



Next generation sequencing elucidates cacao badnavirus diversity and reveals the existence of more than ten viral species



E. Muller^{a,b,*}, S. Ravel^{a,b}, C. Agret^{c,d}, F. Abrokwah^e, H. Dzahini-Obiatey^f, I. Galyuon^g, K. Kouakou^h, E.C. Jeyaseelanⁱ, J. Allainguillaume^j, A. Wetten^{k,1}

^a CIRAD, UMR BGPI, 34398, Montpellier, France

^b BGPI, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France

^c CIRAD, UMR AGAP, 34398, Montpellier, France

^d AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France

^e Department of Biochemistry, University of Cape Coast, Ghana

^f Cocoa Research Institute of Ghana, P. O. Box 8, Akim Tafo, Ghana

^g Department of Molecular Biology and Biotechnology, University of Cape Coast, Ghana

^h World Cocoa Foundation, Abidjan, Cote d'Ivoire

ⁱ University of Jaffna, Faculty of Science, Department of Botany, Jaffna, Sri Lanka

^j University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol, BS16 1QY, United Kingdom

^k School of Agriculture, Policy and Development, University of Reading, Whiteknights, Reading, RG6 7BE, United Kingdom

ARTICLE INFO

Keywords:

Cacao swollen shoot virus
Complete genomes
Cacao
Phylogeny
Illumina sequencing

ABSTRACT

Cacao swollen shoot virus is a member of the family *Caulimoviridae*, genus *Badnavirus* and is naturally transmitted to *Theobroma cacao* (L.) by several mealybug species. CSSV populations in West African countries are highly variable and genetically structured into several different groups based on the diversity in the first part of ORF3 which encodes the movement protein. To unravel the extent of isolate diversity and address the problems of low titer and mixed viral sequences in samples, we used Illumina MiSeq and HiSeq technology. We were able to reconstruct *de novo* 20 new complete genomes from cacao samples collected in the Cocoa Research Institute of Ghana (CRIG) Museum and from the field samples collected in Côte d'Ivoire or Ghana. Based on the 20% threshold of nucleotide divergence in the reverse transcriptase/ribonuclease H (RT/RNase H) region which denotes species demarcation, we conclude there exist seven new species associated with the cacao swollen shoot disease. These new species along with the three already described leads to ten, the total number of the complex of viral species associated with the disease. A sample from Sri Lanka exhibiting similar leaf symptomatology to West African CSSD-affected plants was also included in the study and the corresponding sequence represents the genome of a new virus named cacao bacilliform SriLanka virus (CBSLV).

1. Introduction

Cacao swollen shoot disease (CSSD) which results from cacao swollen shoot virus (CSSV) infection is now regarded as the major viral disease affecting cacao and has been recognized as one of the most important diseases in West Africa limiting cacao production. CSSD was first described in Ghana at Effiduase in the New Juabeng district of the Eastern region in 1936 (Steven) although the disease was probably present in the nearby Nankese township of Ghana from 1922 (Paine,

1945). The disease subsequently appeared in all major cacao growing areas in West Africa with CSSD reported in Côte d'Ivoire in 1943 (Burle, 1961; Mangenot et al., 1946), in Nigeria in 1944 (Thresh, 1959), in Togo in 1949 (Partiot et al., 1978) and in Sierra Leone in 1958 (Attafuah et al., 1963). In addition, West African Amelonado cacao, planted uniformly throughout West Africa, appeared to be highly susceptible and sensitive to CSSV and has favored the rapid spread of the disease. CSSD has always been described as a disease endemic to West Africa, as it has never been reported in South America, the cacao tree's

Abbreviations: CSSV, cacao swollen shoot virus; CSSCDV, cacao swollen shoot CD virus; CSSTAV, cacao swollen shoot Togo A virus; CSSD, cacao swollen shoot disease; PCR, polymerase chain reaction; RT, reverse transcriptase; RNase H, ribonuclease H; ORF, open reading frame; ³H-methionyl transfer RNA; CRIG, Cocoa Research Institute of Ghana; ICTV, International Committee on Taxonomy of Viruses; CBSLV, cacao bacilliform SriLanka virus; NGS, next generation sequencing

* Corresponding author at: CIRAD, UMR BGPI, TA A – 54/K, Campus international de Baillarguet, 34398, Montpellier Cedex 5, France.

E-mail address: emmanuelle.muller@cirad.fr (E. Muller).

¹ Present address: University of the West of England, Frenchay Campus, Clodharbour lane, Bristol, BS16 1QY, United Kingdom.

<https://doi.org/10.1016/j.virusres.2017.11.019>

Received 7 September 2017; Received in revised form 15 November 2017; Accepted 18 November 2017

Available online 21 November 2017

0168-1702/ © 2017 Elsevier B.V. All rights reserved.