Virus Evolution, 2023, **9(2)**, 1–17
DOI: https://doi.org/10.1093/ve/vead049
Advance access publication date: 21 August 2023
Research Article

Dynamics of the rice yellow mottle disease in western Burkina Faso: Epidemic monitoring, spatio-temporal variation of viral diversity, and pathogenicity in a disease hotspot

Estelle Billard,[1,‡] Mariam Barro,[1,2,§] Drissa Sérémé,[3,**] Martine Bangratz,[1] Issa Wonni,[2,††] Moustapha Koala,[3,‡‡] Abalo Itolou Kassankogno,[2,§§] Eugénie Hébrard,[1,***] Gaël Thébaud,[1,†††] Christophe Brugidou,[1] Nils Poulicard,[1,†,*,‡‡‡] and Charlotte Tollenaere[1,†,*,§§§]

[1]PHIM, Plant Health Institute of Montpellier, Univ. Montpellier, IRD, CIRAD, INRAE, Institute Agro, Montpellier, France, [2]INERA, Institut de l'Environnement et de Recherches Agricoles, Laboratoire de Phytopathologie, Bobo-Dioulasso, Burkina Faso and [3]INERA, Institut de l'Environnement et de Recherches Agricoles, Laboratoire de Virologie et de Biologie Végétale, Kamboinsé, Burkina Faso
[†]Equal contribution.
[‡]https://orcid.org/0009-0008-9745-0169
[§]https://orcid.org/0000-0002-3171-6496
[**]https://orcid.org/0000-0002-2439-5735
[††]https://orcid.org/0000-0003-1891-5447
[‡‡]https://orcid.org/0000-0003-0832-909X
[§§]https://orcid.org/0000-0003-3168-1972
[***]https://orcid.org/0000-0002-2128-1302
[†††]https://orcid.org/0000-0002-2987-4997
[‡‡‡]https://orcid.org/0000-0003-3969-2131
[§§§]https://orcid.org/0000-0003-2045-2824
*Corresponding authors: E-mail: charlotte.tollenaere@ird.fr; nils.poulicard@ird.fr

Abstract

The rice yellow mottle virus (RYMV) is a model in plant virus molecular epidemiology, with the reconstruction of historical introduction routes at the scale of the African continent. However, information on patterns of viral prevalence and viral diversity over multiple years at a local scale remains scarce, in spite of potential implications for crop protection. Here, we describe a 5-year (2015–9) monitoring of RYMV prevalence in six sites from western Burkina Faso (geographic areas of Bama, Banzon, and Karfiguela). It confirmed one irrigated site as a disease hotspot and also found one rainfed lowland (RL) site with occasional high prevalence levels. Within the studied fields, a pattern of disease aggregation was evidenced at a 5-m distance, as expected for a mechanically transmitted virus. Next, we monitored RYMV genetic diversity in the irrigated disease hotspot site, revealing a high viral diversity, with the current coexistence of various distinct genetic groups at the site scale (ca. 520 ha) and also within various specific fields (25 m side). One genetic lineage, named S1bzn, is the most recently emerged group and increased in frequency over the studied period (from 20 per cent or less in 2015–6 to more than 65 per cent in 2019). Its genome results from a recombination between two other lineages (S1wa and S1ca). Finally, experimental work revealed that three rice varieties commonly cultivated in Burkina Faso were not different in terms of resistance level, and we also found no significant effect of RYMV genetic groups on symptom expression and viral load. We found, however, that infection outcome depended on the specific RYMV isolate, with two isolates from the lineage S1bzn accumulating at the highest level at early infections. Overall, this study documents a case of high viral prevalence, high viral diversity, and co-occurrence of divergent genetic lineages at a small geographic scale. A recently emerged lineage, which comprises viral isolates inducing severe symptoms and high accumulation under controlled conditions, could be recently rising through natural selection. Following up the monitoring of RYMV diversity is required to confirm this trend and further understand the factors driving the local maintenance of viral diversity.

Keywords: rice; RYMV; crop virus; epidemiology; spatio-temporal dynamics; diversity; recombination.

Introduction

Biotic constraints drastically affect crop production, with pathogens and pests being responsible for an estimated 17–30 per cent of global yield losses (Savary et al. 2019). Food-deficit regions with fast-growing populations are disproportionately affected, notably Sub-Saharan Africa (Savary et al. 2019), where production losses were estimated to reach over one billion dollars and threaten food security (Sileshi and Gebeyehu 2021). Viruses

cause almost half of the reported plant-emerging infectious diseases (Anderson et al. 2004) and constitute a major challenge to agriculture, with an annual global cost of more than 30 billion dollars (Sastry and Zitter 2014). Such virus-induced yield losses may even increase in the next decades, as a consequence of climate change and agricultural intensification, urging to improve the scientific understanding of epidemiological dynamics of crop viral diseases (Jones and Naidu 2019).

Spatio-temporal studies of crop virus dynamics are of major importance to better understand the drivers of epidemics and to develop appropriate control strategies (McLeish, Fraile, and García-Arenal 2021). In this respect, several studies have successfully used knowledge of small-scale crop virus dynamics to guide disease management methods, as exemplified by the long-term study of viruses infecting cucurbit and solanaceous crops and weeds in southern France (Desbiez et al. 2009, 2020). Virus detection, genetic and pathogenic diversity (including in neighboring weeds), and spatial structure analyses inform for each considered virus the respective key prophylactic and cultural measures to undertake control of insect vectors, weeding, deployment of specific resistance genes, and sanitation of human-assisted plant material exchanges (Desbiez et al. 2009, 2020). For two tomato begomoviruses in Brazil as well, Macedo et al. (2019) revealed that primary spread was the most important mechanism for epidemics of the two common whitefly-transmitted viruses, so that preventing the arrival of viruliferous insects to production areas is the key to manage golden mosaic and yellowing diseases. Finally, a 21-year monitoring of pepper-infecting tobamoviruses in southeastern Spain documented the rapid genetic viral diversification related to the deployment of pepper resistant cultivars (Fraile et al. 2011). Overall, spatial and temporal variation in prevalence and diversity allows us to identify disease hotspots, inoculum sources, and origin of new pathogen variants, to characterize transmission roads, and to identify newly adapted pathogen genotypes that are likely to increase in frequency and spread. All this information is mandatory for science-based deployment of control methods to efficiently reduce virus epidemics and emergence risks (McDonald and Linde 2002; Burdon and Thrall 2008; Picard et al. 2017; Desbiez et al. 2020; Saubin et al. 2022).

Rice (*Oryza spp.*) is of critical importance regarding present and future food security, with more than 3.5 billion people now relying on it for more than 20 per cent of their daily diet (Ahmadi and Bouman 2013). In West Africa, rice consumption experienced a recent 8 per cent increase each year between 2009 and 2019 (Soullier et al. 2020). To face this growing demand, proactive policies led to a rapid increase in rice production with an average annual growth higher than 10 per cent between 2009 and 2019 (Soullier et al. 2020). This increase in rice production is the consequence of an expansion of rice cultivation areas and agricultural intensification. However, intensification of rice cropping may render rice production more vulnerable to disease and pest outbreaks in West Africa. While as much as sixteen viruses were reported worldwide (Cartwright et al. 2018), only two rice viruses have been reported to date in West Africa: *Rice stripe necrosis virus* and, most importantly in terms of yield impact, *Rice yellow mottle virus* (RYMV). RYMV is a *Sobemovirus* (family *Solemoviridae*) endemic of Africa (see Hébrard, Poulicard, and Rakotomalala 2021 for review), mechanically transmitted by insects (mainly species of the Chrysomelidae family; Bakker 1970) or other animals, or during human-mediated agricultural manipulations (Traoré et al. 2009). RYMV causes severe yield losses to rice production (Hébrard, Poulicard, and Rakotomalala 2021) and was consequently listed within the fourteen emerging infectious pathogens threatening food security (Savary et al. 2019) and economies in Africa (Sileshi and Gebeyehu 2021) and in the top ten list of economically important plant viruses (Rybicki 2015).

RYMV is also one of the major models for plant virus molecular epidemiology (Trovão et al. 2015; Pagán, Fraile, and García-Arenal 2016). Continental-scale molecular epidemiology studies were undertaken (Pinel-Galzi et al. 2009, 2015) and compared to RYMV field epidemiology (Rakotomalala et al. 2019; Issaka et al. 2021), revealing a strong geographic structuration of genetic diversity (Pinel-Galzi et al. 2015). Various viral strains may, however, overlap in some regions, as exemplified by the coexistence of several RYMV strains in Côte d'Ivoire (N'Guessan et al. 2000) or at the local scale in Tanzania (Kanyeka et al. 2007). Description of recombinant isolates in East Africa (Pinel-Galzi et al. 2009; Ochola et al. 2015; Adego et al. 2018) suggests a substantial level of viral strain co-occurrence and co-infection, but this remains poorly documented. In addition, compared to country- or continental-scale studies, no study finely described local-scale RYMV epidemics in terms of disease prevalence and genetic diversity (but see Awoderu 1991; Traore et al. 2005). Such local monitoring is, however, the relevant scale for the efficient deployment of resistant cultivars.

In Burkina Faso, yellow mottle disease was first reported in 1981 (John, Thottapilly, and Awoderu 1984; Salaudeen et al. 2010) and research on this virus began with an intensive sampling in 1994–5 (Konate, Traore, and Coulibaly 1997). Yellow mottle disease was known by the farmers from western Burkina Faso: reported as the most serious rice disease by farmers from ten villages of the Cascades Region (Kam et al. 2013) and as a major concern for rice cultivation in the irrigated perimeter of Banzon (Traoré et al. 2015). Burkina Faso is located at the crossroads of RYMV dispersal roads in West Africa (Dellicour et al. 2018) and is one of the countries where various RYMV strains have been reported (S1wa, S1ca, and S2; Pinel-Galzi et al. 2015), with potential contrasting resistance-breaking capacities (considering the S1ca pathotype T' able to evolve and infect resistant cultivars under controlled conditions and reported in neighboring countries; Hébrard et al. 2018). In western Burkina Faso, Barro et al. (2021a) monitored four major rice diseases in six study sites (three irrigated areas and neighboring RLs) over 4 years (2016–9), following an initial survey of viral and bacterial diseases in a selection of sites in 2015 (Tollenaere et al. 2017). Over the 2016–9 period, yellow mottle disease symptoms were observed in fifteen-four fields over the 179 observations (30.2 per cent of visited rice fields). The interaction between the rice production system (irrigated area vs RL) and the geographical zone (three zones distant from each other from 40 to 90 km) was highly significant to explain yellow mottle occurrence. Indeed, the irrigated area of Banzon and, to a lesser extent, the RL of Bama (Badala village) were identified as yellow mottle disease hotspots, while the four other sites presented much fewer symptomatic plants (Barro et al. 2021a). Banzon is a 520-ha irrigated perimeter located approximately 60 km west of the city of Bobo-Dioulasso. As a consequence of the creation of the irrigation scheme (initiated in 1973 and finished in 1981), the population has increased tenfold from 1975 to 1985 and the practice of fallow disappeared (Toe 1992).

The present study complements the previous study of multiple rice diseases in western Burkina Faso by Barro et al. (2021a), based on symptom observation only. We focus here specifically on the rice yellow mottle disease and use serological detection and molecular work to address the following questions that have never been addressed for RYMV: what is the spatio-temporal dynamics of the disease at a regional scale and can we identify drivers of infection at local (field) and regional scales? What is the level of

viral genetic diversity at the local scale and does it vary through time? Do various viral strains coexist within the same irrigated area or even the same field? If so, do these strains differ in terms of pathogenicity on susceptible rice host? Globally, this study aims to monitor the dynamics of epidemics and characterize the drivers of infection and genetic diversity for the major viral threat of rice, a staple food far from self-sufficiency in West Africa.

Material and methods

Longitudinal sampling in selected farmer's fields in western Burkina Faso

Symptom observations and sampling were performed between 2015 and 2019 in six sites of western Burkina Faso (Bobo-Dioulasso area). Only two of these sites (Karfiguela and Banzon) were studied in 2015. Then, annually from 2016 to 2019, we visited three irrigated areas and three RLs located nearby the irrigated sampled sites (within three geographical zones: Bama, Banzon, and Karfiguela; see the map in Fig. 1). The sampling scheme thus presents six sites in total and a total of 203 (24–50 per year) observations over the 2015–9 period (Barro et al. 2021a). These visits were always performed at the maximum tillering or heading initiation, growth stage (September–early December). In each selected field (25 × 25 m), a 4 × 4 grid was delimited and symptoms associated with multiple pathogens were assessed visually in the four cells of the diagonal of the grid (Barro et al. 2021a). Yellow mottle symptoms were scored from 0 (no symptom observed) to 100 per cent (all plants symptomatic) in the four diagonal cells of each studied field. In addition, the sixteen plants located at the grid nodes were inspected for symptoms and sampled (three leaves per plant put in an envelope and dried using silica gel). In every case, we obtained permission from the farmers to work (observations and leaves sampling) in their fields. We interviewed most farmers to get data on the cultivar and agricultural practices for each studied field, as we also asked questions to document constraints to rice production, farmer's knowledge of rice diseases, and control methods. These farmer's interview data were analyzed in Barro et al. (2021a) mostly to contrast the major agricultural practices used in each rice growing system. Further details on the methodology and all the data (symptom observation as well as farmer's interviews) for each field studied every year are available in DataSuds (https://doi.org/10.23708/8FDWIE).

Enzyme-linked immunosorbent assay detection of RYMV in collected samples

In all cases of observation of specific yellow mottle disease symptoms in a field, the sixteen sampled plants were analyzed by RYMV-specific serological detection. Each plant sample consisted on one 1-cm long leaf sample from each of the three different leaves sampled (3 cm equivalent). Serological detection was performed on these samples, using a previously described Double Antibody Sandwich Enzyme Linked Immunosorbent Assay (DAS-ELISA) protocol with the RYMV-Mg antiserum (Ndjiondjop et al. 1999; Traoré et al. 2008). Absorbance at 405 nm was read using a Metertech Σ960 microplate reader in the PathoBios detection platform (Kamboinse, Burkina Faso) for 2015–7 samples or a Tecan Spark Multimode plate Reader in the IRD-Genetrop platform (Montpellier, France) for 2018–9 samples. Following Traoré et al. (2008), samples were considered as positive if the optical density (OD) obtained was higher than the average of negative samples plus three times the standard deviation. A constant threshold of OD = 0.3 was also used to compare the obtained results to a more conservative analysis (see also N'Guessan et al. 2000).

RYMV-specific ELISA data analysis

We used the R software (R Core Team 2018) and the package 'ggplot2' (Wickham 2016) to visualize prevalence data (frequency of infection within fields) in the different sites and years. The relationship between prevalence estimates based on DAS-ELISA data and symptom-based prevalence data was tested using a Spearman correlation test (cor.test function).

Generalized Linear Mixed Models (GLMMs) were performed on DAS-ELISA data using the R package 'lme4' (Bates et al. 2015). Because our sampling design corresponds to visiting the same fields annually, we included the particular site in each year as a random effect in the models. Model selection was performed using Akaike information criterion criteria, and final models were checked for overdispersion and using the package 'DHARMa' (Hartig 2017). Post hoc tests were performed using the package 'lsmeans' (Lenth 2016).

First, we tested for an effect of the rice growing system, the geographical zone, and the interaction between these two variables (that correspond to each particular site) on RYMV prevalence. The response variable was the prevalence based on serological detection (binomial variable of the number of ELISA-positive samples and number of ELISA-negative samples in each field). Next, we tested for an effect of agricultural practices on yellow mottle disease prevalence. Because previous analyses revealed a very strong effect of the site, we performed this analysis on a subset of data: all observations performed within the irrigated perimeter of Banzon (Banzon-IR), revealed as a yellow mottle disease hotspot by the aforementioned analyses. Three agricultural practices presented variability within the Banzon-IR dataset: rice cultivar (four cultivars), off-season cultivation (rice or nothing), and the number of urea applications (no or only one application, compared to two or more applications). However, off-season cultivation and the number of urea applications were highly correlated, so off-season cultivation was only kept in the model. GLMMs were again performed on the binomial variable of 'RYMV prevalence', with the agricultural variables ('rice cultivar' and 'off-season cultivation') and their interaction as fixed explanatory variables.

Finally, we tested whether RYMV presence/absence data were spatially aggregated at the within-field (25 × 25 m) level. This analysis was performed for the RYMV symptom observation dataset, as well as for the DAS-ELISA detection dataset. The same methodology as described in Thébaud et al. (2005) was used. The test compares the observed distribution of distances between infected plants to the simulation under H_0 (random distribution of infected plants, i.e. no spatial aggregation). This test was performed over the fields presenting at least two infected plants (twenty-nine cases for the symptoms and thirty-three cases for DAS-ELISA).

Sanger sequencing of coat protein gene and viral genome

Genetic analysis of collected samples was performed in the irrigated area of Banzon only and particularly for samples collected within five fields (distant from ca. 400 m to 2.2 km), chosen for their high RYMV prevalence over the five sampling years.

Amplification of the coat protein (CP) gene (*ORF4*) was performed by reverse transcription (RT)-Polymerase Chain Reaction (PCR) as described earlier (Pinel et al. 2000; Fargette et al. 2004), using the primer RYMV-M 5′-CGCTCAACATCCTTTTCAGGGTAG-3′for RT and the primer pair RYMV-III 5′-CAAAGATGGCCAGGAA-3′ and RYMV-M for the PCR. Amplified products were sent to Genewiz (Leipzig, Germany) for Sanger sequencing with the same primer pair used for the PCR. A 720-bp sequence was obtained

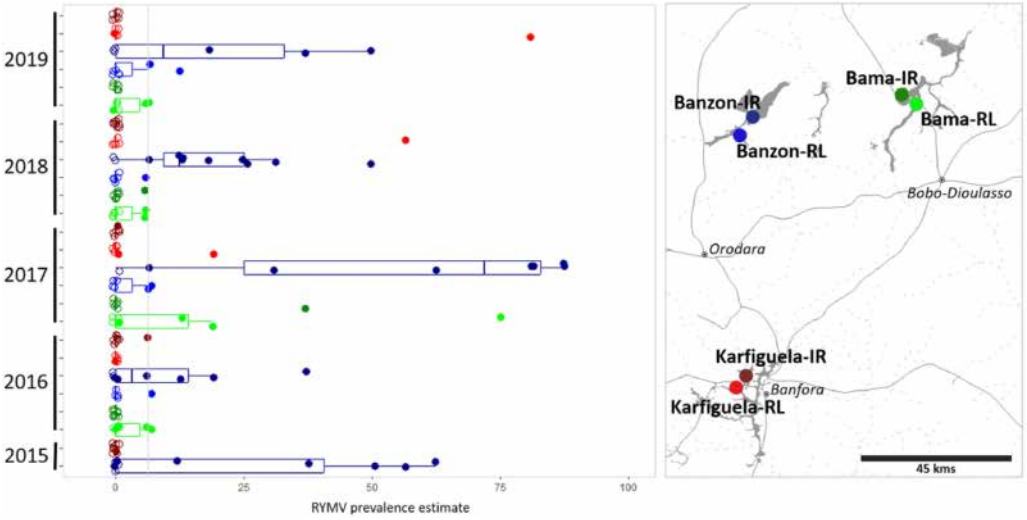

Figure 1. RYMV prevalence data over the different sites in western Burkina Faso within the 2015–9 period. All filled dots correspond to RYMV prevalence estimates derived from specific serological detection of RYMV in sixteen plants analyzed in each field where yellow mottle disease symptoms were observed. Empty dots correspond to the fields where no yellow mottle disease symptoms could be observed, ELISA tests were not performed in these cases, and the prevalence was set to zero. The function 'geom_jitter' (adding a small amount of random variation to the location of each point) was used to decrease the overlap between points. A light gray line shows the case of one positive plant found out of sixteen (6.25 per cent prevalence).

for 132 isolates, and alignment was performed using MEGA-X software.

We performed full-length genome sequencing for three representative isolates, using the approach described in Fargette et al. (2004). Briefly, a RT was performed using the primer RYMV-II 5′-CTCCCCCACCCATCCCGAGAATT-3′ and two PCRs were performed with the two pairs of primers: RYMV-R1 5′-CAATTGAAGCTAGGAAAGGAG-3′ with RYMV-14bis 5′-ACTTCG-CCGGTTTCGCAAAGGATT-3′ (expecting product size of 2,399 pb) and RYMV-2136 5′-CATGCTGGGAAAAGTGTCTG-3′ with RYMV-II (expecting product size of 2,315 pb). Four pairs of additional primers were used for sequencing four genome fragments (Fargette et al. 2004), using the Sanger technique (Genewiz, NJ, USA).

The obtained sequences were deposited in GenBank (accession numbers from OQ716828 to OQ716971 for ORF4 and OQ858571, OQ858572, and OQ858573 for complete genomes).

Genetic diversity, phylogenetic analyses, and detection of recombination events

The ORF4 sequence dataset and the three full-length genome sequences obtained from Banzon (MP1401, MP1493, and MP3582) were compared to those available in National Center for Biotechnology Information (cf. Issaka et al. 2021), i.e. 261 CP gene sequences and thirty-seven complete genome sequences from isolates collected in West and Central Africa.

Multiple sequences were aligned using MUSCLE (Edgar 2004) implemented in SEAVIEW v4.7 (Gouy, Guindon, and Gascuel 2010). We reconstructed Maximum-Likelihood (ML) phylogenetic trees with SEAVIEW using the best-fitting nucleotide substitution models (K2 + G + I and GTR + G + I for the CP gene and the complete genome datasets, respectively) determined with MEGA-X (Kumar et al. 2018) and 100 bootstrap replicates. Phylogenetic trees were drawn using FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). In addition, we used the median-joining network method implemented in the Network software (version 4.611; www.fluxus-engineering.com) (Bandlet, Forster, and Röhl 1999) to reconstruct the minimum spanning network (MSN) connecting all the CP sequences (at the amino acid level) obtained in this study.

Genetic diversity of the CP gene dataset was estimated per year, field, and rice cultivar using the K2 + G + I substitution model, with standard errors of each measure based on 100 bootstrap replicates, as implemented in MEGA-X. The difference in nucleotide diversity of the virus populations among groups (i.e. year, field and rice cultivars) was tested by analysis of molecular variance (AMOVA), as implemented in Arlequin v. 5.3.1.2 (Excoffier, Laval, and Schneider 2005). AMOVA calculates the F_{ST} index explaining the between-group fraction of total genetic diversity. Significance of these differences was obtained by performing 1,000 permutations.

Potential recombination signals from complete RYMV genome sequences or the CP gene dataset were searched using the pairwise homoplasy test (Pairwise Homoplasy Index test) using SplitsTree v4.18.3 (Huson and Bryant 2006) and the seven algorithms implemented in the RDP program v4.97 (Martin et al. 2015). Only recombination events detected by five or more methods and with *P*-values above 10^{-5} were considered.

Bayesian evolutionary inferences and discrete phylogeographic analyses

In order to examine the degree of temporal signals in the CP gene dataset (261 CP gene sequences from West and Central Africa and available in public database in January 2023 and 132 from this study), we first used an exploratory linear regression approach (Duchêne et al. 2015; Murray et al. 2016). Based on the ML phylogenetic tree reconstructed from the CP gene dataset, the temporal signal of this sequence dataset was visualized and tested in PhyloStemS (Doizy et al. 2023) to regress phylogenetic root-to-tip distances against the sampling date using the root that minimized the residual mean squares. In addition, the significance of the temporal signal was evaluated by a date-randomization test. Thus, the mean rate and its 95 per cent highest probability density (HPD) estimated with the observed sampled dates using the Bayesian Evolutionary Analysis Sampling Trees (BEAST) v1.10.4

package (Suchard et al. 2018, see later) were compared with a null distribution obtained by randomly permutating the tip dates ten times (Firth et al. 2010). As previously described (Duchêne et al. 2015; Murray et al. 2016), the criterion for a significant temporal signal was that the 95 per cent HPD for the rate estimate obtained with the observed sampled dates should not overlap with the 95 per cent HPD for the estimate obtained with randomized sampling times.

The discrete phylogeographic reconstructions were generated on the CP gene dataset using a Bayesian statistical framework implemented in BEAST v1.10.4 (Suchard et al. 2018) and the Broad-platform Evolutionary Analysis General Likelihood Evaluator library (Ayres et al. 2012) to improve computational performance. BEAST uses Markov chain Monte Carlo (MCMC) integration to average over all plausible evolutionary histories for the data, as reflected by the posterior probability. The nucleotide substitution process was reconstructed using the HKY85 + G substitution model, the lognormal relaxed clock model, and an initial substitution rate of 10^{-3} substitutions/site/year as previously determined for RYMV (Trovão et al. 2015; Dellicour et al. 2018; Issaka et al. 2021).

Among the RYMV sequences from the literature, only six are from the same geographic area, western Burkina Faso (collection date from 1990 to 2000), and were compiled with the sequences obtained in this study. Discrete phylogeographic inferences were estimated at the geographical level (western Burkina Faso 138 sequences, vs the rest of West and West-Central Africa, 255 sequences) using the continuous-time Markov chain process (Lemey et al. 2009) and with a Bayesian stochastic search variable selection. This method reconstructs the dispersal history between discrete locations and infers a posterior distribution of trees whose internal nodes are associated with an estimated ancestral location. MCMC analyses were run for one billion iterations, sampling every 100,000th and 10 per cent iterations discarded as the chain burn-in. The maximum clade credibility (MCC) tree was obtained using TreeAnnotator v1.10.4 (BEAST package), and convergence and mixing properties (e.g. based on effective sample sizes >200 for the parameters) were inspected using Tracer v1.7.1 (http://tree.bio.ed.ac.uk/software/tracer).

Statistical analysis of the spatio-temporal distribution of genetic groups

Contingency tables of genotype frequencies were built using the R software, and pie charts were mapped on an aerial view of the irrigated perimeter of Banzon using QGIS (QGIS 3.28, Geographic Information System; QGIS Association, http://www.qgis.org). To test whether the distribution of genotypes (frequencies of each genetic group) varied spatio-temporally (over the five sampling years and five fields), we carried out multinomial regressions with the 'nnet' package (Venables and Ripley 2002). We tested the interaction between the factor 'year' and the factor 'field' by comparing (using a likelihood ratio test in the 'Anova' function of package 'car'; Fox and Weisberg 2019) the complete model (including the main effects and their interaction) with the model without the interaction and then tested each explanatory factor.

Phenotypic evaluation of RYMV isolates under controlled conditions

Two experiments were conducted under controlled greenhouse conditions (IRD Montpellier, France) with cycles of 12 h of light at 28°C and 80 per cent relative humidity (RH) and 12 h of dark at 26°C and 70 per cent RH. Five plants per pot were grown in pots, filled with potting soil and appropriate fertilizers; soil moisture was checked throughout the experiments. RYMV isolates used in these experiments were sequenced for the CP protein (as described earlier), and construction of the ML phylogenetic tree (as earlier) allowed them to be assigned to the previously identified genetic lineages.

The first experiment (Experiment 1) aimed to assess the behavior of four viral isolates, each representing one identified genetic group, on various rice cultivars. Three rice cultivars (*Oryza sativa* subsp. *indica*) commonly grown in Banzon were selected: FKR19, FKR62N, and FKR64 (Barro et al. 2021a). We also included in the experiment the reference cultivars: IR64 (*O. sativa* subsp. *indica*) and Azucena (*O. sativa* subsp. *japonica*), as, respectively, susceptible (Albar et al. 2006) and partially resistant (Ndjiondjop et al. 1999). Three viral isolates collected at a local scale in Southwestern Burkina Faso from 2015 to 2019 were chosen within each of the identified genetic lineages of RYMV based on CP gene sequencing: BF706, BF710, and BF711. The isolate BF1, from the genetic lineage S2, was included as a reference severe isolate (Pinel et al. 2000). Plants were grown in 1-l pots. and the experiment was performed using forty-five plants per rice cultivar for each of the five conditions (four viral isolates tested and the virus-free control). Pots were arranged in two randomized blocks, and the whole experiment was subdivided into three successive monthly replicates.

The second experiment (Experiment 2) aimed to test the behavior of thirteen viral isolates (including those used in the first experiment) from the identified genetic group on susceptible rice cultivar IR64. We used the same isolates as in Experiment 1 (BF706, BF710, and BF711) and also selected three additional viral isolates per genetic group. These nine isolates were sampled during other surveys conducted in Burkina Faso in 2014 (Tollenaere et al. 2017), 2016, and 2021 (unpublished): BF708, BF709, EF0738, BF712, EF0333, EF0744, BF705, EF0619, and EF0780. The BF1 isolate (Pinel et al. 2000) was also included in this experiment. Plants were grown in 3-l pots, and we used forty IR64 plants for each of the fourteen conditions (thirteen viral isolates and the control). Pots were arranged in two randomized blocks, and all the inoculations were performed at a single date.

All viral isolates were first multiplied in the susceptible rice cultivar IR64 to provide sufficient viral particles to be used as inoculum in the experiments. Then, for both experiments, viral isolates were inoculated under controlled conditions at ca. 2.10^{11} copies of viral genomes per plant. For this purpose, each inoculum was prepared by grinding infected leaves in 0.1 M phosphate buffer (pH 7.2) (Pinel-Galzi et al. 2018). Then, RYMV-specific quantitative RT-PCR (qRT-PCR) assay was used to adjust the viral loads for each inoculum, as described in Poulicard et al. (2010) with the following modifications: amplifications were carried out using 10 µl of Takyon™ No Rox 2X SYBR MasterMix blue dTTP (Eurogentec, Angers, France) and performed on a LightCycler 480™ Roche (Roche Diagnostics, Meylan, France). Finally, plants were mechanically inoculated 2 weeks after sowing, using 50 µl of inoculum per leaves for each RYMV isolate, or inoculated with phosphate buffer alone (virus-free control), as described in Pinel-Galzi et al. (2018).

Infected plants were monitored for visual symptoms beginning at 9 days post-inoculation (dpi) and every 3–4 days until 24 dpi. Yellow mottle symptoms appeared on newly developed leaves after 10 days at the earliest and were scored according to the symptom severity scale described in Pinel-Galzi et al. (2018). In order to estimate the intra-plant viral load for each condition, we sampled meristem from one infected plant per pot early in the infection process, at 7 dpi (see also Tollenaere et al. 2017). This specific time

point was selected as it aligns with the rapid systemic spread of the virus (Poulicard et al. 2010) and allows optimal detection of variations in viral accumulation between isolates personal communication (pers. com.). Then, we performed a qRT-PCR assay as described previously. The number of RYMV copies was estimated using a standard curve obtained with RNA from purified virus (BF5 isolate, S1wa strain).

Experimental infection data analysis

The overall aggressiveness of viral isolates on plants was evaluated by considering the severity of symptoms and viral load. For each inoculated plant, symptoms at 9, 13, 16, 21, and 24 dpi were used to compute the area under the disease progression curve (AUDPC) using the R package 'agricolae' (De Mendiburu and Yaseen 2020).

Results obtained for each of the three replicates of the first experiment were analyzed jointly for statistical analysis of AUDPC and viral load. The R package 'ggplot2' was used to visualize the effect of the virus on symptom expression (AUDPC) and viral load (RYMV copy number at 7 dpi expressed in logarithm) for each RYMV isolate on the different rice cultivars.

Both experiments were analyzed with GLMM using R package 'lme4'. For the first experiment, the effects of the variables 'RYMV isolates', 'rice cultivars', and their interaction were tested independently on the AUDPC and the viral load (number of RYMV copies at 7 dpi). Inoculation date was included in the models as a random factor. For the second experiment, we tested the effect of the variables 'RYMV isolates' and 'genetic group' on the variables AUDPC and viral load, including the variable 'block' as a random factor. Post hoc tests were performed using the R package 'lsmeans'.

Results

Yellow mottle disease prevalence varies spatio-temporally in western Burkina Faso

A total of 203 observations were performed between 2015 and 2019, with the number of studied fields per year varying between twenty-four and fifty over in the six sites (Table 1). Over these 203 observations, the total number of cases where yellow mottle disease symptoms were observed was sixteen two, representing a frequency of 30.5 per cent over the whole area and period. In each studied site and year, the frequency of fields where yellow mottle disease symptoms were found varied from 0 to 87.5 per cent (Table 1), with a maximum of seven fields with symptoms out of eight studied fields in Banzon in 2017. Detailed results obtained for yellow mottle disease symptoms, observed in the four cells of the diagonal of the grid, on the one hand, and in the sixteen plants of grid nodes, on the other hand, are available on DataSuds (https://doi.org/10.23708/GZCM1O).

Serological analysis was performed for the sixteen plants collected in the sixteen-two fields where yellow mottle disease symptoms were observed, resulting in a dataset of 990 samples (62 × 16 and two missing data). We observed a total of 228 positive plants (global seroprevalence: 23.0 per cent) when applying a constant threshold (OD = 0.3) (see also N'Guessan et al. 2000), while 353 plants (global seroprevalence: 35.7 per cent) were DAS-ELISA-positive when applying a threshold following Traoré et al. (2008). Comparing serological results with the presence/absence of observed symptoms led us to choose the constant threshold for subsequent analysis (all figures and main tables). Indeed, with the approach of Traoré et al. (2008), we obtained 107 DAS-ELISA-positive plant where no symptom could be observed (30.3 per cent of DAS-ELISA-positive plants) and sixteen negative symptomatic plants (1.6 per cent of total analyzed plants). On the other hand, with a constant threshold, we obtained only twenty-eight ELISA-positive plants where no symptom could be observed (12.3 per cent of DAS-ELISA-positive plants) and twenty negative symptomatic plants (2.0 per cent of total analyzed plants). Statistical analyses were also, however, performed with the approach of Traoré et al. (2008), and the general conclusions did not change (only a few differences in contrasts analyses, see Supplementary Table S1).

In the fields where we did not observe any yellow mottle disease symptom, we assumed that the disease was absent and set prevalence to zero. The distribution of RYMV prevalence based on DAS-ELISA results over the six studied sites and the 5-year period is presented in Table 1 and Fig. 1. The relationship between prevalence estimates based on DAS-ELISA data and symptom-based prevalence data is shown in Supplementary Fig. S1. It revealed highly significant correlations, both when symptoms are observed on the same sixteen plants (cor = 0.97; $P < 0.001$) and when symptoms are observed in the four diagonal grid cells (cor = 0.78; $P < 0.001$). The interaction between the rice production system and the geographical zone, which reflect the particular site (among six sampled sites), had a highly significant effect ($P < 0.0001$) on detection results (Supplementary Table S1).

For all five studied years, the site presenting the highest RYMV prevalence is the irrigated perimeter of Banzon (Banzon-IR, Fig. 1, Table 1, and Supplementary Fig. S2). In the Banzon-IR site, 36.9 per cent of studied fields presented yellow mottle symptoms, with the average prevalence up to 13.3 per cent, over the 2015–9 period (Table 1). It is the only site with median prevalence different from zero (in 4 years out of 5, see Table 1), and it differs significantly from all other sites, except the RL of the Bama zone (Bama-RL, Supplementary Table S1). RYMV prevalence in Bama-RL is significantly different only from the irrigated perimeter of Karfiguela (Supplementary Table S1a). This site occasionally presents quite high frequency of infected fields and average prevalences (Table 1 and Supplementary Fig. S2), with an opposite effect of the rice growing system in this geographical zone of Bama (more epidemics in RL), while the opposite was found for the Banzon geographical zone (Supplementary Table S1a).

Agricultural practices affect yellow mottle disease prevalence in a disease hotspot

GLMM performed to explain the RYMV prevalence within the irrigated perimeter of Banzon revealed a significant effect of the interaction between the rice cultivar and off-season cultivation ($P = 0.016$, Supplementary Table S1b). Contrast analysis revealed that in the fields where rice was grown in the previous year, highest RYMV prevalence was found for cultivar FKR34 compared to FKR62N, which has higher prevalence than FKR64 (Supplementary Fig. S3, Supplementary Table S1b). Both the double rice season (compared to only one rice cultivation per year) and various fertilization applications (two or more, as compared to one or less) correspond to higher RYMV prevalence (Supplementary Fig. S3).

Yellow mottle disease aggregation is observed within small rice fields

Within the small (20 × 20 m) studied fields, there were 15 per cent more pairs of infected plants in the first distance class (5 m) than expected under the null hypothesis of random distribution of the infection, both for observed yellow mottle disease symptoms

Table 1. Frequency of studied fields presenting yellow mottle disease symptoms and median and average prevalence in the six sites studied from 2015 to 2019 in western Burkina Faso.

	Number of fields with yellow mottle disease symptoms observed (%) Median and average prevalences based on DAS-ELISA results					
Site	2015	2016	2017	2018	2019	All years (2015-19)
Bama-RL	NA	4/6 (66.7%)	4/8 (50.0%)	2/7 (28.6%)	3/6 (50.0%)	13/27 (48.1%)
(Badala)		0–2.1%	0–13.3%	0–1.8%	0–2.1%	0–5.3%
Bama-IR	NA	0/10 (0%)	1/8 (12.5%)	1/9 (11.1%)	0/7 (0%)	2/34 (5.9%)
(Bama)		0–0%	0–4.7%	0–0.7%	0–0%	0–1.3%
Banzon-RL	NA	1/5 (20.0%)	2/7 (28.6%)	1/7 (14.3%)	2/7 (28.6%)	6/26 (23.1)
(Senzon)		0–1.3%	0–1.8%	0–0.9%	0–2.7%	0–1.7%
Banzon-IR	4/12 (58.3%)	6/8 (75.0%)	7/8 (87.5%)	9/11 (81.8%)	3/6 (50%)	32/45 (71.1%)
(Banzon)	0–18.2%	3.1–9.4%	71.9–54.7%	12.5–17.6%	9.4–17.7%	12.5–22.9%
Karfiguela-RL	NA	1/5 (20.0%)	2/7 (28.6%)	1/7 (14.3%)	2/6 (33.3%)	6/25 (24.0%)
(Tengrela)		0–0%	0–2.7%	0–8.0%	0–13.5%	0–6.25%
Karfiguela-IR	1/12 (8.3%)	1/9 (11.1%)	1/8 (12.5%)	0/9 (0%)	0/8 (0%)	3/46 (6.5%)
(Karfiguela)	0–0%	0–0.7%	0–0%	0–0%	0–0%	0–0.1%
Total	8/24 (33.3%)	13/43 (30.2%)	17/46 (36.9%)	14/50 (28%)	10/40 (25%)	62/203 (30.5%)
(all sites)	0–9.1%	0–2.3%	0–13.3%	0–5.5%	0–5.5%	0–7.0%

The number of fields where yellow mottle disease was observed are mentioned over the total number of rice fields where symptom observations were performed for each year and site. Median and average prevalences are based on the RYMV-specific DAS-ELISA test in all studied fields (including both fields with and without yellow mottle disease symptoms observed). The different sites are grouped by geographical zones, see Fig. 1. IR: irrigated areas; NA: Not applicable.

Table 2. List of the five fields studied from 2015 to 2019 within the irrigated area of Banzon with the estimated RYMV prevalence and number of plants for which CP sequence was obtained.

	RYMV prevalences Number of plant samples with CP sequence in each field and year					
Field	2015	2016	2017	2018	2019	Total (average all years)
BZ02	62.5%	6.2%	0%	12.5%	0%	16.2%
	5	3	0	4	0	12
BZ04	12.5%	0%	87.5%	25.0%	50.0%	35.0%
	4	0	6	4	16	30
BZ09	37.5%	37.5%	87.5%	50.0%	0%	42.5%
	6	5	5	9	0	25
BZ10	56.2%	18.7%	81.2%	12.5%	18.7%	37.5%
	10	3	4	10	10	37
BZ11	50.0%	0%	81.2%	6.2%	37.5%	35.0%
	4	4	4	1	15	28
Total (all fields)	18.2%	9.4%	54.7%	17.6%	17.7%	22.9%
	29	15	19	28	41	132

RYMV prevalences are based on the RYMV-specific serological test (DAS-ELISA). The last line (Total, all fields) refers to all studied fields in the irrigated area of Banzon (not only the five fields shown here).

($P < 0.001$) and for serological RYMV detection ($P < 0.001$). This pattern of aggregation between infected plants remains significant ($P < 0.05$) even within a 10-m radius.

The disease hotspot harbors high genetic diversity

A total of 132 CP gene sequences (clear chromatograms) were generated from samples collected in Banzon between 2015 and 2019, with the number of samples per field and year varying between zero and sixteen (average: 5.3 ± 4.3; Table 2). In addition, six samples (4.3 per cent of the dataset) were amplified but resulted in unclear chromatograms, possibly as a result of mixed infection by various isolates.

Interestingly, the genetic diversity of the 132 CP gene sequences observed in Banzon ($d = 0.0391 \pm 0.0051$ nucleotide substitutions per site) is similar to those estimated for the whole Burkina Faso ($d = 0.0449 \pm 0.0054$ sub/site) and is higher than those from most countries in West- and West-Central Africa (Supplementary Fig. S4), which could reflect a complex or old epidemiological history of RYMV in this region.

Various historical introductions shaped RYMV viral diversity in western regions of Burkina Faso

Based on the CP dataset collected in Banzon in this study (132 CP) and in West- and West-Central Africa in the literature (261CP, including some sequences from several regions of Burkina Faso), we reconstructed the evolutionary history of RYMV to gain information on the epidemic history of this virus in southwestern Burkina Faso. First, we established that no recombination event ($P = 0.93$) was detected within the CP dataset (393 sequences in total). Then, the strength of the temporal signal was examined on this CP dataset by linear regression exploration of root-to-tip distances and by date-permutation tests. The linear regression exploration showed a low but significant temporal signal in the 393 CP dataset ($R^2 = 0.087$; $P = 1.62 \times 10^{-9}$; Supplementary Fig. S5). The presence of temporal signals was

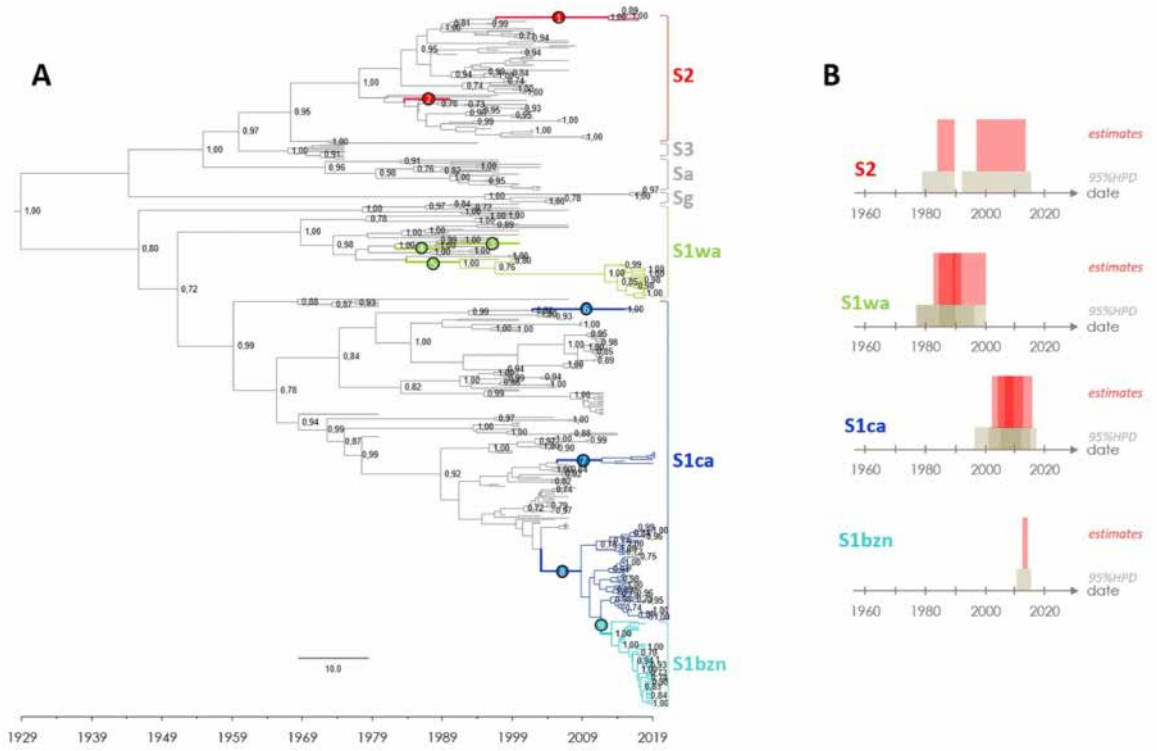

Figure 2. (A) Time-calibrated MCC tree reconstructed by continuous Bayesian evolutionary inference with the 393 CP gene sequence dataset. Branches corresponding to the isolates collected in western Burkina Faso are colored according to the RYMV strains (S1wa: green; S1ca: blue; S1wa × S1ca recombinant line named S1bzn: light blue; S2: red). The values on the nodes correspond to their posterior probabilities (values above 0.70 are indicated). The bottom axis gives the timeframe (by date) of the RYMV diversification. Nine branches of interest are labeled on the MCC tree (from 1 to 9), and their relative nodes (1a, 1b, 2, 3, 4, 5a, 5b, 6a, 6b, 7a, 7b, 8a, 8b, 9a, 9b) were annotated with their time to the most common ancestor and their 95 per cent HPD intervals (Table 3). (B) Dates (in red) and the 95 per cent HPD intervals (in gray) of the introduction/emergence of the S2, S1wa, S1ca, and recombinant line S1bzn in western Burkina Faso estimated as the intervals between nodes of interest (Table 3). The darker areas represent the overlaps between the estimates.

confirmed with the date-randomization tests (tip-date randomization: Supplementary Fig. S6A and B; clustered tip randomizations: Supplementary Fig. 6C and D) using BEAST, as none of the estimates of the permutations overlapped with the ones of the real (non-permutated) 393 CP dataset (Supplementary Fig. S6).

We reconstructed the discrete dispersal histories of the RYMV in West- and West-Central Africa by Bayesian inferences. We obtained a MCC tree where the tip locations (western Burkina Faso vs the rest of West- and West-Central Africa) were mapped (Fig. 2A). Eight independent introductions of RYMV to western Burkina Faso were estimated ($MarkovJumps_{Introduction} = 8.11$ [8.00; 9.00]; branches 1–8 indicated in Fig. 2A), two corresponding to S2 strain introductions (branches 1 and 2; posterior probabilities $PP_{introduction1} = 0.99$ and $PP_{introduction2} = 1.00$), three to S1wa strains (branches 3–5; $PP_{introduction3} = 0.99$, $PP_{introduction4} = 1.00$, $PP_{introduction5} = 0.98$), and three to S1ca strains (branches 6–8; $PP_{introduction6} = 1.00$, $PP_{introduction7} = 1.00$, $PP_{introduction8} = 1.00$). We also distinguish Branch 9 that corresponds to a group named S1bzn. The date of emergence of each RYMV strain in western Burkina Faso could be estimated based on the date intervals of the nodes of branches of interest. The introductions of S2 and S1wa strains in this region were estimated between the late 1980s and the 2000s/2010s (Fig. 2B, Table 3). The introduction of the S1ca strain probably occurred later, from the 2000s to the mid-2010s (Fig. 2B, Table 3). Finally, the emergence of the group S1bzn was estimated to be in the beginning of the 2010s (Fig. 2B, Table 3). As there is no regular and continuous sampling in time and space of RYMV isolates in Burkina Faso, we cannot exclude that the emergence of Lineage 9 corresponds to an introduction from another region of Burkina Faso instead of a recombination event that occurred within western Burkina Faso following the co-circulation of S1wa and S1ca strains.

Note that no migration of RYMV populations from western Burkina Faso to other regions of West- and West-Central Africa was reported using this dataset ($MarkovJumps_{Release} = 0.11$ [0.00; 1.00]). Among S2 and S1wa genetic groups, some lineages may have gone extinct; successful introductions in western Burkina Faso are attested by their detection in Banzon during the 2015–9 period only for Branch 1 within S2 lineage and for Branch 5 in S1wa lineage (Fig. 2).

Co-circulation of viral genetic groups, with temporal variation throughout the sampling period

The MSN connecting all the CP sequences (at the amino acid level) from Banzon is presented in Fig. 3. While the genetic differentiation is low between fields ($F_{ST(field)} = 0.0562$; $P = 0.0134$) and rice cultivars ($F_{ST(rice)} = 0.0474$; $P = 0.0098$), it is highly significant between years ($F_{ST(years)} = 0.2796$; $P < 0.0001$, Fig. 2). Notably, the genomes identified in 2018 and 2019 were significantly different from the genomes sampled in the other years, suggesting wide variations on RYMV populations circulating in Banzon during the last two sampling years.

Table 3. Introduction dates of RYMV strains in western Burkina Faso.

Node number	Time	Time 95%HPD	Date	Date 95%HPD	Genetic group
1a	22.71	[27.79–18.18]	1997.12	[1992.04–2011.65]	S2
1b	6.59	[9.05–4.53]	2013.24	[2010.78–2015.30]	
2	35.86	[40.97–31.29]	1983.97	[1978.86–1988.54]	
3	30.94	[35.24–26.59]	1988.89	[1984.59–1993.24]	S1wa
4	37.04	[42.40–32.51]	1982.79	[1977.43–1987.32]	
5a	35.32	[43.07–27.82]	1984.51	[1976.76–1992.01]	
5b	27.88	[33.22–23.48]	1991.95	[1986.61–1996.35]	
6a	17.90	[23.34–14.29]	2001.93	[1996.49–2005.54]	S1ca
6b	4.48	[6.85–3.05]	2015.35	[2012.98–2016.78]	
7a	13.74	[15.02–12.92]	2006.09	[2004.81–2006.91]	
7b	7.53	[10.07–5.13]	2012.30	[2009.76–2014.70]	
8a	15.77	[18.97–13.43]	2004.06	[2000.86–2006.40]	
8b	10.65	[13.80–8.23]	2009.18	[2006.03–2011.60]	
9a	7.56	[9.64–6.07]	2012.27	[2010.19–2013.76]	S1bzn
9b	5.91	[7.39–4.79]	2013.92	[2012.44–2015.04]	

Date intervals for nine branches of interest (labeled from 1 to 9 in Fig. 2A) and their relative nodes (1a, 1b, 2, 3, 4, 5a, 5b, 6a, 6b, 7a, 7b, 8a, 8b, 9a, 9b), annotated with their time to the most common ancestor and their 95 per cent HPD intervals. The columns "Time" indicate the time before present (2019,83).

Three genetic groups are distinguished on the network (Fig. 3) and named based on the phylogenetic tree including reference isolates from West- and West-Central Africa (Fig. 2A): S1ca, S1wa, and S2. A fourth group, corresponding to Lineage 9 in the phylogenetic tree, is labeled S1bzn. This group presents a clear temporal structure with genotypes that progressively diverge from the rest of the network through time (Fig. 3).

Whole-genome sequencing was performed for one isolate from each of the three major genetic groups (S1wa, S1ca, and S1bzn), and the three newly obtained complete sequences were compared to the thirty-seven RYMV genomes described from samples collected in West- and West-Central Africa (Supplementary Fig. S7). For the samples of group S1bzn (2017MP1493), the software RDP identified one putative recombination event, with a CP from the S1ca group, while the rest of the genome corresponds to the S1wa group (Supplementary Fig. S8). In accordance with this detected recombination event, we obtained contrasted phylogenies for the CP sequences, as compared to the rest of the genome (Supplementary Fig. S7).

Contingency tables of the four genetic groups are presented in Supplementary Table S2, and pie charts are represented in each field and year (Fig. 4). Various genetic groups were found within the same field, with up to three genetic groups coexisting at the same sampling date (Fig. 4 and Supplementary Table S2). Multinomial regressions showed no effect of the interaction between the year and the field, and no geographical structure within the irrigated perimeter (factor 'field' not significant). On the other hand, a strong effect of the factor 'year' is evidenced ($\chi^2 = 95.2$; $df = 12$; $P < 0.001$). Temporal variations in the frequency of genetic groups are presented in Fig. 5. The overall frequencies over the 5 years are 44.6 per cent for the genetic group S1ca, 37.1 per cent for S1bzn, 15.2 per cent for S1wa and only 3.0 per cent for S2. The three groups S1wa, S1ca, and S1bzn coexist throughout the 5-year period with fluctuations in time, while the S2 strain was only found in 2015 and 2017. The S1wa strain was rare in 2015 and absent in 2016, 2017, and 2019; however, it was the most frequent lineage in 2018 (64 per cent). A decrease in S1ca frequency is observed over the period (from 76 per cent in 2015 to 32 per cent in 2019), whereas S1bzn was rare in 2015 and 2016 (14–20 per cent) but seemed to rise in 2019 (up to 68 per cent). Finally, the overall (2015–9) genetic diversity in the studied site is 0.039 ± 0.005 subsite and varies in time with a maximum reached in 2017 (0.047 ± 0.005 subsite) and 2018 (0.042 ± 0.005 subsite) and a minimum in 2019 (0.015 ± 0.003 subsite; Fig. 5).

Symptoms and viral load are affected by the rice cultivar, the viral isolate, and their interactions

Detailed results obtained for all experimental infections performed are available on DataSuds (https://doi.org/10.23708/MDQIUE). The phylogenetic positions of the different isolates used in experiments are presented in Supplementary Fig. S9 and Supplementary Table S3: isolates BF706, EF0738, BF708, and BF709 belong to the S1wa strain. BF710, EF0744, BF712, and EF0333 are identified as S1ca, and EF0619, EF0780, BF705, and BF711 as S1bzn.

First, in order to characterize the pathogenicity of the major viral genetic groups identified, we evaluated the symptom severity and intra-plant viral accumulation of four RYMV isolates from each genetic group identified (S2, S1wa, S1ca, and S1bzn) on the rice cultivars commonly grown in Burkina Faso (FKR19, FKR62N, and FKR64) and two control varieties (susceptible rice cultivar IR64 and partially resistant cultivar Azucena) (Supplementary Fig. S10).

GLMM performed to explain the symptom severity, on the one hand, and viral load, on the other hand, revealed in both cases an effect of the rice cultivar, the RYMV isolate, and the interaction between these two factors (Table 4). Tukey post hoc tests revealed that IR64 and Azucena differed from other rice cultivars for both symptom severity and viral load ($P < 0.0001$; Supplementary Table S4). Highest symptoms and viral load were observed for the four RYMV isolates inoculated on IR64, while the lowest symptoms and viral load were observed on the rice cultivar Azucena. In contrast, a common intermediate pattern was observed for the three cultivars from Burkina Faso (FKR19, FKR62N, and FKR64), for both symptoms and viral load. Indeed, Tukey post hoc tests revealed no significant differences between the three rice cultivars for symptoms and viral load (Supplementary Table S4).

For the five rice cultivars tested, we observed similar patterns of variability in aggressiveness according to RYMV isolates (Supplementary Figure S10). First, the BF1 isolate (S2 strain) induced strong symptoms as isolate BF706 (S1wa strain) and BF710 (S1ca strain) and accumulated to the highest level in all rice cultivars. In contrast, isolate BF711 (from lineage S1bzn) was the least aggressive isolate, with mildest symptoms and

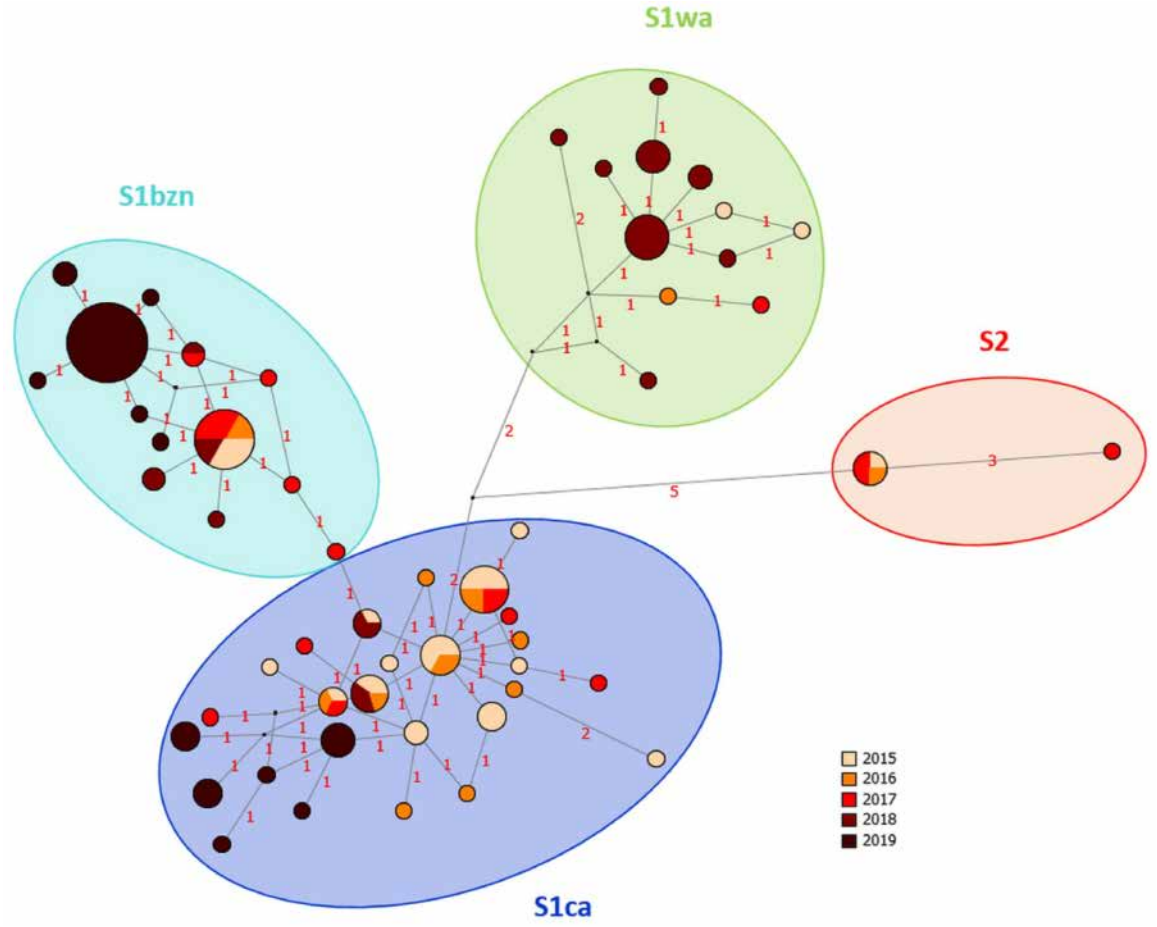

Figure 3. MSN of the 132 CP sequences (at the amino acid level) identified in Banzon. Each CP sequence is indicated by one node in which the circle area is proportional to the number of individual sequences for this particular sequence, and the year of sampling of these sequences is represented as a proportional pie chart. The number of amino acid substitutions between sequences is indicated on the branches. Seven putative intermediate sequences connecting the CP sequences correspond to the small red dots.

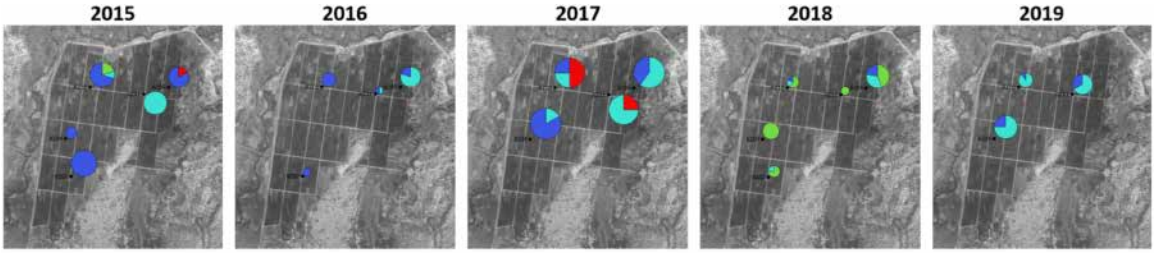

Figure 4. RYMV prevalence and genetic diversity (CP sequences) follow-up data over the five selected quadrats within the irrigated area of Banzon between 2015 and 2019. The colors represent the four identified genetic groups: S2 in red, S1wa in green, S1ca in dark blue, and the recombinant S1bzn in light blue (see also Supplementary Table S2). The surface of each pie chart represents the RYMV prevalence estimate based on serological data (see also Table 2). The only exception is for BZ11 in 2016, where not any of the sixteen plants analyzed were positive in DAS-ELISA, while a few symptomatic plants were observed and sequenced (Table 2), and we set the prevalence to five in this field to be able to represent the genetic groups identified.

a low level of accumulation in all rice cultivars. Tukey honest significant difference (HSD) tests, all presented in Supplementary Table S4, showed highly significant differences between isolate BF711 and all isolates, for both symptoms and viral load. Although no differences between isolate BF1 (S2), BF706 (S1wa), and BF710 (S1ca) were observed for symptoms, an intermediate level of accumulation in all rice cultivars was observed for isolate BF706 (S1wa) and BF710 (S1ca), which were different from BF1 (S2) and BF711 (S1bzn) (see Tukey HSD tests; Supplementary Fig. S4).

Isolates from the S1bzn group induced severe symptoms and high viral load

A second experiment was performed to evaluate potential variability in pathogenicity within genetic groups. For this purpose, we evaluated on the susceptible rice cultivar IR64 the experimental behavior of three additional isolates for each of three genetic groups (S1wa, S1ca, and S1bzn). GLMM results revealed that the genetic group was not a determinant of either symptom severity or viral load; only the specific RYMV isolate has an effect on symptom severity and viral load (Table 4). Fig. 6 shows the ranking of all

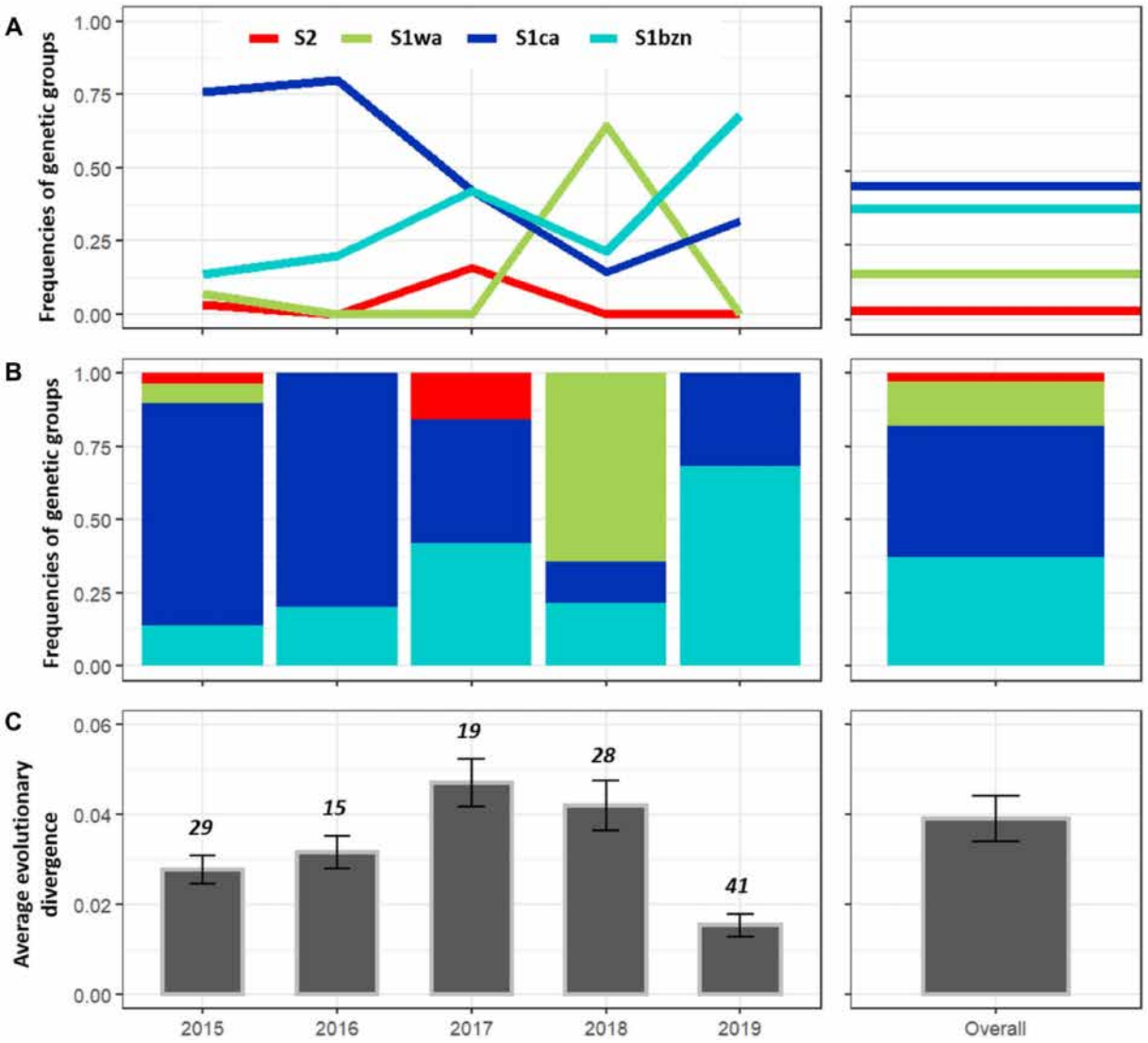

Figure 5. Temporal variation of RYMV genetic diversity (CP region) over the 2015–9 period in the irrigated perimeter of Banzon. (A and B) Overall frequencies of the four genetic groups in Banzon for the years 2015, 2017, 2018, and 2019. The four genetic groups are represented by four different colors with S2 in red, S1wa in green, S1ca in dark blue, and S1bzn in light blue. (C) Estimates of average evolutionary divergence over sequence pairs within groups. The number of base substitutions per site from averaging over all sequence pairs within each group is shown. Standard error estimate(s) are shown above the diagonal. Analyses were conducted using the Kimura two-parameter model. The right panel labeled 'Overall' represents the whole dataset of 132 CP sequences sampled over the 2015–9 period.

Table 4. Statistical analyses of the experimental infection results: estimations of symptom severity and viral load in a set of four isolates and five cultivars (Experiment 1) and in a set of thirteen isolates on rice cultivar IR64 (Experiment 2).

	AUDPC			Viral load		
Model (GLMM)	Chisq	*df*	Pr(>Chisq)	Chisq	*df*	Pr(>Chisq)
Experiment 1						
Rice cultivar	362.732	4	$<2.2 \times 10^{-16}$	627.336	4	$<2.2 \times 10^{-16}$
Isolate	34.637	3	1.5×10^{-7}	91.508	3	$<2.2 \times 10^{-16}$
Rice cultivar:Isolate	25.792	12	0.011	38.776	12	0.0001
Experiment 2						
Genetic group	6.293	3	0.098	7.598	3	0.055
Genetic group:Isolate	1245.945	9	$<2.0 \times 10^{-16}$	215.785	9	$<2.2 \times 10^{-16}$

Results of the GLMM for Experiment 1 including 'Inoculation date', 'Block', 'Rice cultivar*Isolate' and for Experiment 2 including 'Block' and 'Genetic group/Isolate' are reported. GLMM results on symptom severity (integrated over the symptom development: AUDPC) are shown on the left and on viral load (estimated from specific qPCR on the meristematic area at 7 dpi) on the right, for each experiment.

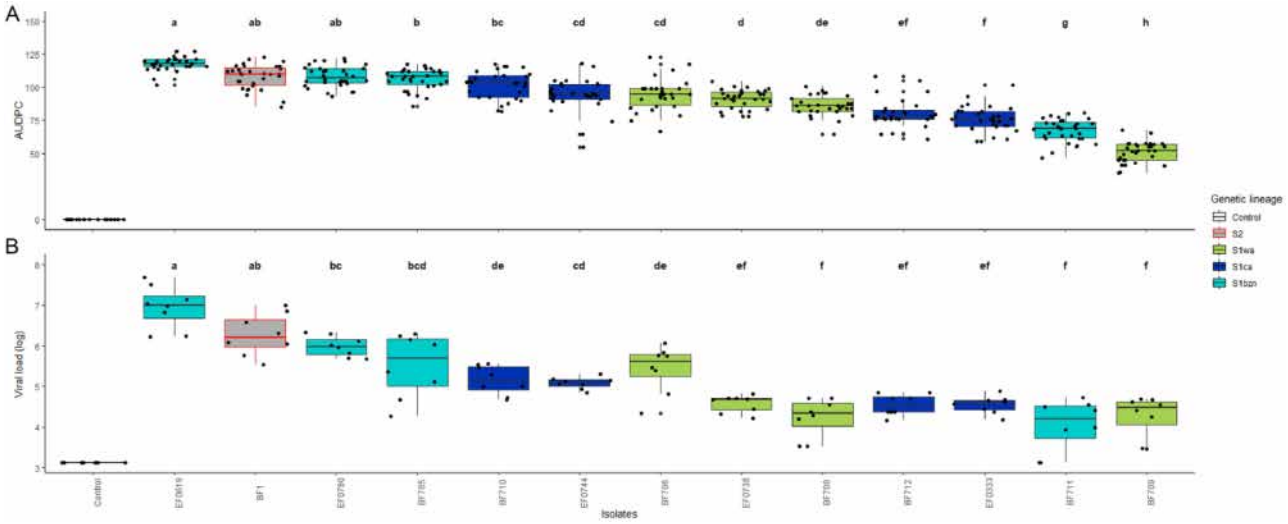

Figure 6. Symptom estimate and viral load of thirteen RYMV isolates, from the three genetic groups identified in Banzon (S1wa, S1ca, and S1bzn), evaluated on a susceptible rice cultivar. (A) The symptom severity estimated by the area under disease progress curve (AUDPC) of thirteen RYMV isolates inoculated on the rice cultivar IR64. AUDPC was determined according to the notation severity scale over five dates (from 9 to 24 dpi), on thirty-two phenotyped plants. (B) The intra-plant viral load of the thirteen RYMV isolates estimated on the rice cultivar IR64. The number of RYMV copies was estimated by RT-qPCR on eight meristematic area collected at 7 dpi, using a standard curve. Isolates are ordered according to their aggressiveness (highest to lowest AUDPC values). The thirteen RYMV isolates are represented by four different colors corresponding to their genetic group, S1wa (green), S1ca (dark blue), S1bzn (light blue), and S2 (red box) corresponding to the reference severe isolate BF1 (gray). Letters indicate the significant differences between each isolate (Tukey honestly significant difference, $P < 0.05$).

tested isolates, in terms of aggressiveness and viral accumulation (also see the contrast analysis in Supplementary Table S4).

For the four isolates tested in the first experiment, we obtained the same pattern of infection outcomes as in Experiment 1: BF1 (S2) being the most aggressive and BF711 (S1bzn) the least aggressive. In addition, this second experiment showed that the isolate causing the highest symptoms was EF0619 from the S1bzn group. Both this isolate and EF780 also from the S1bzn group accumulated in IR64 at the highest level than isolates from other groups (see Tukey HSD tests; Supplementary Table S4). The genetic group S1bzn also contains another aggressive isolate (BF705), so that out of four isolates tested from this group, three induce strong symptoms and accumulate at high levels.

Discussion

Pathogens offer the unique opportunity to study contemporaneous evolution and monitor 'evolution-in-action' (McDonald 2010), a fascinating context for evolutionary biology. Understanding pathogen evolution is also a priority for public health, as well as food security, especially for virus diseases affecting subsistence agriculture in developing countries (Jones and Naidu 2019). In this context, we document here the spatio-temporal variation in prevalence, as well as genetic and pathogenic diversity, of the rice yellow mottle disease at a local scale in western Burkina Faso.

Serological detection of RYMV performed in all the studied fields where yellow mottle disease symptoms were observed, in a survey of six sites over 5 years, confirmed previous results based on symptoms: the rice growing system is not a structuring factor for the rice yellow mottle disease in western Burkina Faso (Barro et al. 2021a). Indeed, the irrigated perimeter of Banzon is identified as the most frequently affected by the disease in this region and where prevalence is higher. It confirms this site as a hotspot of yellow mottle disease as previously noticed (Traoré et al. 2015; Barro et al. 2021a). In addition, one RL site (Bama-RL) occasionally experiences quite a high frequency of infected fields and prevalences, but much less than Banzon-IR. This geographical pattern of yellow mottle disease is consistent with farmers' interviews data. Indeed, rice diseases were mentioned among the major threats to rice production in twelve cases out of eighty-three (14.5 per cent) over the six study sites, seven of which (58 per cent) were from the irrigated perimeter of Banzon. Among the diseases known by the farmers, yellow mottle disease was cited in twenty-five cases out of eighty-two (30.5 per cent), eighteen (72 per cent) of which were from the irrigated perimeter of Banzon. Consequently, these results confirm that the site where disease burden is highest corresponds to the site where the farmers recognize yellow mottle disease, as also shown in a previous study of RYMV in Burkina Faso (Traoré et al. 2015). The disease control methods mentioned by the farmers in the irrigated perimeter of Banzon are (1) uprooting of symptomatic plants, (2) avoiding entering a rice field if symptoms are observed to avoid further plant-to-plant transmission by contact, and (3) shifting rice season later, by sowing in August or even September instead of July (the most likely control mechanism being the desynchronization between rice cultivation and virus sources such as reservoir plants or vectors). Also, various farmers state their incapacity to fight the disease. More work is required to specifically address agronomic and sociological questions related to rice yellow mottle disease in Banzon.

At the within-field level, we found small-scale aggregation in DAS-ELISA-positive plants, reflecting field symptom-based observations (pers. com.). This is congruent with expected major within-field secondary transmission, mechanically, through (1) direct leaf-to-leaf contact in crowded plants; (2) contaminated hands or tools, particularly when transplanting, and applying fertilizers (these agricultural practices also enhancing direct leaf-to-leaf contact); and/or (3) small-scale vector transmission, mostly by insect's biting and chewing mouthparts (Banwo, Alegbejo, and Abo 2004; Traore et al. 2006; Traoré et al. 2009).

Our results show that yellow mottle virus prevalence is affected by agronomic factors, as well as rice cultivars. First, the latter suggests potential differences in terms of resistance in the cultivars used in Banzon. Rice cultivars from Burkina Faso are moderately tolerant to various RYMV isolates experimentally. Three cultivars from Burkina Faso (FKR19, FKR62N, and FKR64), commonly grown in the study sites (Barro et al. 2021a), were found to be susceptible, but less than IR64, in our experiments, which is congruent with the results obtained in the field for FKR19 and FKR64 (Sérémé et al. 2016a) and in semi-controlled experiments for FKR19 and FKR62N (Sérémé et al. 2016b). On the other hand, FKR34, commonly grown in Banzon in 2017 only, is associated with higher RYMV prevalence, which corresponds to the temporal pattern of highest disease burden in 2017. Because FKR34 was not included in our experimental infections, we cannot rule out any of the two following hypotheses: (1) FKR34 is more susceptible than other cultivars grown in Burkina Faso or (2) other factors led to a higher disease pressure in 2017, resulting in spurious association between FKR34 cultivar and disease prevalence. Finally, field data associating cultivar names and RYMV prevalence have to be taken with caution, as they rely on cultivar names given by the farmers, which do not always correspond to genetically well-defined cultivars, according to a recent study performed in the same study sites (Barro et al. 2021b).

Second, we tested for an effect of agronomic factors, namely, off-season cultivation and the number of urea applications, that were actually highly correlated, allowing us to distinguish intensively cultivated fields (various urea applications and two rice seasons) from low (one urea application or less, and only one rice season). We showed that this factor, in interaction with the rice cultivar, affects yellow mottle disease prevalences. This relates to the effect of fertilization on yellow mottle disease (nitrogen effect, Bouet et al. 2012) and the likely increased viral transmission associated with longer rice cultivation period (although yellow mottle disease is much less observed during the dry season; Issaka et al. 2012).

A high genetic diversity of RYMV was found in the studied site, with the coexistence of at least four genetic lineages. Over the whole period, we have three common groups (S1ca: 39.4 per cent, S1wa: 15.2 per cent, and S1bzn: 42.4 per cent) and a rare strain: S2 (3.0 per cent). The genetic diversity in the studied site (2015–9) is 0.039 sub/site. This is a very high level of diversity considering the small geographical scale of this study (the irrigated perimeter of Banzon covers ca. 520 ha), compared to the estimates for the whole country (Burkina Faso: 0.045 sub/site) and for West Africa (0.063 sub/site). Various West African countries have lower levels of diversity than this peculiar site, and although some sampling bias (in terms of not only the number of samples but also the geographic distribution of sampling within the countries) may limit this comparison, this reflects the impressive RYMV diversity at a local scale in this hotspot site.

A recent study showed a relationship between the genetic composition of pathogen populations and disease intensity, with higher infection rates in wild plant populations infected by fungal pathogen populations harboring higher strain diversity (Eck et al. 2022). Following this, the high prevalence of yellow mottle disease in Banzon fields may somehow be a consequence of this very high RYMV viral diversity, but this remains speculative, especially as we have no information on the viral genetic diversity in the five other sites with lower disease pressure. The conditions that lead to yellow mottle disease outbreaks are still poorly understood (Traoré et al. 2009) and deserve further research.

Interestingly, the various lineages coexist at the rice field scale, with up to three distinct lineages detected simultaneously in delimited areas of around 625 m^2. Co-occurrence of various lineages at a very local scale is particularly interesting as it is a favorable context for mixed infection (multiple infection of a rice plant by various RYMV genetic lineages, see later), a prerequisite for recombination.

Our results show that at least eight introductions of RYMV occurred in western Burkina Faso including the irrigated perimeter of Banzon, enhancing the viral diversity locally. Indeed, phylogeographic reconstructions suggest that strains S1wa and S2 were introduced first in this region, for some lineages as early as the 1980s and globally before 2000. These two groups were known in the region (Pinel-Galzi et al. 2015). On the other hand, the group S1ca was known more in the West (Pinel-Galzi et al. 2015). Although frequently found in the beginning of our survey (75 per cent or more of the samples in 2015–6), the S1ca group was most likely introduced later than S2 and S1wa, probably after 2001. The irrigated perimeter of Banzon was built in early 1980s (Toe 1992), as well as others in the region (Nebie 1993). In the last decades, Burkina Faso experienced rice cultivation intensification (yield doubled between the 1960s and the 1980s, Food and Agriculture Organization (FAO 1997)), followed by an increase in areas cultivated in rice (fourfold increase between the period before and after 2008, (FAO 1997)). Increased exchanges of material, and possibly highest probability of new introduction, are likely associated with these recent changes in rice cultivation. Within one irrigated site, we found the coexistence of genetic groups introduced various decades ago (S2 and S1wa) and others more recently emerged. Specifically, a recombination event between S1wa and S1ca involving the CP coding region, either occurring locally (in the Banzon site) or elsewhere (but most presumably in the area of western Burkina Faso), led to the recent emergence of a lineage (S1bzn), maintained over the 5-year period. Other studies on crop viruses have documented the emergence of recombinant groups with variable outcomes. For example, a recombinant strain of pepino mosaic virus found in 2004 may be extinct as it was not detected again since 2005 (Alcaide et al. 2020), while for tomato yellow leaf curl virus (TYLC), a recombinant strain displaced its parental viruses in southern Morocco (Belabess et al. 2015). The success of the recombinant strain of TYLC in Morocco is likely the consequence of a selective advantage, as it exhibits greater within-host viral accumulation in a resistant cultivar that had been recently deployed at the time of emergence (Belabess et al. 2016).

The source of these introductions may be the neighboring RLs. However, we notice that the studied site in the neighborhood (Banzon-RL = Senzon village) experiments a low level of virus circulation. Alternatively, long-distance migration events may explain observed patterns, in accordance with the average dispersal rate of 13 km/year estimated for RYMV in West Africa (Trovão et al. 2015). Putative transmission events may involve insect vectors flying long distances or the transport of infective tissue (leaf fragments or empty rice spikelets) associated with farmers' seed exchanges (Banwo, Alegbejo, and Abo 2004). A significant level of exchange of infected material between farmers from distant areas in Burkina Faso is expected considering that other rice pathogens exhibit a low level of spatial structure in this country (Kaboré et al. 2022). Control measures would include a more careful control of rice material imports to prevent introduction of new pathogen strains. The high genetic diversity of RYMV in Banzon illustrates the increased viral circulation from the 1980s that gradually blurs the strong geographical structure

of RYMV documented at the continental level (Traore et al. 2005; Dellicour et al. 2018).

Although little geographic structure was found within the studied irrigated perimeter, we detected genetic differentiation between the 5 sampling years and temporal variation in frequencies of the genetic groups. A recently introduced group, S1bzn, seems to increase in frequency over the studied period, from 14 to 20 per cent in 2015–6 to 68 per cent in 2019. On the other hand, the strain S1ca was particularly frequent at the beginning of the survey (76–80 per cent in 2015–6) and decreases in frequency to only 14 per cent in 2018 and 32 per cent in 2019. The two other lineages (S1wa and S2), which correspond to the lineages introduced the earliest, do not exhibit any trend (either regular increase or decrease) over this 5-year survey. Instead, the S1wa lineage was only found in 2015 (but very rare, 3 per cent of the samples) and 2018, where it was surprisingly the most common group, totaling 64 per cent of the samples. Although this group was not found in 2019, we hypothesize that it is not extinct as it was also absent from 2016 and 2017 surveys but then very common in 2018. The S2 strain is very rare over the whole survey (3 per cent of the samples) and was not found in the last two surveyed years, either because it was not sampled by chance or because it was actually absent these years and maybe extinct.

The turnover rate of RYMV genetic groups over the study period raises questions on the underlying processes, including annual founding effects or selective pressures. To investigate the selection hypothesis, we characterized as a fitness proxy the outcome of infection by viral isolates belonging to the three genetic groups (S1wa, S1ca, and S1bzn) under controlled conditions. The experimental results showed that induced symptoms and viral load varied between isolates, while the genetic group was not a determinant of infection outcome. These results are congruent with other studies in plant viruses, where the genetic groups are not homogeneous in terms of pathogenicity (N'Guessan et al. 2000; Fraile et al. 2011).

However, interestingly, we found that two isolates from S1bzn genetic lineage resulted in highest viral loads, compared to all isolates from other groups tested. A higher viral load would be in agreement with the hypothesis of a selective advantage of the lineage S1bzn. Such an effect of natural selection is supported by the overall increase in this lineage through years (from 20 per cent or less in 2015–6, to more than 65 per cent in 2019). However, this has to be taken with caution considering the absence of the effect of the genetic group on the infection outcome. In addition, our experimental fitness estimates may have been too simple and considering other fitness components such as the outcome of mixed infections by various RYMV strains would be interesting in further experiments. Indeed, in Ivory Coast, although no difference between two strains could be detected in mono-infections, the S2 strain of RYMV from Ivory Coast dominates the S1 strain in experimental mixed infection (N'Guessan et al. 2000). The S2 strain is very prevalent in other neighboring countries such as Ivory Coast (N'Guessan et al. 2000) and Ghana (Omiat et al. 2023), compared to our study in Burkina Faso, so that we could speculate that environmental factors differentially affect RYMV genetic groups in these coastal countries compared to Sahelian areas. Again connected with the results in N'Guessan et al. (2000), we notice that multiple RYMV infections within the same plant may be quite common in western Burkina Faso, as we obtained mixed Sanger chromatograms for 4.3 per cent of sequenced samples. Long-term coexistence of two viral strains may occur in spite of differential fitness in single infection, provided that the same strain is not favored both in mono-infections and in mixed infections, as exemplified for TYLC in La Reunion (Péréfarres et al. 2014). As a result, multiple infections by distinct viral genetic groups can significantly alter the outcomes of infections and dynamics of disease epidemics and deserve to be considered in further research on this pathosystem.

Innovations in science and technology, including outbreak forecasting, constitute a promise for improving the management of viral epidemics, but this requires detailed understanding of small-scale epidemiology. For the RYMV, Hébrard et al. (2018) built a resistance-breaking risk map at the scale of the African continent, based on the geographical repartition of the capacity of isolates to evolve and infect cultivars with resistance genes and alleles under controlled conditions. This work also identified a hypervirulent pathotype within the S1ca group, able to overcome all the sources of high resistance identified in rice. Our work focusing on the local scale (region, site, and field level) in western Burkina Faso illustrates the complexity of durable resistance deployment, as multiple genetic groups, likely presenting different resistance-breaking abilities, coexist in the same site. This highlights the importance to track the different RYMV pathotypes with innovative and effective detection tools. Furthermore, it urges to continue the development of new resistances, including (1) the combination of multiple resistance genes into a single plant genotype (pyramiding) and (2) the introduction of innovative ways to create new resistance. This monitoring will be followed up as they evidence the relevance of this pathosystem to address major current questions for the ecology and evolution of plant viruses, including the impact of multiple infections and recombination in viral evolution (Lefeuvre et al. 2019).

Supplementary data

Supplementary data is available at Virus Evolution online.

Acknowledgements

The authors thank the facilities of the 'International joint Laboratory LMI PathoBios: Observatory of plant pathogens in West Africa: biodiversity and biosafety' (www.pathobios.com; twitter.com/PathoBios). Sylvain Zougrana, Kader Guigma, Edouard Kabore, Yacouba Kone, Martial Kabore, Jean-Noël Ouattara, Moumouni Traoré, and Roméo Dabiré contributed to the sampling in western Burkina Faso. We are grateful to the rice farmers from all the study sites (Badala, Bama, Banzon, Karfiguela, Senzon, and Tengrela) for their collaboration. We also thank Laurence Albar and Agnès Pinel-Galzi for constructive discussions, as well as Emilie Thomas, Manon Perez, and Pascaline Boyer for help in the experimental work.

Funding

This work was publicly funded through ANR (the French National Research Agency) under the program 'Investissements d'avenir' with the reference ANR-10-LABX-001-01 Labex Agro (VIRBIARF, E-Space and RiPaBIOME projects), coordinated by Agropolis Fondation under the frame of I-SITE MUSE (ANR-16-IDEX-006), the project EVCOPAR (ANR-20-CE35-0004-01), as well as the CGIAR Research Program on Rice Agri-food Systems (RICE).

Conflict of interest: None declared.

References

Adego, A. K. et al. (2018) 'Full-Length Genome Sequences of Recombinant and Nonrecombinant Sympatric Strains of Rice Yellow Mottle Virus from Western Kenya', *Genome Announcements*, 6: e01508–17.

Ahmadi, N., and Bouman, B. (2013) 'Riz Et Rizicultures, Enjeux Économiques, Écologiques Et Scientifiques', *Cahiers Agricultures*, 22: 333–5.

Albar, L. et al. (2006) 'Mutations in the eIF(iso)4G Translation Initiation Factor Confer High Resistance of Rice to Rice Yellow Mottle Virus', *The Plant Journal : For Cell and Molecular Biology*, 47: 417–26.

Alcaide, C. et al. (2020) 'Long-Term Cocirculation of Two Strains of Pepino Mosaic Virus in Tomato Crops and Its Effect on Population Genetic Variability', *Phytopathology*, 110: 49–57.

Anderson, P. K. et al. (2004) 'Emerging Infectious Diseases of Plants: Pathogen Pollution, Climate Change and Agrotechnology Drivers', *Trends in Ecology and Evolution*, 19: 535–44.

Awoderu, V. A. (1991) 'The Rice Yellow Mottle Virus Situation in West Africa', *Journal of Basic Microbiology*, 31: 91–9.

Ayres, D. L. et al. (2012) 'BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics', *Systematic Biology*, 61: 170–3.

Bakker, W. (1970) 'Rice Yellow Mottle Virus a Mechanically Transmissible Virus Disease of Rice in Kenya', *Netherlands Journal of Plant Pathology*, 76: 53–63.

Bandlet, H.J., Forster, P., and Röhl, A. (1999) 'Median-joining Networks for Inferring Intraspecific Phylogenies', *Molecular Biology and Evolution*, 16: 37–48.

Banwo, O. O., Alegbejo, M. D., and Abo, M. E. (2004) 'Rice Yellow Mottle Virus Genus Sobemovirus: A Continental Problem in Africa', *Plant Protection Science*, 40: 26–36.

Barro, M. et al. (2021a) 'Spatiotemporal Survey of Multiple Rice Diseases in Irrigated Areas Compared to Rainfed Lowlands in the Western Burkina Faso', *Plant Disease*, 105: 3889–99.

——— et al. (2021b) 'Assessment of Genetic Diversity of Rice in Registered Cultivars and Farmers' Fields in Burkina Faso', *Crops*, 1: 129–40.

Bates, D. et al., 2015. 'Fitting Linear Mixed-Effects Models Using lme4', *Journal of Statistical Software*, 67: 1–48.

Belabess, Z. et al. (2015) 'Monitoring the Dynamics of Emergence of a Non-canonical Recombinant of Tomato Yellow Leaf Curl Virus and Displacement of Its Parental Viruses in Tomato', *Virology*, 486: 291–306.

——— et al. (2016) 'The Non-canonical Tomato Yellow Leaf Curl Virus Recombinant that Displaced Its Parental Viruses in Southern Morocco Exhibits a High Selective Advantage in Experimental Conditions', *Journal of General Virology*, 97: 3433–45.

Bouet, A. et al. (2012) 'Effet de la Fertilisation Azotée Et Phosphorée Sur le Développement de la Panachure Jaune En Riziculture Aquatique En Côte d'Ivoire', *International Journal of Biological and Chemical Sciences*, 6: 4071–9.

Burdon, J. J., and Thrall, P. H. (2008) 'Pathogen Evolution across the Agro-ecological Interface: Implications for Disease Management', *Evolutionary Applications*, 1: 57–65.

Cartwright, R. D. et al. (2018) *Compendium of Rice Diseases and Pests*, 2nd edn. Diseases and Pests Compendium Series.

Dellicour, S. et al. (2018) 'On the Importance of Negative Controls in Viral Landscape Phylogeography', *Virus Evolution*, 4: vey023.

De Mendiburu, F., and Yaseen, M. (2020), *Agricolae: Statistical Procedures for Agricultural Research* <https://cran.r-project.org/package=agricolae>.

Desbiez, C. et al. (2009) 'Emergence of New Strains of Watermelon Mosaic Virus in South-eastern France: Evidence for Limited Spread but Rapid Local Population Shift', *Virus research*, 141: 201–8.

——— et al. (2020) 'Distribution and Evolution of the Major Viruses Infecting Cucurbitaceous and Solanaceous Crops in the French Mediterranean Area', *Virus Research*, 286: 198042.

Doizy, A. et al. (2023) 'Phylostems: A New Graphical Tool to Investigate Temporal Signal of Heterochronous Sequences at Various Evolutionary Scales', *Bioinformatics Advances*, 3: 1.

Duchêne, S. et al. (2015) 'The Performance of the Date-Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data', *Molecular Biology and Evolution*, 32: 1895–906.

Eck, J. L. et al. (2022) 'Strain Diversity and Spatial Distribution are Linked to Epidemic Dynamics in Host Populations', *American Naturalist*, 199: 59–74.

Edgar, R. C. (2004) 'MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput', *Nucleic Acids Research*, 32: 1792–7.

Excoffier, L., Laval, G., and Schneider, S. (2005) 'Arlequin (Version 3.0): Integrated Software Package Population Genetic', *Evolutionary Bioinformatics*, 1: 47–50.

FAO, Food and Agriculture Organization of the United Nations (1997), *FAOSTAT Statistical Database* (Rome) <http://www.fao.org/faostat>.

Fargette, D. et al. (2004) 'Inferring the Evolutionary History of Rice Yellow Mottle Virus from Genomic, Phylogenetic, and Phylogeographic Studies', *Journal of Virology*, 78: 3252–61.

Firth, C. et al. (2010) 'Using Time-Structured Data to Estimate Evolutionary Rates of Double-Stranded DNA Viruses', *Molecular Biology and Evolution*, 27: 2038–51.

Fox, J., and Weisberg, S. (2019) *An R Companion to Applied Regression*, 3 rd edn. Sage Publications Thousand Oaks: CA.

Fraile, A. et al. (2011) 'Rapid Genetic Diversification and High Fitness Penalties Associated with Pathogenicity Evolution in a Plant Virus', *Molecular Biology and Evolution*, 28: 1425–37.

Gouy, M., Guindon, S., and Gascuel, O. (2010) 'SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building', *Molecular Biology and Evolution*, 27: 221–4.

Hartig, F. (2017), *DHARMa: Residual Diagnostics for Hierarchical (Multi-level/mixed) Regression Models* (R package) <http://florianhartig.github.io/DHARMa/>.

Hébrard, E. et al. (2018) 'Identification of A Hypervirulent Pathotype of Rice Yellow Mottle Virus: A Threat to Genetic Resistance Deployment in West-Central Africa', *Phytopathology®*, 108: 299–307.

Hébrard, E., Poulicard, N., and Rakotomalala, M. (2021) 'Rice Yellow Mottle Virus', in Bamford, D., and Zuckerman, M. (eds.) *Encyclopedia of Virology*, 4 th edn. Vol. 4. pp. 485–90. Oxford: Elsevier.

Huson, D. H., and Bryant, D. (2006) 'Application of Phylogenetic Networks in Evolutionary Studies', *Molecular Biology and Evolution*, 23: 254–67.

Issaka, S. et al. (2012) 'Diagnosis and Importance of Rice Yellow Mottle Disease Epidemics in Niger Republic', *Journal of Applied Biosciences*, 50: 3501–11.

——— et al. (2021) 'Rivers and Landscape Ecology of a Plant Virus, Rice Yellow Mottle Virus along the Niger Valley', *Virus Evolution*, 7: veab072.

John, V. T., Thottapilly, G., and Awoderu, V. A. (1984) 'Occurrence of Rice Yellow Mottle Virus in Some Sahelian Countries in West Africa', *FAO Plant Protection Bulletin*, 32: 86–7.

Jones, R. A. C., and Naidu, R. A. (2019) 'Global Dimensions of Plant Virus Diseases: Current Status and Future Perspectives', *Annual Review of Virology*, 6: 387–409.

Kaboré, K. H. et al. (2022) 'Genetic Diversity and Structure of Bipolaris Oryzae and Exserohilum Rostratum Populations Causing Brown Spot of Rice in Burkina Faso Based on Genotyping-by-sequencing', *Frontiers of Plant Science*, 13: 1022348.

Kam, H. et al. (2013) 'Rice Traits Preferred by Farmers and Their Perceptions of Rice Yellow Mottle Virus (RYMV) Disease in Cascades Region of Burkina Faso', *African Journal of Agricultural Research*, 8: 2703–12.

Kanyeka, Z. I. et al. (2007) 'Distribution and Diversity of Local Strains of Rice Yellow Motile Virus in Tanzania', *African Crop Science Journal*, 15: 201–9.

Konate, G., Traore, O., and Coulibaly, M. M. (1997) 'Characterization of Rice Yellow Mottle Virus Isolatesin Sudano-Sahelian Areas', *Archives of Virology*, 142: 1117–24.

Kumar, S. et al. (2018) 'MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms', *Molecular Biology and Evolution*, 35: 1547–9.

Lefeuvre, P. et al. (2019) 'Evolution and Ecology of Plant Viruses', *Nature Reviews. Microbiology*, 17: 632–44.

Lemey, P. et al. (2009) 'Bayesian Phylogeography Finds Its Roots', *PLoS Computational Biology*, 5: e1000520.

Lenth, R. V. (2016) 'Least-squares Means: The R Package Lsmeans', *Journal of Statistical Software*, 69: 1–33.

Macedo, M. A. et al. (2019) 'Temporal and Spatial Progress of the Diseases Caused by the Crinivirus Tomato Chlorosis Virus and the Begomovirus Tomato Severe Rugose Virus in Tomatoes in Brazil', *Plant Pathol*, 68: 72–84.

Martin, D. P. et al. (2015) 'RDP4: Detection and Analysis of Recombination Patterns in Virus Genomes', *Virus Evolution*, 1: vev003.

McDonald, B. (2010) 'How Can We Achieve Durable Disease Resistance in Agricultural Ecosystems?', *The New Phytologist*, 185: 3–5.

McDonald, B. A., and Linde, C. (2002) 'Pathogen Population Genetics, Evolutionary Potential, and Durable Resistance', *Annual Review of Phytopathology*, 40: 349–79.

McLeish, M. J., Fraile, A., and García-Arenal, F. (2021) 'Population Genomics of Plant Viruses: The Ecology and Evolution of Virus Emergence', *Phytopathology*, 111: 32–9.

Murray, G. G. R. et al. (2016) 'The Effect of Genetic Structure on Molecular Dating and Tests for Temporal Signal', *Methods in Ecology and Evolution*, 7: 80–9.

Ndjiondjop, M. N. et al. (1999) 'The Genetic Basis of High Resistance to Rice Yellow Mottle Virus (RYMV) in Cultivars of Two Cultivated Rice Species', *Plant Disease*, 83: 931–5.

Nebie, O. (1993) 'Les Aménagements Hydro-agricoles Au Burkina Faso: Analyse Et Bilan Critiques', *Journal of Transport Geography*, 83: 123–40.

N'Guessan, P. et al. (2000) 'Evidence of the Presence of Two Serotypes of Rice Yellow Mottle Sobemovirus in Côte d'Ivoire', *European Journal of Plant Pathology*, 106: 167–78.

Ochola, D. et al. (2015) 'Emergence of Rice Yellow Mottle Virus in Eastern Uganda: Recent and Singular Interplay between Strains in East Africa and in Madagascar', *Virus Research*, 195: 64–72.

Omiat, E. G. et al. (2023) 'Genetic Diversity and Epidemic Histories of Rice Yellow Mottle Virus in Ghana', *Virus Research*, 329: 199106.

Pagán, I., Fraile, A., and García-Arenal, F. (2016) 'Evolution of the Interactions of Viruses with Their Plant Hosts', *Virus Evolution Current Research Future Dir*, 127–54.

Péréfarres, F. et al. (2014) 'Frequency-dependent Assistance as a Way Out of Competitive Exclusion between Two Strains of an Emerging Virus', *Proceedings of the Royal Society B: Biological Sciences*, 281: 20133374.

Picard, C. et al. (2017) 'Exploiting Genetic Information to Trace Plant Virus Dispersal in Landscapes', *Annual Review of Phytopathology*, 55: 139–60.

Pinel, A. et al. (2000) 'Molecular Variability of Geographically Distinct Isolates of Rice Yellow Mottle Virus in Africa', *Archives of Virology*, 145: 1621–38.

Pinel-Galzi, A. et al. (2018) 'Protocol for RYMV Inoculation and Resistance Evaluation in Rice Seedlings', *Bio-protocol*, 8: e2863–e2863.

——— et al. (2009) 'Recombination, Selection and Clock-like Evolution of Rice Yellow Mottle Virus', *Virology*, 394: 164–72.

——— et al. (2015) 'The Biogeography of Viral Emergence: Rice Yellow Mottle Virus as a Case Study', *Current Opinion in Virology*, 10: 7–13.

Poulicard, N. et al. (2010) 'Why Rice Yellow Mottle Virus, a Rapidly Evolving RNA Plant Virus, Is Not Efficient at Breaking Rymv1-2 Resistance', *Molecular Plant Pathology*, 11: 145–54.

Rakotomalala, M. et al. (2019) 'Comparing Patterns and Scales of Plant Virus Phylogeography: Rice Yellow Mottle Virus in Madagascar and in Continental Africa', *Virus Evolution*, 5: vez023.

R Core Team (2018), *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria) <https://www.R-project.org/>.

Rybicki, E. P. (2015) 'A Top Ten List for Economically Important Plant Viruses', *Archives of Virology*, 160: 17–20.

Salaudeen, M. T. et al. (2010) 'Current Status of Research on Rice Yellow Mottle Sobemovirus', *Archives Of Phytopathology And Plant Protection*, 43: 562–72.

Sastry, K. S., and Zitter, T. A. (2014) 'Management of Virus and Viroid Diseases of Crops in the Tropics', in Sastry, K. S. (ed.) *Plant Virus and Viroid Diseases in the Tropics*, pp. 149–480. Springer.

Saubin, M. et al. (2022) 'Improving Sustainable Crop Protection Using Population Genetics Concepts', *Molecular Ecology*, 32: 2461–2471.

Savary, S. et al. (2019) 'The Global Burden of Pathogens and Pests on Major Food Crops', *Nature Ecology and Evolution*, 3: 430–9.

Sérémé, D. et al. (2016a) 'Screening Improved Rice Cultivars (*Oryza Spp*) for Their Resistance/Tolerance to Rice Yellow Mottle Virus in West Africa', *International Journal of Agriculture Innovations and Research*, 2319–1473.

——— et al. (2016b) 'Assessment of Yield Losses Due to Rice Yellow Mottle Virus under Field Conditions in Burkina Faso', *International Journal of Current Advanced Research*, 5: 1522–8.

Sileshi, G. W., and Gebeyehu, S. (2021) 'Emerging Infectious Diseases Threatening Food Security and Economies in Africa', *Global Food Security*, 28: 100479.

Soullier, G. et al. (2020) 'The State of Rice Value Chain Upgrading in West Africa', *Global Food Security*, 25: 100365.

Suchard, M. A. et al. (2018) 'Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10', *Virus Evolution*, 4: vey016.

Thébaud, G. et al. (2005) 'Investigating Disease Spread between Two Assessment Dates with Permutation Tests on a Lattice', *Phytopathology*, 95: 1453–61.

Toe. (1992) 'Les Incidences Des Aménagements Hydro-agricoles Sur L'agriculture Traditionnelle: Cas de la Plaine Aménagée de Banzon', Master thesis, University of Ouagadougou, Burkina Faso.

Tollenaere, C. et al. (2017) 'Virus-Bacteria Rice Co-Infection in Africa: Field Estimation, Reciprocal Effects, Molecular Mechanisms, and Evolutionary Implications', *Frontiers of Plant Science*, 8: 645.

Traore, O. et al. (2005) 'Processes of Diversification and Dispersion of Rice Yellow Mottle Virus Inferred from Large-scale and High-resolution Phylogeographical Studies', *Molecular Ecology*, 14: 2097–110.

Traore, O. et al. (2006) 'Rice Seedbeds as a Source of Primary Infection by Rice Yellow Mottle Virus', *European Journal of Plant Pathology*, 115: 181–6.

Traoré, O. et al. (2008) 'Diagnostic Sérologique Des Isolats Soudano-sahéliens du Virus de la Panachure Jaune du Riz (Rice Yellow Mottle Virus, RYMV)', *Tropicultura*, 26: 74–7.

Traoré, O. et al. (2009) 'A Reassessment of the Epidemiology of Rice Yellow Mottle Virus following Recent Advances in Field and Molecular Studies', *Virus Research*, 141: 258–67.

Traoré, V. S. E. et al. (2015) 'Farmers' Perception and Impact of Rice Yellow Mottle Disease on Rice Yields in Burkina Faso', *Journal of Agricultural Science*, 06: 943.

Trovão, N. S. et al. (2015) 'Host Ecology Determines the Dispersal Patterns of a Plant Virus', *Virus Evolution*, 1: vev016.

Venables, W. N., and Ripley, B. D. (2002) *Modern Applied Statistics with S*, 4th edn. New York: Springer.

Wickham, H. (2016) 'Data Analysis', in *Ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer.