Bulletin of Insectology ***64*** *(Supplement): S13-S14, 2011*
ISSN 1721-8861

'Flavescence dorée' phytoplasma genome: a metabolism oriented towards glycolysis and protein degradation

Patricia CARLE[1,2], Sylvie MALEMBIC-MAHER[1,2], Nathalie ARRICAU-BOUVERY[1,2], Delphine DESQUÉ[1,2], Sandrine EVEILLARD[1,2], Sébastien CARRÈRE[3], Xavier FOISSAC[1,2]
[1]*INRA, UMR1332 Fruit Biology and Pathology (BFP), Villenave d'Ornon, France*
[2]*University of Bordeaux, UMR1332 Fruit Biology and Pathology (BFP), Villenave d'Ornon, France*
[3]*INRA, UMR441, Laboratoire des Interactions Plantes-Microorganismes (LIPM), Castanet-Tolosan, France*

Abstract

The 670 kbp chromosome of 'flavescence dorée' phytoplasma line FD92 was partially sequenced by pyrosequencing and SOLEXA. More than 94% of the chromosome could be assembled and the 22 largest contigs representing 85% of the chromosome were annotated under the semi-automatic annotation platform iANT. Out of 464 chromosomal coding sequences (CDS), 174 CDS (38%) were involved in information transfer (DNA replication, protein production, RNA modification and regulation), 88 CDS (19%) were encoding metabolic enzymes, 40 CDS (9%) corresponded to transporters, 8 CDS (1%) corresponded to cellular processes, whereas 145 CDS (31%) remained cryptic. At this stage of incomplete assembly, repeated sequences were underestimated and transposon and phage-related CDS (2%) could not yet be precisely evaluated. FD92 phytoplasma possesses a complete glycolytic pathway and has a prominent system for proteolysis possibly resulting from the adaptation to its woody hosts.

Key words: Phloem-restricted bacteria, plant pathology, bacterial genomics, ATP-dependent zinc protease, HflB.

Introduction

Vineyards of southern Europe are affected by the 'flavescence dorée' phytoplasma (FD), inducing an epidemic and quarantine disease (Boudon-Padieu, 2002). Whereas important progress has been made in phytoplasma classification and ecology, little is known about mechanisms of phytoplasma physiology, phytopathogenicity and transmission by insects. These research areas should benefit from the comparative analysis of phytoplasma genomes. A physical and genetic map of the FD92 line chromosome (670 kbp) was established (Malembic-Maher *et al.*, 2008). To decipher FD phytoplasma genome, enriched DNA of FD line FD92 was prepared from broad beans by repeated bis-benzymide-CsCl density gradient centrifugations. A preliminary set of shotgun sequences were homologous to known phytoplasma sequences in 27% of the cases. The sequencing of the FD phytoplasma genome could therefore be undertaken.

Materials and methods

FD92 line was transmitted to *Vicia faba* var. Aqua Dulce using *Scaphoideus titanus* collected in infected vineyards of South-West France in 1992, and maintained since then by transmission to *V. faba* by *Euscelidius variegatus*. Total nucleic acids were extracted from stems and leaf midribs by the CTAB procedure (Murray *et al.*, 1980) followed by 4 repeated bis-benzymide Cesium chloride gradients (Kollar *et al.*, 1989).

Pyrosequencing of 8 µg of FD92 DNA enriched fraction was performed on a 454 GLX (Roche) by GATC BIOTECH (Germany). This produced 528,000 flowgrams totalling 122 Mb with an average of 230 bp read length. After assembly with Newbler (Roche), 1,067 contigs containing at least ten reads were compared to the genome of '*Candidatus* Phytoplasma asteris' line OY-M and to the non redundant database (nr) with BLASTX and BLASTN considering significant alignments below a cut-off E value of 10^{-10}. Out of the 1,067 contigs, positive contigs were therefore selected as they showed homology to phytoplasmas or other bacteria. Plant DNA was screened out from the assembly. The 72 selected contigs were composed of 107,000 reads with an average coverage of 39 X. Additional contigs showing no homology but having coverage in the range of 20-100X were also selected after PCR performed on healthy and FD-infected *V. faba*. Phytoplasma contigs were further extended by genome walking (BD Bioscience, USA). Finally, 85 contigs totalling 629 kbp (94% of the chromosome) resulted from the final assembly with Phred-Phap-Consed Package. To overcome the errors on homopolymers falsely introduced by 454GFLX reads, Solexa sequencing was performed by GATC Biotech (Germany). About 100 Mb of 35 bp reads allowed to correct the homopolymers errors.

The 22 largest contigs, ranging from 11 to 58 kbp (85% of the chromosome), were annotated under the semi-automatic annotation platform iANT developed by LIPM at INRA Toulouse.

Results

The FD92 chromosome has 21.1% G+C, possesses 2 rRNA operons, 27 tRNAs as well as 464 coding sequences (CDS) (table 1). Most of the sequenced CDS (38%) participate to the information transfer, *i.e.* replication, transcription and production of the proteins. Regarding to DNA rearrangement, FD92 should be capable of homologous recombination as it has a complete *rec*A gene, *ruv*A and *ruv*B genes encoding DNA

helicases and RuvX, a putative holiday junction resolvase. In addition, 5 putative phage recombinases of *xer*C family and phage-related integrases could be evidenced among the 464 CDS of FD92 chromosome.

Table 1. Structural RNAs and functional distribution CDS from partial FD92 chromosomal sequences.

Functional classification	Number of CDS
16S-23S-5S rRNA operons	2
tRNAs	27
CDS	464
Information transfer	174
Protein translation	98
DNA replication and modification	43
RNA modification	30
Regulation	3
Metabolism	88
Protein degradation	24
Glycolysis and energy metabolism	10
Nucleic acid metabolism	10
Lipids and phospholipids metabolism	9
Cofactors	3
Amino acids metabolism	2
Others	30
Transport	40
ABC transporters	25
Protein secretion	6
Cation dependent P-ATPase	4
2-hydroxycarboxylate transporters	2
Others	3
Cellular processes	8
Phage related	8
Transposon related	1
Cryptic	145

Metabolic genes account for 19% of the sequenced CDS out of which 27% participate to protein degradation. It comprises 15 copies often truncated of *Hfl*B/*Fts*H encoding ATP-dependent zinc proteases as well as various other proteases. '*Ca.* P. asteris' line OY-M possesses 20 copies of *Hfl*B/*Fts*H genes obviously not all functional (Oshima *et al.*, 2004). As most of phytoplasma genomes sequenced to date, FD92 chromosome encodes the complete glycolytic pathways. Nine CDS are controlling lipids and phospholipids metabolism, whilst 10 CDS are involved in amino acid synthesis. Finally, the precise role of 30 other CDS metabolic pathway could not precisely determined.

Most of the transporters were ABC transporters (62%) aimed to transport maltose, spermidin/putrescine, oligopeptides, methionine, cobalt and manganese/zinc. Other transporters were assigned to protein secretion (sec system), to cation dependent P-ATPase and two 2-hydroxycarboxylate (malate/citrate) cation transporters.

As long as the remaining 15% repeated part of the genome are not sequenced and annotated, it is too early to state about the presence of complete potential mobiles units homologous to '*Ca*. P. asteris' PMU (Bai *et al.*, 2006). However, *dna*D, *ins*K (transposase), *ssb* and *hup* (putative bacterial nucleoid binding protein) were scarcely and randomly distributed into contigs. *Hfl*B, were not adjacent to CDS homologous to PMU genes.

As virulence and defence factors, FD92 has a type III haemolysin and a superoxide dismutase (*sod)* that allow phytoplasma to resist to oxidative stress.

Discussion

FD phytoplasms complete genome deciphering is not yet achieved. However, guidelines can already be drawn: FD phytoplasma metabolism is oriented towards protein degradation and transport. Its carbon, oxydoreductive and energetic metabolism seems to rely exclusively on glycolysis, maltose and 2-hydroxycarboxylate import and degradation, and P-ATPases. These findings will help a better understanding of FD phytoplasma specialisation in woody hosts.

Acknowledgements

We greatly acknowledge the regional council of Aquitaine and CIVB for their financial support. We also thank K. Guionnaud and J-L. Danet for greenhouse and insectarium maintenance and M. Fontanieu for excellent technical assistance.

References

BAI X. D., ZHANG J. H., EWING A., MILLER S. A., RADEK A. J., SHEVCHENKO D. V., TSUKERMAN K., WALUNAS T., LAPIDUS A., CAMPBELL J. W. AND 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(10): 3682-3696.

BOUDON-PADIEU E., 2002.- Flavescence dorée of the grapevine: knowledge and new developments in epidemiology, etiology and diagnosis.- *Atti Giornate Fitopatologiche*, 1: 15-34.

KOLLAR A., SEEMÜLLER E., 1989.- Bases composition of the DNA of mycoplasma-like organisms associated with various plant diseases.- *Journal of Phytopathology*, 127: 177-186.

MALEMBIC-MAHER S., CONSTABLE F. E., CIMERMAN A., ARNAUD G., CARLE P., FOISSAC X., BOUDON PADIEU E., 2008.- A chromosome map of the flavescence dorée phytoplasma.- *Microbiology*, 154: 1454-1463.

MURRAY M. G, THOMPSON W. F., 1980.- Rapid isolation of high molecular weight plant DNA.- *Nucleic Acids Reseach*, 8: 4321-4326.

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(1): 27-29.

Corresponding author: Xavier FOISSAC (e-mail: foissac@bordeaux.inra.fr), INRA and University of Bordeaux UMR1332 Fruit Biology and Pathology, 71 avenue Edouard Bourlaux, BP81, F-33883, Villenave d'Ornon, France.