

SHORT REPORT

Open Access

Evolution of *African cassava mosaic virus* by recombination between bipartite and monopartite begomoviruses

Fidèle Tiendrébéogo^{1,2,3}, Pierre Lefeuvre², Murielle Hoareau², Mireille A Harimalala^{2,4,5}, Alexandre De Bruyn², Julie Villemot², Valentin SE Traoré¹, Gnissa Konaté¹, Alfred S Traoré³, Nicolas Barro³, Bernard Reynaud², Oumar Traoré¹ and Jean-Michel Lett^{2*}

Abstract

Background: Cassava mosaic disease (CMD) is a major constraint on cassava cultivation in Africa. The disease is endemic and is caused by seven distinct cassava mosaic geminiviruses (CMGs), some of them including several variants.

Findings: From cassava leaf samples presenting CMD symptoms collected in Burkina Faso, four DNA-A begomovirus components were cloned and sequenced, showing 99.9% nucleotide identity among them. These isolates are most closely related to *African cassava mosaic virus* (ACMV) but share less than 89% nucleotide identity (taxonomic threshold) with any previously described begomovirus. A DNA-B genomic component, sharing 93% nucleotide identity with DNA-B of ACMV, was also characterized. Since all genomic components have a typical genome organization of Old World bipartite begomoviruses, this new species was provisionally named African cassava mosaic Burkina Faso virus (ACMBFV). Recombination analysis of the new virus demonstrated an interspecies recombinant origin, with major parents related to West African isolates of ACMV, and minor parents related to *Tomato leaf curl Cameroon virus* and *Cotton leaf curl Gezira virus*.

Conclusion: This is the first report of an ACMV-like recombinant begomovirus arisen by interspecific recombination between bipartite and monopartite African begomoviruses.

Keywords: Cassava mosaic disease, *African cassava mosaic virus*, Begomovirus, Evolution, Recombination

Findings

Cassava (*Manihot esculenta* Crantz), the primary food crop in sub-Saharan Africa, is severely affected by cassava mosaic disease (CMD) with losses estimated to up to 19 millions tons (1.9-2.7 billions dollars) annually [1]. CMD is caused by distinct cassava mosaic geminiviruses (CMGs) (family *Geminiviridae*: Genus *Begomovirus*) which are transmitted by the whitefly *Bemisia tabaci* (Genn.) [1]. Whereas begomoviruses are either bipartite with two genomic components (DNA-A and DNA-B) or monopartite with only a DNA-A-like component, all described CMGs are bipartite. Their two circular single

stranded DNA molecules (~2.8 kb each) encode a total of eight viral proteins responsible, among other functions, for replication, transcription enhancement, encapsidation and viral movement (see [2] for a review).

CMD is endemic in Africa and seven distinct species are found associated with the disease: *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), *East African cassava mosaic Cameroon virus* (EACMCV), *East African cassava mosaic Kenya virus* (EACMKV), *East African cassava mosaic Malawi virus* (EACMMV), *East African cassava mosaic Zanzibar virus* (EACMZV) and *South African cassava mosaic virus* (SACMV) [3]. Several strains of the EACMV species are also described throughout Africa [2], many of which resulting from inter- or intra-species recombination [2]. Indeed, it was established that begomoviruses

* Correspondence: jean-michel.lett@cirad.fr

²CIRAD, UMR PVBMT, Pôle de Protection des Plantes, Saint-Pierre, Ile de La Réunion 97410, France

Full list of author information is available at the end of the article

are subject to frequent recombination which provide the viruses with a large sequence space and may promote the adaptation to environmental changes and new ecological niches [4-7]. In contrast to other CMGs, including the EACMV-like viruses (EACMV, EACMCV, EACMKV, EACMMV, EACMZV), SACMV, *Sri Lanka cassava mosaic virus* (SLCMV) and *Indian cassava mosaic virus* (ICMV), the genomes of ACMV isolates present little genetic variation and do not display evidence of recombination, suggesting that ACMV has evolved as a distinct lineage (reviewed by [2]).

In West Africa, cassava is frequently cultivated in association with several other crops such as vegetables (tomato, pepper, okra), tubers (cassava, sweet potato, yam), and cereals. In August/September 2008, leaves from cassava plants presenting CMD symptoms were collected from such a field around Ouagadougou (12° 23'31.42"N; 1°30'19.40"W; Burkina Faso). Total DNA was extracted using DNeasy[®] Plant Minikit (Qiagen) and tested for the presence of begomoviruses using polymerase chain reaction (PCR) with degenerate primers representative of either DNA-A [8] or DNA-B [9].

Ten samples showing mosaic symptoms were found to be positive for the presence of begomovirus DNA using PCR. The primary results showed the presence of ACMV and EACMV-UG in seven and one samples respectively [10]. From two of those samples (BF127 and BF128), full-length viral genomes were successfully amplified by rolling-circle amplification (RCA) [11]. The amplified DNAs were digested with endonucleases *Bam*HI, *Pst*I, *Apa*I or *Nco*I and the DNA fragments of the expected size (~2.8 kb) were cloned into pGEM[®] vectors (Promega Biotech). Cloned genome components were sequenced by Macrogen Inc. (South Korea) by primer walking and genomic components were assembled using the DNAMAN software (version 5.2.9; Lynnon, Quebec, Canada). Four full-length DNA-A nucleotide sequences, one from sample BF127 (BF:Oua:BF127:08) and three from sample BF128 (BF:Oua:BF128A:08; BF:Oua:BF128B:08; BF:Oua:BF128C:08), as well as one full-length DNA-B nucleotide sequence from sample BF127 (BF:BF127:08), were obtained (Table 1). The sequences

were deposited in the EMBL database under the accession numbers HE616777 to HE616781.

The cloned genomes showed an organization typical of bipartite begomoviruses from the Old World, *i.e.* DNA-A components (2768 to 2770 nt each) with two ORFs on the viral strand (AV2/MP and AV1/CP), and four on the complementary-sense strand (AC1/Rep, AC2/TrAP, AC3/REn and AC4) and a DNA-B component (2737 nt) with one ORF on the viral-sense (BV1/MP), and one on the complementary-sense (BC1/NSP). Sequences were aligned with representative isolates of begomoviruses using the Muscle [12] and ClustalW [13] alignment tools implemented in the MEGA 5 software [14]. Pairwise sequence comparisons performed with MEGA 5 (with pairwise deletion of gaps) showed that the four genomic DNA-A components are genetically related with 99.9% identity between one another. The four DNA-A sequences displayed highest nucleotide identity (87%) with ACMV-[Côte d'Ivoire:1999] (Accession n° AF259894) and ACMV-[Nigeria:1990] (X17095). The DNA-B sequence displayed 93% nucleotide identity with ACMV-[Nigeria:Ogo:2002] (AJ427911). Based on the species demarcation criterion of the genus *Begomovirus* (< 89% nucleotide identity for DNA-A) [15], the begomovirus molecularly characterised in this work could be considered as a novel species, provisionally named African cassava mosaic Burkina Faso virus (ACMBFV).

The putative Rep protein binding site of the four ACMBFV isolates included a single core binding site (GGGGT) with a potential inverted repeat (GGACC), and the corresponding iteron-related domain (IRD) was identified to be MAPPRKFRIN. These features are different of those from previously described ACMV isolates (Iteron: GGAGA and IRD: MRTPRFRIQ or MRTPRFRVQ), but are similar to those found in isolates of *Tomato leaf curl Cameroon virus* (ToLCCMV) and *Cotton leaf curl Gezira virus* (CLCuGV), two monopartite begomoviruses [16,17]. Interestingly, the presence of the same iteron (GGGGT) in the intergenic region of the ACMBFV DNA-B is consistent with its trans-replication by the Rep protein encoded by the ACMBFV DNA-A component.

Table 1 Name, acronym and accession numbers of begomovirus isolates reported in this study

Name	Acronym	DNA molecule, size (nt) and accession numbers	
		DNA-A	DNA-B
African cassava mosaic Burkina Faso virus-[Burkina Faso:Ouagadougou:BF127:2008]	ACMBFV-[BF:Oua:BF127:08]	2770 (HE616777)	2737 (HE616778)
African cassava mosaic Burkina Faso virus-[Burkina Faso:Ouagadougou:BF128A:2008]	ACMBFV-[BF:Oua:BF128A:08]	2768 (HE616779)	
African cassava mosaic Burkina Faso virus-[Burkina Faso:Ouagadougou:BF128B:2008]	ACMBFV-[BF:Oua:BF128B:08]	2768 (HE616780)	
African cassava mosaic Burkina Faso virus-[Burkina Faso:Ouagadougou:BF128C:2008]	ACMBFV-[BF:Oua:BF128C:08]	2768 (HE616781)	

As these results suggested that the ACMBFV isolates could have a recombinant origin, a recombination analysis was performed on a set of 234 sequences, with 222 sequences representing the whole Old World begomovirus diversity, ten *Curtovirus* and two *Topocuvirus* sequences. The data set was retrieved from public sequence databases using taxonomy browser (<http://www.ncbi.nlm.nih.gov>) on September 2011 and aligned as described above with our sequences. Detection of potential recombinant sequences, identification of likely parental sequences, and localization of possible recombination breakpoints was carried out using RDP [18], GENECONV [19], BOOTSCAN [20], MAXIMUM CHI SQUARE [21], CHIMAERA [20], SISTER SCAN [22] and 3Seq [23] recombination detection methods as implemented in RDP3 [24]. The analysis was performed so that only triplets of sequences involving at least two CMG sequences were considered. Default settings for the different detection methods and a Bonferroni corrected *P*-value cut-off of 0.05 were used. Only events detected with 3 methods or more were accepted.

Three distinct recombination events (a, b and c; Figure 1) were detected within the genomic DNA-A of ACMBFV, whereas no recombination was detected within the DNA-B. Recombination event (a) involves a major parent related to the Ivory Coast isolate of the bipartite ACMV (ACMV-[Côte d'Ivoire:1999]), and a minor parent related to the Cameroon isolate of the monopartite ToLCCMV (ToLCCMV-[Cameroon:Buea:Okra:2008]). Event (b) detected in the intergenic region occurred between ACMV-[Nigeria] as a major parent and CLCuGV-[Cameroon:2008], another monopartite begomovirus described in Cameroon on okra as a minor

parent. The event (c) was detected into the event (a) between the same major parent and an unknown minor parent (undescribed or already extinct virus).

Maximum likelihood (ML) phylogenetic trees (with 100 bootstrap replicates) were constructed from full DNA components as well as alignments of both parts of the recombination event (a) using PHYML v3.0. [25] and the model of sequence evolution selection procedure implemented in RDP3 [24]. Trees were visualized using FigTree v1.3.1 software (available at <http://tree.bio.ed.ac.uk/software/figtree/>). Consistently, phylogenetic analyses performed with the non-recombinant portion of the genome (nt 27-1727 relative to ACMBFV-[BF:Oua:BF127:08]; covering AV2, AV1, AC2 and AC3) placed ACMBFV in the cluster of ACMV (Figure 2A). With the recombinant portion (nt 1728-26 relative to ACMBFV-[BF:Oua:BF127:08]; covering 90% of 3' portion of Rep, entire AC4 and IR) the same analysis placed ACMBFV in a cluster with ToLCCMV-[CM:Bue:Ok:08] (Figure 2B). Taken together, these results confirm the recombinant nature of the ACMBFV, with the majority of the genome (61.5%) derived from a major ACMV parent, while the rest of the genome (38.5%) is related to ToLCCMV-[CM:Bue:Ok:08] and CLCuGV-[CM:08] minor parents. As expected, phylogenetic reconstruction performed with the DNA-A sequences, using GTR + I + G4 as the best fit model of evolution, placed ACMBFV outside the cluster of ACMV isolates, due to its recombinant nature (Figure 3). A similar analysis performed with the DNA-B sequence of ACMBFV under GTR + G4 evolutionary model placed it within the cluster of ACMV isolates (Figure 4). In order to assess its biological features, the infectivity of our

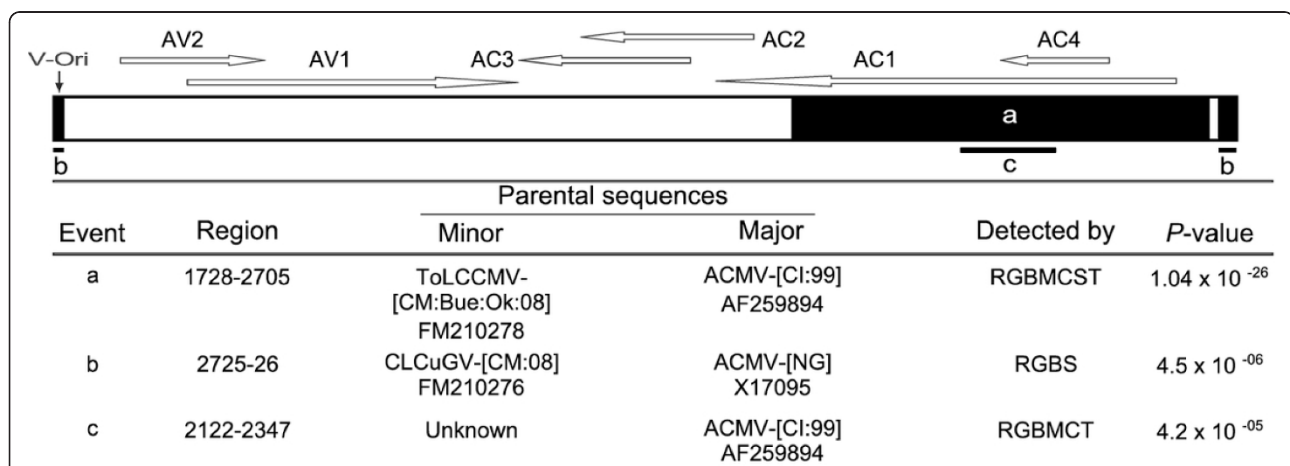


Figure 1 Recombinant regions (a, b and c) detected within the isolates of *African cassava mosaic Burkina Faso virus* using RDP3. The genome at the top of the figure corresponds to the schematic representation of sequence below. Region coordinates are nucleotide positions of detected recombination breakpoints. Wherever, possible parental sequences are identified. "Major" and "Minor" parents are sequences that were used, along with the indicated recombinant sequence to identify recombination. Whereas for each identified event the minor parent is apparently the contributor of the sequence within the indicated region, the major parent is the apparent contributor of the rest of the sequence. Note that the identified "parental sequences" are not the actual parents but are simply those sequences most similar to the actual parents in the analysed dataset. Recombinant regions and parental viruses were identified using the RDP (R), GENECONV (G), BOOTSCAN (B), MAXIMUM CHI SQUARE (M), CHIMAERA (C), SISTER SCAN (S) and 3Seq (T) methods.

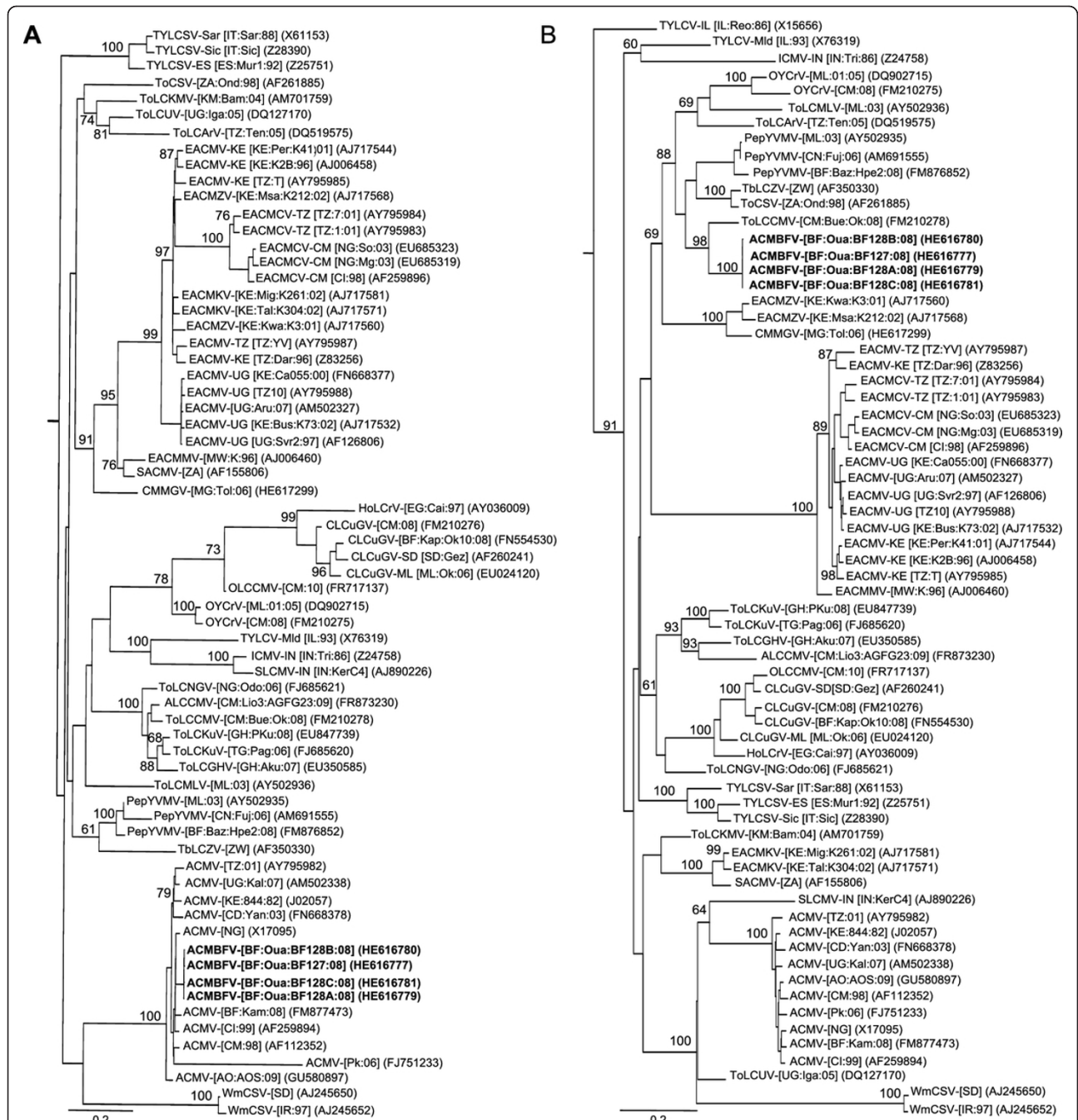


Figure 2 Maximum likelihood trees based on the alignment of the sequences of four African cassava mosaic Burkina Faso virus isolates (in bold; see Table 1 for isolates names and acronyms), plus additional DNA-A and DNA-A-like sequences originated from Africa and the Mediterranean region. In (A) the non-recombinant portion (ORFs: AV2, AV1, AC2 and AC3; see Figure 1) of the genome and in (B) the recombinant portion (covering 90% of 3' portion of Rep, entire AC4 ORF and IR; see Figure 1). Begomovirus acronyms used are described in the Additional file 1: Table S1 according to Fauquet *et al.* [15] isolate descriptors.

recombinant DNA-A and the need of a cognate DNA-B of ACMV type should be investigated in the future. In conclusion, we describe for the first time an ACMV-like recombinant begomovirus arisen by interspecific recombination between bipartite and monopartite

African begomoviruses. Although current data suggest that ACMV has evolved as a distinct lineage, our results prove that the ACMV genotype is also prone to recombination. Our study confirms that West Africa is likely a centre of begomovirus diversity [17], where

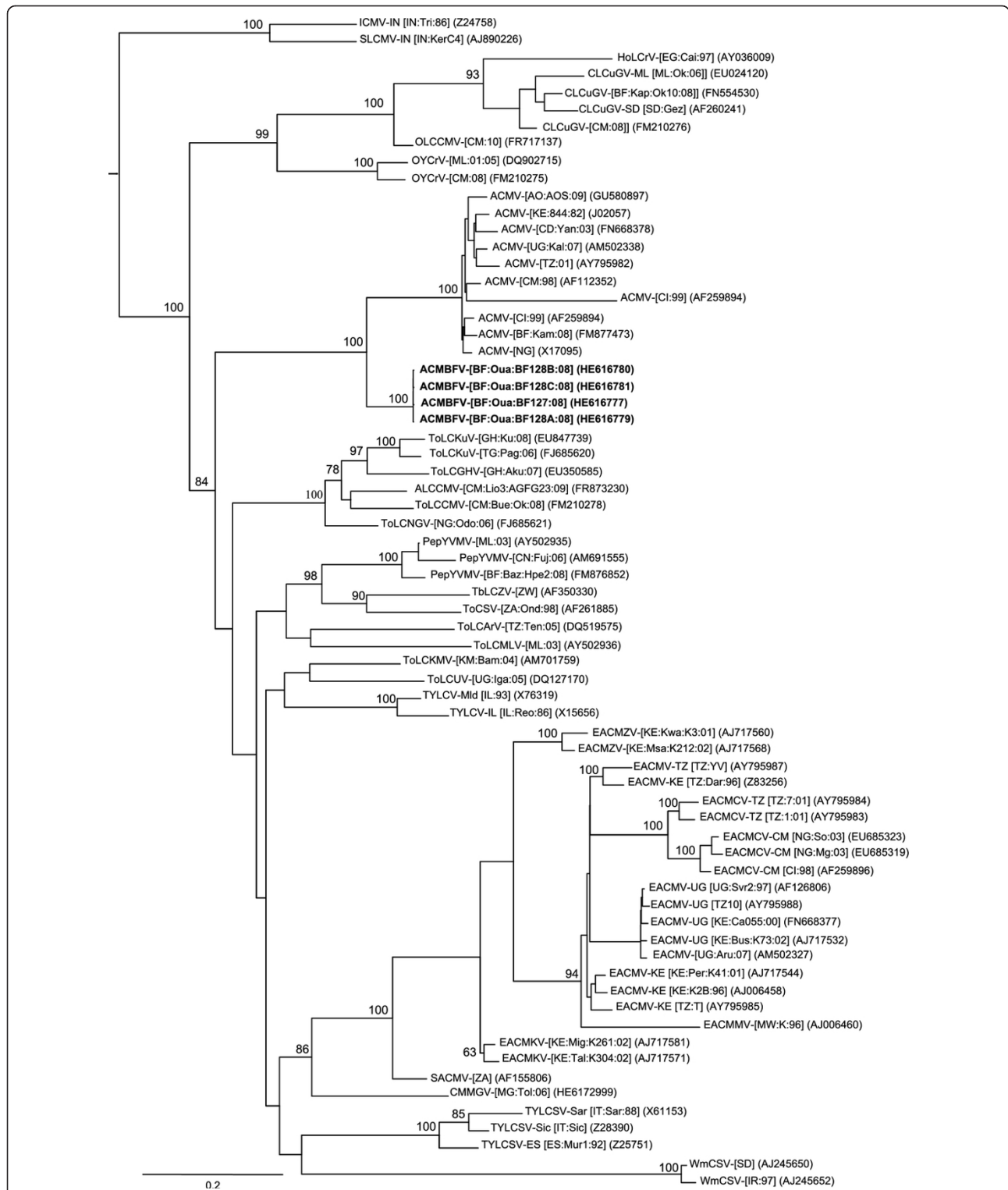
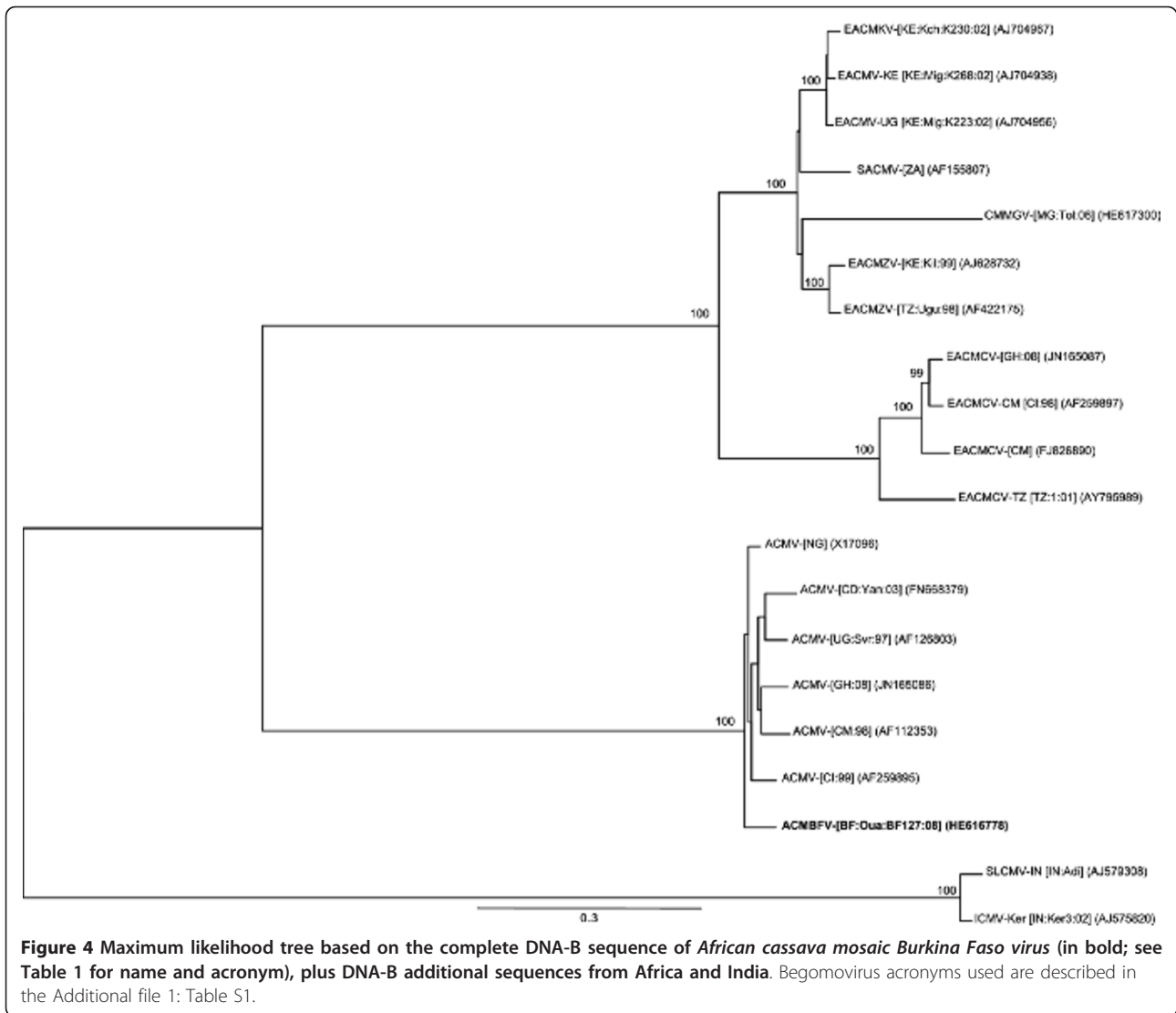


Figure 3 Maximum likelihood tree based on the complete DNA-A sequences of four African cassava mosaic Burkina Faso virus isolates (in bold; see Table 1 for isolates name and acronyms), plus additional DNA-A and DNA-A-like sequences originated from Africa and the Mediterranean region. Begomovirus acronyms used are described in the Additional file 1: Table S1.



recombination seems to play a major role in the evolution of food and vegetable crops infecting begomoviruses.

Additional material

Additional file 1: Table S1. Description of begomovirus acronyms used in the phylogenetic trees.

Acknowledgements

This study was supported by the following institutions: AIRES-Sud: a programme from the French Ministry of Foreign and European Affairs implemented by the *Institut de Recherche pour le Développement* (IRD-DSF), CRSBAN/UFR-SVT (University of Ouagadougou), CIRAD, *Conseil Régional de La Réunion* and European Union (FEDER) and France Embassy in Burkina Faso. We express our gratitude to Mr François GAUDEAU of the *Service de Coopération et d'Action Culturelle* (SCAC) of the France Embassy in Ouagadougou (BF). The authors wish to thank Nathalie Becker for her

comments, critical reviewing, and for revising the English version of the manuscript. All authors read and approved the final manuscript.

Author details

¹Institut de l'Environnement et de Recherches Agricoles (INERA), 01 BP 476, Ouagadougou 01, Burkina Faso. ²CIRAD, UMR PVBMT, Pôle de Protection des Plantes, Saint-Pierre, Ile de La Réunion 97410, France. ³CRSBAN/UFR/SVT, Université de Ouagadougou, 03 BP 7021, Ouagadougou 03, Burkina Faso. ⁴Faculté des Sciences, Université d'Antananarivo, Antananarivo BP 906, Madagascar. ⁵Université de La Réunion, UMR PVBMT, Pôle de Protection des Plantes, Saint-Pierre, Ile de La Réunion 97410, France.

Authors' contributions

FT, VSET and OT collected samples. FT, MH, MAH, ADB and JV cloned and sequenced the viruses; FT, PL and JML analysed the data and prepared the manuscript. JML, OT, NB, GK, AST, VSET and BR secured funding for the project's execution, and provided ideas and comments during manuscript preparation. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 16 December 2011 Accepted: 14 March 2012
Published: 14 March 2012

References

1. Legg JP, Fauquet CM: Cassava mosaic geminiviruses in Africa. *Plant Mol Biol* 2004, **56**:585-599.
2. Patil BL, Fauquet CM: Cassava mosaic geminiviruses: actual knowledge and perspectives. *Mol Plant Pathol* 2009, **10**:685-701.
3. Bull SE, Briddon RW, Sserubombwe WS, Ngugi K, Markham PG, Stanley J: Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. *J Gen Virol* 2006, **87**:3053-3065.
4. Pita JS, Fondong VN, Sangare A, Otim-Nape GW, Ogwal S, Fauquet CM: Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J Gen Virol* 2001, **82**:655-665.
5. Owor EB, Martin DP, Shepherd DN, Edema R, Monjane AL, Rybicki EP, Thomson JA, Varsani A: Genetic analysis of maize streak virus isolates from Uganda reveals widespread distribution of a recombinant variant. *J Gen Virol* 2007, **88**:3154-3165.
6. Varsani A, Shepherd DN, Dent K, Monjane AL, Rybicki EP, Martin DP: A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. *Virology* 2009, **6**:36.
7. Lefevre P, Lett JM, Varsani A, Martin DP: Widely conserved recombination patterns among single-stranded DNA viruses. *J Virol* 2009, **83**:2697-2707.
8. Delatte H, Martin DP, Naze F, Golbach RW, Reynaud B, Peterschmitt M, Lett JM: South West Indian Ocean Islands tomato begomovirus populations represent a new major monopartite begomovirus group. *J Gen Virol* 2005, **86**:1533-1542.
9. Rojas MR, Gilbertson RL, Russel DR, Maxwell DP: Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminivirus. *Plant Dis* 1993, **77**:340-347.
10. Tiendrébéogo F, Lefevre P, Hoareau M, Traoré VSE, Barro N, Reynaud B, Traoré AS, Konaté G, Traoré O, Lett JM: Occurrence of East African cassava mosaic virus-Uganda (EACMV-UG) in Burkina Faso. *Plant Pathol* 2009, **58**:783.
11. Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T: A simple method for cloning the complete begomovirus genome using the bacteriophage ϕ 29 DNA polymerase. *J Virol Methods* 2004, **116**:209-211.
12. Edgar RC: MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004, **32**:1792-1797.
13. Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994, **22**:4673-4680.
14. Tamura K, Peterson D, Stecher G, Nei M, Kumar S: MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011, **28**:2731-2739.
15. Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M, Zhou X: Geminivirus strain demarcation and nomenclature. *Arch Virol* 2008, **153**:783-821.
16. Argüello-Astorga GR, Ruiz-Medrano R: An iteron-related domain is associated to motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. *Arch Virol* 2001, **146**:1465-1485.
17. Kon T, Gilbertson RL: Two genetically related begomoviruses causing tomato leaf curl disease in Togo and Nigeria differ in virulence and host range but do not require a betasatellite for induction of disease symptoms. *Arch Virol*, DOI 10.1007/s00705-011-1139-0.
18. Martin D, Rybicki E: RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 2000, **16**:562-563.
19. Padidam M, Sawyer S, Fauquet CM: Possible emergence of new geminiviruses by frequent recombination. *Virology* 1999, **265**:218-225.
20. Martin DP, Posada D, Grandall KA, Williamson C: A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retroviruses* 2005, **21**:98-102.
21. Maynard SJ: Analyzing the mosaic structure of genes. *J Mol Evol* 1992, **34**:126-129.
22. Gibbs MJ, Armstrong JS, Gibbs AJ: Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 2000, **16**:573-582.
23. Boni MF, Posada D, Feldman MW: An Exact Nonparametric Method for Inferring Mosaic Structure in Sequence Triplets. *Genetics* 2007, **176**:1035-1047.
24. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefevre P: RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 2010, **26**:2462-2463.
25. Guindon S, Lethiec F, Duroux P, Gascuel O: PHYML Online-a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res* 2005, **33**:557-559.

doi:10.1186/1743-422X-9-67

Cite this article as: Tiendrébéogo *et al.*: Evolution of African cassava mosaic virus by recombination between bipartite and monopartite begomoviruses. *Virology Journal* 2012 **9**:67.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

