INSECT-TRANSMITTED PROCARYOTES

Spiroplasma citri: From Functional Genomics to . . . Genomics!*


ABSTRACT. Spiroplasma citri, the causal agent of citrus stubborn disease, was the first mollicute of plant origin to have been cultured as early as 1970. Even today, none of the many phytoplasmas (formerly MLOs) have been obtained in culture. S. citri was also the first representative of a new genus in the Mollicutes, the genus Spiroplasma with unique properties: motility and helical morphology. Being available in culture, many cellular and molecular properties of S. citri could be studied in the ensuing 20 yr: cell growth and division, cell polarity, spiroplasma proteins, extrachromosomal DNA (viruses and plasmids), DNA and RNA polymerases, insensitivity to rifampin, cloning and sequencing of protein genes (such as spiralin, fibril, etc.), heterologous expression of spiralin in Escherichia coli, sequences of viral genomes and identification of UGA as a tryptophane codon, sequences of UGA and UGG tryptophane tRNAs, physical and genomic map of the S. citri genome, 16S rDNA based taxonomy and phylogeny, cloning and sequencing of the replication origin (oriC) of the spiroplasmal chromosome.

As interesting as the above studies might have been, they threw no light on the most characteristic features of the spiroplasma: multiplication in, and transmission by, leafhopper vectors, mechanisms of phytopathogenicity, helical morphology and motility. To approach these problems, molecular genetics had to be used, and the required tools had first to be developed: transformation (by electroporation or PEG-mediated) or transfection (with virus SpV1 replicative form), development of S. citri oriC artificial plasmids such as shuttle plasmid pBOT1 used, for instance, to complement S. citri mutants with wild type genes, transposon Tn4001 mutagenesis, and production of mutants by gene disruption through homologous recombination. These tools have made it possible to identify genes involved in the mechanism of phytopathogenicity (fructose operon), motility, morphology, and helicity (gene scm1, fibril gen, gene mreB1), leafhopper transmission (gene sc76), and the regulation of spiroplasma operons (genes fruR and treR).

In fact, the preceding studies are part of what is called “functional genomics” (what are the genes involved in a function?). They were undertaken before the era of “genomics” or the sequencing of full genomes. Sequencing and annotation of the S. citri genome is however underway. This will greatly accelerate functional genomics, and render possible comparisons with sequenced genomes of other phytopathogenic bacteria of the apoplast, the phloem or the xylem. The S. citri genome should be available in 2003. Many more exciting years are ahead of us.

The first 15 yr of research on Spiroplasma citri, the causal agent of citrus stubborn disease, were reviewed at the 10th congress of the IOCV covering references up to 1988 (3). The major points covered in the 1986 review will be summarized and updated in this introduction. The essential part of the present review will be devoted to functional genomics, a field developed more recently. Sequencing of the S. citri genome is also underway, and will be briefly mentioned.

The first plant mollicutes, now called phytoplasmas, were discovered in 1967 (11), but have not yet been obtained in culture. S. citri was seen in 1970 in the sieve tubes of stubborn affected sweet orange leaves, and could be cultured as early as 1971. Hence, S. citri was the first mollicute cultured from plants. The organism had an optimal growth temperature of 32°C. All plant mollicutes, i.e. phytoplasmas and spiroplasmas, are restricted to the phloem sieve tubes. The stubborn agent was characterized and described as a new mollicute with motility and helical morphology in 1973, but the mechanisms involved in these features begin to be understood only today, as will be seen later. The organism was

*Invited presentation.
shown to be the causal agent of stub-
born disease in 1974. The leafhopper
Circulifer tenellus was found to be
the major vector of S. citri in the
U.S.A. in 1976. C. haematoceps,
absent from the U.S.A., transmits the
spiroplasma in the Old World as
found in 1986, but C. tenellus is also
present. Growth and division of this
unique, helical mollicute in liquid
medium was studied in the early
1980s. The smallest viable cell of
S. citri was found to be a two-turn
helix (elementary helix). The elemen-
tary helices grow into longer parent-
ial cells, which then divide by
constriction to yield elementary heli-
ces. The helical morphology is con-
served during this process. The
discovery of S. citri led to the iso-
lation of many other spiroplasmas,
mainly from insects and ticks. There
were 30 different species in 1986,
almost 50 in 2001. In addition to
S. citri, there are only two other spiro-
plasmas of plant origin: S. kunkelii,
a corn pathogen, and S. phoeniceum,
isolated from naturally infected peri-
winkle (Catharanthus roseus) plants
in Syria (3).

The molecular and cellular biol-
y of spiroplasmas was reviewed in
1989 (5). Spiroplasmas as infectious
agents of plants, arthropods and ver-
tebrates were reviewed in 1997 (4).
S. citri is serologically related to, and
shares DNA homology with the two
other plant pathogenic spiroplas-
mas, as well as with S. melliferum, a
honeybee pathogen. By 1987, three
S. citri viruses (SpV1, SpV2, and
SpV3), and one S. melliferum virus
(SpV4) had been discovered. SpV1 is
a filamentous non-lytic Plectrovirus
within the Inoviridae, while SpV4 is
an icosahedral lytic Spiromicrovirus
within the Microviridae (9, 28). The
single-stranded, circular DNA
genomes of SpV4 and SpV1 have been
sequenced in 1987 and 1990,
respectively (28).

Much data on molecular biology
and gene organization have been
acquired over the years. Some will
be mentioned here. The mollicutes
are phylogenetically related to
Gram positive bacteria with low
G+C DNA (33). This explains many
of the properties of S. citri. The G+C
content of S. citri DNA is 26%. The
genome size varies from 1,650 to
1,910 kbp according to the S. citri
strain considered. (36). The physical
and genomic map of the S. citri
genome was established in 1992
(37). Viral SpV1 DNA sequences are
inserted at several sites within the
genome (1). Features for DNA tran-
scription (RNA polymerase, pro-
moter and terminator) are similar to
those of Gram positive bacteria.
However, contrary to bacteria, S.
citri and all other mollicutes, are
insensitive to rifampin. It was
shown in 1996 that this property is
due to the fact that amino acid 526
in the β subunit of RNA polymerase
is asparagin (18). In spiroplasmas
and other mollicutes, UGA is not a
stop codon, but codes for tryp-
tophan. Spiroplasmas have thus two
tryptophan codons: UGG (regular
codon) and UGA. The UGG-tRNA
and the UGA-tRNA have been iso-
lated and sequenced (10). Spiralin is
the major membrane protein of S.
citri, and the spiralin gene was the
first mollicute gene to be expressed
in Escherichia coli. Many other mol-
llicute genes could not be expressed
in this way. When the spiralin gene
was eventually sequenced (7, 25), it
was found that there were no tryp-
tophan codons, and especially no
UGA codons, in the gene, in agree-
ment with the absence of tryp-
tophan in the protein, and hence the
E. coli ribosomes were able to trans-
late the whole gene. However, with
mollicute genes containing UGA-
tryptophan codons, the bacterial
ribosomes stop at the first UGA
codon encountered, as UGA means
“stop” to them, and no complete,
functional protein is obtained.

**FUNCTIONAL GENOMICS**

As interesting as the above stud-
ies might have been, they threw no
light on the most characteristic features of the spiroplasma; multiplication in, and transmission by, leafhopper vectors, mechanisms of phytopathogenicity, helical morphology and motility. To approach these problems, molecular genetics had to be used, and the required tools had first to be developed.

**Tools for Molecular Genetics of S. citri.** Electroploration rather than the polyethylene glycol method has been mainly used to promote penetration of DNA into S. citri cells (17). However, efficiency of transformation is also strain dependent. Strain GII3 was readily transformed with plasmid pBOT1, while strain R8A2 was very poorly transformed (X. Foissac, unpublished). This is one reason why GII3 was used in most of the work reported here, another reason being the fact that GII3 is easily transmitted to plants by leafhoppers.

Artificial plasmids have been developed to introduce, into S. citri cells, antibiotic resistance markers and other genes such as wild type genes to complement mutants. These plasmids contain the origin of DNA replication of the S. citri chromosome (oriC) (38). One such plasmid is pBOT1 (27, 29). This plasmid combines the 2kbp oriC region from S. citri strain R8A2 (quasi identical to that of strain GII3) with the tetracyclin resistance gene tetM from transposon Tn916, and linearized E. coli plasmid pBS containing a colE1 origin of replication. Because of its two origins of replication, oriC and colE1, pBOT1 is able to shuttle between S. citri and E. coli. When introduced into S. citri, pBOT1 replicates first as a free extrachromosomal element. However later, the plasmid integrates into the spiroplasmal chromosome during passaging of the spiroplasmal transformants. Plasmid integration was shown to occur by homologous recombination involving one crossover at the oriC region. Once integrated into the host chromosome, the whole plasmid is stably maintained. Plasmids derived from pBOT1 and containing S. citri wild type genes have been used, as will be seen later, to complement various S. citri mutants.

The production of S. citri mutants has been obtained by random insertion of transposon Tn4001 into genes causing their inactivation (14). Following random mutagenesis, mutants with the expected phenotype were then selected. pBOT1-derived plasmids have been used to produce mutants by disrupting targeted genes (fructose operon, motility gene scm1 through homologous recombination involving one crossover (12, 20).

**Mechanism of Phytopathogenicity.** In addition to citrus, S. citri has many other natural or experimental hosts (6). Periwinkle is a natural host in countries where citrus stubborn disease occurs, and has been used throughout the present studies. Transmission of S. citri strains or mutants to young periwinkle plants (6-8 leaf stage) was by the leafhopper C. haematoceps reared in the greenhouse (15, 22). About 0.1 µl of S. citri culture of known titer (10^8-10^9 cells/ml) was microinjected intra-abdominally into adult female leafhoppers. The injected insects were caged on the plants (10 insects per plant, two plants per S. citri strain or mutant) for a period of 2 weeks at 30°C (transmission period). After this period, the leafhoppers were killed with insecticide (dichlorvos), and the plants were kept in the greenhouse at 30°C for symptom development (observation period) and S. citri analyses.

Periwinkle plants infected with the wild type S. citri strain GII3 (w.t.GII3) showed symptoms during the first week of the observation period, and symptoms became soon severe. Among 257 mutants obtained by random insertion of transposon Tn4001 into the chromosome of w.t.GII3, one mutant, GMT553, began to induce symptoms only during the fourth week of the observa-
tion period, and thereafter symptoms remained mild for several weeks, until eventual reversion of the mutant to wild type \textit{S. citri} by excision of the transposon (16). It was found that in mutant GMT553, transposon Tn4001 was inserted in the first gene, \textit{fruR}, of the fructose operon. This gene is coding for an activator protein of the fructose operon (see below). The second gene of the operon, \textit{fruA}, codes for fructose permease which enables uptake of fructose by the spiroplasma cells, and its concomitant phosphorylation to fructose-1-phosphate. The third gene, \textit{fruK}, codes for 1-phosphofructokinase which phosphorylates fructose-1-phosphate into fructose-1,6-bis-phosphate. In mutant GMT553, transcription and expression of the fructose operon is abolished, and hence, the mutant cannot utilize fructose as carbon and energy sources for growth (19). To confirm the involvement of the fructose operon in pathogenicity and fructose utilization, mutant GMT553 was functionally complemented with various combinations of the fructose operon genes carried by pBOT-derived plasmids. These experiments demonstrated that both fructose utilization and phytopathogenicity were restored with \textit{fruR-fruA-fruK}, \textit{fruA-fruK} or \textit{fruA}, but not \textit{fruR} or \textit{fruR-fruA}. Additional mutants of the fructose operon have been produced by gene disruption using pBOT-derived plasmids, or by selection of spontaneous xylitol-resistant mutants (20). Results obtained with these additional mutants were very similar to those seen with mutant GMT553 (21). To sum up, plants infected with spiroplasmas capable of using fructose (fructose\textsuperscript{*}) showed symptoms as early as one week after the transmission period, and symptoms became always severe, while spiroplasmas unable to utilize fructose (fructose\textsuperscript{-}) induced symptoms much later, and symptoms remained mild until revertants, with a fructose\textsuperscript{*} phenotype, eventually occurred. It should be pointed out that both fructose\textsuperscript{*} and fructose\textsuperscript{-} spiroplasmas are able to use glucose.

\textit{S. citri} cells are restricted to the phloem sieve tubes. Strangely, the concentration of fructose and glucose in the sieve tubes seems to be very low (39) or undetectable (8, 13). Sucrose is the major carbohydrate in sieve tubes. Loading of sucrose into a sieve tube requires the action of companion cells. Fructose, in combination with UDP-glucose, is needed by the companion cell for sucrose loading (Fig. 1). If fructose\textsuperscript{*} \textit{S. citri} cells are present in the sieve tubes, they also use fructose, and drain it away from the companion cell. Thus, the companion cell and the spiroplasmas compete for fructose. Fructose utilization by the spiroplasmas will impair sucrose loading into the sieve tubes by the companion cells, and result in accumulation of carbohydrates in “source” leaves and depletion of carbon sources in “sink” tissues. The lower level of carbohydrates in sink organs could lead to growth impairment as observed for young leaves and roots. Chlorosis of older leaves might be a consequence of sugar accumulation in source leaves. Preliminary studies in healthy and \textit{S. citri}-infected periwinkle plants show a modification in soluble sugar concentration in source and sink leaves. These results seem to indicate that pathogenicity of sieve tube-restricted plant mollicutes may involve a novel mechanism in which fructose utilization by the pathogen interferes with the normal physiology of the plant. However, even though the fructose strains induce only mild symptoms, they are still pathogenic, indicating that mechanisms of pathogenicity other than fructose utilization, are probably involved.

\section*{Regulation of the fructose and trehalose operons of \textit{S. citri}.}

In the course of these studies, the fructose operon of \textit{S. citri} has been characterized functionally (19). In particular, the role of the first gene of
In vivo transcription of the operon is greatly enhanced by the presence of fructose in the growth medium while glucose has no effect. When fruR is not expressed (fruR mutants), transcription of the operon is not stimulated by fructose, and fructose fermentation is decreased, indicating that FruR, the protein product of fruR, is an activator of the fructose operon (22). Trehalose is the major sugar in leafhoppers and other insects. The trehalose operon of S. citri has a gene organization very similar to that of the fructose operon, and the first gene of the trehalose operon, treR, also codes for a transcriptional activator of the operon (W. Maccheroni, unpublished). This is the first time that transcriptional regulation of mollicute operons has been reported (22).

**Motility and Pathogenicity.**

*S. citri*, when growing on low-agar medium, forms fuzzy colonies with occasional surrounding satellite colonies due to the ability of the spiroplasmal cells to move through the agar matrix. In liquid medium, the helical cells flex, twist, and rotate rapidly. By using transposon Tn4001 insertional mutagenesis on w.t. *S. citri GII3*, a motility mutant, G540, was isolated on the basis of its nondiffuse, sharp-edged colonies. The mutant flexed at a low frequency and had lost its ability to rotate about the helix axis. The transposon was found to be inserted into an open reading frame coding for a putative polypeptide of 409 amino acids showing no significant homology with known proteins. The corresponding gene, *scm1*, was recovered from the w.t. strain, and introduced into mutant G540 by using plasmid pBOT1 as the vector. The appearance of fuzzy colonies and the observation of active rotary and flexional movements showed the motile phenotype to be restored. This functional complementation of the motility mutant 540 proves the *scm1* gene product to be involved in the motility mechanism of *S. citri* (23). In these studies, mutant G540 was obtained from a
high-passage culture of *S. citri* GII3 that had lost its ability to be transmitted to plants by the leafhopper vector. Therefore, a motility mutant was produced by disruption of gene *scm1* in a low-passage, insect-transmissible isolate of *S. citri* GII3. The *scm1*-disrupted motility mutant multiplied efficiently in the leafhoppers, and when introduced into periwinkle plants, produced symptoms indistinguishable from those obtained with w.t. strain GII3. These experiments show that motility of *S. citri* is not essential for its pathogenicity. They also indicate that gene disruption through homologous recombination with pBOT1-derived plasmids, is effective (12).

**Motility and Helicity: Role of the Cytoskeleton and associated Proteins.** Mollicutes are among bacteria with an internal cytoskeleton. The spiroplasmal cytoskeleton is a flat, monolayered, membrane-bound ribbon which follows the shortest helical line on the cellular coil (Fig. 2). The ribbon is composed of several, well-ordered protein fibrils. The fibrils in spiroplasmas were first discovered in 1974 (34). The nucleotide sequence of the *S.citri* fibril protein gene was established in 1991 (35). The mass of the fibril protein was deduced from the gene sequence and found to be 59 kDa. The cellular and molecular organization of the *S. melliferum* cytoskeleton has been studied very recently (31). The fibril protein forms tetramer subunits which assemble into flat fibrils. The subunits in the fibrils undergo conformational changes from circular to elliptical (Fig. 3), which results in shortening of the fibrils and helix contraction, or from elliptical to circular leading to length increase of the fibrils and cell helix. The cytoskeleton being bound to the spiroplasmal membrane over its entire length, acts as a scaffold, and controls the dynamic helical shape of the cell. The cytoskeleton is involved in motility through its linear contractibility and interactions with the cell membrane. The contractile cytoskeleton can thus be seen as a “linear motor” in contrast to the common bacterial “rotary motor” which is part of the flagellar apparatus.

The 59 kDa fibril protein consistently copurifies with other polypeptides (31). The 26 kDa spiralin is among these peptides, but also, most interestingly, the 45 kDa *scm1* gene product shown to be indispensable for motility (see above), and the product of gene *mreB1*, an actin-like protein (32). Mutations in *mreB* affect cell morphology (24). *MreB* genes are present in rod-shaped, filamentous and helical bacteria, but not in coccoid and pleiomorphic bacteria. Interestingly, in mollicutes, they are present in the helical spiroplasmas, but not in the pleiomorphic mycoplasmas. *S. citri* contains five homologues of *B. subtilis* *mreB* genes (W. Maccheroni, J. Renaudin...
et al., unpublished). Four of them are organized in tandem and transcribed as two separate operons. One, mreB1, is not part of an operon, and has the highest transcription level. Disruption of mreB1 resulted in reduced size colonies with sharp contours, characteristic of altered motility, and much longer cells with increased aggregation.

**Leafhopper transmission.** As seen above, the fructose-phytopathogenicity mutants induce symptoms in periwinkle plants much later than the wild type spiroplasmas, and these symptoms remain mild as long as no fructose+ revertants occur. If the number of wild type *S. citri* cells injected by a leafhopper (*C. haematoceps*) into a periwinkle plant is supposed to be 100%, the number of cells corresponding to a putative mutant impaired in its ability to be transmitted by the leafhoppers, would be expected to be much lower. If so, it would take some time for the mutant in the plant to reach titers high enough for symptoms to appear, and thus symptoms would occur late, but contrary to a fructose-phytopathogenicity mutant, symptoms would be severe. Tn4001 transpositional mutant G76 is precisely a mutant with such properties (A. Boutareaud and C. Saillard, unpublished). Plants infected with mutant G76 show symptoms 4 to 5 weeks later than those infected with wild type strain GII3, but symptoms induced are severe. Mutant G76 multiplies and leafhoppers as efficiently as the wild type strain. However, under conditions where leafhoppers injected with the wild type spiroplasma transmit the spiroplasma to 100% of the plants exposed to transmission, those injected with mutant G76 infect only one plant out of two. This suggests that mutant G76 is injected into plants by the leafhoppers less efficiently than the wild type strain. To check this possibility, the number of spiroplasma cells injected by a leafhopper through a parafilm membrane into culture medium, was determined. As expected, 20 times less mutant cells were transmitted through parafim membranes than were wild type cells. This result was confirmed by measuring the number of spiroplasma cells (colony forming units) present in the salivary glands of injected leafhoppers. This number was 20 times higher for the wild type spiroplasma cells than for cells of mutant G76. These results suggest that mutant G76 is injected into plants by the leafhoppers less efficiently than the wild type strain.

Finally, transmission of *S. citri* by the leafhopper vector must involve adherence to and invasion of insect host cells. A putative *S. citri* gene inactivated by insertion of Tn4001 in mutant G76, gene sc76, was sequenced. The gene product contains 466 amino acids (51.8 kDa), has 3 transmembrane α helices, a cystein at position 24, a signal peptide characteristic of bacterial lipoproteins, and 12 tryptophane codons (1UGG, 11UGA). The protein has sequence homology with lipoprotein MG040 of *Mycoplasma genitalium*, but the role of this protein is unknown.
adhesion related protein (SARP1) has recently been purified (2).

GENOMICS: 
THE SPIROPLASMA CITRI 
GENOME PROJECT

The physical and genetic map of the S. citri genome was established in 1992 (36, 37, 38). Over 100 genetic markers could be localized on the map. The first approach to sequence the S. citri genome was based on this map. Three large restriction fragments totaling the whole genome were purified by pulsed field gel electrophoresis. The fragments were subjected to partial digestion with restriction enzyme Sau3A. Eleven libraries were constructed which yielded 6,000 sequence reads. To obtain further sequences, a shotgun approach was decided. The spiroplasma DNA was mechanically sheared, and two plasmid libraries were constructed. In one, the plasmid inserts were in the range of 3.5 kbp and yielded 7,000 reads. In the second, the 1.5 kbp inserts, not yet sequenced, will lead to 3000 reads. A 20 kbp insert miniBAC library, for long distance assembly, was obtained from DNA partially digested with Sau3A, and will also lead to 1,400 reads.

At this time, 1,262 kbp (69.4%) of the spiroplasma genome have been assembled, each base having been sequenced 3.4 times (n = 3.4). Additional 114 kbp (6.3%) occur as singlets (n = 1). The genome should be entirely sequenced and annotated by 2003.

LITERATURE CITED


