

Phytoplasma PMU1 exists as linear chromosomal and circular extrachromosomal elements and has enhanced expression in insect vectors compared with plant hosts

Tania Y. Toruño,^{1†} Martina Šeruga Musić,²
Silvia Simi,^{1‡} Mogens Nicolaisen³ and
Saskia A. Hogenhout^{1,4*}

¹Department of Entomology, The Ohio State University-Ohio Agricultural Research and Development Center (OARDC), Wooster, OH 44691, USA

²Department of Biology, Faculty of Science, University of Zagreb, HR-10000 Zagreb, Croatia

³Aarhus University, Department of Integrated Pest Management, DK-4200 Slagelse, Denmark

⁴Department of Disease and Stress Biology, The John Innes Centre, Norwich Research Park, Colney Lane, Norwich NR4 7UH, UK.

Summary

Phytoplasmas replicate intracellularly in plants and insects and are dependent on both hosts for dissemination in nature. Phytoplasmas have small genomes lacking genes for major metabolic pathways. Nevertheless, their genomes harbour multicopy gene clusters that were named potential mobile units (PMUs). PMU1 is the largest most complete repeat among the PMUs in the genome of Aster Yellows phytoplasma strain Witches' Broom (AY-WB). PMU1 is *c.* 20 kb in size and contains 21 genes encoding DNA replication and predicted membrane-targeted proteins. Here we show that AY-WB has a chromosomal linear PMU1 (L-PMU1) and an extrachromosomal circular PMU1 (C-PMU1). The C-PMU1 copy number was consistently higher by in average approximately fivefold in insects compared with plants and PMU1 gene expression levels were also considerably higher in insects indicating that C-PMU1 synthesis and expression are regulated. We found that the majority of AY-WB virulence genes lie on chromosomal PMU regions that have similar gene content and organization as PMU1

providing evidence that PMUs contribute to phytoplasma host adaptation and have integrated into the AY-WB chromosome.

Introduction

Phytoplasmas are plant pathogenic bacteria that are transmitted by phloem-feeding insects of the order Hemiptera, including leafhoppers, planthoppers and psyllids (Weintraub and Beanland, 2006). In plants, phytoplasma replication is mainly restricted to phloem sieve cells, whereas they invade the entire insect vector. They replicate in various tissues of the insect including the salivary glands from which they are introduced into the phloem during feeding (Hogenhout *et al.*, 2008). Four phytoplasma genomes have been completely sequenced: Onion Yellows phytoplasma strain M (OY-M) (Oshima *et al.*, 2004) and Aster Yellows phytoplasma strain Witches' Broom (AY-WB) (Bai *et al.*, 2006) that belong to different subgroups of *Candidatus* Phytoplasma asteris; *Ca.* Phytoplasma australiense strain Australian grapevine yellows (AUSGY) (Tran-Nguyen *et al.*, 2008); and the apple proliferation (AP) disease agent *Ca.* Phytoplasma mali strain AT (Kube *et al.*, 2008). The genomes are greatly reduced in the number of genes involved in metabolism and other basic processes consistent with phytoplasmas having obligate associations with their hosts (Moran and Plague, 2004). Phytoplasmas have not been cultured in cell-free medium. Genome comparisons indicated that the gene composition and content of phytoplasmas are different from those of other members of the class Mollicutes, such as spiroplasmas and mycoplasmas, some of which can be cultured (Oshima *et al.*, 2004; Hogenhout and Seruga Music, 2010).

Despite their small genomes, phytoplasma chromosomes contain a substantial number of open reading frames (ORFs) that are present as multiple copies (Oshima *et al.*, 2004; Bai *et al.*, 2006; Tran-Nguyen *et al.*, 2008). These ORFs are frequently organized in clusters, which we named potential mobile units (PMUs), because they resemble composite transposons (Bai *et al.*, 2006). The largest PMU on the AY-WB chromosome, named PMU1, is 20 kb in length and contains 21 ORFs. PMU1 is

Accepted 5 July, 2010. *For correspondence. E-mail saskia.hogenhout@bbsrc.ac.uk; Tel. (+44) 1603 450393; Fax (+44) 1603 450045. Present addresses: †University of Nebraska – Lincoln, N311 Beadle Center, Lincoln, NE 68588-0660, USA; ‡Institute of Biology, State University of Campinas, UNICAMP, Campinas, São Paulo, Brazil.

flanked at one border by a *tra5* gene, which is predicted to encode a full-length transposase, and a 327 bp repeat. The other border of PMU1 is flanked by a truncated *tra5* sequence and another 327 bp repeat (Bai *et al.*, 2006). Between these borders, PMU1 contains predicted ORFs for DNA replication (*ssb*, *dnaB* and *dnaG*) and synthesis (*tmk*), recombination (*himA*) proteins, predicted membrane-targeted and secreted proteins, and proteins with unknown functions (Bai *et al.*, 2006; Hogenhout and Seruga Music, 2010). All ORFs have the same transcriptional orientation on the genomic strand with the exception of the truncated *tra5* sequence. Furthermore, the 327 bp repeats are inverted relative to one another. Other PMUs on the AY-WB chromosome, such as PMU2, 3 and 4, have parallel gene compositions as PMU1, but are smaller because some of the ORFs are truncated (predicted pseudogenes) or are not present (Bai *et al.*, 2006). Furthermore, PMU3 and 4 are flanked by the 327 bp repeat at only one border that is adjacent to the full-length *tra5* genes, but not at the other border, and PMU2 lacks these 327 bp repeats completely (Bai *et al.*, 2006). Hence, PMUs 2, 3 and 4 appear to be degenerated versions of PMU1.

In addition to these four PMUs, the AY-WB and OY-M chromosomes harbour several other apparently degenerated PMUs that together account for 10.2% and 14.1% of the total number of predicted ORFs of AY-WB and OY-M respectively (Bai *et al.*, 2006). The PMUs exhibit a non-random distribution across the genome and tend to congregate as tandem or multiple repeats (Bai *et al.*, 2006; Hogenhout and Seruga Music, 2010). PMU-like sequences are also prevalent in the AUSGY chromosome (Tran-Nguyen *et al.*, 2008), and the AP chromosome harbours a sequence similar to PMU1 and a PMU remnant (Kube *et al.*, 2008). The multicopy genes in the sequenced and other phytoplasma genomes have also been named sequence-variable mosaics (SVMs) (Jomantiene and Davis, 2006; Jomantiene *et al.*, 2007), and it was suggested that the SVMs were created from attacks of phages of the order *Caudovirales* (Wei *et al.*, 2008). However, no phage core genes typically used for identification of phages were identified in the phytoplasma genomes (Wei *et al.*, 2008). Hence, it remains unclear whether the multicopy genes are derived from phage attacks.

Because AY-WB PMU1 resembles a composite transposon that contains genes for DNA replication and is flanked by large 327 bp inverted repeats, we hypothesized that PMU1 can excise from the chromosome and replicate independently (Bai *et al.*, 2006; Hogenhout *et al.*, 2008). Here we provide evidence that PMU1 exists in two forms, a linear PMU1 (L-PMU1) located in the AY-WB chromosome and an extrachromosomal circular PMU1 (C-PMU1). We found that C-PMU1 abundance and

PMU1 transcript levels are higher in insects than in plants suggesting that genes within this region may encode proteins relevant to phytoplasma adaptation to an insect host.

Results

Evidence for the presence of an extrachromosomal PMU1 unit by PCR

As a first step to determine whether AY-WB has an extrachromosomal circular PMU1 (C-PMU1), we designed primers p1 and p2 (Table S1) that point outward towards the borders of the chromosomal L-PMU1 (Fig. 1A and B) (GenBank Accession No.: NC_007716). These primers generated a product of ~1700 bp from DNA isolated from AY-WB-infected China aster and *Macrostes quadrilineatus*, which is the insect vector of AY-WB (Fig. 1C). To provide independent confirmation of this result we designed primers p5 and p6 (Table S1) that similarly point outward but are located more internally in the L-PMU1 region (Fig. 1A and B). These primers generated a product of approximately 4000 bp from DNA of AY-WB-infected China aster and leafhoppers (Fig. 1C). In both cases, these results suggest that the left and right ends of L-PMU1 connect to form a circular DNA molecule.

We sequenced the p1–p2 PCR products originating from AY-WB-infected plants and insects (Fig. 2A). Sequence analyses confirmed that the left and right ends of L-PMU1 join in C-PMU1; in C-PMU1 the *sigF* gene at the right border of L-PMU1 was located immediately downstream of the 328 bp repeat at the left border of L-PMU1 (Fig. 1A and B). However, C-PMU1 lacks the truncated *tra5* gene and the 327 bp repeat at the right border of L-PMU1 (Fig. 1A and B, area indicated with Δ , Fig. 2A). Except for this deletion, the L-PMU1 and C-PMU1 sequences were identical, but have nucleotide sequence differences with other PMUs in the AY-WB genome. The C-PMU1 sequences from China aster and leafhoppers were also identical except for a two-nucleotide difference at the site where sequences corresponding to the left and right ends of L-PMU1 connect in C-PMU1 (Fig. 2A). We also sequenced the ends of the p5–p6 PCR product and this confirmed that the PCR product contained the predicted regions of C-PMU1 (Fig. 1B) and the sequences were identical to L-PMU1 (NC_007716). These results suggest that C-PMU1 is derived from L-PMU1 and not from the other PMUs in the AY-WB genome, and that both forms are present in insects and plants.

Evidence for the presence of extrachromosomal PMU1 by Southern blot hybridizations

To provide further evidence for the presence of L-PMU1 and C-PMU1, we carried out Southern blot hybridization

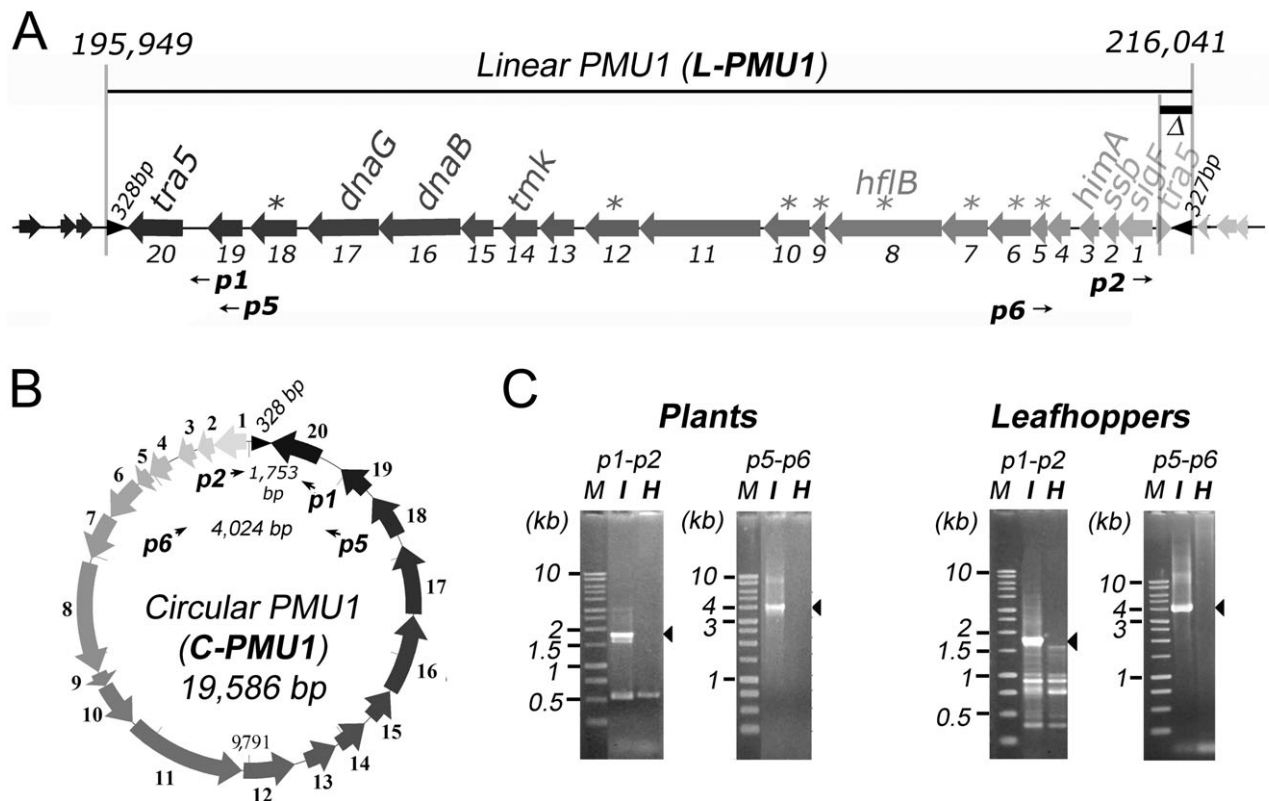


Fig. 1. Evidence for the presence of a circular extrachromosomal PMU1 by PCR.

A. Linear PMU1 (L-PMU1) in the AY-WB chromosome between nucleotides 195 949 and 216 041 bp as indicated. L-PMU1 is flanked by 327/328 bp inverted repeats (arrowheads) and contains 21 predicted ORFs of which 20 appear full-length (numbered block arrows) oriented in the same direction on the negative sense chromosomal strand (Bai *et al.*, 2006; Hogenhout *et al.*, 2008). The identity of genes with similarities to genes with known functions in GenBank are specified above the block arrows. ORFs encoding predicted *trans*-membrane or secreted proteins are indicated with asterisks. Primers p1, p2, p5 and p6 used in PCRs (C) are shown.

B. The proposed extrachromosomal circular PMU1 (C-PMU1) with a predicted size of 19 586 bp. The grey scales of block arrows and ORF numbers correspond to those in L-PMU1 (A). The location of primers p1, p2, p5 and p6 and the predicted sizes of PCR products are indicated.

C. Agarose gels showing the PCR products of ~1.7 kb of primer pair p1 and p2, and ~4 kb of primer pair p5 and p6 amplified from total DNA isolated from AY-WB-infected China aster plants and leafhopper vector *Macrosteleus quadrilineatus* (lanes indicated with I and arrowheads at right of gels). Total DNA extracted from uninfected plants and leafhoppers was used as controls in the PCRs (lanes indicated with H). The sizes of the marker DNA ladder (M) in kb are indicated to the left of the gels. The p1–p2 PCR products were sequenced, confirming the predicted configuration of C-PMU1 (B) and revealing identical sequence as L-PMU1 except for the region indicated with Δ (A).

experiments. The first experiment was conducted to confirm that PMU1 exists as one copy and not as tandem repeats in the AY-WB chromosome. Whole genomic DNA isolated from AY-WB-infected and healthy *Arabidopsis thaliana* plants was digested with BglII and PciI, which recognize restriction sites external to the inverted repeats of PMU1 in the AY-WB chromosome (Fig. 3A). In the presence of one PMU1 copy, BglII would be expected to generate a fragment of 25 925 bp, and PciI a fragment of 29 129 bp (Fig. 3A). In the event PMU1 exists as tandem repeats, BglII would generate a fragment of no less than 45 481 bp or greater in the event that there are multiple PMU1 repeats. Similarly, in the event of more than one L-PMU1 repeat, PciI digestion would generate a fragment of at least 48 685 bp. The digested genomic DNA was size-separated by pulsed-field gel electrophoresis and the

Southern blot was hybridized to DIG-labelled probe 14S generated by a pair of primers 14Fs3/14R (Table S1) that anneal to genes *AYWB_188* (ORF 7 on PMU1, Fig. 1) and *AYWB_189* (ORF 6 on PMU1, Fig. 1) respectively. ORFs 6 and 7 are uniquely present in PMU1 in the AY-WB genome (Bai *et al.*, 2006) and therefore probe 14S3 should hybridize to PMU1 fragments only. Indeed, a BLASTN search of the probe 14S3 nucleotide sequence against the entire AY-WB genome did not yield additional sequences besides the target area in PMU1 (data not shown). The hybridized Southern blots showed fragments of ~26 kb for BglII and ~29 kb for PciI (Fig. 3B). Thus, there is one copy of PMU1 located on the AY-WB chromosome.

The whole genomic DNA preparations isolated from AY-WB-infected plants and insects were digested with

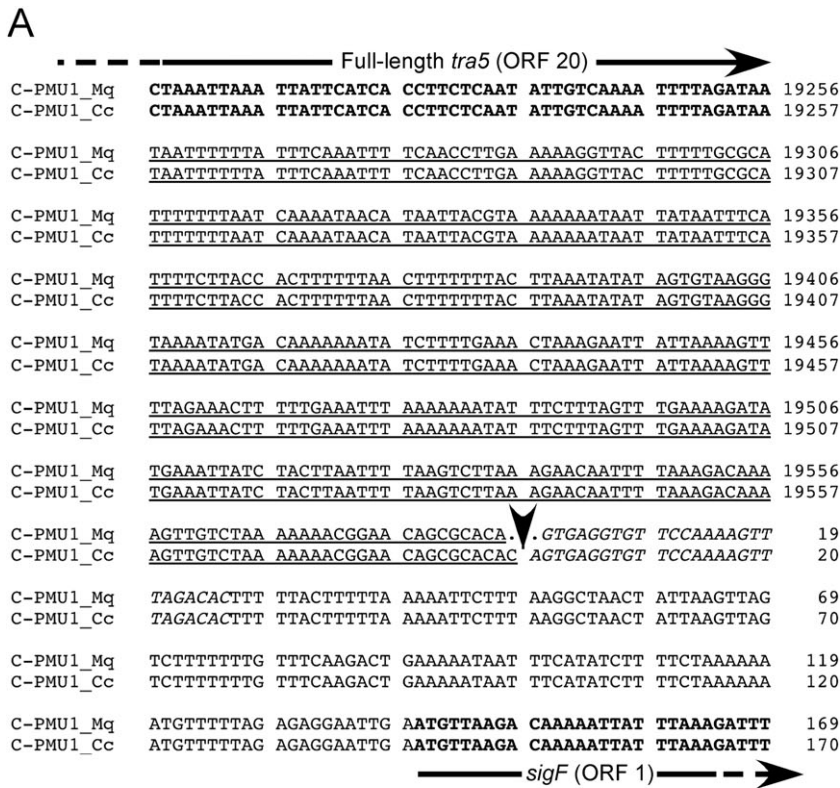
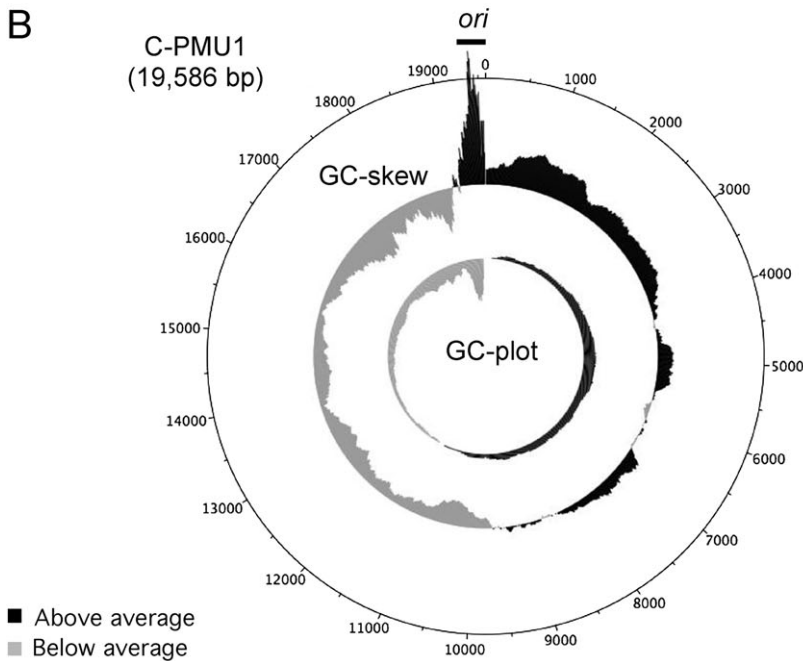
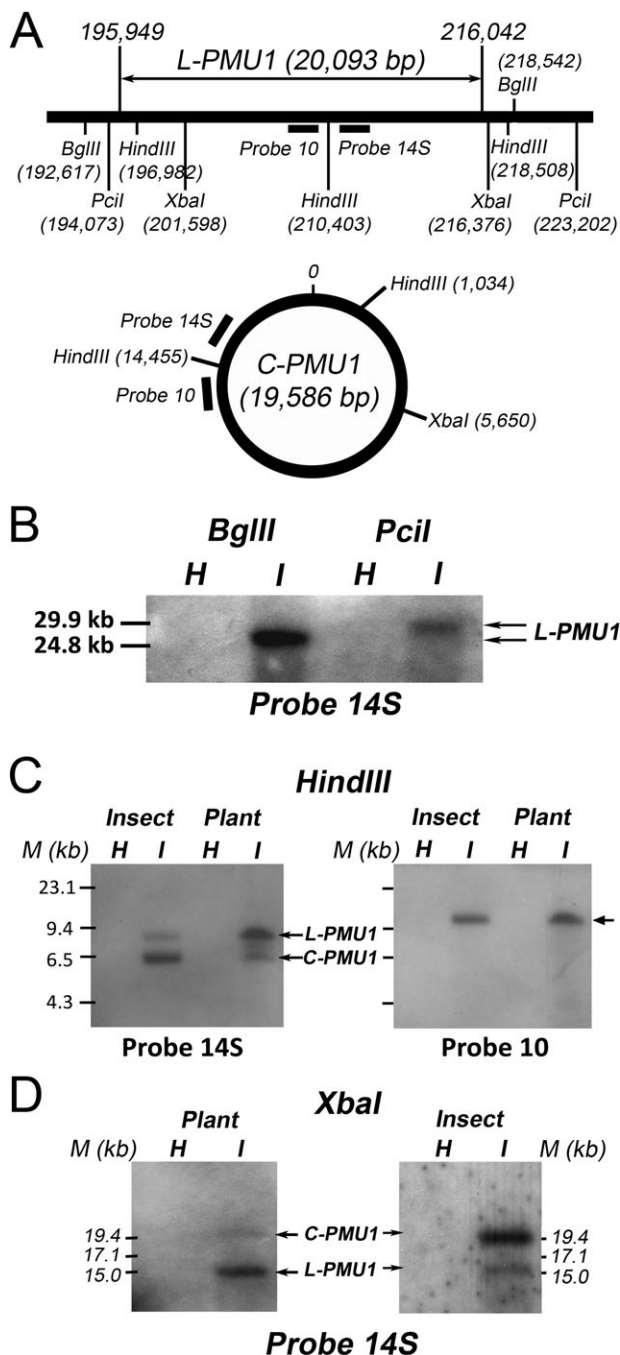


Fig. 2. The 328 bp region is the predicted origin of replication (*ori*) of C-PMU1. A. C-PMU1 contains one 328 bp repeat. The p1–p2 PCR products shown in Fig. 1C were sequenced from aster (Cc) and leafhoppers (Mq) and sequences were aligned. The C-PMU1 sequence between the two full-length ORFs (indicated as bold font) of *tra5* (ORF 20 in Fig. 1B) and *sigF* (ORF 1 in Fig. 1B) is shown. In C-PMU1, the 328 bp repeat (underlined) lying to the left of L-PMU1 (Fig. 1A) is connected (arrow head) to part of the truncated *tra5* sequence (indicated in italics font) to the right of L-PMU1 (Fig. 1A), resulting in the absence of the 327 bp repeat and a large portion of the truncated *tra5* sequence present in L-PMU1 (shown as Δ in Fig. 1A). The aster and leafhopper C-PMU1 sequences show a two-nucleotide difference at the junction site (arrow head). B. Cumulative GC-skew and GC-plot predicted based on the sequences shown in (A) and of L-PMU1. The 328 bp repeat is the switch point from below to above average GC-skew and is also more AT-rich (GC-plot) than surrounding sequences. These are typical features of *ori* regions.



HindIII and XbaI. These enzymes were selected, because they recognize sites internally and externally to the chromosomal L-PMU1 inverted repeats and hence can be used to differentiate between the L-PMU1 and C-PMU1 forms (Fig. 3A). Southern blot hybridizations with DIG-labelled probe 14S3 showed two bands for HindIII in the

plant and insect genomic DNA (Fig. 3C, left panel). The larger band is ~8 kb (Fig. 3C), corresponding to the expected 8105 bp HindIII chromosomal L-PMU1 fragment (Fig. 3A), and the smaller band is ~6 kb (Fig. 3C) corresponding to the expected 6165 bp HindIII extrachromosomal C-PMU1 fragment (Fig. 3A). A Southern blot of



HindIII-digested DNA was also hybridized to DIG-labelled probe 10. This probe was generated from the PCR product of primer pair 10F1/R1 (Table S1) that anneals to *AYWB_186* (ORF 9 on PMU1, Fig. 1), which is unique to PMU1 (Bai *et al.*, 2006) as confirmed by a BLASTN search of the probe nucleotide sequence against the entire AY-WB genome sequence (data not shown). Hybridizations with probe 10 generated one ~13 kb band for plant and insect DNA (Fig. 3C, second panel) consistent with the predicted size of the 13 421 bp HindIII–HindIII frag-

Fig. 3. L-PMU1 exists as a single copy in the AY-WB chromosome and C-PMU1 is present at higher concentrations relative to L-PMU1 in insects compared with plants.

A. Schematic illustrations of L-PMU1 and C-PMU1 showing the positions of restriction sites and probes 10 and 14S, which correspond to sequences unique in PMU1 of AY-WB (Bai *et al.*, 2006).

B. L-PMU1 exists as a single repeat in the AY-WB chromosome. Total DNA of AY-WB-infected (I) and non-infected plants (H) was digested with BglIII and PciI that recognize restriction sites externally to the borders of PMU1. Fragments were size-separated on agarose gels by pulsed-field gel electrophoresis and Southern blots were hybridized with DIG-labelled probe 14S. The fragments that hybridized are similar in size to the predicted 25 925 bp (BglIII) and 29 129 bp (PciI) fragments corresponding to a single PMU1 copy.

C and D. C-PMU1 is present at higher concentrations in insects compared with plants. HindIII and XbaI digest internally and externally to the borders of PMU1 and hence can distinguish fragments derived from L-PMU1 and C-PMU1 as indicated. Total DNA isolated from AY-WB infected (I) and uninfected (H) plants and insects were size-separated by conventional (HindIII) or pulsed-field gel (XbaI) electrophoresis and Southern blots were hybridized with DIG-labelled probes 10 and 14S. Fragments corresponding to L-PMU1 and C-PMU1 are indicated. XbaI cuts once in C-PMU1, generating the entire linear C-PMU1 of the predicted size of ±19.5 kb. The HindIII fragments of L-PMU1 and C-PMU1 detected with probe 10 are similar in size (arrow at right). In C, the sizes of marker DNA bands (M) for both blots are shown in kb at the left side of the blot at left.

ment in L-PMU1 and C-PMU1 (Fig. 3A). Southern blot hybridizations of the XbaI digests with probe 14S3 showed a band of ~15 kb (Fig. 3D) corresponding to the expected 14 778 bp XbaI–XbaI fragment of chromosomal L-PMU1. The second ~20 kb band (Fig. 3D) corresponds to the entire extrachromosomal C-PMU1 of a predicted size of 19 586 bp that contains one predicted XbaI restriction site (Fig. 3A). Together these Southern blot hybridization results indicate that the chromosomal L-PMU1 and extrachromosomal C-PMU1 have similar sequences.

C-PMU1 and PMU1 transcripts are more abundant in insects than in plants

Southern blot hybridizations of the HindIII- and XbaI-digested DNA revealed that C-PMU1 has a higher copy number in insects compared with plants (Fig. 3C and D). This was confirmed by quantitative PCR (qPCR) experiments; C-PMU1 levels were 4.6, 5.9 and 5.8 times higher in *M. quadrilineatus* compared with *Arabidopsis*, *China aster* and *lettuce* respectively.

To determine how the presence of L-PMU1 and C-PMU1 relates to PMU1 gene expression levels, transcript levels for each of the PMU1 genes were measured. Quantitative reverse transcriptase-PCR (qRT-PCR) experiments performed on total RNA of AY-WB-infected and healthy *M. quadrilineatus* and infected and healthy *Arabidopsis* and *Nicotiana benthamiana* demonstrated that the vast majority of the PMU1 genes are significantly

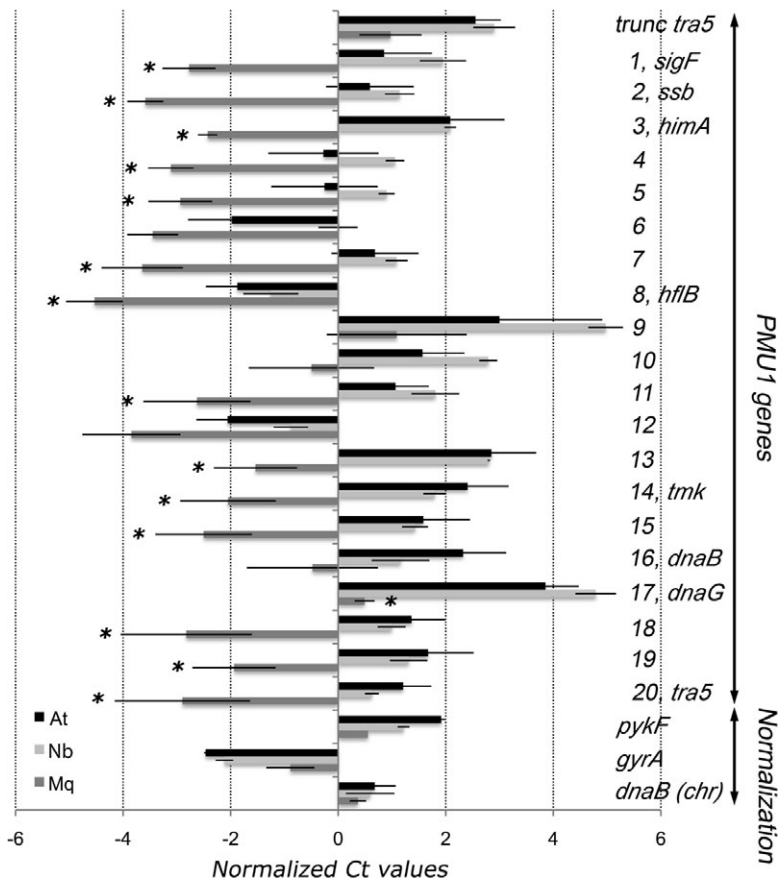


Fig. 4. The AY-WB PMU1 genes have higher expression levels in insects than in plants. Primer pairs were designed for all PMU1 genes for qRT-PCR reactions of total RNA isolated from AY-WB-infected *A. thaliana* (At), *N. benthamiana* (Nb) and the leafhopper vector *M. quadrilineatus* (Mq). The Ct values were normalized against the average expression levels of genes encoding pyruvate kinase (*pykF*), DNA gyrase subunit A (*gyrA*) and chromosomal replicative DNA replicase (*dnaB_chr*) that are not encoded on PMUs (Arashida *et al.*, 2008). PMU1 genes are numbered as in Fig. 1, and annotations are indicated when available. Negative Ct levels indicate higher expression levels compared with those of *pykF*, *gyrA* and *dnaB_chr*, whereas positive Ct levels indicate lower expression levels relatively to these genes. Bars indicate the averages of three independent biological replicates, each consisting of two replicate reactions. Lines on the bars indicate $1 \times$ SEM values. All PMU1 genes were more highly expressed in *M. quadrilineatus* compared with both *A. thaliana* and *N. benthamiana*. The majority of these gene expression differences were significant as indicated with asterisks (P -value < 0.05, Student's *t*-test), and in all these cases PMU1 gene expression levels were not significantly different between *A. thaliana* and *N. benthamiana*.

more abundantly expressed in insects than in plants (Fig. 4). Thus, AY-WB generates more C-PMU1 during insect infection, which is correlated with a higher expression level of PMU1 genes.

The observation that C-PMU1 is more abundant in insects compared with plants and the presence of genes involved in DNA replication (*dnaB*, *dnaG*, *ssb*) on PMU1 make it more likely that C-PMU1 is a self-replicating plasmid. To obtain additional evidence for this, we conducted GC-skew ($[(G+C)]/[G+C]$) and %GC-plot analyses on the circularized L-PMU1 that has the C-PMU1 sequence with one 328 bp repeat (Fig. 2A). This showed above average GC-skew and low GC content at the 328 bp sequence and GC-skew and GC-content switch points at the 328 bp sequence and directly opposite of this sequence at *c.* 10 kb in C-PMU1 (Fig. 2B). These results point to a bidirectional replication strategy for C-PMU1 with the 328 bp sequence as the origin of replication (*ori*) and directly opposite, at *c.* 10 kb, the predicted replication terminus (*ter*) (Arakawa and Tomita, 2007). The accurate initiation of bidirectional DNA replication from *ori* is governed by the interaction of DnaB and DnaG in *Escherichia coli* (Hiasa and Marians, 1999), and both proteins are encoded on PMU1 (Fig. 1).

Discussion

In this work, we demonstrate that there are two forms of PMU1, L-PMU1 and C-PMU1, in the AY-WB genome. The first is integrated in the AY-WB chromosome, whereas the latter is a circular extrachromosomal DNA. Sequence analyses of the p1-p2 PCR product revealed that C-PMU1 and L-PMU1 are identical in the overlapping regions (Fig. 2A), and Southern blot hybridizations of digested DNA based on the L-PMU1 restriction map indicated that C-PMU1 and L-PMU1 are highly similar in sequence. Phytoplasmas in insects have higher concentrations of C-PMU1 and higher expression levels of PMU1 genes relatively to phytoplasmas in plants (Fig. 3B and C). This finding is particularly interesting given that C-PMU1 harbours one inverted repeat that is located immediately upstream of the first gene of PMU1, *sigF*, and has the opposite orientation in C-PMU1 relative to the inverted repeat upstream of *sigF* in L-PMU1 (Fig. 1A and B). We hypothesize that the upstream *sigF* region in C-PMU1 may contain promoters and enhancers that drive the expression of PMU1 genes.

Our finding that the C-PMU1 copy number consistently increases during AY-WB infection of insects compared with

that of plants (Figs 3 and 4) indicates that C-PMU1 can be amplified implying the presence of an active regulatory mechanism. This is consistent with PMU1 containing genes involved in DNA replication, including *dnaB*, *dnaG* and *ssb* (del Solar *et al.*, 1998) (Fig. 1). Furthermore, GC-skew and GC-plot analyses predict a bidirectional replication of the proposed C-PMU1 with an *ori* at the 328 bp sequence between *tra5* and *sigF* genes, and a site directly opposite of the *ori* at c. 10 kb where the replication is predicted to end (Fig. 2B) (Hiasa and Mariani, 1999; Arakawa and Tomita, 2007). The GC-skew is also apparent in the L-PMU1 region of the AY-WB chromosome (Bai *et al.*, 2006). Based on these observations, we propose that C-PMU1 is a self-replicating unit (plasmid), which increases in copy numbers when AY-WB is in the insect environment. Since C-PMU1 does not contain all the genes required for initiation of DNA replication, including for instance *dnaA* (del Solar *et al.*, 1998), such an increase in copy number would, at least partly, be regulated by the AY-WB chromosome.

All sequenced phytoplasmas chromosomes contain PMUs in a range of forms, sizes and copy numbers. This suggests that two events may have occurred. First, extra-chromosomal PMUs appear to have reintegrated into phytoplasma chromosomes followed by sequence diversification and gene degeneration. C-PMU1 reintegration is consistent with PMU *tra5* being a member of the IS3 family, IS150 subgroup insertion elements that tend to integrate into their own inverted repeats (Mahillon and Chandler, 1998). This hypothesis is also consistent with the observation that PMUs of OY-M, AY-WB and AUSGY often cluster together in certain regions of the phytoplasma chromosomes (Bai *et al.*, 2006; Tran-Nguyen *et al.*, 2008), and the occurrence of multiple (and degenerate) versions of PMUs in phytoplasmas, including PMUs 2, 3 and 4 in AY-WB (Bai *et al.*, 2006). Second, phytoplasmas may acquire PMUs by horizontal gene transfer, probably from other phytoplasmas. We note that these two events (i.e. reintegration and horizontal gene transfer) are not mutually exclusive. Indeed, there is evidence for horizontal DNA exchange of the *mgs1* gene and flanking *tra5* sequences of AY-WB PMU3 sequences among plant-associated mollicutes (Bai *et al.*, 2004; Hogenhout *et al.*, 2008).

PMU1 encodes a series of membrane-targeted proteins (protein IDs AYWB_177, AYWB_183 and AYWB_185-190 corresponding to ORFs 18, 12 and 10 through 5 in Fig. 1A) (Bai *et al.*, 2006). One possibility is that these proteins associate into membrane-associated structures that allow for horizontal exchange of C-PMU1 to other phytoplasmas through conjugative transfer or a natural transformation process (Lovell *et al.*, 2009) for which the insect might provide a better environment than the plant. An alternative possibility is that PMU1 is involved in

AY-WB adaptation to the insect environment. The PMU1-encoded membrane proteins may change the phytoplasma membrane surface composition to facilitate insect cell infection, such as adhering to and invasion of insect cells, and avoiding or counteracting insect immune defence responses. It is striking that, in AY-WB, PMU regions contain the majority of predicted secreted proteins that are putative effectors involved in manipulation of plant and insect hosts (Bai *et al.*, 2009). Previously, 56 Secreted AY-WB Proteins (SAPs) were identified from the AY-WB genome sequence (Bai *et al.*, 2009). Herein, we found that 34 SAPs lie on PMU regions. Also, the AY-WB phytoplasma effector protein SAP11, which targets plant cell nuclei and affects plant development, is encoded on a PMU-like region along with other candidate effectors (Hogenhout *et al.*, 2008; Bai *et al.*, 2009; Strauss, 2009). Thus, PMUs are apparently involved in phytoplasma virulence and may be retained in the genome because they enhance phytoplasma fitness.

Whereas integration of PMUs into the chromosome probably occurs, that of C-PMU1 should have happened very recently as both sequence and Southern blot results revealed that L-PMU1 and C-PMU1 are almost identical. However, the chance of detection of a recent integration event is very low. Therefore, we should also consider the alternative possibility that C-PMU1 is being synthesized *de novo* from L-PMU1. This is possible when considering that the proliferation of insertion sequences of the IS3 family occurs through a copy-and-paste mechanism involving covalently closed circles that is different from the cut-and-paste mechanism used by many other insertion sequences and mobile elements (Mahillon and Chandler, 1998; Duval-Valentin *et al.*, 2004). This is best exemplified by IS911, which is an IS3 family member (Mahillon and Chandler, 1998). In the first step of IS911 proliferation, a single DNA strand at the inverted repeat of one end is cleaved and transferred to the same strand of DNA at the other end leading to figure-eight structures when IS911 is located on a plasmid (Duval-Valentin *et al.*, 2004). These figure-eight structures are precursors for replication, which starts at the inverted repeats and lead to the regeneration of the parental transposon concomitantly with the production of double-stranded covalently closed circles (Duval-Valentin *et al.*, 2004). Such circles have also been observed for IS2, IS3 and IS150 (Duval-Valentin *et al.*, 2004). Interestingly, *tra5* encodes a full-length OrfAB transposase (Bai *et al.*, 2006). Thus, C-PMU1 might replicate from a L-PMU1 intermediate in which the two 327/328 bp repeats at each end are joined. Because OrfAB is absolutely essential to catalyse the formation of the intermediate structure and replication initiation (Duval-Valentin *et al.*, 2004), C-PMU1 might be synthesized solely when *tra5* is expressed. This is in agreement with our finding that the upregulation of *tra5* coincides with a greater copy

number of C-PMU1 in insects (Figs 3 and 4). However, both inverted repeats (IRL and IRR) are present in the covalent IS911 circles, whereas we have demonstrated that C-PMU1 contains only a single 328 bp repeat at the left site of L-PMU1 (Fig. 1A and B). Our observation of nucleotide differences at the junction of the left and right ends between C-PMU1 copies (Fig. 2A) suggests that there may be another way the L-PMU1 left and right ends associate to initiate replication. The possibility that C-PMU1 might replicate independently following synthesis from L-PMU1 has to be considered as well. Future research should reveal whether C-PMU1 is a self-replicating plasmid or is being generated *de novo* from L-PMU1. Nonetheless, our data demonstrate that PMUs exist as extrachromosomal DNAs.

Experimental procedures

Phytoplasma source and maintenance

AY-WB strain was originally collected from diseased lettuce plants in Celeryville, Ohio in 1998 (Zhang *et al.*, 2004) and sequenced in 2006 (Bai *et al.*, 2006). It is maintained in greenhouse and growth chambers by serial transmission to China aster (*Callistephus chinensis*) plants by aster leafhoppers (*M. quadrilineatus* L.). For all experiments, AY-WB-infected plants were generated by exposure of young plants to AY-WB-infected *M. quadrilineatus* for 1 week. Plant materials used for DNA/RNA extraction were harvested 2–3 weeks after the first day of leafhopper exposure as soon as plants were fully symptomatic. For generating AY-WB-infected *M. quadrilineatus* leafhoppers, healthy fourth-instar nymphs were exposed to AY-WB-infected plants for at least 1 week. These leafhoppers were transferred to healthy plants for at least 2 weeks and then harvested for DNA and RNA extraction. Controls were young plants exposed to non-infected leafhoppers and fourth-instar leafhoppers exposed to non-infected plants.

PCR and sequencing

Genomic DNA was isolated from AY-WB phytoplasma-infected China aster (*C. chinensis*) and leafhoppers (*M. quadrilineatus*) by DNeasy plant maxi kit (Qiagen, Valencia, CA). Healthy China aster and leafhoppers DNA was also isolated and used as controls. Two primer pairs, p1–p2 and p5–p6 (Table S1), located on the edges of the linear form of PMU1 and pointing outward were designed to confirm the circular form of PMU1. BLAST searches of primer sequences against the AY-WB genome (GenBank Accession No.: NC_007716) showed that primer p1 hybridizes to the chromosomal region between PMU1 AYWB-175 and AYWB_176 (between ORFs 19 and 20, respectively, in Fig. 1A; nucleotides 197 332–197 318) and four other regions in the AY-WB chromosome adjacent to *tra5* sequences (nucleotides 56 311–56 325, 314 885–314 871, 383 076–383 090 and 635 870–635 856), whereas primer p2 hybridizes solely to PMU1 AYWB_194 (ORF 1 in Fig. 1A; nucleotides 215 166–

215 186). Also, primer p5 hybridizes to PMU1 AYWB_176 (ORF 19 in Fig. 1A; nucleotides 197 454–197 431) and two other regions in the AY-WB genome (nucleotides 382 954–382 977 and 635 992–635 969), whereas primer p6 hybridizes solely to one region of PMU1 corresponding to AYWB_189 and AYWB_190 (ORFs 5 and 6 in Fig. 1A; nucleotides 213 017–213 036). Typical thermal cycler conditions for generating the p1–p2 amplification product were 95°C for 2 min, 35 cycles of 95°C for 1 min, 56°C for 1 min and 68°C for 2 min, followed by a final extension of 68°C for 10 min. Those for generating the p5–p6 amplification product were 95°C for 2 min, 35 cycles of 95°C for 1 min, 50°C for 1 min and 68°C for 15 min, followed by a final extension of 72°C for 10 min. The products were size-separated on 1% agarose gels. The p1–p2 PCR products were gel-purified by Qiaquick Gel Extraction Kit (Qiagen) following manufacturer's instructions. Fragments were sequenced to completion from both directions at the Molecular and Cell Imaging Center, Wooster, Ohio using an ABI Prism 3100xl genetic analyser (Applied Biosystems, Foster City, CA).

Assessment of cumulative GC-skew and GC-plot

Cumulative GC-skew and GC-plot for C-PMU1 were analysed with DNAPlotter::Release 1.3 software (Wellcome Trust Sanger Institute, Cambridge, UK) (Carver *et al.*, 2009). Settings were 5000 bp Window Size and 25 bp Step Size for the GC-skew ($[GC]/[G+C]$), and 10 000 bp Window Size and 50 bp Step Size for the %GC-plot.

Determination of C-PMU1 copy number

Relative C-PMU1 copy numbers in insects versus plants were determined by qPCR using an AYWB189 assay (unique gene on PMU1) and normalizing using AYWB007 and AYWB064 assays (not located on PMUs) (Table S1). DNA was extracted from 50 mg of material from infected plants and insects using DNeasy (Qiagen, Hilden, Germany). qPCR were performed as for qRT-PCRs (described below) except that DNA was used as template directly in the qPCR. AYWB189 Ct values were normalized by subtracting an average of Ct AYWB007 and Ct AYWB064 values (ΔCt). Finally, normalized AYWB189 Ct values from insects were subtracted from plant values to give $\Delta\Delta Ct$. Assuming a PCR efficiency of 100% the relative copy number in insects was calculated using $2^{\Delta\Delta Ct}$.

Southern blot hybridization experiments

Genomic DNA of up to 100 kb was extracted from AY-WB phytoplasma-infected *Arabidopsis thaliana* plants and adult *M. quadrilineatus* with OmniPrep™ Kit (G-Biosciences, Genotech, Maryland Heights, MO, USA) according to manufacturer's instructions. One hundred milligrams of plant tissue and 10 insects per sample were used. Healthy plants and insects were also included in experiments as controls.

Three micrograms of high genomic DNA were digested separately with BglII, PciI, XbaI and HindIII. BglII, PciI and XbaI digests were separated by pulsed-field gel electrophoresis on 1% gel in 0.5× TBE buffer chilled to 14°C, with 18 h run

time and 0.1–9 s pulse switch, at the 6 V cm⁻¹ (CHEF-DR II system, Bio-Rad Laboratories, Hercules, CA, USA). HindIII digests were separated by conventional electrophoresis on 0.8% agarose gel in 1× TBE buffer for 3 h at 60 V. All gels were subjected to depurination, denaturation and neutralization according to standard protocols, as described in Sambrook *et al.* (1989). DNA fragments were transferred to Hybond-N+ membranes (GE Healthcare Life Sciences, Amersham, UK) overnight by capillary transfer and fixed to the membrane with the Stratilinker UV cross-linker (Stratagene, Agilent Technologies, La Jolla, CA, USA). Labelling of the probes was conducted with the DIG System with dTTP to DIG-11-dUTP ratio of 4:1 and primer pairs 14Fs3/14R and 10F1/10R1 (Table S1) to produce probes 14S and 10, respectively, and the following PCR conditions: denaturation at 95°C for 2 min; 34 cycles of denaturation at 95°C for 1 min, annealing 58°C for 1 min, 68°C for 2 min; 10 min 68°C. BLAST searches of the primer sequences against the AY-WB genome sequences (GenBank Accession No.: NC_007716) showed that the full-length primers 10F1 and 10R1 hybridize solely to PMU1 *AYWB_186* (ORF 9 in Fig. 1A; nucleotides 208 967–208 989 and 209 210–209 188 respectively) on the AY-WB chromosome. The full-length primers 14Fs3 and 14FR hybridize solely to *AYWB_188* and *AYWB_189*, respectively (ORFs 7 and 6 in Fig. 1A; nucleotides 212 207–212 230 and 212 419–212 399 respectively) on the AY-WB chromosome. Probes were hybridized to the blots following the manufacturer's instructions (Roche Diagnostics GmbH, Roche Applied Science, Mannheim, Germany). Probe-target hybrids were detected by immunological detection using anti-DIG antibodies conjugated to alkaline phosphatase, and a chemiluminescent substrate CDP-Star (Roche Diagnostics GmbH, Roche Applied Science), followed by exposure to X-ray film (Fuji medical X-ray film Super RX, Fujifilm UK, Bedfordshire, UK).

Expression analysis by qRT-PCR

Total RNA was extracted from infected and uninfected leaves of *A. thaliana*, *N. benthamiana* and from *M. quadrilineatus* adults using TRIzol (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol. The RNA was treated with DNase I (Invitrogen) according to the manufacturer's instructions. cDNA was prepared from 2 µg of total RNA using the High Capacity RNA-to-cDNA Kit (Applied Biosystems, Foster City, CA, USA). qRT-PCR was performed using 2 µl of template in 1× Power SYBR Green Master Mix and 0.9 µM of each primer in a total volume of 15 µl using an ABI PRISM 7900 HT sequence detection system (Applied Biosystems). Primers designed for AY-WB are listed in Table S1. After completion of the PCR, a dissociation curve was run to confirm amplification of a single PCR product. As a control for DNA contamination, DNase-treated RNA that had not been reverse transcribed was included in one gene assay. Results were expressed as threshold cycle (Ct) values. Two replicate reactions were run for each sample and their Ct values averaged. Gene expression was normalized to three genes involved in chromosome replication (*dnaB_chr* and *gyrA*) and basic metabolism (*pyfk*). The averaged Ct value of the controls was subtracted from the Ct value of the PMU1 genes under investigation to give ΔCt. Difference in expression

between plant and insect samples was calculated by subtracting the respective ΔCt values to give a ΔΔCt value.

Acknowledgements

S.A.H. is supported by The John Innes Centre and The Gatsby Charitable Foundation. The John Innes Centre is grant-aided by the Biotechnology and Biological Sciences Research Council (BBSRC). This research was partly funded by EMBO Short Term Fellowship ASTF141.00-2008 awarded to M.S.M., and by the National Research Institute of the USDA Cooperative State Research, Education, and Extension Service, Grant 2002-35600-12752 awarded to S.A.H., and the Ohio Agricultural Research and Development Center competitive grants programme. We thank Angela Strock, Heather Kingdom and The John Innes Centre Insectary for generating the leafhopper and plant materials used in experiments, and Sir David Hopwood, David Hodgson, John Mansfield and Allyson MacLean for helpful discussions.

References

- Arakawa, K., and Tomita, M. (2007) The GC skew index: a measure of genomic compositional asymmetry and the degree of replicational selection. *Evol Bioinform Online* **3**: 159–168.
- Arashida, R., Kakizawa, S., Hoshi, A., Ishii, Y., Jung, H.Y., Kagiwada, S., *et al.* (2008) Heterogeneous dynamics of the structures of multiple gene clusters in two pathogenetically different lines originating from the same phytoplasma. *DNA Cell Biol* **27**: 209–217.
- Bai, X., Zhang, J., Holford, I.R., and Hogenhout, S.A. (2004) Comparative genomics identifies genes shared by distantly related insect-transmitted plant pathogenic mollicutes. *FEMS Microbiol Lett* **235**: 249–258.
- Bai, X., Zhang, J., Ewing, A., Miller, S.A., Jancso Radek, A., Shevchenko, D.V., *et al.* (2006) Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. *J Bacteriol* **188**: 3682–3696.
- Bai, X., Correa, V.R., Toruño, T.Y., Ammar, E.D., Kamoun, S., and Hogenhout, S.A. (2009) AY-WB phytoplasma secretes a protein that targets plant cell nuclei. *Mol Plant Microbe Interact* **22**: 18–30.
- Carver, T., Thomson, N., Bleasby, A., Berriman, M., and Parkhill, J. (2009) DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* **25**: 119–120.
- Duval-Valentin, G., Marty-Cointin, B., and Chandler, M. (2004) Requirement of IS911 replication before integration defines a new bacterial transposition pathway. *EMBO J* **23**: 3897–3906.
- Hiasa, H., and Marians, K.J. (1999) Initiation of bidirectional replication at the chromosomal origin is directed by the interaction between helicase and primase. *J Biol Chem* **274**: 27244–27248.
- Hogenhout, S.A., and Seruga Music, M. (2010) Phytoplasma genomics, from sequencing to comparative and functional genomics – what have we learnt? In *Phytoplasmas – Genomes, Plant Hosts and Vectors*. Weintraub, P.G., and Jones, P. (eds). Wallingford: CABI, pp. 19–36.

- Hogenhout, S.A., Oshima, K., Ammar, E.D., Kakizawa, S., Kingdom, H.N., and Namba, S. (2008) Phytoplasmas: bacteria that manipulate plants and insects. *Mol Plant Pathol* **9**: 403–423.
- Jomantiene, R., and Davis, R.E. (2006) Clusters of diverse genes existing as multiple, sequence-variable mosaics in a phytoplasma genome. *FEMS Microbiol Lett* **255**: 59–65.
- Jomantiene, R., Zhao, Y., and Davis, R.E. (2007) Sequence-variable mosaics: composites of recurrent transposition characterizing the genomes of phylogenetically diverse phytoplasmas. *DNA Cell Biol* **26**: 557–564.
- Kube, M., Schneider, B., Kuhl, H., Dandekar, T., Heitmann, K., Migdoll, A.M., *et al.* (2008) The linear chromosome of the plant-pathogenic mycoplasma 'Candidatus Phytoplasma mali'. *BMC Genomics* **9**: 306.
- Lovell, H.C., Mansfield, J.W., Godfrey, S.A., Jackson, R.W., Hancock, J.T., and Arnold, D.L. (2009) Bacterial evolution by genomic island transfer occurs via DNA transformation *in planta*. *Curr Biol* **19**: 1586–1590.
- Mahillon, J., and Chandler, M. (1998) Insertion sequences. *Microbiol Mol Biol Rev* **62**: 725–774.
- Moran, N.A., and Plague, G.R. (2004) Genomic changes following host restriction in bacteria. *Curr Opin Genet Dev* **14**: 627–633.
- Oshima, K., Kakizawa, S., Nishigawa, H., Jung, H.Y., Wei, W., Suzuki, S., *et al.* (2004) Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. *Nat Genet* **36**: 27–29.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M., and Diaz-Orejas, R. (1998) Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**: 434–464.
- Straus, E. (2009) Phytoplasma research begins to bloom. *Science* **325**: 388–390.
- Tran-Nguyen, L.T., Kube, M., Schneider, B., Reinhardt, R., and Gibb, K.S. (2008) Comparative genome analysis of 'Candidatus Phytoplasma australiense' (subgroup *tuf*-Australia I; *rp-A*) and 'Ca. Phytoplasma asteris' Strains OY-M and AY-WB. *J Bacteriol* **190**: 3979–3991.
- Wei, W., Davis, R.E., Jomantiene, R., and Zhao, Y. (2008) Ancient, recurrent phage attacks and recombination shaped dynamic sequence-variable mosaics at the root of phytoplasma genome evolution. *Proc Natl Acad Sci USA* **105**: 11827–11832.
- Weintraub, P.G., and Beanland, L. (2006) Insect vectors of phytoplasmas. *Annu Rev Entomol* **51**: 91–111.
- Zhang, J., Hogenhout, S.A., Nault, L.R., Hoy, C.W., and Miller, S.A. (2004) Molecular and symptom analyses of phytoplasma strains from lettuce reveal a diverse population. *Phytopathology* **94**: 842–849.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.