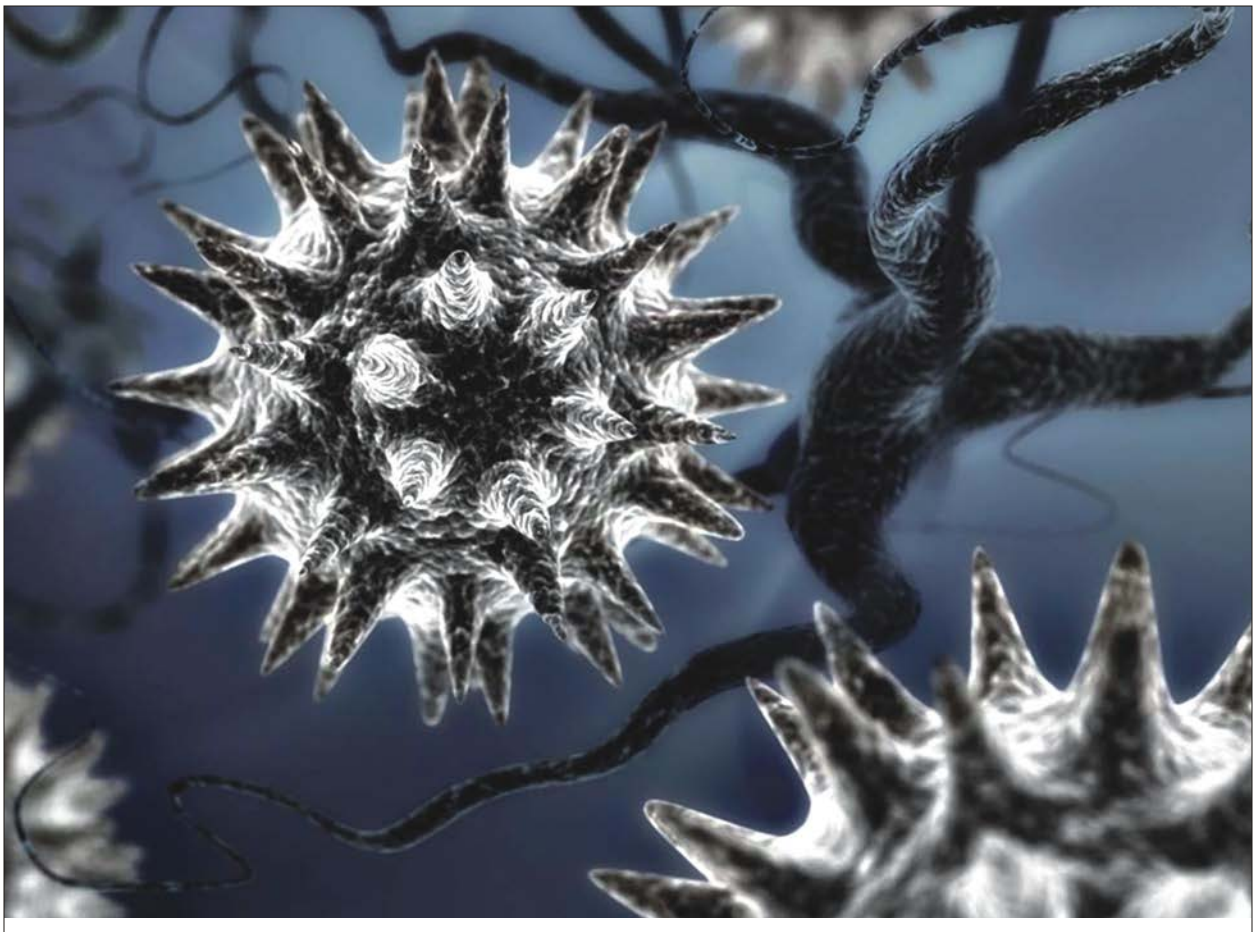

International Journal of Virology

ISSN 1816-4900

Academic Journals Inc.

www.academicjournals.com

OPEN ACCESS International Journal of Virology

ISSN 1816-4900
DOI: 10.3923/ijv.2017.80.89

Research Article
Genetic Diversity and Molecular Epidemiology of *Rice Yellow Mottle Virus* in Central African Republic

[1,2]R.D.S. Longué, [3]A. Galzi-Pinel, [1]I. Zinga, [1]S. Semballa, [3]D. Fargette, [4]N. Barro and [2]O. Traoré

[1]Laboratoire des Sciences Biologiques et Agronomiques Pour le Développement (LASBAD), Université de Bangui, BP 908 Bangui, Central African Republic
[2]Institut de l'Environnement et de Recherches Agricoles (INERA), 01 BP 476 Ouagadougou 01, Burkina Faso
[3]Institut de Recherche Pour le Développement, UMR IPME (IRD-Cirad-Université Montpellier), 911 Avenue Agropolis 34394 Montpellier, France
[4]Université de Ouagadougou, UFR/SVT, Département de Biochimie-Microbiologie, BP 7021 Ouagadougou, Burkina Faso

Abstract

Background: *Rice yellow mottle virus* (RYMV) of the genus *Sobemovirus* is the most important viral disease of rice in Africa, mainly occurring in the lowland and irrigated ecologies. **Materials and Methods:** The coat protein gene of forty six isolates of RYMV collected between 2011 and 2014 in Central African Republic was amplified and sequenced. **Results:** Analysis of sequences revealed that the average nucleotide diversity among isolates was low, 2.0% in nucleotide and 0.9% in amino acid. The ratio of non-synonymous over synonymous nucleotide substitutions per site was 0.07, indicating a virus diversification under a high conservative selective pressure. All isolates shared the amino acids specific of the serotype Ser1, a Val_{115} involved in the response with monoclonal A and a Thr_{115} which accounts for the lack of reaction with monoclonal D. Phylogenetic analyses showed that isolates of the south of Central African Republic belong to two sister monophyletic groups related to the S1ca strain, a strain which gathers all isolates from the East of West Africa to the West of Central Africa. Molecular clock dating of the age of each of the two groups and of their common ancestor suggests that RYMV was introduced a few decades ago in Central African Republic from the West of Central Africa. **Conclusion:** RYMV remained at a low level and undetected in wild hosts and in traditional rice fields until recently. RYMV emergence occurred less than ten years ago, likely favored by the on-going rice intensification and resulted in the current epidemics. This is the first study of the molecular diversity of RYMV in Central African Republic. It indicates that the isolates of the South of the country belonged to the S1 strain.

Key words: *Rice yellow mottle virus*, rice cultivation, molecular epidemiology, coat protein, diversity index

Received: December 13, 2016 Accepted: February 07, 2017 Published: March 15, 2017

Citation: R.D.S. Longué, A. Galzi-Pinel, I. Zinga, S. Semballa, D. Fargette, N. Barro and O. Traoré, 2017. Genetic diversity and molecular epidemiology of *Rice yellow mottle virus* in Central African Republic. Int. J. Virol., 13: 80-89.

Corresponding Author: R.D.S. Longué, INERA, Laboratoire de Virologie et de Biotechnologie Végétale (LVBV), 01 BP 476 Ouagadougou 01, Burkina Faso Tel: +226 77 37 72 99

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

First detected in Kenya[1] in 1966, *Rice yellow mottle virus* (RYMV) of the genus *Sobemovirus* is the most important virus of rice in Africa. Only found in Africa and in Madagascar, RYMV is widespread in most countries where rice is grown[2]. The RYMV was reported in Cameroon and Chad[3] in 2001, countries bordering the Central African Republic at the West and at the North, respectively. In 2012, RYMV was detected in the south of the Central African Republic[4].

The diversity among RYMV isolates in Africa was first assessed using polyclonal and monoclonal antibodies raised against an isolate of Madagascar[5-8]. A set of six Mabs (A, B, C, D, E and G) was used. Five serotypes were delineated through the pattern of detection by this set of monoclonal antibodies and named Ser1 to Ser5. The RYMV diversity was further characterized by sequencing the highly variable 720 nucteotides-long Coat Protein (CP) gene. Comparatively, CP sequencing is more discriminating than immunological typing since several sequence variants belong to the same serotype[9]. Seven major strains-named S1-S6 and Sa were determined. They had contrasting geographical distribution. Four strains were found in West and Central Africa (S1-S3 and Sa) and three strains in East Africa (S4-S6). Molecular epidemiology of RYMV indicated a center of origin in East Africa and a secondary center of diversification in West Africa[10]. Furthermore, genetic differences were found between isolates from Savanna and forest regions of West Africa[11].

In this study, we collected and sequenced the coat protein gene of 46 isolates from the South of the Central African Republic, the part of the country where most of the rice is grown. The genetic diversity of the isolates was assessed, their phylogenetic relationships with representative isolates of the main strains in Africa were determined and the serospecific amino-acids were identified. Clock-like diversification of RYMV allows molecular dating[12]. Accordingly, the Time of the Most Recent Common Ancestor (TMRCA) of the isolates of the Central African Republic was estimated, providing information on virus introduction and disease emergence in the country. Then, a scenario for the epidemiology of RYMV in the Central African Republic is proposed.

MATERIALS AND METHODS

Virus isolates: Symptomatic leaf samples were collected between 2011 and 2014 in large irrigated rice production schemes around 30 km West of Bangui, the capital of the Central African Republic and in lowland subsistence fields in Bangui outskirts. Isolates were propagated by mechanical inoculation on the susceptible cultivar IR64 to increase the viral concentration and to get sufficient material for RNA extraction. Symptomatic leaves were collected 2 weeks after inoculation and used for molecular typing.

Extraction of total RNA from rice leaves: Total RNA was extracted from infected rice leaves using TRIzol reagent (Invitrogen, USA) according to the manufacturer's instruction. The mains steps were as follows: the mortars were first sterilized 2 h at 120°C and stored overnight at -80°C. The frozen leaves were ground to a fine powder and transferred into sterile 2 mL Eppendorf tubes. The powder was mixed with TRIzol reagent (about 1 mL) and the mixture was vortexed. Then, chloroform was added and the tubes were placed on ice for 5 min to separate RNA into cellular proteins and DNA. After centrifugation 12,000 rpm for 15 min at 4°C, the aqueous phase was transferred into sterile 1.5 mL Eppendorf tubes. To precipitate total RNAs, 550 µL of isopropanol were added and left for 30 min at -20°C. After centrifugation at 12,000 rpm for 10 min, total RNAs (pellet) were then washed with 500 µL of alcohol 75%. The resulting total RNAs were air dried, resuspended in 30 µL of sterile water and stored at -20°C.

RT-PCR of the coat protein gene: The first strand of the CP gene cDNA was synthesized using M-MLV reverse transcriptase (Promega corp, USA). A volume of 5 µL of total RNAs was denatured in presence of 10 µM of reverse primer RYMV II or P052 (5'CTCCCCCACCCATCCCGAGAATT3') for 5 min at 70°C. The RT reaction mixture contained 5 x RT buffer, dNTP (10 mM), M-MLV-RT (200 U), RNase inhibitor (40 U) and sterile water (8 µL) to a total volume of 25 µL and followed by incubation at 42°C for 1 h.

The PCR amplification was performed as described earlier[13]. Specific primers P013 5'CAAAGATGGCCAGGAA3' (forward) and P052 5'CTCCCCCACCCATCCCGAGAATT3' (reverse) were used to cover the full length of the CP gene. The thermal conditions were as follows: Initial denaturation at 94°C for 5 min and 30 cycles of denaturation at 94°C for 1 min, hybridization at 55°C for 30 sec, extension at 72°C for 1 min, followed by final extension 10 min at 72°C and then held at 15°C. All RT-PCR reactions were conducted in Applied Biosystems 2720 Thermal Cycler (Life Technologies, France). The amplified fragments were subjected to electrophoresis in 1% (p/v) agarose gels containing ethidium bromide and were visualized using a UV transilluminator in order to check the size of the amplified fragments according to the covering region. Positives results were sent to Beckman Coulter Genomics (France) for sequencing. Forty sequences were obtained.

Sequence analysis: The sequences were edited and aligned using EditSeq and MegAlign of DNASTAR Lasergene software version 7 followed by similarity assessment of pairwise sequences. Multiple nucleotide sequences alignment was made using Clustal W with default parameters[14]. The aligned sequences were checked and corrected manually where necessary. Molecular diversity among the isolates was estimated using the Kimura's two parameter method[15] for nucleotide and Jones-Taylor-Thornton (JTT) model for amino-acid, as implemented in MEGA-6 software[16]. The ratio between non synonymous (π_s) and synonymous (π_a) substitutions per site ($\omega = \pi_a/\pi_s$) was calculated to assess the selection pressure acting on the CP gene. Then, a Z-test was computed to validate the type of selection pressure. A p-value less than 0.05 was considered significant at the 5% level. All tests of natural selection were computed using the Nei-Gojobori method as implemented in MEGA-6 software[16].

Phylogenetic relationships between the Central African Republic isolates (named CF isolates) and between CF isolates and isolates of the main strains in Africa were determined using the Maximum-Likelihood (ML) method implemented in MEGA-6 (Table 1). The Tamura 3-parameter model with rate variation across sites estimated from a gamma distribution (T92+G) was selected as the best fitting model. The significance of the internal branches was evaluated using 1000 bootstrap replications. The phylogenetic tree was visualized and edited using the MEGA-6 software.

Molecular clock dating: The reference dataset used for molecular clock dating consists of 300 sequences that were sampled between 1966 and 2012 in 20 countries across East, West and Central Africa[17] to which were added the CF sequences. The TMRCA of the nodes of interest and their associated 95% highest density probability (HPDs) intervals were estimated within a Bayesian coalescent framework by a Markov Chain Monte Carlo (MCMC) implemented in BEAST v 1.8.2[18]. A relaxed (uncorrelated lognormal) clock model, a HKY2+G4 model and the flexible Bayesian skygrid model as prior were selected. In all cases, the BEAST analyses were run until all relevant parameters converged (ESS>200) with 20% of the MCMC chains discarded as burn-in. Statistical confidence in the parameter estimates was represented by values for 95% HPDs intervals around the marginal posterior parameter means.

RESULTS

Agarose gel visualization of RT-PCR products: Forty isolates were analyzed by RT-PCR using specific primers that amplified a zone covering the CP gene. The amplified fragments were about 1000 bp long, corresponding to the expected size according to the genetic marker (Fig. 1). Then, 40 sequences were obtained.

Molecular variability among Bangui (CAR) isolates: The 40 CP sequences of isolates of Central African Republic together with 6 sequences of the same country previously published[4] were analyzed. The total number of nucleotide sites was 720 nucleotides, encoding 239 amino acid residue. Analyses of the polymorphic sites among sequences revealed 55 and 12 variable sites in nucleotide and amino acid, respectively. Indeed, 75% of the mutations occurred at the 3rd nucleotide position of the codons and none of these mutations result in an amino acid change. Two hundred and thirty nine amino acid residues were mostly hydrophobic amino acids (Table 2). Similarity among the isolates ranged from 94.7-100% in nucleotide whereas the average divergence was 2.0% in nucleotide and 1.1% in amino acid. Divergence among sequences ranged from 0.0-4.5 in nucleotide and from 0.0-2.5% in amino acid. The average number of nucleotide synonymous substitutions per site was low (π_s = 0.06), yet 15 times higher than the number of non-synonymous substitutions per site (π_a = 0.004), i.e., a ω ratio (π_a/π_s) of 0.07 (Table 3). The maximum of non-synonymous and of synonymous divergence between

Table 1: Name, origin, reference and sequence accession No. of the RYMV isolates used in this study

Isolates	Country	Region	Strain	References	Sequences accession No.
CF1a	Central African Republic	Central Africa	S1ca	Longue *et al.*[4]	KF054740
CF2a	Central African Republic	Central Africa	S1ca	Longue *et al.*[4]	KF054741
CF4a	Central African Republic	Central Africa	S1ca	Longue *et al.*[4]	KF054743
CF26a	Central African Republic	Central Africa	S1ca	Longue *et al.*[4]	KF054744
CF31a	Central African Republic	Central Africa	S1ca	Longue *et al.*[4]	KF054745
CF32a	Central African Republic	Central Africa	S1ca	Longue *et al.*[4]	KF054742
CF17	Central African Republic	Central Africa	S1ca	This study	
CF28a	Central African Republic	Central Africa	S1ca	This study	
CF31b	Central African Republic	Central Africa	S1ca	This study	
CF46	Central African Republic	Central Africa	S1ca	This study	

Table 1: Continue

Isolates	Country	Region	Strain	References	Sequences accession No.
CF63	Central African Republic	Central Africa	S1ca	This study	
CF93	Central African Republic	Central Africa	S1ca	This study	
CF170	Central African Republic	Central Africa	S1ca	This study	
CF175	Central African Republic	Central Africa	S1ca	This study	
CF177	Central African Republic	Central Africa	S1ca	This study	
CF181	Central African Republic	Central Africa	S1ca	This study	
CF183	Central African Republic	Central Africa	S1ca	This study	
CF5	Central African Republic	Central Africa	S1ca	This study	
CF13a	Central African Republic	Central Africa	S1ca	This study	
CF14a	Central African Republic	Central Africa	S1ca	This study	
CF22a	Central African Republic	Central Africa	S1ca	This study	
CF16	Central African Republic	Central Africa	S1ca	This study	
CF19	Central African Republic	Central Africa	S1ca	This study	
CF20	Central African Republic	Central Africa	S1ca	This study	
CF22b	Central African Republic	Central Africa	S1ca	This study	
CF26b	Central African Republic	Central Africa	S1ca	This study	
CF28b	Central African Republic	Central Africa	S1ca	This study	
CF30	Central African Republic	Central Africa	S1ca	This study	
CF32b	Central African Republic	Central Africa	S1ca	This study	
CF37	Central African Republic	Central Africa	S1ca	This study	
CF47	Central African Republic	Central Africa	S1ca	This study	
CF53	Central African Republic	Central Africa	S1ca	This study	
CF58	Central African Republic	Central Africa	S1ca	This study	
CF59	Central African Republic	Central Africa	S1ca	This study	
CF61	Central African Republic	Central Africa	S1ca	This study	
CF64	Central African Republic	Central Africa	S1ca	This study	
CF72	Central African Republic	Central Africa	S1ca	This study	
CF73	Central African Republic	Central Africa	S1ca	This study	
CF76	Central African Republic	Central Africa	S1ca	This study	
CF78	Central African Republic	Central Africa	S1ca	This study	
CF79	Central African Republic	Central Africa	S1ca	This study	
CF98	Central African Republic	Central Africa	S1ca	This study	
CF120	Central African Republic	Central Africa	S1ca	This study	
CF122	Central African Republic	Central Africa	S1ca	This study	
CF163	Central African Republic	Central Africa	S1ca	This study	
CF176	Central African Republic	Central Africa	S1ca	This study	
Ca4	Cameroon	Central Africa	S1ca	Traore *et al.*[3]	AJ317951
Tc1	Chad	Central Africa	S1ca	Traore *et al.*[3]	AJ317952
Tc3	Chad	Central Africa	S1ca	Traore *et al.*[3]	AJ317954
Ng4	Niger	West Africa	S1ca	GenBank	AJ885148
Ng18	Niger	West Africa	S1ca	GenBank	FN432841
Ni1	Nigeria	West Africa	S1ca	Pinel *et al.*[13]	AJ279932
Ni2	Nigeria	West Africa	S1ca	Pinel *et al.*[13]	AJ279933
CI1	Côte d'Ivoire	West Africa	S1wa	Pinel *et al.*[13]	AJ279902
BF5	Burkina Faso	West Africa	S1wa	Pinel *et al.*[13]	AJ885090
Ma2	Mali	West Africa	S1wa	Pinel *et al.*[13]	AJ279927
Ma77	Mali	West Africa	Sa	Pinel *et al.*[13]	AJ608210
Ma145	Mali	West Africa	Sa	Pinel *et al.*[13]	AJ885137
BF1	Burkina Faso	West Africa	S2	Pinel *et al.*[13]	AJ279901
CI10	Côte d'Ivoire	West Africa	S2	Pinel *et al.*[13]	AJ279911
Ma4	Mali	West Africa	S2	Pinel *et al.*[13]	AJ279929
CI17	Côte d'Ivoire	West Africa	S3	Pinel *et al.*[13]	AJ279918
SL2	Sierra Leone	West Africa	S3	Pinel *et al.*[13]	AJ279936
SL4	Sierra Leone	West Africa	S3	Fargette *et al.*[10]	AJ608215
Tz1	Tanzania	East Africa	S4	Pinel *et al.*[13]	AJ279938
Mg2	Madagascar	East Africa	S4	Pinel *et al.*[13]	AJ279922
Mg3	Madagascar	East Africa	S4	Pinel *et al.*[13]	AJ279923
Tz2	Tanzania	East Africa	S5	Pinel *et al.*[13]	AJ279939
Tz3	Tanzania	East Africa	S5	Pinel *et al.*[13]	AJ279940
Tz11	Tanzania	East Africa	S6	Abubakar *et al.*[20]	AJ511800
Ke12	Kenya	East Africa	S6	GenBank	FN432839
Tz18	Tanzania	East Africa	S6	Abubakar *et al.*[20]	AJ885159

S1-S6 and Sa indicate the different strains of RYMV, the isolate code represent the country of origin

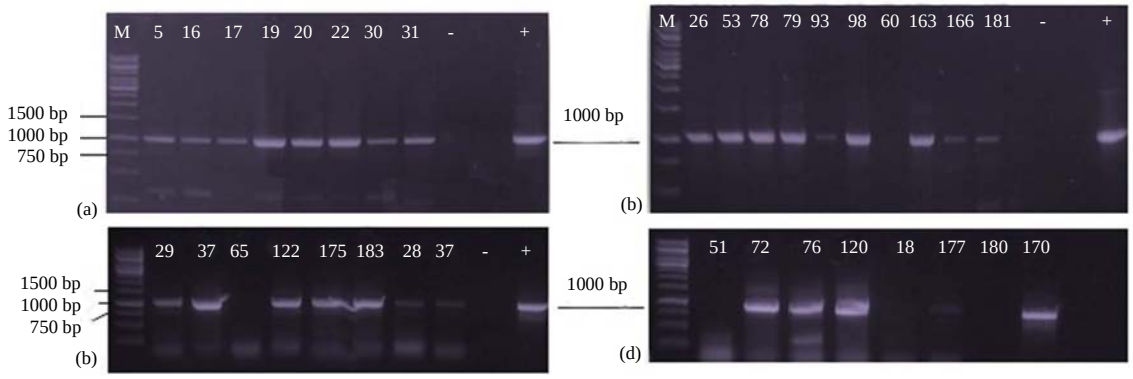

Fig. 1(a-d): Agarose gel (1%) electrophoresis of amplified fragments of RYMV using primers for the coat protein gene. (a-d) Represent different RT-PCR tests. The molecular weight was estimated using a DNA ladder (M, 1 kb)

Table 2: Amino acid composition of the coat protein

Amino acids*																			
Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Total
Number																				
20	5	12	5	5	16	2	7	12	15	7	8	18	6	18	31	16	23	6	7	239

*Amino acids from left to right are as follows: Alanine, cysteine, aspartic acid, glutamine, phenylalanine, glycine, histidine, isoleucine, lysine, leucine, methionine, asparagine, proline, glutamine, arginine, serine, threonine, valine, tryptophan and tyrosine

Table 3: Diversity index, based on the coat protein gene nucleotide sequences, of Group 1 and Group 2 made of the isolates from the Central African Republic and of the combined groups 1 and 2

	Diversity index			
	π	π_a	π_s	ω (π_a/π_s)
Group 1	0.008	0.005	0.013	0.38
Group 2	0.005	0.002	0.013	0.15
Groups 1 and 2	0.02	0.004	0.06	0.07

π: Average of nucleotide diversity (No. of substitutions per site), π_a: No. of non-synonymous substitutions per site, π_s: No. of synonymous substitutions per site

two sequences was 1.2 and 13%, respectively. As $\omega<1$ is indicative of a purifying selection pressure and as the p-value of the Z-test was highly significant ($p<0.001$), the CP gene of the CF isolates diversified under a strong purifying selection. The transition-to-transversion ratio was estimated to 7.3.

Phylogenetic relationships: The phylogenetic reconstruction from the nucleotides sequences based on the maximum likelihood method indicates that all isolates from the Central African Republic clustered into two monophyletic sister groups with high bootstrap support which were subsequently designated group 1 and group 2 (Fig. 2). Group 1 contained 17 isolates whereas group 2 included 29 isolates. In some instances, isolates collected in the same rice field were quite similar, sometimes with identical sequences. However, isolates collected in the same rice field may differ. In the Ngola locality in particular, isolates of both groups were found in the same field. Within each group, there were variants with a high bootstrap support. The diversity within and between group 1 and group 2 was assessed. Nucleotide substitution per site between isolates of the two groups ranged from 2.8-4.5% with an average of 3.5%. However, within each group, the average nucleotide diversity was low: 0.8% in group 1 and 0.5% in group 2, with a maximum divergence between two isolates of 1.6 and 1.5%, respectively.

To assess the phylogenetic relationship between isolates of the Central African Republic and isolates representative of the main strains in Africa[11], phylogenetic tree was reconstructed using the ML method as described earlier (Fig. 3). The CF isolates belong to the strain S1 (bootstrap support 92%) made of the S1wa and the S1ca sister strains. The S1wa strain included isolates from the West of West Africa only. The CF isolates clustered within the S1ca strain (node support 95%) referred to as the West-Central African strain, a strain which gathers all isolates from the east of West Africa (Benin, Togo, Niger and Nigeria) to the West of Central Africa (Chad and Cameroun). Consistently, genetic distance between CF isolates and S1ca isolates was up to 4.4% in nucleotide, lower than 5.9% between CF and S1wa isolates (Table 4). Interestingly, one isolate from Chad (Tc3) was close to group 2 and formed a monophyletic group not including group 1 with a high bootstrap support (78%).

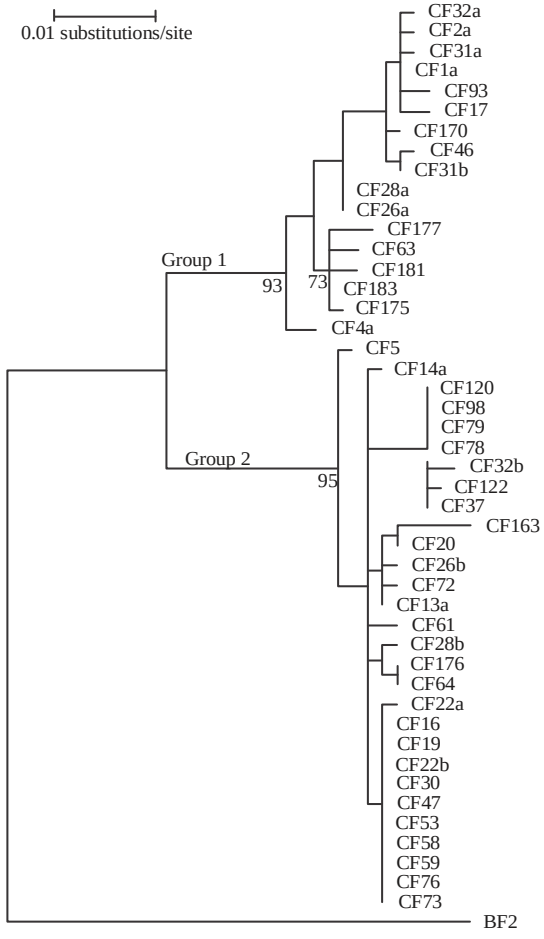

Fig. 2: Phylogenetic analysis of 46 RYMV isolates from the Central African Republic based on the neucleotide sequences of the coat protein gene. The isolate BF2 from Burkina Faso, which belonged to the S1wa strain, was used as an outgroup. The phylogenetic tree was constructed using the maximum likelihood method with the Tamura 3-parameter model implemented in MEGA-6. The bootstrap supports of the nodes (>70%) of the main clusters are indicated as the percentage of 1000 replications. CF is the code name for isolates of the Central African Republic

Comparison of molecular and serological typing: Although the phylogenetic analyses reconstructed from the nucleotide and amino acid sequences displayed an overall similar pattern, there was yet a noticeable difference. Whereas isolates CF63, CF175, CF177, CF181 CF183 and CF4a belonged to group 1 at the nucleotides level, they did not cluster with the other isolates of this group at the aa level (Fig. 4). This difference is due to specific amino acid changes (which distinguish groups) at positions 68, 80 and 217 corresponding to Lys/Thr, Asp/Asn

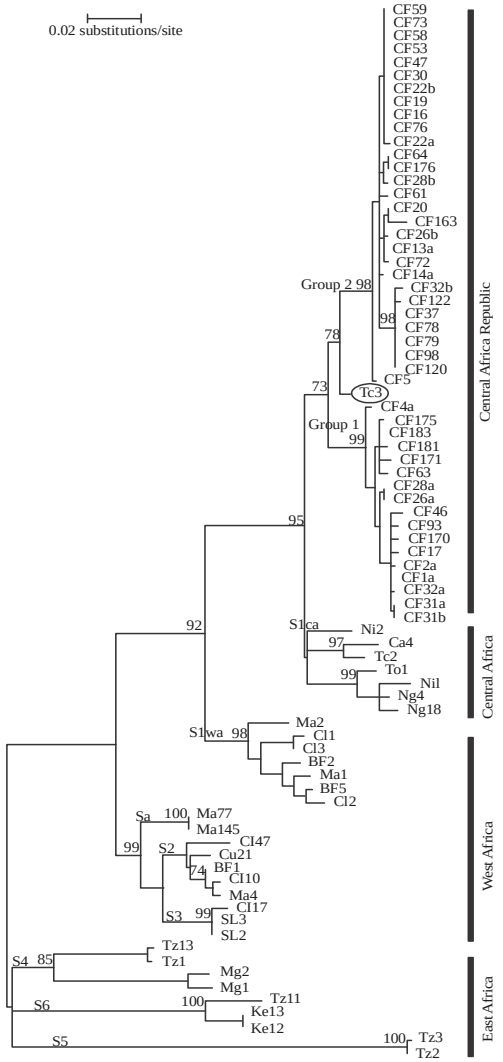

Fig. 3: Phylogenetic relationship between the isolates of the Central African Republic and the main strains of RYMV in Africa. The phylogenetic tree was reconstructed from the coat protein nt sequences by the maximum likelihood method. The bootstrap values for the node of the main strains (>70%) were computed using 1000 random replications. A total of 80 isolates was used in the phylogenetic reconstruction: 3 from Burkina Faso (BF), 1 from Cameroon (Ca), 46 from the Central African Republic (CF), 1 from Guinea (Gu), 6 from Ivory Coast (CI), 2 from Kenya (Ke), 2 from Madagascar (Mg), 5 from Mali (Ma), 2 from Niger (Ng), 2 from Nigeria (Ni), 2 from Sierra Leone (SL), 2 from Chad (Tc), 5 from Tanzania (Tz) and 1 from Togo (To). The framed isolate Tc3 refers to an isolate of Chad which groups with some of the CF isolates. The geographical origins of isolates are shown by vertical bars to the right of the figure. The strains (S1-S6 and Sa) are indicated above their matching branches

Table 4: Amino acid and nucleotide pairwise divergence (%) between group 1 and group 2 of isolates of the Central African Republic and of the main strains of RYMV

	Group 1	Group 2	S1ca	S1wa	Sa	S2	S3	S4	S5	S6
Group 1	-	1.4	1.8	3.1	5.0	4.0	6.4	8.8	8.6	8.9
Group 2	3.5	-	2.1	3.0	5.6	4.2	6.7	9.2	9.2	8.9
S1ca	4.0	4.4	-	3.1	4.5	3.9	6.0	8.2	8.8	8.7
S1wa	5.8	5.9	6.1	-	6.0	4.6	7.1	10.0	10.6	10.9
Sa	8.1	7.9	7.6	6.7	-	2.4	3.2	6.7	8.7	8.1
S2	6.9	8.5	7.6	6.3	3.9	-	3.4	7.6	8.6	8.3
S3	8.7	9.1	8.3	7.8	3.1	3.6	-	7.3	9.8	9.2
S4	10.8	11.5	10.4	10.6	9.2	8.9	9.4	-	8.4	7.8
S5	10.2	11.0	11.4	11.7	12.7	11.8	12.4	12.2	-	8.5
S6	12.4	12.3	11.6	12.5	10.2	9.4	10.0	9.6	12.1	-

The average of nucleotide pairwise divergence is shown above the diagonal and below the diagonal for the amino acid pairwise diversities

Table 5: Time of the Most Recent Common Ancestor (TMRCA) of *Rice yellow mottle virus* in the South of the Central African Republic and of the S1ca strain

Strain	TMRCA	95% HPD	Date	95% HPD
Group 1*	8	5-12	2004	2000-2007
Group 2*	6	4-9	2006	2003-2008
Groups 1 and 2**	18	8-28	1994	1984-2004
S1ca strain**	65	51-81	1947	1931-1961

*All isolates collected in the South of the Central African Republic belong to group 1 or group 2, **Group 1 and 2 belong to the S1ca strain, S1ca strain gathered all isolates of West-Central Africa

and Ala/Thr residue changes, respectively (Fig. 5). These positions also allowed to distinguish CF isolates from Ser1 isolates i.e., when group 1 consensus sequence was compared to that of Ser1 (Fig. 5). In addition, isolates CF28b, CF64 and CF176 from group 2 had an hydrophobic amino acid Leu (TTA) at position 164 instead of a polar residue Ser (TCA). Also, seven other isolates from group 2 had an Arg residue versus a Lys, both basic amino acids, at position 18.

Isolates of the serotype Ser1 are characterized by a detection by Mab A and a lack of recognition by Mab D[9]. Moreover, a Val at position 191 characterizes the reaction with Mab A, whereas a Thr at position 115 accounts for the lack of reaction with Mab D. Sero-specific amino acids of the coat protein of the two groups of CF isolates were identified. Any isolate from the Central African Republic; be it of group 1 or of group 2 had a Val at position 191 and a Thr at position 115 (Fig. 5). This result suggests that, despite their differences in amino acid sequences in the coat protein, isolates of group 1 and 2 shared the serotype Ser1.

Molecular dating: Molecular clock dating provides an estimate of the TMRCA of group 1 and of group 2 and of their common ancestor (Table 5). The diversification of each of the two groups was recent, less than 10 years ago, i.e., after year 2000 (Table 5). The TMRCA was estimated to 8 years (i.e., year 2004) and 6 years (i.e., year 2006) for group 1 and 2, respectively. The overlap of their 95% HPDs intervals

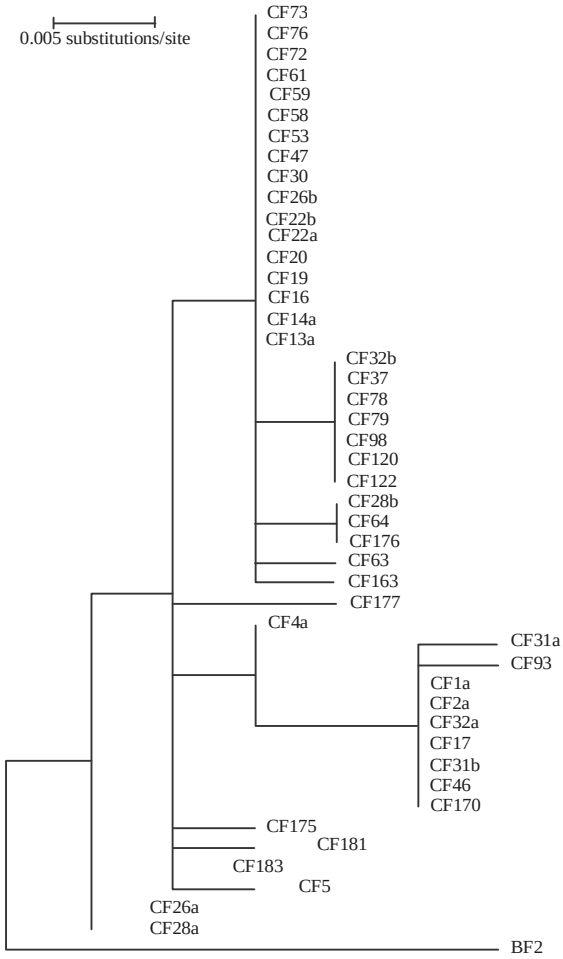

Fig. 4: Phylogenetic tree constructed using aa sequences of the coat protein. The maximum likelihood method was applied. The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site

indicates a non-significant difference between the two TMRCAs. The TMRCA of the common ancestor of the two sister group 1 and 2 was substantially older, 18 years

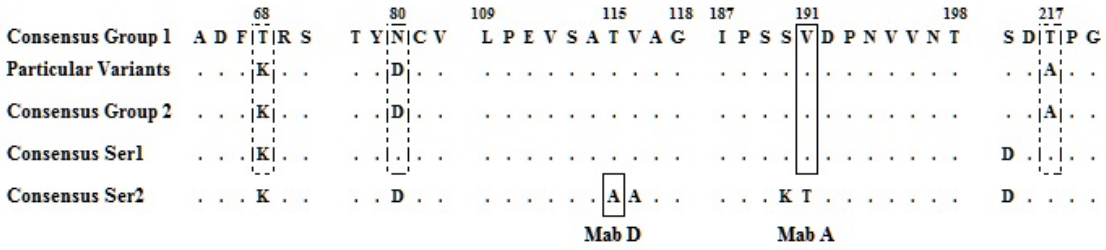

Fig. 5: Amino acid residues of the antigenic regions of the coat protein which reacted with Mab A or Mab D and the main positions which distinguished groups and variants (boxed in dashed lines). The consensus sequence of isolates from the Central African Republic is given in the top line as reference. Dots indicated an amino acid identical to the consensus sequence. The amino acid residues associated to the response with Mabs are boxed in full line (Mab A above at right and Mab D below at left). The antigenic regions are highlighted in black in the top line. Particular variant represents sequence consensus of variants of group 1 (level nucleutide) which are distinguished from other at amino acid level

old (i.e., year 1994). The TMRCA of all isolates from strain S1ca, including the CF isolates was 65 years (i.e., year 1947).

DISCUSSION

Isolates from the Central African Republic were specific as they belong to two phylogenetic groups not found in other countries, even in the neighboring countries Cameroun and Chad. Yet, they are genetically related to isolates of the West-Central strain S1ca as described by Traore *et al.*[3]. This result jointly indicates that RYMV in the Central African Republic belongs to the S1ca strain and that the S1ca strain had a geographical distribution more extended at the East than earlier reported, by at least 500 km[19]. Interestingly, isolates of the S4-S6 strains of East Africa and at the East of Central Africa (Burundi, Rwanda) and at the east of the Democratic Republic of the Congo, a country bordering the Central African Republic at the South were not found in the Central African Republic. Conversely, isolates of the S1ca strain were neither found at the East of Central Africa nor in East Africa. This clear-cut split between the geographical distribution of the S1ca strain and that of the S4-S6 strains suggests that the large forests of the Democratic Republic of the Congo with scarce rice cultivation make an efficient and long-lasting barrier to RYMV propagation across them[12].

The isolates of the Central African Republic belonged to two distinct genetic groups. There was a small genetic distance between the two groups and a low genetic diversity within each group. Similar diversity (π = 1.0%) was found within isolates from Sierra Leone at the far West of Africa and originating from mangrove habitats, according to Abubakar *et al.*[20]. Based on these low variabilities between sequences, we suppose that diversification of the two groups in the South of the Central African Republic is relatively recent. This genetic diversity is also characterized by a low ratio of non-synonymous over synonymous substitution per site, indicative of a strong purifying selective pressure acting on the CP gene and a high transition to transversion ratio, two characteristics shared with isolates of the main strains of RYMV[11]. The results confirm those showed by Pinel *et al.*[13] suggesting a conservative evolution pressure over the CP gene and explains that variation is less at amino acid than at nucleotide levels.

According to N'Guessan *et al.*[8] and Fargette *et al.*[9], amino acid sequences of the coat protein gene suggested that isolates from the Central African Republic shared the Ser1 serotype. This is consistent with serological studies which showed that isolates from the Central African Republic were all recognized by Mab A but were undetected by Mab D, indicative of the Ser1 serotype[4,21].

Molecular clock dating indicates a recent diversification of each group (<10 years). This recent and simultaneous emergence of RYMV in the South of the Central African Republic is consistent with farmer's opinion who reported that disease emergence coincided with the recent rice intensification in the country and occurred not only in traditional rice areas grown since the 1960s but also in new rice cropping areas. Detection of isolates of each group in proximate fields, sometimes even within the same fields, indicates a current and rapid spread of isolates of the two strains within the area and probably beyond. The inclusion of each of the two groups within the S1ca strain and the relationship with isolates from Cameroon and Chad, suggest introduction(s) of the virus, a few decades ago, from these

countries located at the West and North of the Central African Republic, respectively, as suggested earlier[17]. According to the previous studies, the Time of the Most Recent Common Ancestor (TMRCA) of the mains strains of West Africa is estimated around 100 years[22]. Whereas, the results showed the TMRCA of S1ca strain is around 65 years suggesting that RYMV remained at a low level and undetected in wild hosts and in traditional rice fields until recently. The RYMV emergence occurred less than ten years ago, likely favored by the on-going rice intensification and resulted in the current epidemics. Such a scenario is consistent with what has been observed in other countries in Africa, but need to be documented from denser and larger surveys in the Central African Republic and neighboring countries.

CONCLUSION

Rice yellow mottle virus emergence in the Central African Republic is recent and coincided with rice intensification in the country. Particular attention should be paid to the current epidemics in order to prevent major crop loss as what occurred in West Africa in 1990s.

This study contributes to the understanding of the genetic relationships of RYMV through Africa. Since these data are from isolates from of the South of the country, it would be interesting to extend the study to isolates from other parts of country. This could result in a better understanding of the virus spread within the country and would also contribute to the understanding of its evolutionary history across Africa.

SIGNIFICANT STATEMENT

- The RYMV has been extensively studied at the molecular level in many countries of East and West Africa. It was found that the virus had spread from East to West
- Very few studies have been conducted in the Central African region
- Several isolates from the South of the Central African Republic were sequenced. They were phylogenetically related to the isolates of Central Africa and of West/Central Africa
- Analysis of molecular epidemiology revealed two groups of isolates. They shared a common ancestor which had emerged ca 18 years ago. The current intensification has played a major role in the recent emergence of the virus in the Central African Republic

ACKNOWLEDGMENTS

This study was supported financially by the International Foundation for Science (IFS) through fellowship No. C/5359-1 to R D Longué Sokpé and SCAC from the French Embassy in Bangui. We also gratefully acknowledged the laboratoire mixte international (LMI-PathoBios), Burkina Faso, for providing technical assistance.

REFERENCES

1. Bakker, W., 1970. Rice yellow mottle, a mechanically transmissible virus disease of rice in Kenya. Neth. J. Plant Pathol., 76: 53-63.
2. Abo, M.E., A.A. Sy and M.D. Alegbejo, 1998. Rice yellow mottle virus (RYMV) in Africa: Evolution, distribution, economic significance on sustainable rice production and management strategies. J. Sustain. Agric., 11: 51-58.
3. Traore, O., A. Pinel, D. Fargette and G. Konate, 2001. First report and characterization of Rice yellow mottle virus in Central Africa. Plant Dis., 85: 920.1-920.1.
4. Longue, R.D.S., A. Galzi-Pinel, S. Semballa, I. Zinga, D. Fargette and O. Traore, 2014. First report of rice yellow mottle virus in rice in the Central African Republic. Plant Dis., 98: 162-162.
5. Fauquet, C. and J.C. Thouvenel, 1977. Isolation of the rice yellow mottle virus in the Ivory Coast. Plant Dis. Rep., 61: 443-446.
6. Mansour, A.N. and K.W. Baillis, 1994. Serological relationships among rice yellow mottle virus isolates. Ann. Applied Biol., 125: 133-140.
7. Konate, G., O. Traore and M. Coulibaly, 1997. Characterization of rice yellow mottle virus isolates in Sudano-Sahalian areas. Arch. Virol., 142: 1117-1124.
8. N'Guessan, P., A. Pinel, M. Caruana, R. Frutos, A. Sy, A. Ghesquiere and D. Fargette, 2000. Evidence of the presence of two serotypes of rice yellow mottle *Sobemovirus* in Cote d'Ivoire. Eur. J. Plant Pathol., 106: 167-178.
9. Fargette, D., A. Pinel, H. Halimi, C. Brugidou, C. Fauquet and M. van Regenmortel, 2002. Comparison of molecular and immunological typing of the isolates of rice yellow mottle virus. Arch. Virol., 147: 583-596.
10. Fargette, D., A. Pinel, Z. Abubakar, O. Traore and C. Brugidou *et al*., 2004. Inferring the evolutionary history of rice yellow mottle virus from genomic, phylogenetic and phylogeographic studies. J. Virol., 78: 3252-3261.
11. Traore, O., F. Sorho, A. Pinel, Z. Abubakar and O. Banwo *et al*., 2005. Processes of diversification and dispersion of rice yellow mottle virus inferred from large-scale and high-resolution phylogeographical studies. Mol. Ecol., 14: 2097-2110.

12. Pinel-Galzi, A., A. Mpunami, E. Sangu, M. Rakotomalala and O. Traore *et al*., 2009. Recombination, selection and clock-like evolution of rice yellow mottle virus. Virology, 394: 164-172.
13. Pinel, A., P. N'Guessan, M. Bousalem and D. Fargette, 2000. Molecular variability of geographically distinct isolates of rice yellow mottle virus in Africa. Arch. Virol., 145: 1621-1638.
14. Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res., 22: 4673-4680.
15. Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol., 16: 111-120.
16. Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar, 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol., 30: 2725-2729.
17. Trovao, N.S., G. Baele, B. Vrancken, F. Bielejec, M.A. Suchard, D. Fargette and P. Lemey, 2015. Host ecology determines the dispersal patterns of a plant virus. Virus Evol., Vol. 1. 10.1093/ve/vev016
18. Drummond, A.J., S.Y.W. Ho, M.J. Phillips and A. Rambaut, 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol., Vol. 4. 10.1371/journal.pbio.0040088
19. Sorho, F., 2006. Rice yellow mottle disease (RYMD) in Ouest Africa: Phylogeography, pathogenicity and durability of natural resistances. Ph.D. Thesis, University of Cocody, Ivory Cost.
20. Abubakar, Z., F. Ali, A. Pinel, O. Traore and P. N'Guessan *et al*., 2003. Phylogeography of rice yellow mottle virus in Africa. J. Gen. Virol., 84: 733-743.
21. Longue, R.D.S., I. Zinga, S. Semballa, N. Barro and O. Traore, 2016. Detection and serological characterization of rice yellow mottle virus in central African republic. Agric. Sci., 7: 911-919.
22. Pinel-Galzi, A., O. Traore, Y. Sere, E. Hebrard D. Fargette, 2015. The biogeography of viral emergence: Rice yellow mottle virus as a case study. Curr. Opin. Virol., 10: 7-13.