Complete Genome Sequence of "*Candidatus* Tremblaya princeps" Strain PCVAL, an Intriguing Translational Machine below the Living-Cell Status

Sergio López-Madrigal,[1,2] Amparo Latorre,[1,2,3] Manuel Porcar,[4] Andrés Moya,[1,2,3] and Rosario Gil[1,2]*

*Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, Apartado Postal 22085, 46071 València, Spain[1];
Departament de Genètica, Universitat de València, Dr. Moliner, 50, 46100 Burjassot (València), Spain[2]; Área de Genómica y Salud,
Centro Superior de Investigación en Salud Pública (CSISP), Avenida de Cataluña 21, 46020 Valencia, Spain[3]; and
Fundació General de la Universitat de València. Apartado Postal 22085, 46071 València, Spain[4]*

The sequence of the genome of "*Candidatus* Tremblaya princeps" strain PCVAL, the primary endosymbiont of the citrus mealybug *Planococcus citri*, has been determined. "*Ca.* Tremblaya princeps" presents an unusual nested endosymbiosis and harbors a gammaproteobacterial symbiont within its cytoplasm in all analyzed mealybugs. The genome sequence reveals that "*Ca.* Tremblaya princeps" cannot be considered an independent organism but that the consortium with its gammaproteobacterial symbiotic associate represents a new composite living being.

Obligate intracellular bacteria are characterized by their genome reduction due to the loss of genes rendered unnecessary in their protected environments. However, they are still able to fulfill the essential cellular functions that allow them to replicate and synthesize their own proteins (7). Mealybugs contain a primary (P) endosymbiont, "*Candidatus* Tremblaya princeps," which harbors a secondary (S) symbiont in its cytoplasm; this is the first described case of a nested bacterial symbiosis (10, 12).

We report the complete genome of "*Ca.* Tremblaya princeps" from the citrus mealybug *Planococcus citri* (Risso, 1813) isolated from cultivated plants in Valencia, Spain. Bacterium-enriched genomic DNA was extracted from adult females by visceral dissection. This genomic DNA was sequenced through a whole-genome shotgun strategy using Roche 454 GS-FLX Titanium pyrosequencing. A total of 19,693 quality reads were assembled and analyzed using the MIRA (http://sourceforge.net/projects/mira-assembler/) and Staden Package software programs (8), providing 41.4× coverage. Screening for RNA genes was performed with ARGORN (4), tRNAscan (5), and Rfam (3), and prediction and annotation of protein-coding genes were performed through BASys (Bacterial Annotation System) (11) and BLAST searches (1). The complete genome is 138,931 kb in length and has a 58.8% G+C content. The genome contains only 110 functional protein-coding genes, 43 of which correspond to ribosomal proteins, and 23 pseudogenes. It contains two 16S-23S rRNA operons, a common feature of other mealybug P-endosymbionts (2). However, only 8 functional tRNA genes are present, no functional aminoacyl-tRNA synthetases have been identified, and the synthesis of some ribosomal proteins and the modification, maturation, and function of ribosomes must depend on the S-symbiont.

Additionally, genes coding for protein translocation machinery, synthesis of nucleotides and cofactors, energy production, transport, and cell envelope biogenesis are absolutely absent, while only part of the replication machinery is preserved. In contrast, "*Ca.* Tremblaya princeps" has retained the ability to synthesize most of the amino acids that are essential for its phloem-feeding insect host. Thus, the genome retains the genes for the synthesis of amino acids of the branched family (valine, leucine, and isoleucine), histidine, threonine, and phenylalanine, and the final step in the synthesis of methionine from homocysteine that, as in other endosymbionts previously analyzed, must be provided by the host (13).

This genome represents another example of the difficulties of maintaining the distinction between autonomous bacterial life and organelles (9). Considering life the emerging property of a system that simultaneously displays homeostasis, self-reproduction, and evolution (6), "*Ca.* Tremblaya princeps" cannot be considered a living organism, although it is not like any organelle previously described. However, the composite organism made up of "*Ca.* Tremblaya princeps" and its so-called S symbiont (genome sequencing is in progress) might resemble a eukaryotic cell, with a compartmentalized living scheme. Similarly to mitochondria, "*Ca.* Tremblaya princeps" retains only some of the genes necessary for symbiotic functions (synthesis of essential amino acids) and vital functions but needs the recruitment of the gene products encoded by the S-symbiont to be exported to the cytoplasm of "*Ca.* Tremblaya princeps," where most translation must take place.

Nucleotide sequence accession number. The complete genome sequence of "*Ca.* Tremblaya princeps" PCVAL was deposited in GenBank under accession number CP002918.

* Corresponding author. Mailing address: Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, Apartado Postal 22085, 46071 València, Spain. Phone: 34 963543824. Fax: 34 963543670. E-mail: rosario.gil@uv.es.

REFERENCES

1. **Altschul, S. F., et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Baumann, L., M. L. Thao, J. M. Hess, M. W. Johnson, and P. Baumann.** 2002. The genetic properties of the primary endosymbionts of mealybugs differ from those of other endosymbionts of plant sap-sucking insects. Appl. Environ. Microbiol. **68:**3198–3205.
3. **Gardner, P. P., et al.** 2009. Rfam: updates to the RNA families database. Nucleic Acids Res. **37:**D136–D140.
4. **Laslett, D., and B. Canback.** 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. **32:** 11–16.
5. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
6. **Luisi, P. L.** 2002. Toward the engineering of minimal living cells. Anat. Rec. **268:**208–214.
7. **Moya, A., J. Pereto, R. Gil, and A. Latorre.** 2008. Learning how to live together: genomic insights into prokaryote-animal symbioses. Nat. Rev. Genet. **9:**218–229.
8. **Staden, R., K. F. Beal, and J. K. Bonfield.** 2000. The Staden package, 1998. Methods Mol. Biol. **132:**115–130.
9. **Tamames, J., et al.** 2007. The frontier between cell and organelle: genome analysis of *Candidatus* Carsonella ruddii. BMC Evol. Biol. **7:**181.
10. **Thao, M. L., P. J. Gullan, and P. Baumann.** 2002. Secondary (gamma-Proteobacteria) endosymbionts infect the primary (beta-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts. Appl. Environ. Microbiol. **68:**3190–3197.
11. **Van Domselaar, G. H., et al.** 2005. BASys: a web server for automated bacterial genome annotation. Nucleic Acids Res. **33:**W455–W459.
12. **von Dohlen, C. D., S. Kohler, S. T. Alsop, and W. R. McManus.** 2001. Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. Nature **412:**433–436.
13. **Wilson, A. C., et al.** 2010. Genomic insight into the amino acid relations of the pea aphid, *Acyrthosiphon pisum*, with its symbiotic bacterium *Buchnera aphidicola*. Insect Mol. Biol. **19**(Suppl 2)**:**249–258.