Bulletin of Insectology ***60*** *(2): 119-120, 2007*
ISSN 1721-8861

Dynamic structures in phytoplasma genomes: sequence-variable mosaics (SVMs) of clustered genes

Robert E. DAVIS[1], Rasa JOMANTIENE[2], Yan ZHAO[1]
[1]*Molecular Plant Pathology Laboratory, USDA-Agricultural Research Service, Beltsville, USA*
[2]*Institute of Botany, Vilnius, Lithuania*

Abstract

Emergence of the phytoplasma clade from an *Acholeplasma*-like ancestor gave rise to an intriguing group of cell wall-less prokaryotes through a remarkable and continuing evolutionary process. In a ceaseless progression, phytoplasmas have evolved reduced genomes, lost biochemical pathways for synthesis of nutrients supplied by hosts, and gained capabilities for transkingdom parasitism and pathogenicity in plants and insects. While continued genome degradation has made phytoplasmas increasingly host dependent, their small, AT-rich genomes have evolved conspicuous flexibility that enables rapid responses to host signals, successful evasion of host surveillance, and adaptation to shifting environments encountered during obligate, transkingdom parasitism. Recent work revealed that multiple, sequence-variable mosaics (SVMs) of clustered genes and repetitive extragenic palindromes are characteristic features of genome architecture in phylogenetically diverse phytoplasma species. SVMs are apparently of ancient origin, while current forms result from dramatic and more recent events. The dynamic nature of SVMs could account for their composite structure and potential for rapid changes significant in phytoplasma-host interactions.

Key words: mobile genetic unit, mobile element, recombination, simple sequence repeats.

Introduction

While the phytoplasma clade experienced evolutionary genome shrinkage resulting in the loss of biosynthetic pathways and causing phytoplasmas to become increasingly host dependent, the small, AT-rich genomes evolved surprising flexibility enabling rapid changes in response to host chemical signals, successful evasion of host surveillance systems, and adaptation to shifting host environments during obligate, transkingdom parasitism of plants and insect vectors. The mechanisms accounting for this flexibility have yet to be fully explored.

Targeted sequencing of segments of the clover phyllody (CPh) phytoplasma genome and comparative analyses with the completely sequenced (Oshima *et al*., 2004) genome of '*Candidatus* Phytoplasma asteris'-related strain OY-M revealed a unique genome architecture consisting of sets of genes clustered in sequence variable mosaics (here termed SVMs) (Jomantiene and Davis, 2006). Potential mobile units (PMUs) described in strain AY-WB (Bai *et al.*, 2006) and SVMs share similar compositions. Based in part on the presence of transposase and phage-related protein gene relics in the SVMs, it was suggested that some SVM genes were probably acquired through horizontal transfer (Jomantiene and Davis, 2006). Interestingly, two research groups independently proposed that sizes and numbers of SVMs could account in part for variations in phytoplasma genome size (Bai *et al.*, 2006; Jomantiene and Davis, 2006).

Further studies revealed that homologous SVMs characterize the genomes of phylogenetically diverse phytoplasma species and thus represent a common feature of phytoplasma genome architecture (Jomantiene *et al.*, 2007). The results indicated that SVMs could function as platforms of genome plasticity, providing *loci* for acquisition of new genes and for targeting of mobile genetic elements to specific regions in phytoplasma chromosomes.

Materials and methods

Phytoplasmas used were members of 16S rDNA RFLP groups 16SrI ('*Ca.* P. asteris'-related strains), 16SrIII (strain SP1), 16SrV ('*Ca*. Phytoplasma ulmi'-related strain EY), 16SrVI ('*Ca*. Phytoplasma trifolii'-related strain PWB), and 16SrXII (strain STOL) (Lee *et al*., 1998). DNA was extracted from plant tissues and used as template in polymerase chain reactions (PCRs). DNA amplification was primed by primer pairs whose annealing sites were located at conserved sequences that flanked a highly variable region of SVMs. PCR products were cloned in *Escherichia coli*, and were sequenced using automated DNA sequencing to achieve at least 3-fold coverage per base position in sequencing both strands. Sequence data for OY-M (GenBank no. NC_005303) and AY-WB (GenBank no. NC_007716) phytoplasmas were obtained from the GenBank database. Nucleotide and amino acid sequence and protein structure analyses were done as previously (Jomantiene and Davis, 2006) and using Artemis release 4: sequence visualization and annotation software (Rutherford *et al.*, 2000).

Results

We hypothesized that SVMs are a common feature of phytoplasma genomes and searched for homologous stretches in phylogenetically diverse phytoplasmas through genome-wide, targeted amplification of SVM segments. Segments included a palindromic sequence

and a CA/CAA-rich stretch possibly having a role analogous with that of short-sequence repeats (SSRs) in transcriptional regulation of downstream genes. In strain OY-M, chromosomal positions of SVM-associated genes were non-randomly distributed, forming clusters homologous with SVM sequences recognized previously (Jomantiene and Davis, 2006). Homologous sequences occur within AY-WB sequence stretches termed potential mobile units (PMUs) by Bai *et al.* (2006).

In each amplified SVM segment, conserved sequences flanked a highly variable region. Results were consistent with the concept that the hyper-variable region is a repeated, length-variable and ORF-variable feature in phytoplasma genomes.

Palindromic DNA sequences were undeviatingly located just upstream of certain conserved sequences, while ORFs occurring immediately upstream of the palindromes varied among different hyper-variable SVM segments.

SVM organization, in which hyper-variable regions are flanked by constant genes, supports the possibility that a joining site for site-specific recombination was located close to the conserved genes. The consistent occurrence of a conserved palindromic or quasi-palindromic DNA sequence at or near this hypothetical site suggested that palindromic DNA may be associated with combinatorial joining sites in phytoplasma genomes.

Discussion

In their descent from Gram-positive walled ancestors, phytoplasmas underwent massive reductions in genome size. Yet, phytoplasma genomes embody considerable plasticity, enabling exceptional flexibility in responses to the shifting landscape of transkingdom parasitism.

The dynamic architecture of phytoplasma genomes is evidenced by recently discovered, unique clusters of diverse genes that exist in multiple, sequence-variable mosaics (SVMs) (Jomantiene and Davis, 2006; Jomantiene *et al*., 2007). The evidence indicated that these unique genomic features formed early in evolution of the phytoplasma clade, after divergence of phytoplasmas and acholeplasmas from a common ancestor.

While SVMs appear to be widespread in phytoplasma genomes, mechanisms responsible for this unique genome architecture remain unexplained. SVMs appear to be composite structures formed in part by multiple events of targeted mobile element attack (Jomantiene and Davis, 2006). The organization of SVMs suggests that palindromic DNA sequences are located at or near combinatorial breakpoints in phytoplasma DNA and may target nearby transposition. Thus, the nonrandomly distributed clusters of repeated DNA could serve as recombination hotspots facilitating genome plasticity, strain variation, and niche adaptation in phytoplasmas.

In phytoplasmas, the occurrence of chromosomal regions composed of arrays of SVM sequences, the possible involvement of palindromic DNA in site-specific recombination in SVMs, probable horizontal acquisition and site-specific assembly of SVM genes, and the composite structure of phytoplasmal SVMs make it tempting to speculate that phytoplasmas possess an unknown mechanism that shares some similarities with the integron/mobile gene cassette system.

Phytoplasmal SVMs appear to have assembled, at given loci, genes encoding disparate functions and including secreted and/or transmembrane, cell surface-interacting proteins, in part through multiple events of mobile element attack recurrently targeted to SVMs. Resulting extensive gene duplication potentially could have enabled repeated excision and reinsertion of integrative sequence stretches into varied positions in SVMs through homologous recombination. This view implies that SVM length and sequence polymorphisms resulted in part from recurrent transpositions of targeted mobile elements, a formative background that conceivably could have yielded the observed, nonrandom chromosomal distribution of composite structures containing clustered, repeated genes.

Acknowledgements

We are grateful to Ellen Dally and Jonathan Shao for excellent technical assistance.

References

BAI X., ZHANG J., EWINGH A., MILLER S. A., RADEK A. J., SCHEVCHENKO D., TSUKERMAN K., WALUNAS T., LAPIDUS A., CAMPBELL J. W., HOGENHOUT S. A., 2006.- Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts.- *Journal of Bacteriology,* 188: 3682-3696.

JOMANTIENE R., DAVIS R. E., 2006.- Clusters of diverse genes existing as multiple, sequence-variable mosaics in the phytoplasma genome.- *FEMS Microbiology Letters,* 255: 59-65.

JOMANTIENE R., ZHAO Y., DAVIS R. E., 2007.- Sequence variable mosaics: Composites of recurrent transposition characterizing the genomes of phylogenetically diverse phytoplasmas.- *DNA and Cell Biology*, in press.

LEE I.-M, GUNDERSEN-RINDAL D. E., DAVIS R. E., BARTOSZYK I., 1998.- Revised classification scheme of phytoplasmas based on RFLP analyses of 16S rRNA and ribosomal protein gene sequences.- *International Journal of Systematic and Evolutionary Microbiology*, 48: 1153-1169.

OSHIMA K., KAKIZAWA S., NISHIGAWA H., JUNG, H.-Y., WEI W., SUZUKI S., ARASHIDA R., NAKATA D., MIYATA S., UGAKI M., NAMBA S., 2004.- Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma.- *Nature Genetics,* 36: 27-29.

RUTHERFORD K., PARKHILL J., CROOK J., HORSNELL T., RICE P., RAJANDREAM M. A., BARRELL B., 2000.- Artemis: sequence visualization and annotation.- *Bioinformatics,* 16: 944-945.

Corresponding author: Robert E. DAVIS (e-mail: robert.davis@ars.usda.gov), USDA-Agricultural Research Service, 10300 Baltimore Avenue, Beltsville, MD USA 20705.