



**HAL**  
open science

## Genetic diversity and epidemic histories of rice yellow mottle virus in Ghana

Emmanuel Gilbert Omiat, Maxwell Darko Asante, Valentin Stanislas Edgar Traoré, Allen Oppong, Beatrice Elohor Ifie, Kirpal Agyemang Ofori, Jamel Aribi, Agnès Pinel-Galzi, Aurore Comte, Denis Fargette, et al.

► **To cite this version:**

Emmanuel Gilbert Omiat, Maxwell Darko Asante, Valentin Stanislas Edgar Traoré, Allen Oppong, Beatrice Elohor Ifie, et al. Genetic diversity and epidemic histories of rice yellow mottle virus in Ghana. *Virus Research*, 2023, 329, pp.199106. 10.1016/j.virusres.2023.199106 . hal-04103331

**HAL Id: hal-04103331**

**<https://hal.science/hal-04103331>**

Submitted on 23 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# Genetic diversity and epidemic histories of rice yellow mottle virus in Ghana

Emmanuel Gilbert Omiat<sup>1,2</sup>, Maxwell Darko Asante<sup>3,4</sup>, Valentin Stanislas Edgar Traoré<sup>5</sup>, Allen Oppong<sup>3</sup>, Beatrice Elohor Ifie<sup>1</sup>, Kirpal Agyemang Ofofu<sup>3</sup>, Jamel Aribi<sup>6</sup>, Agnès Pinel-Galzi<sup>6</sup>, Aurore Comte<sup>6</sup>, Denis Fargette<sup>6</sup>, Eugénie Hébrard<sup>6</sup>, Oumar Traoré<sup>5</sup>, Samuel Kwame Offei<sup>1</sup>, Eric Yirenkyi Danquah<sup>1</sup>, Nils Poulicard<sup>6+</sup>

<sup>1</sup>West Africa Centre for Crop Improvement (WACCI), College of Basic and Applied Sciences, University of Ghana, PMB 30, Legon, Accra, Ghana

<sup>2</sup>National Agricultural Research Organisation (NARO), Buginyanya Zonal Agricultural Research and Development Institute, P. O. Box 1356, Mbale, Uganda

<sup>3</sup>Council for Scientific and Industrial Research – Crops Research Institute (CSIR- CRI), Fumesua, P. O. Box 3785, Kumasi, Ghana

<sup>4</sup>Department of Plant Resources Development, CSIR College of Science and Technology, Fumesua-Kumasi, Ghana

<sup>5</sup>INERA, LNB, Ouagadougou, Burkina Faso

<sup>6</sup>PHIM, Université Montpellier, IRD, INRAE, Cirad, Institut Agro, Montpellier, France

+ corresponding author: [nils.poulicard@ird.fr](mailto:nils.poulicard@ird.fr)

## Abstract

Rice yellow mottle virus (RYMV) has persisted as a major biotic constraint to rice production in Africa. However, no data on RYMV epidemics were available in Ghana, although it is an intensive rice-producing country. Surveys were performed from 2010 to 2020 in eleven rice-growing regions of Ghana. Symptom observations and serological detections confirmed that RYMV is circulating in most of these regions. Coat protein gene and complete genome sequencings revealed that RYMV in Ghana almost exclusively belongs to the strain S2, one of the strains covering the largest area in West Africa. We also detected the presence of the S1ca strain which is being reported for the first time outside its area of origin. These results suggested a complex

30 epidemiological history of RYMV in Ghana and a recent expansion of S1ca to West Africa.  
31 Phylogeographic analyses reconstructed at least five independent RYMV introductions in Ghana  
32 for the last 40 years, probably due to rice cultivation intensification in West Africa leading to a  
33 better circulation of RYMV. In addition to identifying some routes of RYMV dispersion in Ghana,  
34 this study contributes to the epidemiological surveillance of RYMV and helps to design disease  
35 management strategies, especially through breeding for rice disease resistance.

36

37 **Keywords:** *Solemoviridae*; Phylogeography; Spatio-temporal dynamics; Epidemiological  
38 surveillance

39

#### 40 **1. Introduction**

41 Rice yellow mottle virus (RYMV) was first identified in East Africa in 1966 (Bakker, 1974) and  
42 subsequently in 1975 in West Africa (Fauquet and Thouvenel, 1977). Since then, this virus has  
43 progressively been reported in almost all rice-producing countries of sub-Saharan Africa and  
44 Madagascar (Abo et al., 1997; Pinel-Galzi et al., 2015). Endemic to Africa, rice yellow mottle virus  
45 (RYMV) has become one of the most important biotic constraints to rice production in Africa  
46 (Kouassi et al., 2005; Savary et al., 2019; Séré et al., 2013), where it causes high yield losses in  
47 most rice-growing agroecosystems (Agnoun et al., 2019; Suvi et al., 2018).

48 RYMV belongs to the *Sobemovirus* genus of the *Solemoviridae* family. It has icosahedral particles  
49 and a single-stranded positive-sense RNA genome (Hébrard et al., 2020). The RYMV has a narrow  
50 host range limited predominantly to rice species and a few wild *Poacea* (Allarangaye et al., 2007;  
51 Bakker, 1974; Traoré et al., 2009). In infected rice plants, RYMV infection shows symptoms that  
52 include mottling and yellowing of the leaves, stunting, reduced tillering, partial panicle exertion,  
53 sterility and, in severe cases, death of the plant (Attere, 1983; Bakker, 1974). RYMV is mainly  
54 transmitted at short-distance by chrysomelid beetles and by contact between plants during  
55 cultural practices (Traoré et al., 2009). No evidence of seed transmission has been found so far  
56 in rice species (Allarangaye et al., 2006; Bakker, 1974; Fauquet and Thouvenel, 1977; Konate et  
57 al., 2001)

58 Based on RYMV genetic data collected from 24 African countries since the 1970's, the genetic  
59 diversity of this virus has been well described (Pinel-Galzi et al., 2015), and its evolutionary history  
60 and its global dispersion across Africa and Madagascar has started to be deciphered (Issaka et al.,  
61 2021; Rakotomalala et al., 2019; Trovão et al., 2015). All these genetic data revealed a spatially  
62 structured genetic diversity, with strains specifically identified in West and East Africa (Fargette  
63 et al., 2004; Pinel-Galzi et al., 2015). This structure of the genetic diversity results from an  
64 independent dynamic of dispersion of RYMV between these two regions (Trovão et al., 2015). Six  
65 strains have been specifically identified in West Africa: S1ca, S1wa, S2, S3, Sa and Sg (Hébrard et  
66 al., 2020).

67 Three major resistance genes (named *rymv1*, *rymv2* and *RYMV3*) were identified in rice, mainly  
68 in African rice species, *Oryza glaberrima* (Orjuela et al., 2013; Pidon et al., 2020; Pidon et al.,  
69 2017; Thiémélé et al., 2010). Note that these resistant genes have not been used for the  
70 development of commercial rice varieties, *i.e.* the deployment of resistant varieties in fields has  
71 not yet widely been performed. Almost all rice varieties cultivated in Africa currently is of the  
72 Asian species, *O. sativa*, which is highly susceptible to RYMV infection.

73 It has been shown in controlled conditions that RYMV is able to adapt and to overcome the  
74 resistance. The molecular mechanism underlying the resistance-breaking (RB) process has been  
75 identified. It involved the emergence of specific point mutations on different regions of the RYMV  
76 genome (Bonnamy et al., 2022; Pinel-Galzi et al., 2016, 2007; Poulicard et al., 2014; Traore et al.,  
77 2010). Interestingly, the RYMV strains present contrasted RB abilities depending on the  
78 resistance gene. Based on these experimental results and considering the spatial distribution of  
79 RYMV strains in Africa, resistance-breaking risk maps were established (Hébrard et al., 2018).  
80 However, the validity and the sustainability of these risk maps are strictly dependent on the  
81 dispersion of the strains in field conditions and on the evolution of the local RYMV genetic  
82 diversity.

83 With more than 323,900 hectares dedicated to rice cultivation, Ghana is an important rice  
84 producer country in Africa (FAOSTAT 2021;  
85 <https://data.un.org/Data.aspx?d=FAO&f=itemCode%3A27>). Distributed across 6 ecological  
86 zones and 16 regions (Fig. 1A), there are mainly three agroecosystems for rice production in

87 Ghana, namely irrigated, rained lowland and upland rice areas. During the last decades, the  
88 acreage of cultivated rice is rapidly increasing and rice cultivation intensified to cope with the  
89 rising demand due to demographic and societal changes. Such agricultural changes surely make  
90 rice cultivation more exposed and vulnerable to pathogen epidemics. However, very little is  
91 known about the distribution and the genetic diversity of rice pathogens in Ghana, especially  
92 RYMV. Actually, only one coat protein (CP) sequence has been reported so far from a single  
93 isolate collected in Ghana in 1995 (Pinel et al., 2000). This CP sequence indicates that the S2 strain  
94 is present in Ghana but, as Ghana is surrounded by countries where high RYMV genetic diversity  
95 have been reported (with the presence of S1wa, S1ca, and S2 strains in Ivory Coast, Burkina Faso  
96 and Togo (Pinel-Galzi et al., 2015), the genetic diversity of RYMV in Ghana could be much higher.

97  
98 Thus, the aim of this study was to assess the geographical distribution of RYMV in the most  
99 important rice cultivation regions of in Ghana and analyze the genetic diversity using the CP and  
100 complete genome sequencing. In addition, this genetic information will allow us to reconstruct  
101 the spatio-temporal dynamics of the RYMV in Ghana and to identify the ways of introduction and  
102 dispersion of the virus within Ghana and between neighboring countries. Our results will greatly  
103 contribute to the epidemiological surveillance of rice in Ghana and will facilitate the development  
104 of efficient and sustainable control strategies based on the deployment in fields of resistant  
105 varieties against RYMV adapted to this viral diversity to minimize the risk of emergence of RB  
106 variants.

## 107 108 **2. Methods**

### 109 *2.1 Field survey and sampling*

110 Surveys to collect symptomatic RYMV leaf samples were conducted in 2010, 2012 and 2020  
111 during the major rice cropping season (July-September) to cover a total of 11 rice growing regions  
112 spread across different agroecological zones of Ghana, notably Ashanti, Ahafo, Bono, Bono East,  
113 Central, Eastern, Greater Accra, Northern, Upper East, Western and Volta regions (Fig. 1A). All  
114 the samples were collected from rainfed lowland fields. The GPS coordinates were noted for leaf  
115 samples collected in each location, and the rice variety plus the age of the plants were recorded

116 during the 2020 survey (Table S1). The symptomatic leaves sampled from each plant were cut  
117 using a pair of scissors and stored in a notebook used as herbarium. The scissors were disinfected  
118 by dipping in 70% ethanol before sampling from the subsequent plants. The sample locations  
119 were reported on a map based on MapChart website (<https://mapchart.net/>).

120

### 121 *2.2 Serological detection and genome sequencing*

122 The DAS-ELISA tests were performed with polyclonal antibodies targeting the RYMV coat protein  
123 (CP) on the leaf samples as described previously (N'Guessan et al., 2001) to confirm the infection  
124 status of symptomatic leaves. To obtain the coat protein (CP) gene (ORF4; 720 nt) or the complete  
125 genome (ca. 4,45kb) sequences, total RNA from infected leaves were first extracted using  
126 GeneJET Plant RNA Purification Kit (ThermoFisher SCIENTIFIC, MA, USA). Then, the CP genes  
127 and/or the complete genomes were specifically amplified by RT-PCR as previously described  
128 (Pinel et al., 2000). The RT-PCR products were sequenced using the Sanger technique (Genewiz,  
129 NJ, USA) and the sequences were deposited in GenBank (Accession Nos. from OQ200610 to  
130 OQ200655 for CP, from OQ225940 to OQ225946 for complete genomes; Table S2).

131

### 132 *2.3 Genetic diversity and phylogenetic analyses*

133 The partial and complete genome sequences obtained from Ghana were compared to those  
134 already deposited in NCBI (July 2021; cf. (Issaka et al., 2021), i.e. 261 CP gene sequences and 37  
135 complete genome sequences from isolates collected in nearly all-rice growing countries from  
136 West Africa. Multiple sequence alignments were performed using MUSCLE (Edgar, 2004)  
137 implemented in SEAVIEW v4.7 (Gouy et al., 2010). A search for potential recombination signals  
138 in the RYMV sequences was performed with the pairwise homoplasy test (PHI test) using  
139 SplitsTree v4.14.6 (Huson and Bryant, 2006) and the seven algorithms implemented in RDP v4.97  
140 (Martin et al., 2015). Recombination events detected by at least 4 methods and with *P*-values  
141 below  $10^{-5}$  were considered.

142 We reconstructed Maximum-likelihood (ML) phylogenetic trees with SEAVIEW using the best-  
143 fitted nucleotide substitution models (Tamura 3-parameter+G+I and GTR+G+I for the CP and the  
144 complete genome datasets, respectively) determined with MEGA6 (Tamura et al., 2013) and 100

145 bootstrap replications. Phylogenetic trees were drawn using FigTree v1.3.1  
146 (<http://tree.bio.ed.ac.uk/software/figtree/>). The genetic diversity of the ORF4 dataset was  
147 estimated using the best-fitted nucleotide substitution model, with standard errors of each  
148 measure based on 1000 replicate bootstraps, as implemented in MEGA6.

149

#### 150 *2.4 Tests of temporal signal*

151 In order to examine the degree of temporal signal in the CP dataset, we first used an exploratory  
152 linear regression approach (Duchêne et al., 2015; Murray et al., 2016). Based on ML phylogenetic  
153 tree reconstructed with the CP dataset, the temporal signal of this sequence dataset was  
154 visualized and tested in PhyloStemS (Doizy et al., 2020) to regress phylogenetic root-to-tip  
155 distances against sampling date using the root that minimized the residual mean squares.

156 Then, the significance of the temporal signal was evaluated by a date-randomization test. Thus,  
157 the mean rate and its 95% highest probability density (HPD) estimated with the observed  
158 sampled dates using the Bayesian Evolutionary Analysis Sampling Trees BEAST v1.10.4 package  
159 (Suchard et al., 2018), see below) were compared with a null distribution obtained by randomly  
160 permutating the tip dates 10 times (Firth et al., 2010). As previously described (Duchêne et al.,  
161 2015; Murray et al., 2016), the criterion for a significant temporal signal was that the 95% HPD  
162 for the rate estimate obtained with the observed sampled dates should not overlap with the 95%  
163 HPD for the estimate obtained with randomized sampling times.

164

#### 165 *2.5 Bayesian evolutionary inferences*

166 We reconstructed time-calibrated epidemic histories of the RYMV in West Africa based on the  
167 CP dataset using a Bayesian statistical framework implemented in BEAST version 1.10.4 software  
168 (Suchard et al., 2018) and the BEAGLE library (Ayres et al., 2012) to improve computational  
169 performance. BEAST program uses Markov chain Monte Carlo (MCMC) integration to average  
170 over all plausible evolutionary histories for the data, as reflected by the posterior probability. The  
171 nucleotide substitution process was reconstructed using HKY85+G substitution model, the  
172 lognormal relaxed clock model and an initial substitution rate of  $10^{-3}$  substitutions/site/year as  
173 previously determined for RYMV (Dellicour et al., 2018; Issaka et al., 2021; Trovão et al., 2015).

174 To study the geographic dispersion of RYMV in continuous time and space, we used a Bayesian  
175 SkyGrid coalescence model (Gill et al., 2013) and a relaxed random walk model for the continuous  
176 location traits (latitude and longitude of each isolate) with a Cauchy probability distribution  
177 (according to the best model identified in (Trovão et al., 2015). Bayesian inference using  
178 continuous diffusion models yield a posterior distribution of phylogenetic trees, each having  
179 ancestral nodes annotated with location estimates.

180 The MCMC analyses were run for 5 billion generations, sampled every 100,000th and discarding  
181 10% as the chain burn-in until adequate effective sample sizes were obtained (ESS>200 for all  
182 statistics). The parameter estimates were inspected using Tracer 1.7.2  
183 (<http://tree.bio.ed.ac.uk/software/tracer>). A maximum clade credibility (MCC) tree was  
184 summarized from the posterior tree distribution using TreeAnnotator v1.10.4 (BEAST v1.10.4  
185 package) and visualized with FigTree v1.3.1. The dates and the 95% HPD intervals of RYMV  
186 movements (introduction or release of RYMV to/from Ghana) were estimated as the interval  
187 between the nodes from the same clade with different geographical locations as identified in the  
188 MCC tree.

189 The visualization of the continuous phylogeographic reconstruction of the dispersal history of  
190 RYMV into West Africa and Ghana were performed with R functions available in the package  
191 “SERAPHIM” (Dellicour et al., 2016).

192

### 193 **3. Results**

#### 194 *3.1 Field samplings and geographic distribution of RYMV in Ghana*

195 A total of 63 samples showing symptoms that could be attributed to a viral infection, were  
196 collected in 11 regions (Table S1). Serological detection by DAS-ELISA confirmed that 49 of these  
197 samples were positive to RYMV. These positive samples were collected in 9 regions (Ahafo,  
198 Ashanti, Bono East, Central, Eastern, Greater Accra, Northern, Volta, Western, Fig. 1B, Table S1).  
199 In addition, RYMV infection was reported in 9 rice varieties, both improved (AGRA, Bouake189,  
200 CRI-AGRA, Jasmine85, Legon1, Togo Marshall) and local varieties (Aifaifa, Anyifila and Samasa)  
201 (Table S1). During that survey, RYMV was not reported in only two regions (Bono and Upper East),



202 although these regions are characterized by rice varieties and ecological zones similar to those  
203 where RYMV has been observed (Fig. 1A-B; Table S1).

204

### 205 *3.2 Genetic diversity of RYMV in Ghana*

206 All positive samples collected in 2010, 2012 and 2020 were used to analyze the genetic diversity  
207 of RYMV in Ghana. The coat protein gene was successfully amplified by RT-PCR and sequenced  
208 for 46 RYMV isolates (Table S2). These sequences were compared to the 261 CP gene sequences  
209 previously published (Issaka et al., 2021) coming from isolates collected in 15 West-African  
210 countries between 1975 and 2018, including one CP sequence from an isolate collected in 1995  
211 in Ghana. This dataset is thereafter referred to as the WA307 dataset. Based on this CP dataset,  
212 phylogenetic analysis showed that almost all RYMV isolates sampled in Ghana clustered within  
213 the S2 strain (Fig. S1, Fig. 1C), a strain which was previously reported in Ghana in 1995 (Gh1  
214 isolate) and detected in several countries surrounding Ghana (Ivory Coast, Guinea, Mali, Burkina  
215 Faso and Benin). The genetic diversity of S2 RYMV isolates from Ghana is similar to that observed  
216 for the S2 isolates from Ivory Coast and Guinea, and significantly higher than those from Mali or  
217 Benin (Table 1). In addition to the S2 strain, we identified that two isolates (out of 5) from the  
218 Northern region and one isolate (out of 5) from the Eastern region belong to the S1ca strain (Fig.  
219 1C, isolates 2020Gh113, 2020Gh139 and 2020Gh140 in Fig. S1 and Table S1). The genetic diversity  
220 of S1ca RYMV isolates from Ghana is similar to that observed in most of West-African countries  
221 (Table 1). The PHI test did not find statistically significant evidence for recombination within the  
222 WA307 dataset ( $p = 0.93$ ).

223 Based on this CP phylogenetic tree, seven isolates representing the genetic diversity of RYMV in  
224 Ghana were selected to be fully-sequenced (S2 strain: 2010Gh47, 2010Gh55, 2020Gh106,  
225 2020Gh109 and 2020Gh112; S1ca strain: 2020Gh113 and 2020Gh140) and then compared to the  
226 37 complete genome sequences from West Africa. This dataset is thereafter referred to as the  
227 WA44 dataset. The reconstruction of the phylogenetic tree by maximum likelihood confirmed  
228 that these isolates from Ghana clustered within the S2 and S1ca strains and demonstrated that  
229 these isolates are not directly genetically related between each other (Fig. 2). No statistically

230 significant evidence for recombination was detected in the RYMV genomes from Ghana (data not  
231 shown).

232 Altogether, these results showed that the genetic diversity of the S2 and S1ca isolates from  
233 Ghana is representative of those circulating in West Africa, which suggests that several  
234 introductions of RYMV to Ghana occurred and account for the genetic diversity in this country.

235

### 236 *3.3 Epidemic histories of RYMV in Ghana*

237 To reconstruct the evolutionary history of RYMV in West Africa and to gain information on the  
238 epidemic histories and trajectories of this virus in Ghana, we first examined the strength of the  
239 temporal signal of the WA307 sequence dataset. The linear regression exploration of root-to-tip  
240 distances as a function of sampling time suggested a weak but consistent temporal signal (Fig.  
241 S2), with a low correlation coefficient (adjusted  $R^2 < 0.1$ ;  $p < 10^{-6}$ ). The presence of temporal signal  
242 was then confirmed consistently using BEAST and date-permutation tests (tip-date and clustered  
243 tip randomizations). Actually, the estimates of the permutations did not overlap those of the real  
244 (non-permuted) WA307 dataset (Fig. S3). As no signal for recombination event was detected  
245 on the WA307 dataset, we reconstructed the evolutionary and the dispersal histories of the  
246 RYMV in West Africa by Bayesian inferences. The generated maximum clade credibility (MCC)  
247 tree strengthened the topology of the phylogenetic tree previously obtained by maximum  
248 likelihood (Fig. 3A). In addition, this MCC tree and the projection of this MCC tree on sequential-  
249 dated maps (Fig. 3B-E; Fig. S4) highlighted that several independent introductions occurred in  
250 Ghana.

251 By extracting information from the MCC tree and by mapping the tip locations of the MCC tree  
252 in geographic space, we identified at least three independent introductions of isolates belonging  
253 to S2 strain (branches labelled 1 to 3 in Fig. 3A) to southern part of Ghana. Interestingly, all the  
254 S2 clusters identified in Ghana are directly genetically related to isolates from Ivory Coast (Fig.  
255 3A), suggesting that S2 strain introductions to Ghana resulted from RYMV movements from the  
256 Ivory Coast. The date intervals between the nodes of branches 1 to 3 provide estimates of the  
257 timing of these introductions. Thus, a first S2 introduction occurred during the early 1990's  
258 (interval between the node 1 and the date of collection of the 1995Gh1 isolate, *i.e.* 1987.8-1995.6

259 with 95%HPD [1983.3-1995.6]; Fig. 3A-C; Table2). Then, a second S2 introduction occurred during  
260 the early 2000's (branch 2, interval between the nodes 2a and 2b, *i.e.* 1999.2-2003.5 [1993.1-  
261 2007.3]; Fig. 3A-C; Table2). Finally, a more recent S2 introduction occurred after the 2000's  
262 (branch 3, interval between node 3 and the date of collection of the 2020Gh133 isolate, *i.e.*  
263 1999.8-2020.6 [1995.0-2020.6]; Fig. 3A-E; Table2). Although S2 isolates circulated in the South  
264 since the early 1990's, it is interesting to note that these isolates dispersed to the Northern Ghana  
265 only from the 2010's (branch 6, interval between nodes 6a and 6b, *i.e.* 2009.9-2013.8 [2007.3-  
266 2016.8]; Fig. 3A-E; Table2). We also identified one S2 dispersion from Ghana to Benin that  
267 occurred between 2008 and 2010 (branch 7, interval between nodes 7a and 7b, *i.e.* 2008.8-  
268 2009.6 [2007.1-2010.5]; Fig. 3A-D; Table2).

269 In parallel with the dispersion of the S2 strain, the phylogeographic reconstruction showed two  
270 independent introductions of RYMV isolates belonging to the S1ca strain in Ghana (branches  
271 labelled 4 and 5 in Fig. 3A). Because there are few isolates of S1ca strain in Ghana, the range of  
272 dating estimates are more extended than for those of S2, but the phylogeographic  
273 reconstructions suggested that these RYMV movements occurred after the 1990's. A first S1ca  
274 introduction to Ghana was identified as the 2020Gh113 isolate (branch 4, interval between the  
275 node 4 and the date of collection of the 2020Gh113 isolate, *i.e.* 1986.9-2020.6 [1970.7-2020.6];  
276 Fig. 3A-E; Table2). This isolate clustered with isolates from Nigeria and Niger, suggesting an RYMV  
277 way of transmission from these countries (or other intermediate ones as Benin or Togo) to the  
278 South-eastern part of Ghana. In addition, an independent S1ca introduction occurred in the  
279 Northern regions of Ghana at the same period (1992.89-2015.02 [1979.46-2019.32]; Fig. 3A-D,  
280 Table2), probably from isolates genetically related to the ones reported in South-East of Burkina  
281 Faso.

282

#### 283 **4. Discussion**

284 The observation of rice plants with RYMV characteristic symptoms and the confirmation of the  
285 infectious status of these plants by DAS-ELISA and coat protein (gene or complete genome  
286 sequencing demonstrated that the virus has been widely distributed in Ghana since 1995, the  
287 date of collection of the first RYMV isolate in this country. Actually, RYMV has been detected in

288 nine regions covering five ecological zones as Guinea savannah (Northern region), forest  
289 savannah transition (Bono East), semi-deciduous rainforest (Ahafo, Ashanti, Central and Eastern  
290 regions), high rainforest (Western region) and coastal savannah (Greater Accra and Volta regions)  
291 (Fig. 1A-B). In addition, RYMV was detected in all improved and local rice varieties surveyed  
292 during this study (Table S1), suggesting that the resistance to RYMV is not yet a “usual selection”  
293 criteria for the industrial rice producers or the small farm holders.

294 Field surveys between 1995 and 2012 only collected isolates from S2 strain in Ghana. The survey  
295 in 2020 covered a wider geographical scale and larger agro-environmental conditions (from 2 to  
296 11 regions) and confirmed that RYMV isolates circulating in this country belong mainly to the S2  
297 strain (Fig. 2, Fig. S1). Surprisingly, the genetic diversity detected with the CP gene sequences of  
298 the S2 isolates in Ghana is high, which is similar to that reported for the whole S2 strain dataset  
299 (96 isolates) from Ivory Coast, Guinea, Mali, Benin and Burkina Faso (Table1). This observation  
300 suggests that the RYMV isolates collected in Ghana are representative of the genetic diversity of  
301 the S2 strain in West Africa, which reflect the complex epidemiological history of RYMV in Ghana.  
302 In addition, the field survey in 2020 identified three isolates (out of 47, i.e. ca. 6%) from the  
303 Northern and the Eastern regions clustering within the S1ca strain. This suggests that Ghana has  
304 faced dynamic RYMV epidemics in the last decades. Although both RYMV strains circulate within  
305 the same agro-environment (Fig. 1A, Table S1), the analysis of complete RYMV genomes of  
306 isolates from Ghana did not reveal any recombination event (data not shown).

307  
308 Based on 307 CP gene sequences of RYMV isolates collected in 15 West African countries since  
309 1975, including 49 isolates from Ghana sampled from 1995 to 2020, the phylogeographic  
310 reconstructions showed that RYMV was first introduced in Ghana in the 1990’s (Fig. 3C). This  
311 introduction date is remarkably later than that in other countries from West and Central Africa,  
312 including neighboring countries of Ghana. Actually, spatio-temporal modeling revealed that  
313 RYMV disseminated efficiently from Mali to Burkina Faso and Ivory Coast during the 1950’s, to  
314 Nigeria during the 1960’s, to Gambia, Sierra Leone and Cameroon during the 1970’s, and then to  
315 Guinea, Niger and Chad during the 1980’s (Fig. S4; Dellicour et al., 2018; Issaka et al., 2021; Trovão  
316 et al., 2015). We identified at least five independent introductions of RYMV into Ghana (Fig. 3).

317 An important way of S2 strain transmission was identified between Ivory Coast and the Southern  
318 Ghana (probably Ahafo, Ashanti and Western regions). In addition to similar environmental  
319 conditions on both sides of this border (high rainforest), these reiterated RYMV introductions  
320 could be favored by agronomical interactions between Ivory Coast and Ghana that could  
321 drastically increase the probabilities of RYMV dispersion by movement of infected biological  
322 material. The most recent introduction in Ghana was apparent through the isolate 2020Gh133  
323 only (branch 3, Fig. 3A-E) collected in 2020 in Western region, *i.e.* at the border with Ivory Coast,  
324 illustrating the continuous introduction of RYMV isolates from this country since the 1990's.  
325 Interestingly, these introductions are paralleled to the regular increase in rice cultivation areas in  
326 Ghana from the 1990's and more intensively from 2008 (FAOSTAT 2021;  
327 <https://data.un.org/Data.aspx?d=FAO&f=itemCode%3A27>). Note that the number of RYMV  
328 introductions in Ghana is certainly underestimated by these epidemiological models as not all  
329 introductions lead to an epidemic success that maintains the virus in time and space. For  
330 instance, our study shows that no RYMV isolates genetically related to the 1995Gh1 isolate  
331 (branch 1, Fig. 3A) were sampled during the 2010, 2012 and 2020 surveys, which suggest that  
332 this RYMV lineage went extinct.

333 While RYMV isolates belonging to the S2 strain were present in Ghana since the 1990's, the virus  
334 spread very recently to the Northern regions (2010-2017; Table 2, Fig. 3E). The delay of RYMV  
335 dispersion to the North could be due to agronomical and topological factors. The discontinuity of  
336 the rice cultivation landscape in Ghana, and especially with less rice cultivation area in the Central  
337 regions (Fig. 1A), could make a barrier against the RYMV expansion between the Southern and  
338 the Northern regions. Actually, the rice connectivity landscape has been suggested (Trovão et al.,  
339 2015), although not statistically confirmed (Dellicour et al., 2018), to be the major determinant  
340 of RYMV spread. The barrier to RYMV dispersion at the North could also be due to  
341 biogeographical factors. As Northern Ghana is characterized by drier conditions, the agricultural  
342 and environmental conditions could be less favorable to RYMV epidemics in this region (different  
343 cultural practices, less alternative hosts, rice landscape more fragmented; Traore et al., 2009).  
344 However, the initial limitation of RYMV to spread from the South to the North certainly lessened  
345 in the last decades in parallel to rice intensification and extensification in Ghana (Demont, 2013);

346 FAOSTAT 2021, <https://data.un.org/Data.aspx?d=FAO&f=itemCode%3A27>). Similarly, the late  
347 introduction of S1ca could also demonstrate the tendency of RYMV to circulate more easily  
348 during the last decades. Actually, while the phylogeographic reconstructions estimate that RYMV  
349 is present in South-East of Burkina Faso since the 1950's (Fig. 3B) and later on in Niger and Nigeria,  
350 we identified two RYMV movements from these countries to the Northern and the South-eastern  
351 regions of Ghana only after the 1990's (branches 4 and 5 in Fig. 3A-E; Table2). Altogether, because  
352 of the recent introductions of RYMV to the Northern regions of Ghana (both S2 and S1ca strains)  
353 and the agro-environmental factors that could be unfavorable to virus transmission or dispersion,  
354 RYMV has probably not reached an epidemic status in the North of Ghana. This could explain why  
355 no RYMV infected plant has been detected in the Upper East region (Fig. 1B).

356 Surprisingly, although S1wa strain is also present in Ivory Coast (Pinel-Galzi et al., 2015), our  
357 results show that only the S2 strain was introduced several times in Ghana from Ivory Coast. This  
358 suggests that S2 strain has advantageous biological or genetic properties to disperse in time and  
359 space compared to S1wa strain. Contrary to the lineage genetically related to Gh1 that probably  
360 got extinct (branch 2, Fig. 3A), we observed that one RYMV lineage (branch 2, Fig. 3A) has been  
361 detected during the 2010, 2012 and 2020 surveys and showed an efficient expansion to several  
362 Southern and Northern regions of Ghana, and also to Benin (branches 6 and 7, respectively; Fig.  
363 3A-E). These observations suggest that, in addition to agronomical and anthropological factors,  
364 some intrinsic genetic factors could favor the epidemic success of specific RYMV lineages or  
365 strains in West Africa. Earlier studies showed that isolates from S2 strain dominated over those  
366 from S1wa strain in experimental conditions (N'Guessan et al., 2000).

367  
368 The phylogeographic reconstructions established only one release of RYMV from Ghana. One  
369 route of RYMV dispersion was identified from Ghana to Benin (branch 7 in Fig. 3A-D) with a date  
370 estimated between 2008 and 2010 (Table2, interval between nodes 7a and 7b). In West Africa  
371 RYMV history, this movement is only the second dispersion event of RYMV to pass across the  
372 Ghana/Burkina Faso axis, *i.e.* dispersion from West Africa to Central Africa. Actually, a first  
373 expansion of RYMV to Central Africa occurred during the 1950's through Burkina Faso and led to  
374 the genetic differentiation and the emergence of the S1ca strain (Dellicour et al., 2018; Issaka et

375 al., 2021; Trovão et al., 2015). Sixty years later, our results show that RYMV dispersed for a second  
376 time to Central Africa through Ghana, which resulted in the introduction of the S2 strain in West-  
377 Central Africa. Altogether, while the dispersion of RYMV from West Africa to Central Africa across  
378 the Burkina Faso/Ghana axis has been very rare during the RYMV history, this study show that  
379 Ghana corresponds to a junction area of RYMV populations that have circulated independently  
380 in West and Central Africa until the 1990's. Such movements will probably occur more frequently  
381 in the next decades because of the drastic intensification and the extensification of rice  
382 cultivation and in the increase of commercial and agronomical exchanges in all this region  
383 (Demont, 2013).

384

## 385 **5. Conclusion**

386 The generalization of the RYMV movements in West and Central Africa will definitely favor the  
387 coexistence of several strains in fields. As each RYMV strain is characterized by specific biological  
388 and pathogenic properties (Bonnamy et al., 2022; Hébrard et al., 2018; Pidon et al., 2020; Pinel-  
389 Galzi et al., 2007; Poulicard et al., 2012; Traore et al., 2010), the overlapping of the distribution  
390 area of RYMV strains could, in the future, lead to difficulties by controlling RYMV through the  
391 local deployment of rice varieties resistant to the virus. Thus, our study underlines the  
392 importance of both epidemiological surveillance and sanitary regulation at the local, regional,  
393 and international levels to limit the RYMV strain circulation in West Africa to avoid the co-  
394 circulation of RYMV genetic diversity that could make more difficult the design of sustainable  
395 control strategies against the virus.

396

## 397 **CRedit authorship contribution statement**

398 **Emmanuel Gilbert Omiat:** Resources, Investigation, Data Curation, Writing - Original Draft.

399 **Maxwell Darko Asante:** Conceptualization, Supervision, Resources, Investigation, Data Curation,

400 Writing - Review & Editing. **Valentin Stanislas Edgar Traoré:** Resources, Investigation, Writing -

401 Review & Editing. **Allen Oppong:** Validation, Investigation, Writing - Review & Editing. **Beatrice**

402 **Elohor Ifie:** Validation, Investigation. **Kirpal Agyemang Ofofu:** Validation, Investigation, Writing

403 - Review & Editing. **Jamel Aribi:** Methodology, Investigation. **Agnès Pinel-Galzi:** Methodology,

404 Investigation. **Aurore Comte:** Software, Data Curation, Visualization. **Denis Fargette:**  
405 Conceptualization, Software, Writing - Review & Editing, Supervision. **Eugénie Hébrard:** Writing  
406 - Review & Editing, Supervision, Funding acquisition. **Oumar Traoré:** Supervision, Project  
407 administration, Funding acquisition. **Samuel Kwame Offei:** Conceptualization, Supervision,  
408 Project administration, Funding acquisition. **Eric Yirenkyi Danquah:** Conceptualization,  
409 Supervision, Project administration, Funding acquisition. **Nils Poulicard:** Conceptualization,  
410 Software, Data Curation, Writing - Original Draft, Visualization, Funding acquisition.

411

## 412 **Acknowledgements**

413 We thank Dr Drissa Silué and the technical staff at rice breeding programme of the CSIR-Crops  
414 Research Institute, Ghana for helping with the collection of samples from across the country. This  
415 study was conducted under the European Union Funded Scientists in Crop Improvement for Food  
416 Security in Africa (SCIFSA) Project. This work was partially funded by the French National Research  
417 Agency as “young researchers” program (ANR JCJC SPADYVA; ANR-20-CE35-0008-01) and CGIAR  
418 Research Program RICE. The funders had no role in the study design, data collection, and  
419 interpretation, or the decision to submit the work for publication. We thank the ISO 9001  
420 certified IRD iTrop HPC (member of the South Green Platform, <http://www.southgreen.fr>) at IRD  
421 Montpellier for providing HPC resources that have contributed to the research results reported  
422 within this article. URL: <https://bioinfo.ird.fr>.

423

## 424 **Figure captions**

425 **Fig. 1.** Rice cultivation in Ghana and location of the rice samples/RYMV isolates.

426 **(A)** Rice production in Ghana. Regions are colored according to the surface (in hectares, ha)  
427 dedicated to rice cultivation. The Upper West (UW), Upper East (UE), North East (NE), Savannah  
428 Region (SR), Northern (NO), Bono (BO), Bono East (BE), Oti (OT), Ahafo (AH), Ashanti (AS), Eastern  
429 (EA), Volta (VO), Western North (WN), Western (WE), Central (CE) and Greater Accra (GA) regions  
430 and the borders of the ecological zones (dotted lines) are indicated. **(B)** Schematic locations of  
431 the RYMV isolates sampled in Ghana in 1995 (yellow), 2010 (orange), 2012 (brown) and 2020  
432 (black). Samples collected in 2020 and negative to RYMV detection by DAS-ELISA are indicated by



433 white spots. (C) Schematic location of the RYMV isolates belonging to S1ca (red) and S2 (blue)  
434 strains.

435

436 **Fig. 2.** Maximum-likelihood trees inferred from the WA44 complete genome sequence dataset.  
437 The names of the RYMV isolates from Ghana (Gh) are indicated in black, the fully-sequenced  
438 isolates from Ghana are highlighted, and the branches are colored according to the six RYMV  
439 strains of West and Central Africa (S1wa, S1ca, S2, S3, Sa, Sg). Number at each node indicates  
440 bootstrap percentages based on 100 replicates (only values above 60 are indicated). The country  
441 of origin of the isolates is mentioned in these phylogenetic trees: Benin (Be), Burkina Faso (BF),  
442 Cameroon (Ca), Central African Republic (Ce), Ivory Coast (CI), Gambia (Ga), Mali (Ma), Niger (Ng),  
443 Nigeria (Ni), Senegal (Se), Sierra Leone (SL), Chad (Tc) and Togo (Tg).

444

445 **Fig. 3.** Phylogeographic reconstructions of RYMV epidemics in Ghana.

446 (A) Time-calibrated maximum clade credibility (MCC) tree reconstructed by continuous Bayesian  
447 evolutionary inference with the 307 coat protein gene sequence dataset (WA307). Branches are  
448 colored according to the six RYMV strains of West Africa (S1wa, S1ca, S2, S3, Sa, Sg). The values  
449 on the nodes correspond to their posterior probabilities (only values above 0.60 are indicated).  
450 The bottom axis gives the timeframe (by date) of the RYMV diversification. Seven branches of  
451 interest are labeled on the MCC tree (from 1 to 7) and their relative nodes (1, 2a, 2b, 3, 4, 5a, 5b,  
452 6a, 6b, 7a, 7b) were annotated with their time to the most common ancestor and their 95%  
453 highest probability density (HPD) intervals (Table 2). (B to D) Reconstruction of the continuous  
454 spatio-temporal dispersion of RYMV in West Africa with snapshots in 1992 (B), 2008 (C) and 2020  
455 (D) that capture introductions and release of RYMV in/out Ghana (delimited by the grey area).  
456 Black lines show a spatial projection of the MCC tree and colored areas represent statistical  
457 uncertainty in the dated estimated locations of RYMV internal nodes (95% HPD; from red to green  
458 to cover the time period from 1900 to 2020). The seven branches of interest in Figure 3A are  
459 indicated.

460

461 **Supplementary Materials**

462 **Fig. S1:** Maximum-likelihood trees inferred from the WA307 CP sequence dataset.  
463 The names of RYMV isolates from Ghana (Gh) are indicated in black, the fully-sequenced isolates  
464 from Ghana are highlighted, and the clades of the six RYMV strains of West Africa (S1wa, S1ca,  
465 S2, S3, Sa, Sg) are indicated. Number at each node indicates bootstrap percentages based on 100  
466 replicates (only values above 60 are indicated). The country of origin of the isolates is mentioned  
467 by their names: Benin (Be), Burkina Faso (BF), Cameroon (Ca), Central African Republic (Ce), Ivory  
468 Coast (Ci), Gambia (Ga), Mali (Ma), Niger (Ng), Nigeria (Ni), Senegal (Se), Sierra Leone (SL), Chad  
469 (Tc) and Togo (Tg).

470

471 **Fig. S2:** Exploration of the temporal signal from the RYMV WA307 dataset by a linear regression  
472 approach.

473 Root-to-tip divergence as a function of sampling time for Maximum Likelihood tree clusters  
474 obtained with 307 non-recombinant RYMV CP sequences. Plot representations in a 1970-2020  
475 **(A)** and a 1800-2020 **(B)** dating range.

476

477 **Fig. S3:** Exploration of the temporal signal from the RYMV WA307 dataset by date-randomization  
478 tests.

479 Violin plots representing the intervals of the substitution rate (substitution/site/year; **A** and **C**)  
480 and the time to most recent common ancestor (years; **B** and **D**) in the RYMV CP sequence dataset  
481 from the sampled dates (WA307) and from the dates after tip (A and B) or cluster (C and D)  
482 permutations (from perm1 to perm10). The blue areas represent the range of the 95% HPD  
483 estimates.

484

485 **Fig. S4:** Reconstruction of the continuous spatio-temporal dispersion of RYMV in West-Central  
486 Africa based on the Time-calibrated maximum clade credibility (MCC) tree reconstructed by  
487 continuous Bayesian evolutionary inference with the 307 coat protein gene sequence dataset  
488 (WA307).

489 Snapshots of the spatio-temporal dispersion of RYMV in 1992 **(A)**, 2008 **(B)**, 2010 **(C)** and 2020  
490 **(D)**. Ghana is delimited by the grey area. Black lines show a spatial projection of the MCC tree

491 and colored areas represent statistical uncertainty in the dated estimated locations of RYMV  
492 internal nodes (95% HPD; from red to green to cover the time period from 1900 to 2020).

493

#### 494 **References**

495 Abo, M.E., Sy, A.A., Alegbejo, M.D., 1997. Rice Yellow Mottle Virus (RYMV) in Africa: Evolution,  
496 Distribution, Economic Significance on Sustainable Rice Production and Management  
497 Strategies. *Journal of Sustainable Agriculture* 11, 85–111.

498 [https://doi.org/10.1300/J064v11n02\\_08](https://doi.org/10.1300/J064v11n02_08)

499 Agnoun, Y., Yelome, I., Sie, M., Albar, L., Ghesquière, A and Silué, D., 2019. Resistance of  
500 selected *Oryza glaberrima* landraces and their intra-specific breeding lines to Beninese  
501 Rice yellow mottle virus isolates. *Crop Protection* 119, 172–176.

502 <https://doi.org/10.1016/j.cropro.2019.01.022>

503 Allarangaye, M.D., Traoré, O., Traoré, E.V.S., Millogo, R.J., Guinko, S., Konaté, G., 2007. Host  
504 Range of Rice yellow mottle virus in Sudano-Sahelian Savannahs. *Pakistan Journal of*  
505 *Biological Sciences* 10, 1414–1421.

506 Allarangaye, M.D., Traoré, O., Traoré, E.V.S., Millogo, R.J., Konaté, G., 2006. Evidence of non-  
507 transmission of Rice yellow mottle virus through seed of wild host species. *Journal of*  
508 *Plant Pathology* 88, 309–315.

509 Attere, A.F., 1983. Reaction of *Oryza glaberrima* Accessions to Rice Yellow Mottle Virus. *Plant*  
510 *Disease* 67, 420. <https://doi.org/10.1094/pd-67-420>

511 Ayres, D.L., Darling, A., Zwickl, D.J., Beerli, P., Holder, M.T., Lewis, P.O., Huelsenbeck, J.P.,  
512 Ronquist, F., Swofford, D.L., Cummings, M.P., Rambaut, A., Suchard, M.A., 2012.

513 BEAGLE: An Application Programming Interface and High-Performance Computing  
514 Library for Statistical Phylogenetics. *Systematic Biology* 61, 170–173.

515 <https://doi.org/10.1093/sysbio/syr100>

516 Bakker, W., 1974. Characterisation and ecological aspects of Rice yellow mottle virus in Kenya.  
517 *Agricultural Research Report (Wageningen)* 829, 1–152.

518 Bonnamy, M., Pinel-Galzi, A., Gorgues, L., Chalvon, V., Hébrard, E., Chéron, S., Nguyen, T.H.,

519 Poulicard, N., Sabot, F., Pidon, H., Champion, A., Césari, S., Kroj, T., Albar, L., 2022. Rapid

520 evolution of an RNA virus to escape recognition by a rice nucleotide-binding and  
521 leucine-rich repeat domain immune receptor. *New Phytologist* *nph.18532*.  
522 <https://doi.org/10.1111/nph.18532>

523 Dellicour, S., Rose, R., Faria, N.R., Lemey, P., Pybus, O.G., 2016. SERAPHIM: studying  
524 environmental rasters and phylogenetically informed movements. *Bioinformatics* *32*,  
525 3204–3206. <https://doi.org/10.1093/bioinformatics/btw384>

526 Dellicour, S., Vrancken, B., Trovão, N.S., Fargette, D., Lemey, P., 2018. On the importance of  
527 negative controls in viral landscape phylogeography. *Virus Evolution* *4*, vey023.  
528 <https://doi.org/10.1093/ve/vey023>

529 Demont, M., 2013. Reversing urban bias in African rice markets: A review of 19 National Rice  
530 Development Strategies. *Global Food Security* *2*, 172–181.  
531 <https://doi.org/10.1016/j.gfs.2013.07.001>

532 Doizy, A., Prin, A., Cornu, G., Chiroleu, F., Rieux, A., 2020. Phylostems: a new graphical tool to  
533 investigate temporal signal of heterochronous sequences at various evolutionary scales  
534 (preprint). *Evolutionary Biology*. <https://doi.org/10.1101/2020.10.19.346429>

535 Duchêne, S., Duchêne, D., Holmes, E.C., Ho, S.Y.W., 2015. The Performance of the Date-  
536 Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data. *Mol Biol*  
537 *Evol* *32*, 1895–1906. <https://doi.org/10.1093/molbev/msv056>

538 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high  
539 throughput. *Nucleic Acids Research* *32*, 1792–1797. <https://doi.org/10.1093/nar/gkh340>

540 Fargette, D., Pinel, A., Abubakar, Z., Traore, O., Brugidou, C., Fatogoma, S., Hébrard, E., Choisy,  
541 M., Sere, Y., Fauquet, C and Konate, G., 2004. Inferring the evolutionary history of Rice  
542 yellow mottle virus from genomic, phylogenetic, and phylogeographic studies. *Journal of*  
543 *Virology* *78*, 3252–3261. <https://doi.org/10.1128/JVI.78.7.3252>

544 Fauquet, C.M., Thouvenel, J.C., 1977. Isolation of the Rice yellow mottle virus in Ivory Coast.  
545 *Plant Disease* *61*, 443–446.

546 Firth, C., Kitchen, A., Shapiro, B., Suchard, M.A., Holmes, E.C., Rambaut, A., 2010. Using Time-  
547 Structured Data to Estimate Evolutionary Rates of Double-Stranded DNA Viruses.

548 Molecular Biology and Evolution 27, 2038–2051.  
549 <https://doi.org/10.1093/molbev/msq088>

550 Gill, M.S., Lemey, P., Faria, N.R., Rambaut, A., Shapiro, B., Suchard, M.A., 2013. Improving  
551 Bayesian Population Dynamics Inference: A Coalescent-Based Model for Multiple Loci.  
552 Molecular Biology and Evolution 30, 713–724. <https://doi.org/10.1093/molbev/mss265>

553 Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView Version 4: A Multiplatform Graphical User  
554 Interface for Sequence Alignment and Phylogenetic Tree Building. Molecular Biology and  
555 Evolution 27, 221–224. <https://doi.org/10.1093/molbev/msp259>

556 Hébrard, E., Pinel-Galzi, A., Oludare, A., Poulicard, N., Aribi, J., Fabre, S., Issaka, S., Mariac, C.,  
557 Dereeper, A., Albar, L., Silué, D., Fargette, D., 2018. Identification of a Hypervirulent  
558 Pathotype of Rice yellow mottle virus : A Threat to Genetic Resistance Deployment in  
559 West-Central Africa. Phytopathology 108, 299–307. [https://doi.org/10.1094/PHYTO-05-](https://doi.org/10.1094/PHYTO-05-17-0190-R)  
560 [17-0190-R](https://doi.org/10.1094/PHYTO-05-17-0190-R)

561 Hébrard, E., Poulicard, N., Rakotomalala, M., 2020. Rice Yellow Mottle Virus (Solemoviridae), in:  
562 Encyclopedia of Virology. Elsevier, p.675–680. ISBN: 978-0-12-814516-6.  
563 <https://doi.org/10.1016/B978-0-12-809633-8.21244-2>

564 Issaka, S., Traoré, O., Longué, R.D.S., Pinel-Galzi, A., Gill, M.S., Dellicour, S., Bastide, P., Guindon,  
565 S., Hébrard, E., Dugué, M.-J., Séré, Y., Semballa, S., Aké, S., Lemey, P., Fargette, D., 2021.  
566 Rivers and landscape ecology of a plant virus, Rice yellow mottle virus along the Niger  
567 Valley. Virus Evolution 7, veab072. <https://doi.org/10.1093/ve/veab072>

568 Konate, G., Sarra, S., Traore, O., 2001. Rice yellow mottle virus is seed-borne but not seed  
569 transmitted in rice seeds. European Journal of Plant Pathology 107, 361–364.

570 Kouassi, N. K., N’Guessan, P., Albar, L., Fauquet, C. M., and Brugidou, C., 2005. Distribution and  
571 characterization of Rice yellow mottle virus: A threat to African farmers. Plant Disease  
572 89, 124–133. <https://doi.org/10.1094/PD-89-0124>

573 Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and  
574 analysis of recombination patterns in virus genomes. Virus Evolution 1.  
575 <https://doi.org/10.1093/ve/vev003>

576 Murray, G.G.R., Wang, F., Harrison, E.M., Paterson, G.K., Mather, A.E., Harris, S.R., Holmes,  
577 M.A., Rambaut, A., Welch, J.J., 2016. The effect of genetic structure on molecular dating  
578 and tests for temporal signal. *Methods Ecol Evol* 7, 80–89.  
579 <https://doi.org/10.1111/2041-210X.12466>

580 N’Guessan, P., Pinel, A., Caruana, M.L., Frutos, R., Sy, A.A., Ghesquière, A., Fargette, D., 2000.  
581 Evidence of the presence of two serotypes of Rice yellow mottle virus Sobemovirus in  
582 Côte d’Ivoire. *European Journal of Plant Pathology* 106, 167–178.  
583 <https://doi.org/10.1023/A:1008792109954>

584 N’Guessan, P., Pinel, A., Sy, A.A., Ghesquière, A., Fargette, D., 2000. Distribution, pathogenicity,  
585 and interactions of two strains of Rice yellow mottle virus in forested and savanna zones  
586 of West Africa. *Plant Disease* 85, 59-64. <https://doi.org/10.1094/PDIS.2001.85.1.59>

587 Orjuela, J., Deless, E.F.T., Kolade, O., Chéron, S., Ghesquière, A., Albar, L., 2013. A Recessive  
588 Resistance to Rice yellow mottle virus Is Associated with a Rice Homolog of the CPR5  
589 Gene, a Regulator of Active Defense Mechanisms. *MPMI* 26, 1455–1463.  
590 <https://doi.org/10.1094/MPMI-05-13-0127-R>

591 Pidon, H., Chéron, S., Ghesquière, A., Albar, L., 2020. Allele mining unlocks the identification of  
592 RYMV resistance genes and alleles in African cultivated rice. *BMC Plant Biol* 20, 222.  
593 <https://doi.org/10.1186/s12870-020-02433-0>

594 Pidon, H., Ghesquière, A., Chéron, S., Issaka, S., Hébrard, E., Sabot, F., Kolade, O., Silué, D., and  
595 Albar, L., 2017. Fine mapping of RYMV3: a new resistance gene to Rice yellow mottle  
596 virus from *Oryza glaberrima*. *Theoretical and Applied Genetics* 130, 807–818.  
597 <https://doi.org/10.1007/s00122-017-2853-0>

598 Pinel, A., N’Guessan, P., Bousalem, M., Fargette, D., 2000. Molecular variability of  
599 geographically distinct isolates of Rice yellow mottle virus in Africa. *Archives of virology*  
600 145, 1621–1638.

601 Pinel-Galzi, A., Dubreuil-Tranchant, C., Hébrard, E., Mariac, C., Ghesquière, A., Albar, L., 2016.  
602 Mutations in Rice yellow mottle virus Polyprotein P2a Involved in RYMV2 Gene  
603 Resistance Breakdown. *Front. Plant Sci.* 7. <https://doi.org/10.3389/fpls.2016.01779>

604 Pinel-Galzi, A., Rakotomalala, M., Sangu, E., Sorho, F., Kanyeka, Z., Traoré, O., Sérémé, D.,  
605 Poulicard, N., Rabenantoandro, Y., Séré, Y., Konaté, G., Ghesquière, A., Hébrard, E.,  
606 Fargette, D., 2007. Theme and Variations in the Evolutionary Pathways to Virulence of  
607 an RNA Plant Virus Species. *PLoS Pathog* 3, e180.  
608 <https://doi.org/10.1371/journal.ppat.0030180>

609 Pinel-Galzi, A., Traoré, O., Séré, Y., Hébrard, E. and Fargette, D., 2015. The biogeography of viral  
610 emergence: Rice yellow mottle virus as a case study. *Current Opinion in Virology* 10, 7–  
611 13. <https://doi.org/10.1016/j.coviro.2014.12.002>

612 Poulicard, N., Pinel-Galzi, A., Fargette, D., Hébrard, E., 2014. Alternative mutational pathways,  
613 outside the VPg, of rice yellow mottle virus to overcome eIF(iso)4G-mediated rice  
614 resistance under strong genetic constraints. *Journal of General Virology* 95, 219–224.  
615 <https://doi.org/10.1099/vir.0.057810-0>

616 Poulicard, N., Pinel-Galzi, A., Traoré, O., Vignols, F., Ghesquière, A., Konaté, G., Hébrard, E.,  
617 Fargette, D., 2012. Historical Contingencies Modulate the Adaptability of Rice Yellow  
618 Mottle Virus. *PLoS Pathog* 8, e1002482. <https://doi.org/10.1371/journal.ppat.1002482>

619 Rakotomalala, M., Vrancken, B., Pinel-Galzi, A., Ramavovololona, P., Hébrard, E.,  
620 Randrianangaly, J.S., Dellicour, S., Lemey, P., Fargette, D., 2019. Comparing patterns and  
621 scales of plant virus phylogeography: Rice yellow mottle virus in Madagascar and in  
622 continental Africa. *Virus Evolution* 5, vez023. <https://doi.org/10.1093/ve/vez023>

623 Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A., 2019. The  
624 global burden of pathogens and pests on major food crops. *Nat Ecol Evol* 3, 430–439.  
625 <https://doi.org/10.1038/s41559-018-0793-y>

626 Séré, Y., Fargette, D., Abo, M.E., Wydra, K., Bimerew, M., Onasanya, A., Akator, S.K., 2013.  
627 Managing the major diseases of rice in Africa., in: Wopereis, M.C.S., Johnson, D.E.,  
628 Ahmadi, N., Tollens, E., Jalloh, A. (Eds.), *Realizing Africa's Rice Promise*. CABI,  
629 Wallingford, pp. 213–228. <https://doi.org/10.1079/9781845938123.0213>

630 Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian  
631 phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* 4.  
632 <https://doi.org/10.1093/ve/vey016>

633 Suvi, W.T., Shimelis, H. and Mark, L., 2018. Breeding rice for rice yellow mottle virus resistance  
634 in Sub-Saharan Africa: a review. *Acta Agriculturae Scandinavica Section B: Soil and Plant*  
635 *Science* 69, 181–188. <https://doi.org/10.1080/09064710.2018.1523454>

636 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular  
637 Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30, 2725–  
638 2729. <https://doi.org/10.1093/molbev/mst197>

639 Thiémélé, D., Boisnard, A., Ndjiondjop, M.N., Chéron, S., Séré, Y., Aké, S., Ghesquière, A. and  
640 Albar, L., 2010. Identification of a second major resistance gene to Rice yellow mottle  
641 virus, RYMV2, in the African cultivated rice species, *O. glaberrima*. *Theoretical and*  
642 *Applied Genetics* 121, 169–179. <https://doi.org/10.1007/s00122-010-1300-2>

643 Traore, O., Pinel-Galzi, A., Issaka, S., Poulicard, N., Aribi, J., Aké, S., Ghesquière, A., Séré, Y.,  
644 Konate, G., Hébrard, E., Fargette, D., 2010. The adaptation of Rice yellow mottle virus to  
645 the eIF(iso)4G-mediated rice resistance. *Virology*.

646 Traore, O., Pinel-Galzi, A., Sorho, F., Sarra, S., Rakotomalala, M., Sangu, E., Kanyeka, Z., Séré, Y.,  
647 Konaté, G., Fargette, D., 2009. A reassessment of the epidemiology of Rice yellow mottle  
648 virus following recent advances in field and molecular studies. *Virus Research* 141, 258–  
649 267. <https://doi.org/10.1016/j.virusres.2009.01.011>

650 Trovão, N.S., Baele, G., Vrancken, B., Bielejec, F., Suchard, M.A., Fargette, D., Lemey, P., 2015.  
651 Host ecology determines the dispersal patterns of a plant virus. *Virus Evol* 1, vev016.  
652 <https://doi.org/10.1093/ve/vev016>



| Strain | Group    | d     | S.E.  |
|--------|----------|-------|-------|
| S2     | All S2   | 0.016 | 0.002 |
|        | CI       | 0.015 | 0.003 |
|        | Gh       | 0.014 | 0.002 |
|        | Gu       | 0.014 | 0.003 |
|        | Ma       | 0.009 | 0.002 |
|        | Be       | 0.001 | 0.001 |
|        | BF       | nd    | nd    |
| S1ca   | All S1ca | 0.042 | 0.006 |
|        | Ni       | 0.039 | 0.006 |
|        | Ca       | 0.039 | 0.008 |
|        | Gh       | 0.031 | 0.003 |
|        | BF       | 0.029 | 0.009 |
|        | Ce       | 0.024 | 0.006 |
|        | Ng       | 0.018 | 0.003 |
|        | Ch       | 0.017 | 0.003 |
|        | Tg       | 0.007 | 0.003 |
|        | Be       | 0.007 | 0.002 |

653

654 **Table 1.** Genetic diversity of the S2 and S1ca strains in West- and West-Central Africa.

655 d: genetic diversity estimated in substitution per site; S.E.: standard error; Benin (Be), Burkina  
656 Faso (BF), Cameroon (Ca), Central African Republic (Ce), Chad (Ch), Ivory Coast (CI), Ghana (Gh),  
657 Guinea (Gu), Mali (Ma), Niger (Ng), Nigeria (Ni), Togo (Tg). S.E.: standard errors; nd: not  
658 determined.

659

| n° node   | <i>Time before present (2020.6)</i> |               | Date   | Date 95%HPD       |
|-----------|-------------------------------------|---------------|--------|-------------------|
|           | Time                                | Time 95%HPD   |        |                   |
| <b>1</b>  | 32.8                                | [37,3 - 28,2] | 1987.8 | [1983,3 - 1992,5] |
| <b>2a</b> | 21.4                                | [27,5 - 18,1] | 1999.2 | [1993,1 - 2002,5] |
| <b>2b</b> | 17.1                                | [21,6 - 13,3] | 2003.5 | [1999,0 - 2007,3] |
| <b>3</b>  | 20.8                                | [25,6 - 18,2] | 1999.8 | [1995,0 - 2002,4] |
| <b>4</b>  | 33.7                                | [50,0 - 16,4] | 1986.9 | [1970,7 - 2004,2] |
| <b>5a</b> | 27.7                                | [41,1 - 16,0] | 1992.9 | [1979,5 - 2004,6] |
| <b>5b</b> | 5.6                                 | [9,8 - 1,3]   | 2015.0 | [2010,8 - 2019,3] |
| <b>6a</b> | 10.7                                | [13,3 - 8,0]  | 2009.9 | [2007,3 - 2012,6] |
| <b>6b</b> | 6.8                                 | [10,7 - 3,8]  | 2013.8 | [2009,9 - 2016,8] |
| <b>7a</b> | 11.8                                | [13,5 - 10,4] | 2008.8 | [2007,1 - 2010,2] |
| <b>7b</b> | 11.0                                | [12,0 - 10,1] | 2009.6 | [2008,6 - 2010,5] |

660

661 **Table 2.** Dating estimates of nodes of the seven branches of interest identified in the maximum  
662 clade credibility (MCC) tree (Figure 3A).

663 The estimates of the time to the most common ancestor (in years before July 2020, *i.e.* 2020.6),  
664 the dates and the 95% highest probability density (HPD) intervals are mentioned.

665

| Samples | Year | Region        | Variety    | Latitude | Longitude | Rice age | DAS-  | RYMV       |           |        |
|---------|------|---------------|------------|----------|-----------|----------|-------|------------|-----------|--------|
|         |      |               |            |          |           | (weeks)  | ELISA | CP Seq.    | Full Seq. | strain |
| Gh1     | 1995 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 1995Gh1    | S2        |        |
| Gh46    | 2010 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2010Gh46   | S2        |        |
| Gh47    | 2010 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2010Gh47   | 2010Gh47  |        |
| Gh48    | 2010 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2010Gh48   | S2        |        |
| Gh49    | 2010 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2010Gh49   | S2        |        |
| Gh50    | 2010 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2010Gh50   | S2        |        |
| Gh51    | 2010 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2010Gh51   | S2        |        |
| Gh52    | 2010 | Greater Accra | nd         | 5.6      | -0.2      | nd       | +     | 2010Gh52   | S2        |        |
| Gh53    | 2010 | Greater Accra | nd         | 5.6      | -0.2      | nd       | +     | 2010Gh53   | S2        |        |
| Gh54    | 2010 | Greater Accra | nd         | 5.6      | -0.2      | nd       | +     | 2010Gh54   | S2        |        |
| Gh55    | 2010 | Greater Accra | nd         | 5.6      | -0.2      | nd       | +     | 2010Gh55   | 2010Gh55  |        |
| Gh3227  | 2012 | Ashanti       | nd         | 6.6      | -1.3      | nd       | +     | 2012Gh3227 | S2        |        |
| Gh3229  | 2012 | Ashanti       | nd         | 6.6      | -1.3      | nd       | +     | 2012Gh3229 | S2        |        |
| Gh3244  | 2012 | Central       | nd         | 5.6      | -1.2      | nd       | +     | 2012Gh3244 | S2        |        |
| Gh3245  | 2012 | Central       | nd         | 5.6      | -1.2      | nd       | +     | 2012Gh3245 | S2        |        |
| Gh3387  | 2012 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2012Gh3387 | S2        |        |
| Gh3397  | 2012 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2012Gh3397 | S2        |        |
| Gh3399  | 2012 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2012Gh3399 | S2        |        |
| Gh3416  | 2012 | Ashanti       | nd         | 6.7      | -1.6      | nd       | +     | 2012Gh3416 | S2        |        |
| Gh101   | 2020 | Ashanti       | CRI-AGRA   | 6.8      | -1.8      | 7        | +     | 2020Gh101  | S2        |        |
| Gh102   | 2020 | Ashanti       | CRI-AGRA   | 6.8      | -1.8      | 7        | +     | 2020Gh102  | S2        |        |
| TOGO    |      |               |            |          |           |          |       |            |           |        |
| Gh103   | 2020 | Ashanti       | MARSHALL   | 6.7      | -1.5      | 7        | +     | 2020Gh103  | S2        |        |
| Gh104   | 2020 | Ashanti       | BOUAKE 189 | nd       | nd        | 7        | -     |            |           |        |
| Gh105   | 2020 | Ashanti       | CRI-AGRA   | 6.6      | -1.3      | 7        | +     | 2020Gh105  | S2        |        |
| Gh106   | 2020 | Ashanti       | CRI-AGRA   | 6.6      | -1.3      | 8        | +     | 2020Gh106  | S2        |        |
| Gh107   | 2020 | Ashanti       | JASMINE 85 | 6.7      | -1.4      | 8        | -     |            |           |        |
| Gh108   | 2020 | Ashanti       | JASMINE 85 | nd       | nd        | 12       | -     |            |           |        |
| Gh109   | 2020 | Ashanti       | CRI-AGRA   | 6.7      | -1.5      | 7        | +     | 2020Gh109  | S2        |        |
| Gh110   | 2020 | Ashanti       | CRI-AGRA   | 6.7      | -1.5      | 8        | -     |            |           |        |
| Gh111   | 2020 | Volta         | CRI-AGRA   | 6.1      | 0.9       | 8        | +     | 2020Gh111  | S2        |        |
| Gh112   | 2020 | Volta         | CRI-AGRA   | 6.1      | 0.9       | 8        | +     | 2020Gh112  | S2        |        |

|       |      |               |                                |      |      |    |   |            |           |      |
|-------|------|---------------|--------------------------------|------|------|----|---|------------|-----------|------|
| Gh113 | 2020 | Eastern       | LEGON1                         | 6.1  | 0.1  | 12 | + | 2020Gh113  | 2020Gh113 | S1ca |
| Gh114 | 2020 | Eastern       | LEGON1                         | 6.1  | 0.1  | 7  | - |            |           |      |
| Gh115 | 2020 | Eastern       | JASMINE 85                     | 6.0  | -1.0 | 8  | + | 2020Gh115  |           | S2   |
| Gh116 | 2020 | Eastern       | JASMINE 85                     | 6.0  | -1.0 | 7  | + | 2020Gh116  |           | S2   |
| Gh117 | 2020 | Eastern       | CRI-AGRA                       | 6.0  | -1.1 | 12 | + | 2020Gh117  |           | S2   |
| Gh118 | 2020 | Eastern       | JASMINE 85                     | 6.5  | -0.7 | 12 | + | 2020Gh118  |           | S2   |
| Gh119 | 2020 | Bono East     | JASMINE 85                     | 7.5  | -1.6 | 11 | + | partial CP |           | S2   |
| Gh120 | 2020 | Bono East     | CRI-AGRA                       | 7.5  | -1.6 | 10 | + | 2020Gh120  |           | S2   |
| Gh121 | 2020 | Bono East     | JASMINE 85                     | 7.5  | -1.6 | 8  | + | 2020Gh121  |           | S2   |
| Gh122 | 2020 | Ahafo         | CRI-AGRA                       | 7.4  | -2.2 | 12 | + | 2020Gh122  |           | S2   |
| Gh123 | 2020 | Bono          | CRI-AGRA                       | 7.3  | -2.9 | 8  | - |            |           |      |
| Gh124 | 2020 | Ahafo         | JASMINE 85                     | 7.1  | -2.7 | 12 | + | 2020Gh124  | 2020Gh124 | S2   |
| Gh125 | 2020 | Ahafo         | JASMINE 85                     | 6.8  | -2.5 | 12 | + | 2020Gh125  |           | S2   |
| Gh126 | 2020 | Ahafo         | JASMINE 85                     | 6.8  | -2.5 | 11 | + | 2020Gh126  |           | S2   |
| Gh127 | 2020 | Central       | JASMINE 85                     | 5.9  | -1.3 | 12 | + | 2020Gh127  |           | S2   |
| Gh128 | 2020 | Central       | CRI-AGRA                       | 5.9  | -1.3 | 14 | + | 2020Gh128  |           | S2   |
| Gh129 | 2020 | Central       | nd                             | 5.9  | -1.3 | 12 | + | 2020Gh129  |           | S2   |
| Gh130 | 2020 | Western       | Local var.                     | 5.7  | -2.0 | 12 | - |            |           |      |
| Gh131 | 2020 | Western       | Local var.                     | 5.8  | -2.0 | 8  | - |            |           |      |
| Gh132 | 2020 | Western       | CRI-AGRA                       | 5.8  | -2.0 | 13 | - |            |           |      |
| Gh133 | 2020 | Western       | CRI-AGRA                       | 5.8  | -2.0 | 12 | + | 2020Gh133  | 2020Gh133 | S2   |
| Gh134 | 2020 | Greater Accra | CRI-AGRA                       | 5.7  | 0.0  | 8  | - |            |           |      |
| Gh135 | 2020 | Greater Accra | CRI-AGRA                       | 5.7  | -0.1 | 7  | - |            |           |      |
| Gh136 | 2020 | Greater Accra | CRI-AGRA                       | 5.8  | 0.1  | 12 | - |            |           |      |
| Gh137 | 2020 | Greater Accra | JET3                           | 5.8  | 0.1  | 12 | - |            |           |      |
| Gh138 | 2020 | Northern      | nd                             | 9.3  | -1.0 | 6  | + | 2020Gh138  |           | S2   |
| Gh139 | 2020 | Northern      | CRI-AGRA                       | 9.4  | -1.0 | 4  | + | 2020Gh139  |           | S1ca |
| Gh140 | 2020 | Northern      | CRI-AGRA                       | 9.3  | -1.0 | 6  | + | 2020Gh140  | 2020Gh140 | S1ca |
| Gh141 | 2020 | Northern      | Salmasaa<br>Local var.<br>TOGO | 9.6  | -1.0 | 5  | + | 2020Gh141  |           | S2   |
| Gh142 | 2020 | Northern      | MARSHALL                       | 9.6  | -1.0 | 6  | + | 2020Gh142  | 2020Gh142 | S2   |
| Gh143 | 2020 | Upper East    | CRI-AGRA                       | 10.8 | -1.1 | 10 | - |            |           |      |
| Gh144 | 2020 | Upper East    | CRI-AGRA                       | 10.8 | -1.1 | 10 | - |            |           |      |

668

669 **Table S1:** Location, metadata and experimental treatments on rice samples collected in 2020 in  
670 Ghana.

671 Only a portion of the CP gene was amplified and sequenced for the 2020Gh 119 isolates, but the  
672 genetic information was sufficient to confirm that this isolate belongs to the S2 strain. (local var.):

673 local rice variety. nd: not determined.

674

| <b>Isolate</b> | <b>Year</b> | <b>Region</b> | <b>CP</b> | <b>Complete genome</b> |
|----------------|-------------|---------------|-----------|------------------------|
| Gh1            | 1995        | Ashanti       | AJ279919  |                        |
| Gh46           | 2010        | Ashanti       | OQ200610  |                        |
| Gh47           | 2010        | Ashanti       | OQ200611  | OQ225940               |
| Gh48           | 2010        | Ashanti       | OQ200612  |                        |
| Gh49           | 2010        | Ashanti       | OQ200613  |                        |
| Gh50           | 2010        | Ashanti       | OQ200614  |                        |
| Gh51           | 2010        | Ashanti       | OQ200615  |                        |
| Gh52           | 2010        | Greater Accra | OQ200616  |                        |
| Gh53           | 2010        | Greater Accra | OQ200617  |                        |
| Gh54           | 2010        | Greater Accra | OQ200618  |                        |
| Gh55           | 2010        | Greater Accra | OQ200619  | OQ225941               |
| Gh3227         | 2012        | Ashanti       | OQ200620  |                        |
| Gh3229         | 2012        | Ashanti       | OQ200621  |                        |
| Gh3244         | 2012        | Central       | OQ200622  |                        |
| Gh3245         | 2012        | Central       | OQ200623  |                        |
| Gh3387         | 2012        | Ashanti       | OQ200624  |                        |
| Gh3397         | 2012        | Ashanti       | OQ200625  |                        |
| Gh3399         | 2012        | Ashanti       | OQ200626  |                        |
| Gh3416         | 2012        | Ashanti       | OQ200627  |                        |
| Gh101          | 2020        | Ashanti       | OQ200628  |                        |
| Gh102          | 2020        | Ashanti       | OQ200629  |                        |
| Gh103          | 2020        | Ashanti       | OQ200630  |                        |
| Gh105          | 2020        | Ashanti       | OQ200631  |                        |
| Gh106          | 2020        | Ashanti       | OQ200632  |                        |
| Gh109          | 2020        | Ashanti       | OQ200633  |                        |
| Gh111          | 2020        | Volta         | OQ200634  |                        |
| Gh112          | 2020        | Volta         | OQ200635  |                        |
| Gh113          | 2020        | Eastern       | OQ200636  | OQ225942               |
| Gh115          | 2020        | Eastern       | OQ200637  |                        |
| Gh116          | 2020        | Eastern       | OQ200638  |                        |
| Gh117          | 2020        | Eastern       | OQ200639  |                        |

|       |      |           |          |          |
|-------|------|-----------|----------|----------|
| Gh118 | 2020 | Eastern   | OQ200640 |          |
| Gh120 | 2020 | Bono East | OQ200641 |          |
| Gh121 | 2020 | Bono East | OQ200642 |          |
| Gh122 | 2020 | Ahafo     | OQ200643 |          |
| Gh124 | 2020 | Ahafo     | OQ200644 | OQ225943 |
| Gh125 | 2020 | Ahafo     | OQ200645 |          |
| Gh126 | 2020 | Ahafo     | OQ200646 |          |
| Gh127 | 2020 | Central   | OQ200647 |          |
| Gh128 | 2020 | Central   | OQ200648 |          |
| Gh129 | 2020 | Central   | OQ200649 |          |
| Gh133 | 2020 | Western   | OQ200650 | OQ225944 |
| Gh138 | 2020 | Northern  | OQ200651 |          |
| Gh139 | 2020 | Northern  | OQ200652 |          |
| Gh140 | 2020 | Northern  | OQ200653 | OQ225945 |
| Gh141 | 2020 | Northern  | OQ200654 |          |
| Gh142 | 2020 | Northern  | OQ200655 | OQ225946 |

675

676 **Table S2:** Sequence dataset from Ghana and accession numbers.

677