb and J. G. Tully. New York, Academic Press**:** 65-79.

Citti, C. and R. Rosengarten (1997). "*Mycoplasma* genetic variation and its implication for pathogenesis." *Wien Klin. Wochenschr.* **109**(14-15): 562-568.

Cole, R. M., J. G. Tully, T. J. Popkin and J. M. Bove (1973). "Morpholoby, ultrastructure, and bacteriophage infection of the helical Mycoplasma-like organism (*Spiroplasma citri* gen. nov., sp. nov.) cultured from "stubborn" disease of citrus." *J. Bacteriol.* **115**(1): 367-386.

Cole, R. M., W. O. Mitchell and C. F. Garon (1977). "Spiroplasmavirus citri 3: propagation, purification, proteins, and nucleic acid." *Science* **198**(4323): 1262-1263.

Dally, E. L., T. S. Barros, Y. Zhao, S. Lin, B. A. Roe and R. E. Davis (2006). "Physical and genetic map of the *Spiroplasma kunkelii* CR2-3x chromosome." *Can. J. Microbiol.* **52**(9): 857-867.

Davis, R. E., J. F. Worley, R. F. Whitcomb, T. Ishijima and R. L. Steere (1972). "Helical filaments produced by a mycoplasma-like organism associated with corn stunt disease." *Science* **176**: 521-523.

Davis, R. E., E. L. Dally, R. Jomantiene, Y. Zhao, B. Roe, S. Lin and J. Shao (2005). "Cryptic plasmid pSKU146 from the wall-less plant pathogen *Spiroplasma kunkelii* encodes an adhesin and components of a type IV translocation-related conjugation system." *Plasmid* **53**(2): 179-190.

De Almeida, L., B. Raccah and M. Klein (1997). "Transmission characteristics of Spiroplasma citri and its effect on leafhopper vectors from the *Circulifer tenellus* complex." *Ann. Appl. Biol.* **130**: 49-59.

Desiere, F., C. Mahanivong, A. J. Hillier, P. S. Chandry, B. E. Davidson and H. Brussow (2001). "Comparative genomics of lactococcal phages: insight from the complete genome sequence of *Lactococcus lactis* phage BK5-T." *Virology* **283**(2): 240-52.

Dickinson, M. J. and R. Townsend (1984). "Integration of a temperate phage infecting *Spiroplasma citri*." *Isr. J. Med. Sci.* **20**(9): 785-787.

Dickinson, M. J., R. Townsend and S. J. Curson (1984). "Characterisation of a virus infecting the wall-free prokaryote *Spiroplasma citri*." *Virology* **135**(22): 524-535.

Duret, S., J. L. Danet, M. Garnier and J. Renaudin (1999). "Gene disruption through homologous recombination in *Spiroplasma citri*: an scm1-disrupted motility mutant is pathogenic." *J. Bacteriol.* **181**(24): 7449-7456.

Duret, S., N. Berho, J. L. Danet, M. Garnier and J. Renaudin (2003). "Spiralin is not essential for helicity, motility, or pathogenicity but is required for efficient transmission of *Spiroplasma citri* by its leafhopper vector *Circulifer haematoceps*." *Appl. Environ. Microbiol.* **69**(10): 6225-6234.

Feldman, R., B. Henrich, V. Kolb-Bachofen and U. Hadding (1992). "Decreased metabolism and viability of *Mycoplasma hominis* induced by monoclonal antibody-mediated agglutination." *Infect. Immun.* **60**(1): 166-174.

Fletcher, J., G. A. Schultz, R. E. Davis, C. E. Eastman and R. E. Goodman (1981). "Brittle root disease of horseradish: evidence for an etiological role of *Spiroplasma citri*." *Phytopathology* **71**: 1073-1080.

Fletcher, J. (1981, unpublished data). Entomology and Plant Pathology, Oklahoma State University.

Fletcher, J., A. Wayadande, U. Melcher and F. Ye (1998). "The phytopathogenic mollicute-insect vector interface: a closer look." *Phytopathology* **88**(12): 1351-1358.

Fletcher, J., U. Melcher and A. Wayadande (2006). *The Phytopathogenic Spiroplasmas*, Springer-Verlag.

Foissac, X., J. L. Danet, P. Saillard, P. Gurivaud, F. Laigret, C. Pare and J. M. Bove (1997). "Mutagenesis by insertion of Tn4001 into the genome of *Spiroplasma citri*: characterization of mutants affected in plant pathogenicity and transmission to the plant by the leafhopper vector *Circulifer haematoceps*." *Mol. Plant Mibrobe Interact.* **10**(4): 454-461.

Fraser, C. M., J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley and J. L. Fritchman (1995). "The minimal gene complement of *Mycoplasma genitalium*." *Science* **270**(5235): 397-403.

Gasparich, G. E., K. J. Hackett, E. A. Clark, J. Renaudin and R. F. Whitcomb (1993). "Occurrence of extrachromosomal deoxyribonucleic acids in spiroplasmas associated with plants, insects, and ticks." *Plasmid* **29**(2): 81-93.

Gasparich, G. E. (2002). "Spiroplasmas: evolution, adaptation and diversity." *Front. Biosci.* **7**(6): 19-40.

Gaurivaud, P., J. L. Danet, F. Laigret, M. Garnier and J. M. Bove (2000). "Fructose utilization and phytopathogenicity of *Spiroplasma citri*." *Mol. Plant Microbe Interact.* **13**(10): 1145-1155.

Gaurivaud, P., F. Laigret, E. Verdine, M. Garnier and J. M. Bove (2000). "Fructose operon mutants of *Spiroplasma citri*." *Microbiology* **146**: 2229-2236.

Gaurivaud, P., F. Laigret, M. Garnier and J. M. Bove (2001). "Characterization of FruR as a putative activator of the fructose operon of *Spiroplasma citri*." *FEMS Microbiol. Lett.* **198**(1): 73-78.

Glass, J. I., E. J. Lefkowitz, J. S. Glass, C. R. Heiner, E. Y. Chen and G. H. Cassell (2000). "The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*." *Nature* **407**(6805): 757-762.

Henrich, B., R. Feldman and U. Hadding (1993). "Cytoadhesins of *Mycoplasma hominis*." *Infect. Immun.* **61**(7): 2945-2951.

Henrich, B., K. Lang, A. Kitzerow, C. MacKenzie and U. Hadding (1998). "Truncation as a novel form of variation of the P50 gene in *Mycoplasma hominis*." *Microbiol.* **144**: 2979-2985.
Himmelreich, R., H. Hilbert, H. Plagens, E. Pirkl, B. C. Li and R. Herrmann (1996). "Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*." *Nucleic Acids Res.* **24**(22): 4420-4449.
Jaffe, J. D., N. Stange-Thomann, C. Smith, D. DeCaprio, S. Fisher, J. Butler, S. Calvo, T. Elkins, M. G. FitzGerald, N. Hafez, C. D. Kodira, J. Major, S. Wang, J. Wilkinson, R. Nicol, C. Nusbaum, B. Birren, H. C. Berg and G. M. Church (2004). "The complete genome and proteome of *Mycoplasma mobile*." *Genome Res.* **14**(8): 1447-1461.
Joshi, B. D. (2002, unpublished data). Biochemistry and Molecular Biology, Oklahoma State University.
Joshi, B. D., M. Berg, J. Rogers, J. Fletcher and U. Melcher (2005). "Sequence comparisons of plasmids pBJS-O of *Spiroplasma citri* and pSKU146 of *S. kunkelii*: implications for plasmid evolution." *B.M.C. Genom.* **6**: 175.
Killiny, N., B. Batailler, X. Foissac and C. Saillard (2006). "Identification of a *Spiroplasma citri* hydrophilic protein associated with insect transmissibility." *Microbiology* **152**(4): 1221-1230.
Kwon, M. O., A. C. Wayadande and J. Fletcher (1999). "Spiroplasma citri movement into the intestines and salivary glands of its leafhopper vector, *Circulifer tenellus*." *Phytopathology* **89**: 1144-1151.
Lartigue, C., S. Duret, M. Garnier and J. Renaudin (2002). "New plasmid vectors for specific gene targeting in *Spiroplasma citri*." *Plasmid* **48**(2): 149-159.
Lee, I. M., R. E. Davis and D. E. Gundersen-Rindal (2000). "Phytoplasma: phytopathogenic mollicutes." *Annu. Rev. Microbiol.* **54**: 221-255.
Liu, S. L., A. Hessel and K. E. Sanderson (1993). "Genomic mapping with I-Ceu I, an intron-encoded endonuclease specific for genes for ribosomal RNA, in *Salmonella* spp., *Escherichia coli*, and other bacteria." *Proc. Natl. Acad. Sci. U. S. A.* **90**(14): 6874-6878.
Marais, A., J. M. Bove and J. Renaudin (1996). "*Spiroplasma citri* virus SpV1-derived cloning vector: deletion formation by illegitimate and homologous recombination in a spiroplasmal host strain which probably lacks a functional recA gene." *J. Bacteriol.* **178**(3): 862-870.
McBride, M. J. (2001). "Bacterial gliding motility: multiple mechanisms for cell movement over surfaces." *Annu. Rev. Microbiol.* **55**: 49-75.
Melcher, U. and J. Fletcher (1999). "Genetic variation in *Spiroplasma citri*." *Eur. J. Plant. Path.* **105**(6): 519-533.
Melcher, U., Y. Sha, F. Ye and J. Fletcher (1999). "Mechanisms of spiroplasma genome variation associated with SpV1-like viral DNA inferred from sequence comparisons." *Microb. Comp. Genomics* **4**(1): 29-46.
Melcher, U. (1999, unpublished data). Biochemistry and Molecular Biology, Oklahoma State University.
Minion, F. C., E. J. Lefkowitz, M. L. Madsen, B. J. Cleary, S. M. Swartzell and G. G. Mahairas (2004). "The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis." *J. Bacteriol.* **186**(21): 7123-7133.

Montenegro, H., V. N. Solferini, L. B. Klaczko and G. D. D. Hurst (2005). "Male-killing *Spiroplasma* naturally infecting *Drosophila melanogaster*." *Insect Mol. Biol.* **14**(3): 281.
Nunan, L. M., D. V. Lightner, M. A. Oduori and G. E. Gasparich (2005). "*Spiroplasma penaei* sp. nov., associated with mortalities in *Penaeus vannamei*, Pacific white shrimp." *Int. J. Syst. Evol. Microbiol.* **55**(6): 2317-2322.
Ohtani, N. and M. Miyata (2006). "Identification of a novel nucleoside triphosphatase from *Mycoplasma mobile*: a prime candidate for motor of gliding motility." *Biochem. J.* **403**(1): 71-77.
Papazisi, L., T. S. Gorton, G. Kutish, P. F. Markham, G. F. Browning, D. K. Nguyen, S. Swartzell, A. Madan, G. Mahairas and S. J. Geary (2003). "The complete genome sequence of the avian pathogen *Mycoplasma gallisepticum* strain R(low)." *Microbiology* **149**(9): 2307-2316.
Razin, S., D. Yogev and Y. Naot (1998). "Molecular biology and pathogenicity of mycoplasmas." *Microbiol. Mol. Biol. Rev.* **62**(4): 1094-1156.
Razin, S. (1999). "Adherence of pathogenic mycoplasmas to host cells." *Biosci. Rep.* **19**(5): 367-372.
Renaudin, J., P. Aullo, J. C. Vignault and J. M. Bove (1990). "Complete nucleotide sequence of the genome of *Spiroplasma citri* virus SpV1-R8A2." *Nucleic Acids Res.* **18**: 1293.
Renaudin, J. and J. M. Bove (1994). "SpV1 and SpV4, spiroplasma viruses with circular, single-stranded DNA genomes, and their contribution to the molecular biology of spiroplasmas." *Adv. Virus Res.* **44**: 429-463.
Roe, B. A., S. P. Lin, H. G. Jia, H. M. Wu, D. Kupfer and R. E. Davis. (2002). "*Spiroplasma kunkelii* Genome Sequencing Project." 2002, from http://www.genome.ou.edu/spiro.html.
Rottem, S. (2003). "Interaction of mycoplasmas with host cells." *Physiol. Rev.* **83**(2): 417-432.
Saglio, P., D. Lafleche, C. Bonissol and J. Bove (1971). "Isolement et culture invitro des mycoplasmes associes au stubborn des argumes et leur observation au microscope electronique." *C. R. Acad. Sci. Paris* **272**: 1387-1390.
Saglio, P., M. L'hospital, D. Lafleche, G. Dupont, J. Bove, J. G. Tully and E. A. Freundt (1973). "*Spiroplasma citri* gen. and sp. n.: a mycoplasma-like organism associated with "stubborn" disease of citrus." *Int. J. Syst. Bacteriol.* **23**: 191-204.
Sasaki, Y., J. Ishikawa, A. Yamashita, K. Oshima, T. Kenri, K. Furuya, C. Yoshino, A. Horino, T. Shiba, T. Sasaki and M. Hattori (2002). "The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans." *Nucleic Acids Res.* **30**(23): 5293-5300.
Sawyer, S. A. (1999). "GENECONV: A computer package for the statistical detection of gene conversion." *Distributed by the author, Department of Mathematics, Washington University in St. Louis, available at http://www.math.wustl.edu/~sawyer.*
Seto, S., A. Uenoyama and M. Miyata (2005). "Identification of a 521-kilodalton protein (Gli521) involved in force generation or force transmission for *Mycoplasma mobile* gliding." *J. Bacteriol.* **187**(10): 3502-3510.

Sha, Y., U. Melcher, R. E. Davis and J. Fletcher (2000). "Common elements of spiroplasma plectroviruses revealed by nucleotide sequence of SVTS2." *Virus genes* **20**(1): 47-56.

Uenoyama, A., A. Kusumoto and M. Miyata (2004). "Identification of a 349-kilodalton protein (Gli349) responsible for cytadherence and glass binding during gliding of *Mycoplasma mobile*." *J. Bacteriol.* **186**(5): 1537-1545.

Uenoyama, A. and M. Miyata (2005). "Identification of a 123-kilodalton protein (Gli123) involved in machinery for gliding motility of *Mycoplasma mobile*." *J. Bacteriol.* **187**(16): 5578-5584.

Wang, W., W. Gu, Z. Ding, Y. Ren, J. Chen and Y. Hou (2005). "A novel *Spiroplasma* pathogen causing systemic infection in the crayfish *Procambarus clarkii* (Crustacea: Decapod), in China." *FEMS Microbiol. Lett.* **249**(1): 131-137.

Wei, W., D. C. Opgenorth, R. E. Davis, C. J. Chang, C. G. Summers and Y. Zhao (2006). "Characterization of a novel adhesin-like gene and design of a real-time PCR for rapid, sensitive, and specific detection of *Spiroplasma kunkelii*." *Plant Dis.* **90**(9): 1233-1238.

Ye, F., F. Laigret, J. C. Whitley, C. Citti, L. R. Finch, P. Carle, J. Renaudin and J. Bove (1992). "A physical and genetic map of the *Spiroplasma citri* genome." *Nucleic Acids Res.* **20**(7): 1559-1565.

Ye, F., U. Melcher, J. E. Rascoe and J. Fletcher (1996). "Extensive chromosome aberrations in *Spiroplasma citri* strain BR3." *Biochem. Genet.* **34**(7/8): 269-285.

Ye, F., U. Melcher and J. Fletcher (1997). "Molecular characterization of a gene encoding a membrane protein of *Spiroplasma citri*." *Gene* **189**(1): 95-100.

Yu, J., A. Wayadande and J. Fletcher (2000). "*Spiroplasma citri* surface protein P89 implicated in adhesion to cells of the vector *Circulifer tenellus*." *Phytopathology* **90**(7): 716-722.

Zhao, Y., R. W. Hammond, R. Jomantiene, E. Dally, I. M. Lee, H. Jia, H. Wu, S. Lin, P. Zhang, S. Kenton, F. Z. Najar, A. Hua, B. A. Roe, J. Fletcher and R. E. Davis (2003). "Gene content and organization of an 85-kb DNA segment from the genome of the phytopathogenic mollicute *Spiroplasma kunkelii*." *Mol. Genet. Genom.* **269**(5): 592-602.

Zhao, Y., H. Wang, R. W. Hammond, R. Jomantiene, Q. Liu, S. Lin, B. A. Roe and R. E. Davis (2004). "Predicted ATP-binding cassette systems in the phytopathogenic mollicute *Spiroplasma kunkelii*." *Mol. Genet. Genom.* **271**: 325-338.

APPENDIX A

EVALUATING TRANSPOSON MUTAGENESIS FOR GENETIC ANALYSIS OF MOLLICUTE (*S. citri*) GENOMIC DNA

Introduction

Mutagenesis is an important method to understand the mechanisms of movement, colonization, and pathogenicity of *S. citri* in its leafhopper vectors as well as in its plant hosts. Random insertion mutagenesis has provided insight into functional genetics in many bacteria in which transformation is problematic. Transposon-mediated mutagenesis is a promising technique, since it provides random insertion and generation of numerous mutants (Koide *et al.* 2004).

Transposon-mediated mutagenesis has not yet been performed successfully in mollicutes, and tested systems for genetic analysis are limited. The use of TN*916* for mycoplasmas (Dybvig and Alderete 1988) and TN*4001* for *S. citri* (Foissac *et al.* 1997) are among the limited reports of genetic manipulation of mollicutes.

The transposome strategy, a transposon mutagenesis system employing a novel construct designated a transposome, was developed by Epicentre (Madison, WI), and a number of prokaryotes have been mutagenized by this method, including at least 14 gram positive and 44 gram negative bacteria (Epicentre 2000). Epicentre describes the EZ::TN system as “a stable complex of the Tn5 transposase a Tn5-derived transposon.”

The EZ::TN transposon, which is hyperactive 19 bp mosaic end (ME) EZ::TN transposase recognition sequences (Goryshin *et al.* 2000) flanking antibiotic resistance gene, was used in this study. An EZ::TN transposon with a dihydrofolate reductase (DHFR) gene conferring trimethoprim resistance was used previously (Mutaqin 2004) to mutagenize *S. citri* BR3-3X. Five mutants were obtained (XTA-10, XTA-96, XTA-238, XTB-1, and XTC-4) by Mutaqin (2004). Attempts at sequencing the regions surrounding the transposon using genomic (Davis 2007, unpublished) or sheared *S. citri* DNA were unsuccessful.

The objective of this work was to finish the evaluation of the use of the transposome system, mediated by electroporation, to transform *S. citri* BR3-3X and to obtain mutants useful for study of its biology and host interactions. This work describes the cloning and sequencing of regions surrounding the transposon of three of the five mutants (XTA-96, XTA-238, and XTC-4).

Materials and Methods

The *S. citri* transposon mutants were grown in LD8 broth (Chen and Davis 1979) in low volumes, as previously described, to a final volume of 20 mL. Cells were centrifuged 15 min. at 16,000 x g. DNA was extracted using a DNeasy kit (Qiagen). For XTA-96, DNA fragments small enough to be easily cloned were created by digesting XTA-96 DNA with the following two enzymes, which generate compatible ends: *Spe*I (New Endland Biolabs) and *Nhe*I (Promega). The vector DNA, pBluescript SK+, was digested with *Spe*I, and then treated with HK phosphatase (Epicentre). Ligation reactions were carried out in a volume of 10 µl, with final concentrations as follows: 6.9×10^{7}

molecules/μl vector, pBluescript SK+, insert concentrations ranging from 7.8 ng/μl to 24 ng/μl, 1x ligation buffer (Invitrogen), and 10 u T4 DNA ligase (Invitrogen).

Transformation reactions were carried out using chemically competent JM109 cells (Promega) according to the manufacturer's directions, except the following: 20 μl competent cells were used per reaction, cells were incubated on ice for 20 min. after DNA was added, and 80 μl SOC medium was added after heat shock. *S. citri* mutant transformation reactions were plated on LB agar with 100 μg/mL trimethoprim, 100 μg/mL ampicillin, 33.3 μg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), and 8.1 μg/mL isopropyl-β-D-thioglactopyranoside (IPTG).

Plasmids of positive clones were prepared using a QIAprep Spin Miniprep kit (Qiagen). The region surrounding the transposon was sequenced using primers RP-1 (5'-GACACTCTGTTATTACAAATCG-3') and FP-1 (5'-GGCGGAAACATTGGATGCGG), which were designed by the manufacturer (Epicentre).

Results

Plasmids of three positive clones, designated herein as pXTA-96, pXTA-238, and pXTC-4, were obtained, and the regions surrounding the transposon were sequenced. BLAST results revealed that in addition to having the DHFR sequence, pXTA-96 had sequence related to a previously sequenced 9.6 kb segment of *S. citri* BR3-3X (Accession U44405). Nucleotide BLAST searches using the sequence generated from the RP-1 primer and FP-1 primers as queries (queries RP-1 and FP-1, respectively) and the non-redundant database revealed an 89% identity and expect value of $2x10^{-134}$ in the P123

region of the fragment (position 6875-7268) using the RP-1 query, and 90 percent identity and an expect value of $7x10^{-93}$ to the C4-like and P18 region of the fragment (8670-8941) using the FP-1 query. pXTA-96 was also similar to a sequence in the *S. kunkelii* database (Roe *et al.* 2002). pXTA-96 FP-1 query was most simiar to contig 153.comp: 62% identical to 112 residues (positions 4615-4725) of the contig and had an expect value was $2x10^{-9}$. The RP-1 query was most similar to contig 178.comp: 72% identity over 180 residues (positions 34860-35039) and had an expect value of $1x10^{-29}$.

The clone pXTA-238 was similar to a contig from the *S. citri* GII3 chromosome, contig 20 (Accession No. AM285321.1). Using the non-redundant database, sequence generated from the RP-1 primer were 97% identical over 1,139 nucleotides (position 61,043 to 62,181) with an expect value of 0.0 (to one hundred significant digits). The pXTA-238 clone was also similar to a sequence in the *S. kunkelii* database. The RP-1 query was 89% identical to contig 186 over 960 nucleotides (position 60452-59468) with an expect value of less than 10^{-99}.

In the non-redundant database, pXTC-4 was similar to a different contig from the *S. citri* GII3 chromosome, contig 3 (Accession No. AM285304). pXTC-4 was 98% identical over 1079 residues (position 26,936-25,831) with an expect value of 0.0 (to one hundred significant digits). In the *S. kunkelii* database, the RP-1 query of pXTC-4, like pXTA-238, had similarity to contig 186; however, the similarity was in a different location on the contig. The RP-1 query was 97% identical to contig 186 over 412 nucleotides (position 14691-14280) with an expect value of less than 10^{-99}.

Discussion

That the three mutants were inserted in different regions of the *S. citri* BR3-3X chromosome suggests that the transposon was inserted randomly. If so, the EZ:TN transposon system may be a useful system for the generation of mutants in mollicutes. Cloning and sequencing of the other two mutants (XTA-10 and XTB-1) would provide further evidence for this inference.

LITERATURE CITED

Chen, T. A. and R. E. Davis (1979). "Cultivation of spiroplasmas". *The Mycoplasmas*. R. F. Whitcomb and J. G. Tully. New York, Academic Press**:** 65-79.

Davis, L. (2007, unpublished). "Analysis of transposome insertion in *Spiroplasma citri* BR3-3X." Oklahoma State University.

Dybvig, K. and J. Alderete (1988). "Transformation of *Mycoplasma pulmonis* and *Mycoplasma hyorhinis*: transposition of Tn916 and formation of cointegrate structures." *Plasmid* **20**: 33-41.

Epicentre. (2000, July 25, 2007). "New Transposon Construction Vector for Constructing a Custom EZ::TN™ Transposon." *Epicentre Forum Newsletter* Retrieved August, 2007, from www.epicentre.com.

Foissac, X., J. L. Danet, P. Saillard, P. Gurivaud, F. Laigret, C. Pare and J. M. Bove (1997). "Mutagenesis by insertion of Tn4001 into the genome of *Spiroplasma citri*: characterization of mutants affected in plant pathogenicity and transmission to the plant by the leafhopper vector *Circulifer haematoceps*." *Mol. Plant Mibrobe Interact.* **10**(4): 454-461.

Goryshin, I. Y., J. Jendrisak, L. M. Hoffman, R. Meis and W. S. Reznikoff (2000). "Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes." *Nat. Biotechnol.* **18**(1): 97-100.

Koide, T., J. F. da Silva Neto, S. L. Gomes and M. V. Marques (2004). "Insertional transposon mutagenesis in the *Xylella fastidiosa* Citrus Variegated Chlorosis strain with transposome." *Curr. Microbiol.* **48**(4): 247-250.

Mutaqin, K. Molecular typing of *Spiroplasma* species and lines using Rep-polymerase chain reaction and transposome mutagenesis and selection of natural non-adherent mutants of *S. citri* [dissertation]. [Stillwater (OK)]: Oklahoma State University; 2004. 108 p.

Roe, B. A., S. P. Lin, H. G. Jia, H. M. Wu, D. Kupfer and R. E. Davis. (2002). "*Spiroplasma kunkelii* Genome Sequencing Project." 2002, from http://www.genome.ou.edu/spiro.html.

VITA

Jana Lynn Comer

Candidate for the Degree of

Master of Science

Thesis: CHARACTERIZATION OF THE *SPIROPLASMA* P58 MULTIGENE FAMILY

Major Field: Biochemistry and Molecular Biology

Biographical:

Education:
Ringwood High School, Ringwood, Oklahoma, 1993, valedictorian. Oklahoma State University, Stillwater, Oklahoma, B. S., Biochemistry and Molecular Biology, 1999, *Summa Cum Laude*. Completed the requirements for the Master of Science or Arts in Biochemistry and Molecular Biology at Oklahoma State University, Stillwater, Oklahoma in December, May 2008.

Experience:
Sir Alexander Fleming Scholar, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma. Freshman Research Scholar, Biochemistry and Molecular Biology, Robert Matts, advisor,1993; Honors Research Project, Microbiology and Molecular Genetics, Robert Burnap, advisor, 1994; Laboratory technician, Biochemistry and Molecular Biology, Linda Yu, advisor, 1997; Senior Research Project, Biochemistry and Molecular Biology, Ulrich Melcher, advisor, 1998; Oklahoma State University, Stillwater, Oklahoma.