OPEN ACCESS
phytobiomesjournal.org
Phytobiomes Journal
A Transdisciplinary Journal of Sustainable Plant Productivity

RESEARCH

Old Foe, New Host: Epidemiology, Genetic Diversity, and Pathogenic Characterization of Maize Streak Virus in Rice Fields from Burkina Faso

Noun Magdy Ibrahim Fouad,[1] Mariam Barro,[2] Martine Bangratz,[1] Drissa Sérémé,[3] Denis Filloux,[1] Emmanuel Fernandez,[1] Charlotte Julian,[1] Nignan Saïbou,[4] Abalo Itolou Kassankogno,[2] Abdoul Kader Guigma,[2] Philippe Roumagnac,[1] Issa Wonni,[2] Charlotte Tollenaere,[1] and Nils Poulicard[1,†]

[1] PHIM Plant Health Institute, Univ Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France
[2] Laboratoire de Phytopathologie, Institut de l'Environnement et de Recherches Agricoles du Burkina Faso, Bobo-Dioulasso, Burkina Faso
[3] Laboratoire de Virologie et de Biologie Végétale, Institut de l'Environnement et de Recherches Agricoles du Burkina Faso, Kamboinsé, Burkina Faso
[4] Laboratoire Mixte International PathoBios, Institut de Recherche pour le Développement, Bobo-Dioulasso, Burkina Faso

Accepted for publication 24 January 2024.

ABSTRACT

Rice is of critical significance regarding food security worldwide, including in Africa. Only two viruses impacting rice production in Africa have been deeply investigated for decades: the rice yellow mottle virus (*Solemoviridae*) and the rice stripe necrosis virus (*Benyviridae*). Using viral metagenomics, we aimed at broadening knowledge on interacting communities associated with plants in rice landscapes and exploring the diversity and epidemiological status of viruses circulating in rice fields from Burkina Faso. We performed an epidemiological survey in this country between 2016 and 2019 involving 57 small farmers' rice fields under two production systems: rainfed lowlands and irrigated areas. More than 2,700 rice samples were collected without regard to disease symptoms following a regular scheme. Wild and cultivated *Poaceae* (maize and sugarcane) growing in nearby rice fields were also collected. Unexpectedly, metagenomics detected maize streak virus (MSV) (*Geminiviridae*) in analyzed rice samples. Further molecular analyses using rolling circle amplification-polymerase chain reaction showed that MSV was widely distributed and highly prevalent in both rainfed lowlands and irrigated rice areas. MSV-A and MSV-G strains were identified. MSV-G, exclusively identified thus far in wild grasses, was the most prevalent strain, whereas MSV-A, known to cause severe symptoms in maize, was sporadically identified. Using infectious clones in experimental conditions, we confirmed the pathogenicity of both MSV strains in rice. Thus, in addition to contributing to the epidemiological surveillance of rice production in Africa, our results illuminate new epidemiological and pathogenic aspects of one of the most studied plant viruses with significant economic consequences in Africa.

Keywords: Africa, epidemiological surveillance, epidemiological survey, rice, viral metagenomics, virus

[†]Corresponding author: N. Poulicard; nils.poulicard@ird.fr

Funding: Support was provided by the French National Research Agency "Investissements d'avenir" (ANR-10-LABX-001-01) and "young researchers" (ANR-20-CE35-0008-01), Agropolis Foundation (ANR-16-IDEX-006), CGIAR Research Program on Rice Agri-Food Systems (RICE), and Cooperation and Cultural Action Department of the French Embassy (SCAC) in Burkina Faso.

Emerging crop diseases, a high proportion of which are caused by viruses, are a significant burden on the food security and economic stability of societies, especially in developing countries (Anderson et al. 2004; Jones 2021). Although the intensification of agriculture has become one of the major priorities for providing food for people, intensification programs are threatened by climate and global

e-Xtra: Supplementary material is available online.

The author(s) declare no conflict of interest.

changes (environmental, demographic, and socioeconomic), which could subsequently favor the emergence of plant pathogens and increase their impact on food production (Anderson et al. 2004; Baker et al. 2022). In this context, we aimed to characterize the diversity, epidemiology, and virulence of viral pathogens threatening rice production in Africa.

Rice is the most important human staple food crop in the world, directly feeding nearly half of the world's population every day. In Africa, rice cultivation has historically involved two species: the African rice *Oryza glaberrima*, which was domesticated in West Africa approximately 3,000 years ago, and the Asian rice *O. sativa*, which has been introduced repeatedly since the fifteenth century (Portères 1970). Beginning in the second half of the twentieth century and intensifying during the last decades, cultivated rice areas have drastically increased in Africa, and this crop has become important as a strategic commodity for food security according to 2022 statistics from the Food and Agriculture Organization Corporate Statistical Database (FAOSTAT) (https://data.un.org/Data.aspx?d=FAO&f=itemCode%3A27) (Demont 2013; Soullier et al. 2020). Such agricultural changes, including the preference for and intensification of Asian rice cultivation (Cubry et al. 2018), could render rice cultivation more exposed to emerging pathogens, particularly viruses (Anderson et al. 2004). Our capacity to ensure the sustainability of production systems against plant diseases strongly depends on our ability to explore the vast diversity of microorganisms in the environment, to understand how they affect the stability and productivity of ecosystems, and to identify pathogens from the early stage of emergence through epidemiological surveillance.

Among the 19 viruses that have thus far been reported to infect rice worldwide (Wang et al. 2022), five have occasionally been identified in Africa (Abo and Sy 1997). Only two of the five viruses detected in Africa have been genetically characterized and deeply investigated: the rice stripe necrosis virus (*Benyvirus*, *Benyviridae*) (Bagayoko et al. 2021) and the rice yellow mottle virus (RYMV) (*Sobemovirus*, *Solemoviridae*) (Hébrard et al. 2021). The three remaining rice viruses reported in Africa that have been poorly characterized thus far are maize streak virus (MSV) strain A (*Mastrevirus*, *Geminiviridae*), African cereal streak virus (genus and family not determined), and rice crinkle disease (etiology, genus, and family not determined) (Abo and Sy 1997).

The recent methodological innovations in high-throughput sequencing and metagenomics allow us to explore the vast genetic and functional diversity of viruses in wild and cultivated environments and to identify putative emergent viruses (Bernardo et al. 2018; Edgar et al. 2022; Greninger 2018; Lefeuvre et al. 2019). However, despite the continuous detection of partial or complete virus metagenome-assembled genomes from wild or anthropized environments, few of these metagenome-assembled genomes have subsequently been subjected to more in-depth analyses, in most cases using Sanger sequencing. In addition, this sequencing validation step has usually not been supplemented with epidemiological or pathogenetic information on the related viruses (Koonin and Dolja 2018). Nevertheless, among the few novel viruses initially discovered using metagenomics-based approaches and further partially or fully biologically characterized, geminiviruses are among the most represented (Moubset et al. 2022).

In this study, we intended to meet these expectations, focusing on viruses circulating in rice fields in Burkina Faso. To this end, an epidemiological survey focusing on two rice production systems (irrigated [IR] perimeters and rainfed lowlands [RLs]) was conducted in this West African country between 2016 and 2019 (Fig. 1). In parallel, agricultural practices and the use or turnover of rice varieties, as well as the diversity of microbiomes associated with rice roots and (known) rice pathogens (viral, bacterial, and fungal) circulating in these fields, were followed (Barro et al. 2021a, b, 2022; Billard et al. 2023; Kaboré et al. 2022). To expand our knowledge of virus communities associated with wild and cultivated plants from rice landscapes in this region, we performed viral metagenomics analyses on rice and wild grass leaf samples collected in 2016 and 2017 (Moubset et al. 2022). Among the thousands of contigs assigned to viral genomes (including RYMV), these analyses revealed unexpected MSV species (*Mastrevirus*, *Geminiviridae*) infecting rice.

Geminiviruses are responsible for a large number of emerging crop diseases in the world, with considerable impact on the yields of several cash and staple crops (maize, cassava, tomato, cotton, beans, and grain legumes). These viruses constitute a major threat to the food security of tropical and subtropical developing countries (Rybicki 2015; Rybicki and Pietersen 1999). Specifically, MSV, one of the most devastating viruses in maize in Africa (Martin and Shepherd 2009; Savary et al. 2019), was first described in South Africa during the 1870s and later in 1896 during a serious outbreak (Fuller 1901; Shepherd et al. 2010). MSV virions consist of a twinned and quasi-icosahedral (geminate) capsid. The monopartite, circular, single-stranded DNA genome of approximately 2.7 kb, replicating by rolling circle, encodes only four proteins (Fiallo-Olivé et al. 2021; Gong et al. 2021). MSV is transmitted naturally by 18 leafhopper species (Fiallo-Olivé et al. 2021), some of which have been observed in rice fields from Burkina Faso and the surrounding regions (N. Poulicard, *personnal communication*; Tra Bi et al. 2020). Alternatively, MSV is also transmitted experimentally by agroinoculation of infectious clones (Boulton et al. 1989; Grimsley et al. 1987). In addition to maize, MSV is characterized by a large monocotyledonous host range, including wild and cultivated *Poaceae* species, such as sugarcane and sorghum (Kraberger et al. 2017).

Restricted to Africa and surrounding islands, 11 strains (A to K) have been identified. MSV-A is the only strain known to cause severe symptoms in cultivated plants, whereas MSV-B, MSV-C, and MSV-E, mainly detected in infected wild grasses, can also produce mild infections in MSV-susceptible maize genotypes (Claverie et al. 2019; Oyeniran et al. 2021; Shepherd et al. 2010). Previous studies showed that, as with other geminiviruses, recombination played a decisive role in the evolution of MSV. In particular, a recombination event that occurred in the mid-nineteenth century between ancestral MSV-B and MSV-G/F strains probably led to the emergence, efficient dispersion across Africa, and progressive adaptation of the MSV-A strain to maize (Harkins et al. 2009; Monjane et al. 2011, 2020; Varsani et al. 2008).

Although the MSV-A strain has recurrently been reported in maize from Burkina Faso (Kraberger et al. 2017), it was identified only once in rice fields in this country (Konaté and Traoré 1992). The MSV-A strain was also experimentally transmitted to rice by viruliferous leafhoppers (Damsteegt 1983; Konaté and Traoré 1992). However, to our knowledge, information on the geographical distribution, prevalence, genetic diversity, and aggressiveness of this virus in rice fields has never been reported.

Thus, based on viral metagenomics, epidemiological, and experimental approaches, the objective of this study was to evaluate the epidemiological and pathogenic status of MSV in rice fields from Burkina Faso. In addition to contributing to the epidemiological surveillance of rice production in Africa, our results illuminate new epidemiological and pathogenic aspects of one of the most studied plant viruses with significant economic consequences in Africa.

MATERIALS AND METHODS

Study area and samplings. The study area was located in western Burkina Faso in a 100 × 100-km region in the Sudanian

bioclimatic area (Fig. 1A). Six sites located within three geographical zones—Bama (BM)/Badala (BL), Banzon (BZ)/Senzon (SZ), and Karfiguela (KA)/Tengrela (TG)—were surveyed between 2016 and 2019. Each geographical zone comprised one IR site and a neighboring RL site (Fig. 1B).

The regular and longitudinal rice leaf sampling between 2016 and 2019 in these six sites involving 57 rice fields was previously described by Barro et al. (2021a). Briefly, observations and samplings were performed at the maximum tillering/heading initiation stages from September to December each year. Each studied field was a square measuring approximately 25 m on each side, with a regular sampling of 16 plants per field over a grid (Supplementary Fig. S1). This sampling approach did not consider potential disease symptoms (regular sampling without regard to disease symptoms), but the rice leaves were inspected when sampled and disease symptoms were recorded when observed (https://doi.org/10.23708/8FDWIE). Parts of these samples were also submitted for specific detection of RYMV via an enzyme-linked immunosorbent assay (ELISA) (https://doi.org/10.23708/GZCM1O).

In addition to these rice samples, the diversity and percentage of surface covering of wild grass species growing in the rice field borders were estimated for six fields in 2017 (IR production system:

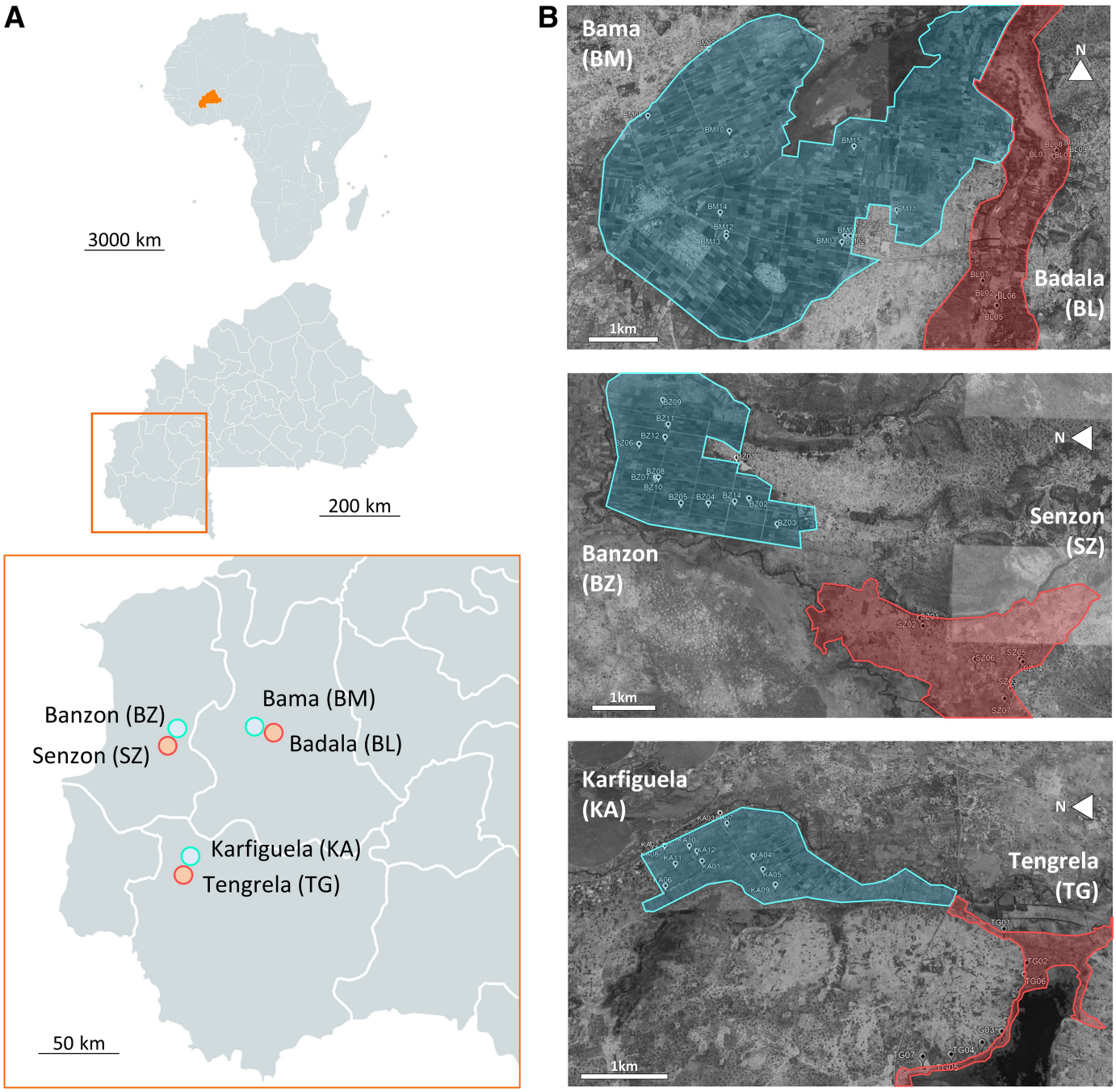

Fig. 1. Study area for the spatiotemporal survey of rice fields in Burkina Faso. **A,** Burkina Faso, Africa, with a focus on western Burkina Faso, where the six sites, Bama (BM), Badala (BL), Banzon (BZ), Senzon (SZ), Karfiguela (KA), and Tengrela (TG), are located. These sites are indicated by colored dots referring to their rice production systems: irrigated (IR), blue, and rainfed lowland (RL), red. Map is based on MapChart (https://mapchart.net/). **B,** Distribution area of IR (blue) and RL (red) rice production systems at the six sites and rice field locations surveyed during this study. Field locations and delimitations are based on Google Earth maps.

BM02, BZ11, KA01; RL production system: BL02, SZ07, TG01) (Supplementary Fig. S1). Details regarding these wild plant observations and samplings are available at https://doi.org/10.23708/1IPJAU. One plant of the five most frequent plant species was randomly collected without regard to symptoms for further analyses (see later discussion). Finally, specific samplings of wild (*Poaceae* species) and cultivated (rice, maize, and sugarcane) plants presenting symptoms putatively related to viral infection (leaf deformations and stripes) were performed within or nearby the rice fields involved in the longitudinal survey. Samples were collected individually in bags or envelopes containing silica gel to dehydrate and then stored at –20°C before further molecular analyses.

Thus, the total number of plant samples analyzed during this study was greater than 2,800, with approximately 2,750 rice plant samples (from 43, 42, 49, and 40 rice fields in 2016, 2017, 2018, and 2019, respectively); 30 wild grasses collected without regard to symptoms; and four, eight, seven, and 15 symptomatic samples of rice, maize, sugarcane, and wild *Poaceae*, respectively. These samples were named according to the year of sampling, the site, the field, and the host plant of collection (e.g., 17TG01 and 17TG01w correspond, respectively, to rice and wild plant samples collected in 2017 in/close to rice field "01" of the TG site).

Viral metagenomics. Detection and identification of both DNA and RNA viruses were performed on rice and wild *Poaceae* samples using a virion-associated nucleic acid (VANA) metagenomics-based approach (Moubset et al. 2022). Specifically, for each rice field surveyed in 2016 and 2017, 16 sampled rice plants were pooled to a 1-g sample that was ground up and prepared for analysis (i.e., 85 pooled rice samples were obtained to represent 85 rice fields, 43 in 2016 and 42 in 2017). Similarly, the five leaf samples of the five most frequent wild plant species collected without regard to symptoms in each border of the rice field were ground up and pooled (i.e., six pooled grass samples were obtained, representing six rice field borders surveyed in 2017).

Each pooled sample of rice or wild grasses was processed using the VANA approach as described by François et al. (2018). Briefly, we isolated viral particles by filtration and ultracentrifugation. The nucleic acids not protected in virus-like particles were further degraded by DNase and RNase, and the total RNA and DNA was then extracted using a NucleoSpin kit (MACHEREY-NAGEL). Reverse transcription was performed with SuperScript III reverse transcriptase (Thermo Fisher Scientific), complementary DNA was purified with a QIAquick PCR Purification Kit (QIAGEN), and complementary strands were synthesized by Klenow DNA polymerase I. Double-stranded DNA was amplified by random polymerase chain reaction (PCR) amplification. Samples were barcoded during reverse transcription and PCR steps using homemade 26-nucleotide Dodeca Linkers and PCR multiplex identifier primers. PCR products were purified using NucleoSpin gel and PCR clean-up (MACHEREY-NAGEL). Finally, libraries were prepared from purified amplicons and sequenced on an Illumina HiSeq to generate 2 × 150-nucleotide paired-end reads (GENEWIZ).

Nucleic acid extraction, rolling circle amplification, and MSV detection by specific PCR. Total DNA extraction of each sample (100 mg of plant material for pooled leaf samples, 20 mg for individual leaf sample) was performed in a lab with no prior exposure to plant material infected with MSV according to the cetyltrimethylammonium bromide protocol. The concentration and quality of the extracted DNA were then assessed with NanoDrop Microvolume Spectrophotometers (Thermo Fisher Scientific).

The circular DNA genomes of MSV were first amplified with a TempliPhi kit (Cytiva) via rolling circle amplification (RCA) according to the manufacturer's protocol. Next, the presence of MSV-G and MSV-A at field or plant level was detected by PCR using 1 µl of RCA products, GoTaq Flexi (Promega) according to the manufacturer's protocol, and primers targeting MSV-G (MSV-F559bp: 5′-GGAGCATGTAAGCTTCGGGA-3′, positions 1,875 to 1,889; MSV-R559bp: 5′-GAGCTCGTTGGTCACTGGAA-3′, positions 2,415 to 2,434; melting temperature [Tm] 57°C; amplification 559 bp) and MSV-A (MSVg-2F: 5′-TCAGCCATGTCCA CGTCCAAG-3′, positions 478 to 498; MSVa-1R: 5′-TCACCAC GAAGCGATGACACA-3′, positions 912 to 932; Tm 55°C; amplification 454 bp). PCR amplification of the 559 and 454 nucleotides was checked on 1× agarose gel, and amplicons were sequenced for further analyses (see later discussion). For inconclusive samples (i.e., included in the uncertainty interval associated with the visualization method), a second PCR was performed using the same protocol as described previously, except for the use of 2 µl of RCA products. The samples still not conclusive after these two PCRs were then considered negative for MSV. For all negative samples, PCRs amplifying the S1 locus area of chromosome 6 of rice were performed as an internal control. It should be noted that the amplicon size was used to discriminate the Asian rice *O. sativa* (935 bp) and the African rice *O. glaberrima* (1,384 bp) (Gnacadja et al. 2018). Differences in the percentage of MSV-positive fields and plants according to sites, rice production systems, and years were assessed based on contingency tables using Fisher's exact test.

Partial and complete genome sequencing. Amplicons obtained by detection PCR were sequenced using the Sanger method (GENEWIZ). We obtained sequences of 491 nucleotides long (i.e. 18.3% of the complete MSV genome), which were used both to identify which MSV strain was present at the field or plant level and to estimate the genetic diversity.

Two approaches were used to obtain complete genome sequences of MSV. For the first approach, the RCA products were digested with *Bam*HI (New England Biolabs), inserted in pGEM-T Easy Vector (Promega), and cloned in JM109 competent cells (Promega) according to the manufacturers' protocols. For the second approach, we performed two overlapping PCRs on the RCA products with GoTaq Flexi according to the manufacturer's protocol to amplify the complete genome of MSV using primers targeting MSV-G (PCR #1: MSVg-2F and MSV-R559bp, Tm 57°C, amplification 1,956 bp; PCR #2: MSV-F559bp and MSVg-2R 5′-AGGCATGT CCGAACCGATGC-3′ at positions 980 to 999, Tm 57°C, amplification 1,811 bp) and MSV-A (PCR #1: MSVg-2F and MSVa-3R 5′-ATTGGCTCCAGCCTAACATCTTCC-3′ at positions 1,898 to 1,921, Tm 55°C, amplification 1,443 bp; PCR #2: MSVa-1F 5′-CGACGATGTAGAGGCTCTGCT-3′ at positions 1,761 to 1,781 and MSVa-1R, Tm 55°C, amplification 1,864 bp). The obtained complete genome sequences were deposited in GenBank (accession numbers OR258386 to OR258402). Two complete genome sequences of representative MSV isolates were used to obtain infectious clones (see later discussion).

Genetic diversity and phylogenetic analyses. Partial and complete genome sequences of MSV obtained during this study were compared with the 885 sequences described by Kraberger et al. (2017). Multiple sequence alignments were performed using MUSCLE (Edgar 2004) implemented in SeaView v4.7 (Gouy et al. 2010). Sequence pairwise identities of a selection of MSV sequences were calculated with SDT v1.2 (Muhire et al. 2014). Maximum likelihood phylogenetic trees were reconstructed with SeaView using the best fit nucleotide substitution models (Tamura three-parameter + G and GTR + G + I for the partial and complete genome data sets, respectively) determined with MEGA X (Kumar et al. 2018) and 100 bootstrap replications. Phylogenetic trees were drawn using FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). The genetic diversity of the partial genome data set was estimated using the best fit nucleotide substitution model, with stan-

dard errors of each measure based on 100 replicate bootstraps, as implemented in MEGAX.

Genetic differentiation of MSV populations according to fields, sites, years, and rice production systems was estimated by analysis of molecular variance obtained by performing 1,000 permutations as implemented in Arlequin v5.3.1.2 (Excoffier et al. 2005). Recombination signals in MSV from Burkina Faso were identified using the seven algorithms implemented in RDP v4.97 (Martin et al. 2015). Recombination events detected by at least five methods and with P values less than 10^{-5} were considered.

Infectious clones and agroinoculations in rice. Three infectious clones of MSV-A and MSV-G were used during this study. First, the clone pBC-KS::MSV-A|R2| was obtained in previous studies and was identified as highly pathogenic in maize (Isnard et al. 1998; Peterschmitt et al. 1996). Next, the clones pCAMBIA0380::MSV-A|53| and pCAMBIA0380::MSV-G|61| were built by gene synthesis (GENEWIZ) of the genome sequences obtained after PCR amplification and cloning of RCA products, respectively. More precisely, the MSV genome sequences were synthesized based on the traditional technique of partial tandem repeats (i.e., with the highly conserved stem-loop region of approximately 50 nucleotides corresponding to the origin of replication repeated on both sides of the genome) (Urbino et al. 2008).

These infectious clones and the empty pCAMBIA0380 plasmid as a negative control (hereafter called pCAMBIA0380::Ø) were introduced into two strains of *Agrobacterium tumefaciens* (C58C1 and EHA105) by electroporation. Transformed *A. tumefaciens* colonies were plated and cultivated at 28°C in Luria-Bertani medium containing 50 µg/ml of kanamycin and 25 µg/ml of rifampicin (and 100 µg/ml of gentamicin for C58C1) and used for agroinoculation as described later after confirming the presence of inserts by PCR in colonies.

Agroinoculations were performed in two rice varieties representing the two cultivated rice species in Africa (*O. sativa indica* 'IR64' and *O. glaberrima* 'Tog5673') and one variety of maize (*Zea mays* 'Golden Bantam'). For each experiment, approximately 40 to 60 plants of each rice variety were inoculated at the two-leaf stage (i.e., 7 to 10 days after seedling) by pricking the apical meristem three times at different levels with the tips of 0.4-mm needles previously dipped in *Agrobacterium* colonies. Two independent assays were performed. Both *A. tumefaciens* strains and pCAMBIA0380::Ø, pBC-KS::MSV-A|R2|, and pCAMBIA0380::MSV-G|61| plasmids were used for the first experiment. As no significant difference in infection was detected between the two *A. tumefaciens* strains ($\chi^2 = 1.101$, $P = 0.294$), we used only EHA105 for the second experiment. In addition, in the second experiment, the pathogenicity of the infectious clone pCAMBIA0380::MSV-A|53| was evaluated. For both experiments, symptom initiation and development were monitored, and the number of leaves and the height of the inoculated plants were measured at 28 days postinoculation (dpi). The fresh weight of inoculated plants was estimated for only the second experiment. Values were expressed for each treatment and cultivar in percentage according to negative controls.

Statistical analyses of these phenotypic measurements were performed using the software STATGRAPHICS Centurion 15.1.02 (StatPoint Technologies). As the distribution of plant size, number of leaves, and fresh biomass was not normal according to the Levene test of equality of error variances, we first analyzed our results using a nonparametric (Kruskal-Wallis) test. However, as the Kruskal-Wallis and analysis of variance (ANOVA) (parametric) tests gave similar results, and because ANOVA is robust to the partial violation of its assumptions and allows the analysis of factor interactions (post hoc least significant difference, significance threshold at $P < 0.05$), whereas Kruskal-Wallis does not, we also presented the results obtained with ANOVA. It should be noted that these statistics were compiled only with the results obtained for *O. glaberrima* 'Tog5673', as the sample size for infected plants was not sufficient for *O. sativa indica* 'IR64'.

RESULTS

Identification of MSV in rice fields from Burkina Faso by viral metagenomics. Eighty-five rice samples were analyzed by a VANA-Illumina HiSeq approach, representing 43 and 42 rice fields surveyed in 2016 and 2017, respectively. The Illumina sequencing produced 288,971,924 raw sequences that were subsequently trimmed and corrected to extract a total of 1,459,838 contig sequences, corresponding to 109.67 Mb (Moubset et al. 2022). BLASTn and BLASTx analyses of these data showed that 1.9% of these contigs matched virus sequences from known (RYMV) and unknown (Moubset et al. 2022) virus species, the rest of the sequences being mainly related to the hosts. Specifically, we identified MSV sequences from two fields surveyed in 2017. More precisely, a total of 1,490 MSV contigs were obtained from field 17BM11 (98 contigs) and field 17TG05 (1,392 contigs), including two long contigs from each field covering the complete genome of MSV. These long contigs respectively shared 99.01 and 98.92% homologies with the genome sequence of an MSV isolate belonging to strain G collected in a wild member of the *Poaceae* family (*Urochloa lata*) from Nigeria (accession number EU628635.1). In addition to rice samples, MSV was identified by the same VANA-Illumina HiSeq approach in one of the six pools of wild plants collected at the borders of one rice field per site from the TG site (17TG01w pool) (https://doi.org/10.23708/1IPJAU) (Supplementary Table S1). The contig obtained from this sample covered the complete genome of this virus, and BLAST analyses revealed 99.37% homology with an MSV-G isolate from Nigeria collected in 2007 from the wild grass *Digitaria horizontalis* (accession number EU628634.1).

Molecular detection of MSV in rice fields. As only 2 years of samplings were analyzed by VANA-Illumina and as the number of fields positive for MSV could be underestimated with this technique, we opted for a targeted MSV detection strategy. For this, we used specific primers to detect by RCA-PCR the presence of the virus in rice fields, with 43 fields in 2016, 40 in 2017, 49 in 2018, and 40 in 2019. Thus, we analyzed a total of 172 rice samples, each representing 16 plants from the same field collected the same year (Supplementary Table S2).

A surprisingly high number ($N = 59$) of rice samples were positive for MSV, suggesting that 34.3% of the surveyed fields were infected with this virus (Table 1). In addition, we noted that MSV was detected in all localities and both production systems (i.e., IR and RL fields), but with no significant difference for the 2016 to 2019 period ($\chi^2 = 6.424$, $P = 0.267$ and $\chi^2 = 0.170$, $P = 0.680$, respectively). However, the percentage of MSV-positive fields was significantly higher in 2018 and 2019 than in 2016 and 2017 ($\chi^2 = 44.269$, $P < 0.001$).

Detection of MSV in wild and cultivated plants. In addition to rice samples, we performed MSV detection PCR on the same pools of wild plants used with the VANA-Illumina approach (Supplementary Table S1). As reported with the VANA-Illumina approach, MSV was detected in the 17TG01w sample. The detection PCR revealed that MSV was also present in the 17TG02w sample. Both pools of samples combined *Poaceae* plant species already described as host plants for MSV; namely, *Setaria pallide-fusca* (Schumach.) Stapf & C.E. Hubb. and *Dactyloctenium aegyptium* Beauv. (Supplementary Table S3) (Kraberger et al. 2017).

Next, we analyzed individually symptomatic wild plants (including the wild rice species *O. longistaminata*) and cultivated *Poaceae* (maize and sugarcane) collected in 2016, 2017, and 2019 around and within the rice fields (Supplementary Table S3). MSV was detected in a large proportion of samples, with all maize (N = 8 [100%]), most sugarcane (N = 7 [87.5%]), and a large fraction of wild plants (N = 15 [40%]) found to be MSV-positive. Overall, among the wild plant species analyzed, MSV was detected in *D. horizontalis*, *Echinochloa colona*, *Eragrostis* spp., and *O. longistaminata*.

Genetic diversity of MSV circulating in rice production areas. To characterize the MSV genetic diversity circulating in rice production areas from western Burkina Faso, we sequenced the amplicons obtained from molecular detection PCRs performed on rice, wild plants, and cultivated *Poaceae* (maize and sugarcane). We obtained 29 sequences (Supplementary Table S2), and phylogenetic reconstructions indicated that most of these sequences belonged to strain G of MSV (27 of 29 [93.1%]), whereas only two (6.9%) (i.e., 16BZ09 and 18TG01) corresponded to strain A (Supplementary Fig. S2; Supplementary Table S2).

In addition to rice, the presence of MSV was tested in wild grasses and cultivated *Poaceae* (maize and sugarcane) collected within or nearby rice fields (Supplementary Table S3). MSV-A was detected in 19 of 30 plants (63.3%) showing symptoms. Interestingly, MSV-G was detected in only one pool of symptomless wild grasses that were randomly collected from one site (17TG01w).

The genetic diversity based on these fragments was estimated at 0.0956 ± 0.0164 substitution (subst.)/site in total (0.0023 ± 0.0007 and 0.0028 ± 0.0017 subst./site for MSV-G and MSV-A, respectively), which is similar to the estimates of genetic diversity of MSV in West Africa in this same portion of the genome (0.0854 ± 0.0101 subst./site). Genetic differentiation of MSV was observed between rice and *Poaceae* ($F_{ST(Rice/Poaceae)} = 0.760, P < 0.001$), reflecting the over-representation of MSV-G in rice and MSV-A in other *Poaceae* species. However, taking each MSV strain separately, no genetic differentiation was revealed between the MSV-G ($F_{ST(Rice/Poaceae)} < 0.001, P > 0.99$) and MSV-A ($F_{ST(Rice/Poaceae)} < 0.001, P = 0.809$) isolates identified in rice and other *Poaceae* species. In addition, genetic differentiation was not detected for MSV isolates according to rice production system (all MSV, $F_{ST(IR/RL)} < 0.001, P > 0.99$, MSV-G, $F_{ST(IR/RL)} < 0.001, P = 0.550$, MSV-A, $F_{ST(IR/RL)} = 0.147, P = 0.124$). Similar results were obtained using other regions of the genome (data not shown).

Complete genomes of MSV. Based on the detection PCR and partial sequencing results, we selected several MSV-positive samples to obtain complete genome sequences of MSV-A and MSV-G from different host species. This effort produced 16 complete MSV genomes: four from rice fields, six from wild grasses, two from maize, and four from sugarcane (Supplementary Table S4). The sequence of 17BM11 MSV isolate (MSV-G|61|) was obtained after the cloning of RCA products. In all other cases, direct sequencing of PCR amplicons was performed. To these 16 sequences, we added the three complete genome sequences obtained by the VANA-Illumina approach (MSV-G in rice 17BM11|Contig1671| and 17TG05|Contig844|, MSV-G in wild grasses 17TG01|Contig711|) for further analyses. It should be noted that the complete MSV-G genome sequences obtained by the Sanger and VANA-Illumina approaches in the same samples were highly similar (11 and one variable sites for 17BM11 and 17TG01w with 99.6 and 100.0% genetic identity, respectively).

The 19 sequences obtained during this study were compared with a complete MSV genome data set that included 885 sequences from all MSV strains, including eight MSV-G sequences from wild grass samples in West Africa (Nigeria and Mali) and Gran Canaria and 695 MSV-A sequences from all over Africa, with five MSV-A sequences collected from maize in 2008 in Burkina Faso (Kraberger et al. 2017). Analyses to detect recombination events in this sequence data set demonstrated that none of the 19 complete genome sequences obtained during this study were recombinant. Phylogenetic analyses showed that the MSV-G isolates identified during this study were closely genetically related (more than 98.9% genetic identity) (Fig. 2). As previously demonstrated with partial genome analyses, no genetic differentiation was observed in MSV collected in rice versus other *Poaceae* species ($F_{ST(Rice/Poaceae)} = 0.362, P < 0.001$) or between rice production modes ($F_{ST(IR/RL)} < 0.001, P > 0.99$). MSV-G isolates from western Burkina Faso were genetically related to those identified in Nigeria in 2007 and Mali in 1987 (more than 99% identity with EU628632.1 and EU628634.1). Similarly, we noted that the MSV-A isolates identified in this study belonged to a clade of isolates exclusively identified in West Africa and were closely genetically related (more than 99.4% identity), with no genetic differentiation with those collected in rice or other *Poaceae* species ($F_{ST(Rice/Poaceae)} < 0.001, P = 0.985$). Surprisingly, these isolates were more genetically related to MSV-A isolates collected in Nigeria in 2015 (e.g., between 99.5 and 99.9% genetic identity with KX787926.1 and KX787927.1, with approximately 1,250-km distance between these isolates and those from our study area) than to MSV-A identified in maize in Burkina Faso in 2008 (between 96.9 and 96.6% genetic identity, with approximately 300-km distance between isolates).

TABLE 1
Detection of maize streak virus (MSV) in rice fields from six sites[a]

	2016		2017		2018		2019		Total				
Site	−	+	−	+	−	+	−	+	−	+	%MSV	%MSV	Prod.
BM	7	3	6	1	5	4	5	2	23	10	30.3		
BZ	6	2	7	0	5	6	5	1	23	9	28.1	33.0	IR
KA	9	0	7	0	2	6	1	7	19	13	40.6		
BL	5	1	7	0	7	0	2	4	21	5	19.2		
SZ	5	0	6	0	0	7	3	4	14	11	44.0	36.0	RL
TG	5	0	5	1	3	4	0	6	13	11	45.8		
Total	37	6	38	2	22	27	16	24	113	59	34.3		
%MSV		14.0		5.0		55.1		60.0		34.3			

[a] Sites (BL = Badala; BM = Bama; BZ = Banzon; KA = Karfiguela; SZ = Senzon; TG = Tengrela) representing two rice production systems (IR = irrigated; RL = rainfed lowland) in 2016, 2017, 2018, and 2019. Percentage of MSV-positive fields (%MSV) is indicated.

MSV prevalence in rice fields. To assess MSV prevalence in rice fields, we analyzed individually the 16 plants of 12 rice fields previously identified as MSV-positive in 2018 by detection PCR (Supplementary Table S2). These fields were selected to represent different sites (BM, BZ, SZ, KA, and TG), genetic diversity of MSV in rice fields (MSV-G versus MSV-A), and rice production system (IR versus RL). The BL site was not included in this analysis, as no MSV-positive fields were identified at this site.

Supplementary Figure S3 and Supplementary Table S5 show the results obtained for the 192 plant samples, among which 135 (70.3%) were annotated as symptomatic and 16 were positive for RYMV according to ELISA assays. We noted that MSV-G was frequently detected (32.1% of plants), with prevalence varying between 25.0 and 62.5% according to the field (Fig. 3) but with no significant difference between fields ($\chi^2 = 16.167$, $P = 0.135$). Interestingly, no significant difference was revealed between the two rice production systems (IR versus RL, $\chi^2 = 0.185$, $P = 0.667$).

The genetic diversity of MSV-G circulating within 18BM13, 18BZ14, 18SZ06, and 18KA08 was estimated at 0.0058 ± 0.0048, 0.0066 ± 0.0035, 0.0060 ± 0.0059, and 0.0035 ± 0.0025 subst./site, respectively, based on the partial genome amplified by detection PCR. We also did not observe genetic differentiation between fields ($F_{ST\text{fields}} < 0.001$, $P > 0.99$).

For MSV-A, in addition to not being frequently detected in rice fields, the prevalence estimated for this strain in the 18TG01 field was drastically lower than that estimated for MSV-G (6.3%) (Fig. 3; Supplementary Fig. S3). No co-infection between MSV-G and MSV-A was observed during this study. Of the 192 individual samples tested, eight plants were positive for both MSV and RYMV, and high levels of co-infection with other diseases, based on symptom recognition, were also reported (Supplementary Table S5).

Experimental validation of MSV pathogenicity in rice. We performed two independent experiments to test the pathogenicity of three MSV infectious clones: two obtained from the isolates identified during this study in rice plants from Burkina Faso (pCAMBIA::MSV-G|61| and pCAMBIA::MSV-A|53|) and one obtained from an isolate identified in a maize plant from Réunion Island (pBC-KS::MSV-A|R2|) (Supplementary Fig. S4). In both ex-

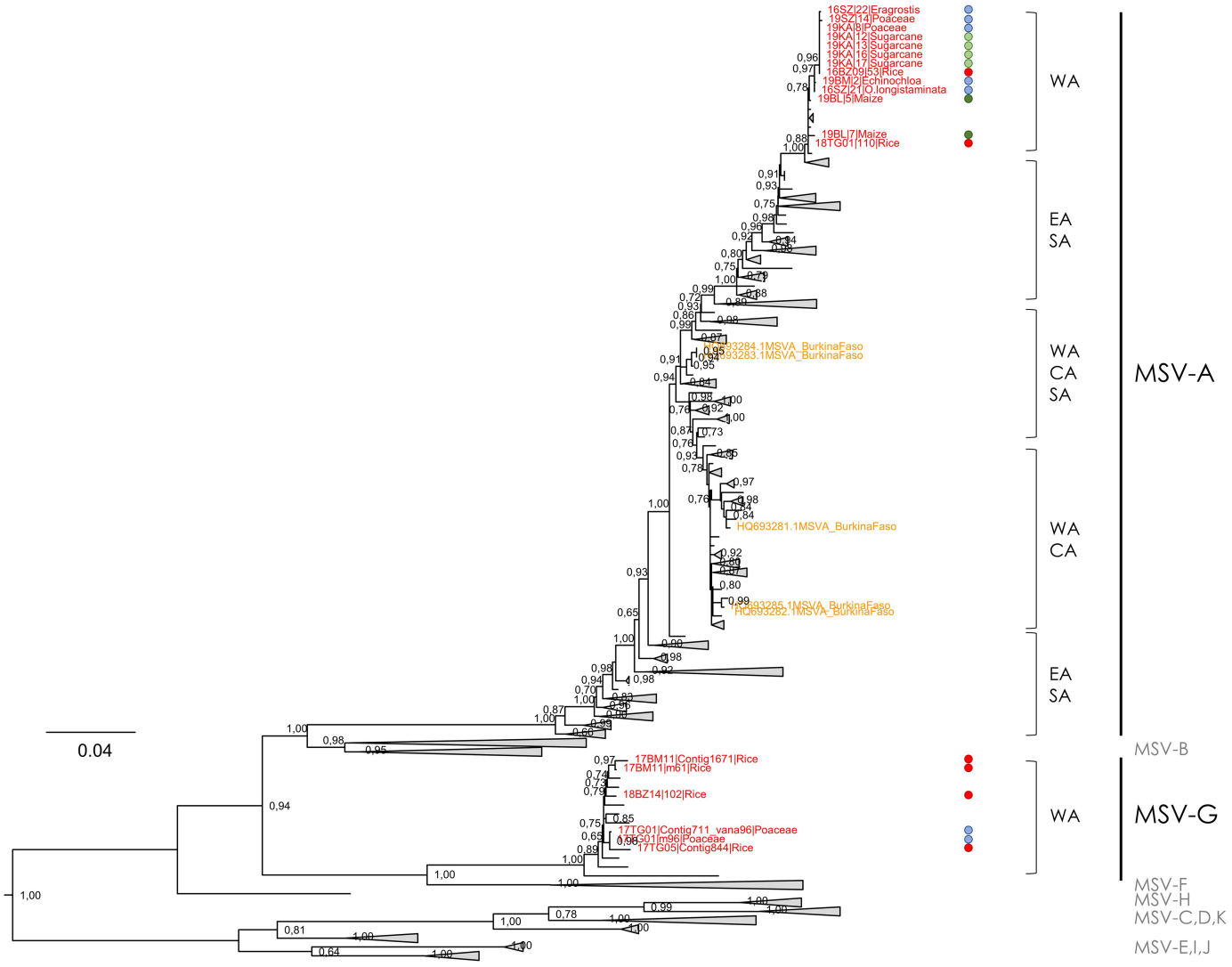

Fig. 2. Condensed phylogenetic tree reconstructed by maximum likelihood based on complete maize streak virus (MSV) genome sequences from western Burkina Faso obtained during this study (names in red) and available from public database (those from Burkina Faso in orange). Numbers at each node correspond to bootstrap values based on 100 replicates (only values above 0.70 are reported). The host plants from which these sequences were identified are indicated by colored circles (red = rice; dark green = maize; light green = sugarcane; blue = wild grasses). The clades corresponding to MSV strains (A to K) and the geographical origin (CA = Central Africa; EA = East Africa; SA = Southern Africa; WA = West Africa) of the sequences are shown.

periments, we observed the appearance of light streaks on the leaves of some rice plants (*O. sativa indica* 'IR64' and *O. glaberrima* 'Tog5673') at 14 dpi. These symptoms then developed drastically into clear and distinct streaks by 28 dpi, both on the leaves where the streaks were first observed and on the systemic and emergent leaves, followed by a drastic reduction in plant growth (Fig. 4). Correlation between symptoms and MSV accumulation in systemic leaves was validated by PCR and Sanger sequencing (data not shown).

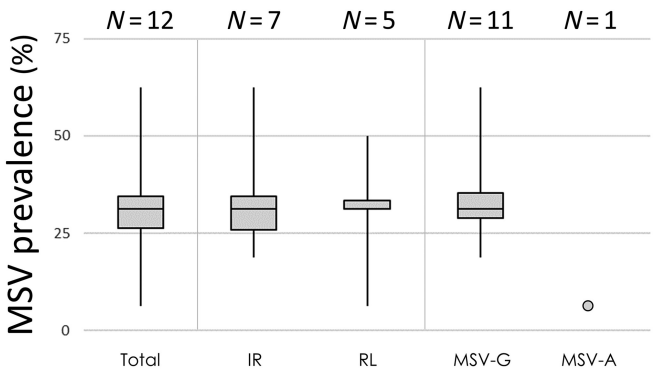

Fig. 3. Prevalence estimation of maize streak virus (MSV)-G and MSV-A in 12 rice fields representing two rice production systems, irrigated (IR) and rainfed lowland (RL), based on 15 to 16 individual plants collected in each field without a priori over a grid in 2018. The numbers of rice fields are indicated.

Unfortunately, most likely due to the variability of the agroinoculation process in rice, only a few cases of successful infection were observed, and not all agroinoculation modalities (i.e., infectious clone versus plant species) succeeded in the same experiment. Across the two experiments, MSV agroinoculation led to successful infection in 46 of 770 plants (6.0%) (Table 2; Supplementary Table S6). No significant difference in MSV transmission efficiency was observed between the three infectious clones in *O. sativa indica* 'IR64' ($\chi^2 = 2.144$, $P = 0.342$) and *O. glaberrima* 'Tog5673' ($\chi^2 = 0.764$, $P = 0.682$). However, MSV was more efficiently transmitted to Tog5673 than to IR64 (42 of 392 plants [10.7%] versus four of 378 plants [1.1%], $\chi^2 = 31.943$, $P < 0.001$) regardless of the infectious clone (pCAMBIA::MSV-G|61|, pCAMBIA::MSV-A|53|, and pBC-KS::MSV-A|R2|, $\chi^2 = 12.968$, $P < 0.001$, $\chi^2 = 17.216$, $P < 0.001$, and $\chi^2 = 2.856$, $P = 0.091$, respectively).

Twenty-eight days after agroinoculation, we measured the height and number of leaves of MSV-infected and noninfected plants. The following statistics were compiled only with the results obtained for *O. glaberrima* 'Tog5673', as the sample size of infected plants was not sufficient for *O. sativa indica* 'IR64'. We noted that the earlier the emergence of symptoms, the greater the impact on plant growth, leading to high between-plant heterogeneity in the data set. Although MSV infection did not systematically affect the number of leaves per plant (*O. glaberrima* 'Tog5673', Kruskal-Wallis,

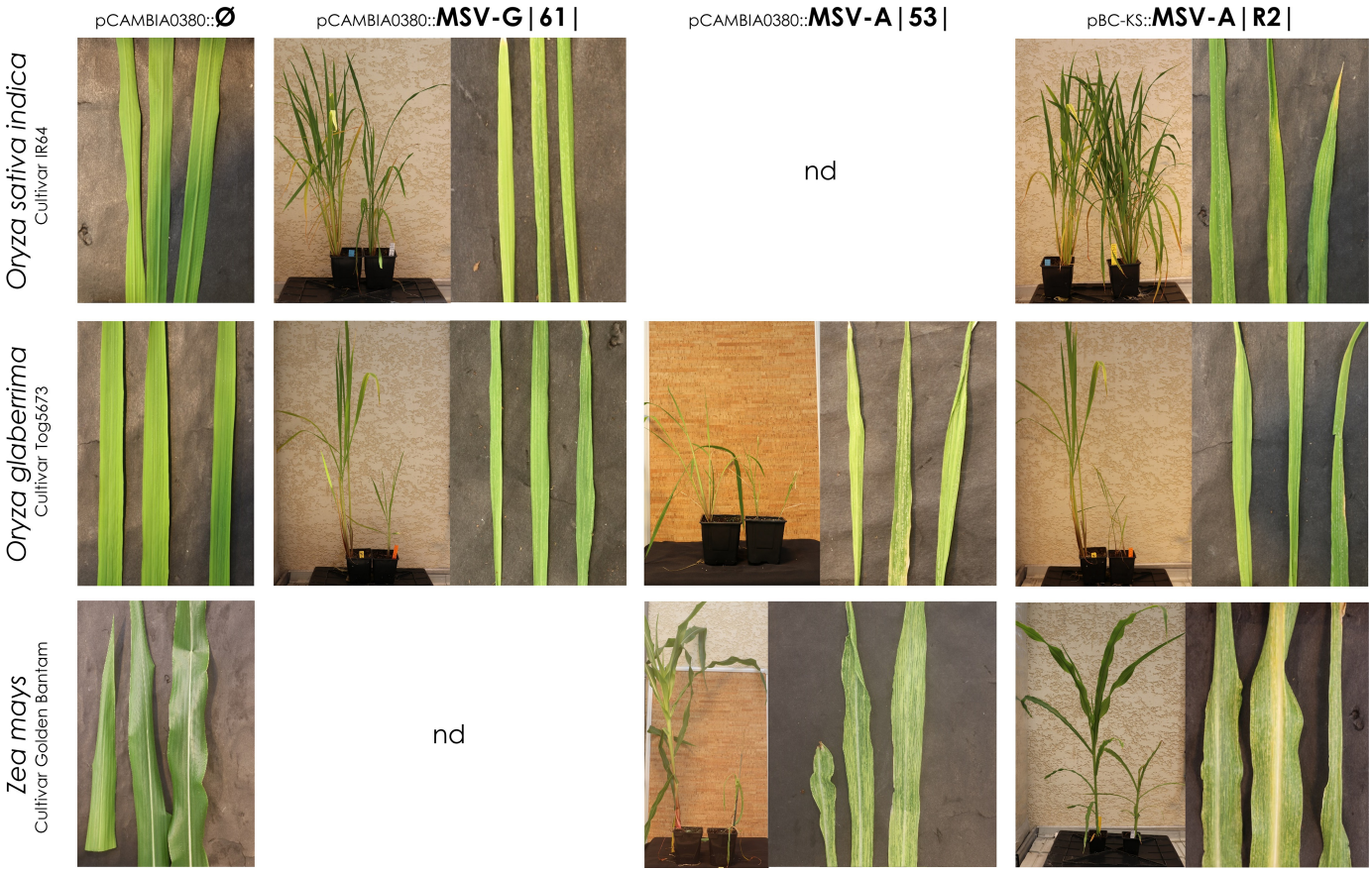

Fig. 4. Symptoms in rice (*Oryza sativa indica* 'IR64' and *O. glaberrima* 'Tog5673') and maize (*Zea mays* 'Goldem Bantam') associated with maize streak virus (MSV) infection after agroinoculations of pCAMBIA0380::MSV-G|61|, pCAMBIA0380::MSV-A|53|, and pBC-KS::MSV-A|R2| infectious clones. Photos of whole plants showing a negative control plant (left) and infected plant (right) were taken 60 days postinoculation (dpi) (28 dpi only for pCAMBIA0380::MSV-A|53| in *O. glaberrima* 'Tog5673'), and photos of systemic and emergent leaves were taken 28 dpi. nd = not determined because no MSV-infected plant was identified.

$F = 26.254$, $P < 0.001$, ANOVA, $F_{3,128} = 13.150$, $P < 0.001$) (Supplementary Fig. S5A), we found that the height of plants was significantly reduced by MSV infection (*O. glaberrima* 'Tog5673', Kruskal-Wallis, $F = 64.973$, $P < 0.001$, ANOVA, $F_{3,128} = 47.880$, $P < 0.001$) (Fig. 5). In addition, the evaluation of plant fresh weight during the second experiment suggested that MSV infection had an impact on biomass production in rice (*O. glaberrima* 'Tog5673', Kruskal-Wallis, $F = 29.633$, $P < 0.001$, ANOVA, $F_{2,116} = 12.060$, $P = 0.001$) (Supplementary Fig. S5B).

In parallel, the infectious clone pBC-KS::MSV-A|R2| was agroinoculated in maize as a positive control, and one symptomatic and MSV-infected plant was obtained (one of five plants) (Fig. 4; Table 2). The emergence of typical symptoms of MSV in maize was observed 9 dpi, and MSV-A|R2| accumulation was validated by PCR and Sanger sequencing (data not shown). By contrast, only one maize plant of 68 (1.5%) agroinoculated with pCAMBIA0380::MSV-A|53| started to develop typical symptoms of MSV 21 dpi (Table 2; Supplementary Table S6). Successful infection was not obtained in maize with pCAMBIA0380::MSV-G|61| agroinoculation.

Field observations of symptoms putatively associated with MSV infection. As phenotypes and symptoms of collected plants in rice fields and pictures of plants were recorded during the sampling process, we tried a posteriori to associate the results of MSV detection with the presence of symptoms that could be associated with MSV infection (based on the results of the experimental MSV agroinoculations). We first noted that "white stripes" were frequently observed in rice plants in 2018 and 2019, whereas this was not reported in 2016 and 2017 (Supplementary Table S2), which could align with the higher frequency of MSV-positive fields the last 2 years of the survey.

No significant association between white stripe phenotypes and MSV-positive samples was observed at the field level ($\chi^2 = 1.442$, $P = 0.230$ for the 2016 to 2019 period and $\chi^2 = 1.199$, $P = 0.274$ and $\chi^2 = 2.063$, $P = 0.151$ for 2018 and 2019, respectively) or at the individual plant level based on plant samples in 2018 from 12 fields used to estimate MSV prevalence, among which six had white stripe symptoms and 51 were positive for MSV ($\chi^2 = 0.919$, $P = 0.340$) (Supplementary Table S5). The only significant association between symptoms and MSV-positive samples was noted in the BL site at the field level (BL, $\chi^2 = 8.864$, $P = 0.003$). Nevertheless, we identified four symptomatic rice plants in fields that we confirmed a posteriori to be infected with MSV-G (detection PCR and Sanger sequencing) (Supplementary Fig. S2).

DISCUSSION

By combining epidemiological field survey of multiple known rice diseases (Barro et al. 2021a), a viral metagenomics approach (VANA-Illumina) (Moubset et al. 2022), molecular epidemiology, and experimental infections, we have described for the first time the epidemiology, genetic diversity, and pathogenicity of MSV in rice (Table 1). Indeed, although MSV has a very large host range, infecting dozens of plant species of the *Poaceae* family (Kraberger

TABLE 2
Total number of asymptomatic (–) and symptomatic (+) plants observed 28 days postinoculation[a]

	Oryza sativa indica 'IR64'			*Oryza glaberrima* 'Tog5673'			*Zea mays* 'Golden Bantam'		
Infectious clone	–	+	% (+)	–	+	% (+)	–	+	% (+)
pCAMBIA0380::Ø	16	0	0.0	20	0	0.0	3	0	0.0
pCAMBIA0380::MSV-G|61|	198	3	1.5	185	20	9.8	84	0	0.0
pCAMBIA0380::MSV-A|53|	127	0	0.0	110	16	12.7	67	1	1.5
pBC-KS::MSV-A|R2|	49	1	2.0	55	6	9.8	4	1	20.0
Total MSV-G/A	374	4	1.1	350	42	10.7	155	2	1.3

[a] pCAMBIA0380::Ø (negative control), pCAMBIA0380::MSV-G|61|, pCAMBIA0380::MSV-A|53|, and pBC-KS::MSV-A|R2| in *O. sativa indica* 'IR64', *O. glaberrima* 'Tog5673', and *Z. mays* 'Golden Bantam'. Percentages of symptomatic plants are indicated with % (+). MSV = maize streak virus.

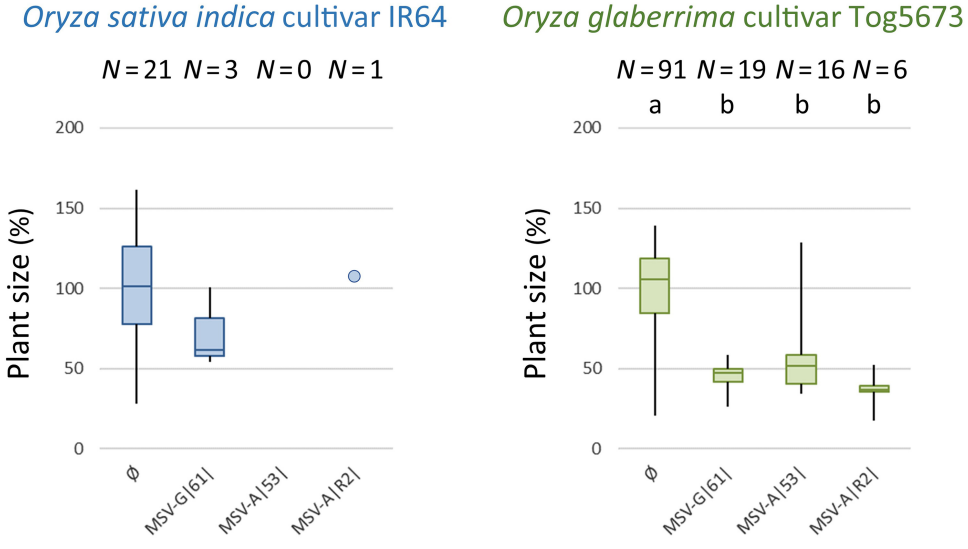

Fig. 5. Maximal height of maize streak virus (MSV)-infected plants (inoculated with pCAMBIA0380::MSV-G|61|, pCAMBIA0380::MSV-A|53|, and pBC-KS::MSV-A|R2|) normalized according to negative control plants (inoculated with pCAMBIA0380::Ø) for *Oryza sativa indica* 'IR64' (blue) and *O. glaberrima* 'Tog5673' (green). The numbers of plants used to obtain these data are indicated. Statistically identical groups are shown (a, b) for assays with *O. glaberrima* 'Tog5673'. The value obtained for the unique IR64 cultivar infected after agroinoculation with pBC-KS::MSV-A|R2| is indicated by a blue dot.

et al. 2017), and has been shown experimentally to be transmitted to rice by leafhoppers (Damsteegt 1983; Konaté and Traoré 1992), no study to date has demonstrated the extent of the MSV epidemic in rice fields from Africa.

Based on rice and wild grass samples collected in 2016 and 2017, sequences sharing identity with MSV strain G were identified from two pools of rice plants and one pool of wild plants using the VANA-Illumina approach. Although the VANA-Illumina approach is extremely valuable for estimating the diversity of virus populations and for expanding the knowledge of virus species circulating within the environment (Moubset et al. 2022), the comparison of MSV detection by VANA-Illumina and RCA-PCR suggested that the detection threshold was lower with RCA-PCR compared with VANA-Illumina. Indeed, although MSV was detected in one rice field by both methods, seven additional rice fields were identified as positive for MSV in 2016 and 2017 by RCA-PCR (Supplementary Table S2). Conversely, we noted that one field was positive for MSV by the VANA-Illumina approach but not by RCA-PCR. As the plant material used for these two approaches was not strictly identical (independent virion or nucleic acid extractions), we can assume that some discrepancy may be generated between extractions and thus MSV detection if the virus is not uniformly distributed within infected plants. In addition, analysis by pooling several plants together may also weaken our ability to consistently detect MSV in rice and wild grass samples. Using both approaches, we can consider that the percentage of rice fields in which MSV was circulating was potentially underestimated in this study.

Despite all of these putative biases of detection, MSV was surprisingly frequent in rice landscapes of western Burkina Faso, especially in 2018 and 2019 (Table 1; Supplementary Table S2). Indeed, the frequency of MSV-positive fields was similar to the frequency of symptoms caused by another well-known virus, RYMV (*Solemoviridae*), in these same fields during the same time period (34.3 versus 30.2%) (Barro et al. 2021a). In addition, similar to RYMV, the production system (IR versus RL) did not have a significant effect on the percentage of MSV-positive fields (Table 1). However, we noted a significant variation in the frequency of MSV-positive fields between years, which could be due to variations in climatic conditions and insect vector populations (as suggested for MSV epidemiology in maize cropping areas from Réunion Island) (Reynaud et al. 2009).

Two MSV strains, MSV-G and MSV-A, were identified in rice fields, and MSV-G was significantly more frequent than MSV-A (Supplementary Fig. S2; Supplementary Table S2). As far as we know, MSV-G has been identified only in wild grasses, whereas MSV-A has been reported in both wild grasses and cultivated *Poaceae*, such as maize and sugarcane (Kraberger et al. 2017). Although specific primers were used in this study to detect MSV-G and MSV-A by RCA-PCR, the MSV genetic diversity assessed in rice landscapes from southwestern Burkina Faso was similar to that estimated in West Africa. Nevertheless, the use of specific primers targeting these two MSV strains could underestimate the genuine diversity of this virus in our study area, even if no other MSV strain has been identified by the VANA-Illumina metagenomics approach.

During this study, MSV-A was identified in only two rice fields (Supplementary Table S2), in which rice and maize were cultivated alongside one another or in rotation (data not shown). Interestingly, we did not detect genetic differentiation between the MSV isolates collected from rice and other wild or cultivated *Poaceae* growing around the rice fields, suggesting that MSV-G and MSV-A circulate similarly between these host plant species. Nevertheless, we noted in our study that MSV-G was detected only in wild *Poaceae* samples regardless of potential symptoms, whereas MSV-A was identified only in plants specifically collected because they displayed symptoms of viral infection (Supplementary Table S3). These results suggest that MSV-A could be more aggressive (i.e., inducing more symptoms) than MSV-G in wild *Poaceae*. In parallel, these results imply that the prevalence and spatial distribution of MSV-G could be underestimated because of more limited symptom induction compared with MSV-A and a lower likelihood of being collected during plant pathology studies, especially within the wild compartment (Lefeuvre et al. 2019).

Based on 12 MSV-positive rice fields surveyed in 2018, we estimated that MSV prevalence within these fields was 32.1% on average and that MSV-A was overall less prevalent than MSV-G (6.3 versus 18.8 to 62.5%) (Fig. 3; Supplementary Table S5). It should be noted that, as previously discussed, the high threshold of MSV detection at the field level (i.e., pool of 16 plant leaves from each field) could imply that only fields with a high proportion of MSV-infected plants were found to be positive by RCA-PCR and thus that the average prevalence may be overestimated. Nevertheless, the analysis of these 12 fields revealed the impressive MSV prevalence in rice fields from Burkina Faso in 2018, with no significant difference between sites or rice production systems. These results suggest that MSV can circulate efficiently within the environment, probably because of the flight performance of the insect vector (Asanzi et al. 1995) and the absence of varietal differentiation between IR and RL rice fields in this area (Barro et al. 2021b).

We used agroinoculations to validate the pathogenicity of MSV-G and MSV-A isolates. Both MSV strains are able to infect both rice species cultivated in Africa (*O. sativa* and *O. glaberrima*). Although the percentage of successful infection was limited (1.1 and 10.7% in *O. sativa indica* 'IR64' and *O. glaberrima* 'Tog5673', respectively) (Table 2; Supplementary Table S6), MSV infection induced severe symptoms (Fig. 4), generally associated with a significant reduction in the size (Fig. 5) and biomass production (Supplementary Fig. S5B) of infected rice plants. The efficiency of agroinoculation in causing infection was similar between MSV-G and MSV-A in each rice variety/species (Table 2). However, the percentage of successful infection with MSV-G, MSV-A, or both strains combined was significantly higher in *O. glaberrima* 'Tog5673' than *O. sativa indica* 'IR64'. Thus, these results suggest that, despite the restricted efficiency of the agroinoculation method to transmit the virus to rice and the restricted number of rice varieties used in this study, the fitness of MSV could be higher in *O. glaberrima* than *O. sativa*. Further studies involving more cultivars of each rice species and insect vectors are, however, required to confirm this hypothesis.

Interestingly, symptoms similar to those induced by experimental MSV agroinoculation were observed in rice fields (Supplementary Tables S2 and S5). However, no significant a posteriori association was identified between these symptoms and MSV detection by PCR at field or plant level. This lack of association could be due to several factors, such as the low detection threshold of MSV by PCR at the field level (see earlier discussion), the rice varieties used in Burkina Faso as well as the environmental conditions that could reduce the intensity of symptoms in fields, or the lack of knowledge of specific symptoms of MSV infection by fieldworkers for rice multipathogen surveys associated with a high prevalence of other diseases (70.3% of individual plants analyzed presented disease symptoms) (Supplementary Table S5) and possible confusion with symptoms induced by other viral infections circulating in these areas (Barro et al. 2021a; Sereme et al. 2014). Nevertheless, the association between symptoms, MSV detection, and MSV sequencing was validated with four plants that were specifically collected because they showed symptoms that could be attributed to viral infection (Supplementary Fig. S2). It should be noted that the BD10 PCR amplifications, which were used as an internal control for DNA extraction quality (see earlier discussion), showed that

all rice samples collected in Burkina Faso for this study exclusively corresponded to *O. sativa* varieties, which is concordant with previous results from the same area (Barro et al. 2021b). In the future, the use of quick and sensitive diagnostic tools for the detection of MSV directly in rice fields, such as loop-mediated isothermal amplification assays (Tembo et al. 2020), will help to unravel the association between MSV infection and symptoms in rice and will improve our understanding of MSV epidemiology in rice fields from Africa.

MSV agroinoculation in maize led to successful infection with pBC-KS::MSV-A|R2| and pCAMBIA0380::MSV-A|53| infectious clones. The isolate MSV-A|R2|, collected in Réunion Island and cloned after serial passages in almost resistant inbred maize lines, was found to be one of the most pathogenic isolates analyzed by Isnard et al. (1998). The isolate MSV-A|53|, collected from rice in Burkina Faso (16BZ09) (Supplementary Table S2), was also able to infect maize. These results demonstrate the ability of these isolates to infect both maize and rice regardless of the host plant of origin, geographical area of origin, or genetic divergence (Supplementary Fig. S4).

Interestingly, we never detected MSV-G infection in maize during our agroinoculation assays (84 plants in total) (Table 2). If this trend is confirmed (especially by assays involving insect vectors to increase MSV transmission efficiency), the asymmetric pattern of infection between MSV-G and MSV-A in rice and maize could shed light on the evolutionary history and host adaptation of MSV. Indeed, MSV-G has thus far been identified only in wild grasses from West Africa, whereas intensive analyses of MSV diversity in cultivated *Poaceae*, such as maize, sorghum, and sugarcane, have been performed (Kraberger et al. 2017). The results obtained in our study suggest the role of rice fields as boosters of MSV-G epidemics in the environment because of the homogeneity, density, and spatial distribution of this crop (Anderson et al. 2004). Thus, as rice and maize are frequently grown within the same fields or areas in Africa, rice could have played an indirect role in the adaptation of MSV to maize by increasing the probability of co-infection with MSV-G and MSV-B, which subsequently led to recombination and the emergence of strain MSV-A (Harkins et al. 2009).

Altogether, in addition to fulfilling Koch's postulates via experimental MSV inoculations in two rice species cultivated in Africa, this study suggests that MSV could be a significant pathogen with regard to rice cultivation in Africa, as it was found to be highly prevalent in rice fields from Burkina Faso and induced severe symptoms in rice plants in controlled conditions. Until our study, MSV-G had been detected only in wild grasses from West Africa, including Nigeria and Mali. As these two countries correspond to the most important rice producers in West Africa according to 2022 statistics from FAOSTAT (https://data.un.org/Data.aspx?d=FAO&f=itemCode%3A27), we can assume that MSV-G could be epidemic in these countries. In addition, we experimentally confirmed that two genetically distinct MSV-A isolates collected from rice in West Africa or maize in East Africa are able to infect both rice species (Fig. 4; Supplementary Fig. S4). Thus, as MSV-A is reported almost everywhere in Africa (Kraberger et al. 2017) where rice is produced on a large scale—for instance, West Africa (Benin, Burkina Faso, Ghana, Nigeria), Central Africa (Cameroon, Central African Republic, Chad), and East Africa (Ethiopia, Kenya, Madagascar, Tanzania, Uganda)—our results suggest lurking MSV epidemics in rice across the continent. Therefore, our study represents a necessary initial step in characterizing a pathogenic virus in rice. This study will allow MSV to be considered by epidemiological surveillance networks composed of producers, agronomists, and phytopathologists dedicated to rice, which will then provide more information on the distribution, diversity, and impact of this virus on rice production in Africa.

ACKNOWLEDGMENTS

This work was performed thanks to the facilities of the "International joint Laboratory LMI PathoBios: Observatory of plant pathogens in West Africa: biodiversity and biosafety" (www.pathobios.com; www.twitter.com/PathoBios). We thank Sylvain Zougrana, Yacouba Kone, Edouard Kabore, Moustapha Koala, Manaka Douanio, Amadou Diallo, Fabrice Nikiema, Bouda Zakaria, Ouedraogo Alassane, Fatoumata Gnacko, Daouda Hema, Drissa Tou, Traoré Momouni, Dabire Roméo, Martial Kabore, and Noël Ouattara for their contributions to the fieldwork in Burkina Faso; the rice farmers from Badala, Bama, Senzon, Banzon, Tengrela, and Karfiguela for their kind collaboration; Michel Peterschmitt, Cica Urbino, and Martine Granier for the valuable discussions and for kindly providing us the infectious clone pBC-KS::MSV-A(R2); and Jamel Aribi for plant care in the greenhouses of l'Institut de Recherche pour le Développement.

LITERATURE CITED

Abo, M. E., and Sy, A. A. 1997. Rice virus diseases: Epidemiology and management strategies. J. Sustain. Agric. 11:113-134.

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., and Daszak, P. 2004. Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. Trends Ecol. Evol. 19:535-544.

Asanzi, C. M., Bosque-Perez, N. A., and Nault, L. R. 1995. Movement of *Cicadulina storeyi* (Homoptera: Cicadellidae) in maize fields and its behaviour in relation to maize growth stage. Int. J. Trop. Insect Sci. 16: 39-44.

Bagayoko, I., Celli, M. G., Romay, G., Poulicard, N., Pinel-Galzi, A., Julian, C., Filloux, D., Roumagnac, P., Sérémé, D., Bragard, C., and Hébrard, E. 2021. Genetic diversity of *Rice stripe necrosis virus* and new insights into evolution of the genus *Benyvirus*. Viruses 13:737.

Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L.-F., Wesolowski, A., and Metcalf, C. J. E. 2022. Infectious disease in an era of global change. Nat. Rev. Microbiol. 20:193-205.

Barro, M., Kassankogno, A. I., Wonni, I., Sérémé, D., Somda, I., Kaboré, H. K., Béna, G., Brugidou, C., Tharreau, D., and Tollenaere, C. 2021a. Spatiotemporal survey of multiple rice diseases in irrigated areas compared to rainfed lowlands in the western Burkina Faso. Plant Dis. 105:3889-3899.

Barro, M., Konate, K. A., Wonni, I., Kassankogno, A. I., Sabot, F., Albar, L., Somda, I., Béna, G., Ghesquière, A., Kam, H., Sié, M., Cubry, P., and Tollenaere, C. 2021b. Assessment of genetic diversity of rice in registered cultivars and farmers' fields in Burkina Faso. Crops 1:129-140.

Barro, M., Wonni, I., Simonin, M., Kassankogno, A. I., Klonowska, A., Moulin, L., Béna, G., Somda, I., Brunel, C., and Tollenaere, C. 2022. The impact of the rice production system (irrigated vs lowland) on root-associated microbiome from farmer's fields in western Burkina Faso. FEMS Microbiol. Ecol. 98:fiac085.

Bernardo, P., Charles-Dominique, T., Barakat, M., Ortet, P., Fernandez, E., Filloux, D., Hartnady, P., Rebelo, T. A., Cousins, S. R., Mesleard, F., Cohez, D., Yavercovski, N., Varsani, A., Harkins, G. W., Peterschmitt, M., Malmstrom, C. M., Martin, D. P., and Roumagnac, P. 2018. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. ISME J. 12:173-184.

Billard, E., Barro, M., Sérémé, D., Bangratz, M., Wonni, I., Koala, M., Kassankogno, A. I., Hébrard, E., Thébaud, G., Brugidou, C., Poulicard, N., and Tollenaere, C. 2023. 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. Virus Evol. 9:vead049.

Boulton, M. I., Buchholz, W. G., Marks, M. S., Markham, P. G., and Davies, J. W. 1989. Specificity of *Agrobacterium*-mediated delivery of maize streak virus DNA to members of the Gramineae. Plant Mol. Biol. 12:31-40.

Claverie, S., Ouattara, A., Hoareau, M., Filloux, D., Varsani, A., Roumagnac, P., Martin, D. P., Lett, J.-M., and Lefeuvre, P. 2019. Exploring the diversity of Poaceae-infecting mastreviruses on Reunion Island using a viral metagenomics-based approach. Sci. Rep. 9:12716.

Cubry, P., Tranchant-Dubreuil, C., Thuillet, A.-C., Monat, C., Ndjiondjop, M.-N., Labadie, K., Cruaud, C., Engelen, S., Scarcelli, N., Rhoné, B.,

Burgarella, C., Dupuy, C., Larmande, P., Wincker, P., François, O., Sabot, F., and Vigouroux, Y. 2018. The rise and fall of African rice cultivation revealed by analysis of 246 new genomes. Curr. Biol. 28:2274-2282.e6.

Damsteegt, V. D. 1983. Maize streak virus: I. Host range and vulnerability of maize germ plasm. Plant Dis. 67:734-737.

Demont, M. 2013. Reversing urban bias in African rice markets: A review of 19 national rice development strategies. Glob. Food Sec. 2:172-181.

Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Edgar, R. C., Taylor, B., Lin, V., Altman, T., Barbera, P., Meleshko, D., Lohr, D., Novakovsky, G., Buchfink, B., Al-Shayeb, B., Banfield, J. F., de la Peña, M., Korobeynikov, A., Chikhi, R., and Babaian, A. 2022. Petabase-scale sequence alignment catalyses viral discovery. Nature 602:142-147.

Excoffier, L., Laval, G., and Schneider, S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol. Bioinform. 1:47-50.

Fiallo-Olivé, E., Lett, J.-M., Martin, D. P., Roumagnac, P., Varsani, A., Zerbini, F. M., and Navas-Castillo, J. 2021. ICTV virus taxonomy profile: *Geminiviridae* 2021. J. Gen. Virol. 102:001696.

François, S., Filloux, D., Frayssinet, M., Roumagnac, P., Martin, D. P., Ogliastro, M., and Froissart, R. 2018. Increase in taxonomic assignment efficiency of viral reads in metagenomic studies. Virus Res. 244:230-234.

Fuller, C. 1901. Mealie variegation. Pages 17-19 in: First Report of the Government Entomologist, 1899-1900. P. Davis & Sons, Government Printers, Pietermaritzburg, South Africa.

Gnacadja, C., Berthouly-Salazar, C., Sall, S. N., Zekraoui, L., Sabot, F., Pegalepo, E., Manneh, B., Vieira-Dalode, G., Moreira, J., Soumanou, M. M., Azokpota, P., and Sie, M. 2018. Phenotypic and genetic characterization of African rice (*Oryza glaberrima* Steud). Int. J. Adv. Res. 6:1389-1398.

Gong, P., Tan, H., Zhao, S., Li, H., Liu, H., Ma, Y., Zhang, X., Rong, J., Fu, X., Lozano-Durán, R., Li, F., and Zhou, X. 2021. Geminiviruses encode additional small proteins with specific subcellular localizations and virulence function. Nat. Commun. 12:4278.

Gouy, M., Guindon, S., and Gascuel, O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27:221-224.

Greninger, A. L. 2018. A decade of RNA virus metagenomics is (not) enough. Virus Res. 244:218-229.

Grimsley, N., Hohn, T., Davies, J. W., and Hohn, B. 1987. *Agrobacterium*-mediated delivery of infectious maize streak virus into maize plants. Nature 325:177-179.

Harkins, G. W., Martin, D. P., Duffy, S., Monjane, A. L., Shepherd, D. N., Windram, O. P., Owor, B. E., Donaldson, L., van Antwerpen, T., Sayed, R. A., Flett, B., Ramusi, M., Rybicki, E. P., Peterschmitt, M., and Varsani, A. 2009. Dating the origins of the maize-adapted strain of maize streak virus, MSV-A. J. Gen. Virol. 90:3066-3074.

Hébrard, E., Poulicard, N., and Rakotomalala, M. 2021. Rice yellow mottle virus (*Solemoviridae*). Pages 675-680 in: Encyclopedia of Virology. Elsevier, Amsterdam, the Netherlands.

Isnard, M., Granier, M., Frutos, R., Reynaud, B., and Peterschmitt, M. 1998. Quasispecies nature of three maize streak virus isolates obtained through different modes of selection from a population used to assess response to infection of maize cultivars. J. Gen. Virol. 79:3091-3099.

Jones, R. A. C. 2021. Global plant virus disease pandemics and epidemics. Plants 10:233.

Kaboré, K. H., Kassankogno, A. I., Adreit, H., Milazzo, J., Guillou, S., Blondin, L., Chopin, L., Ravel, S., Charriat, F., Barro, M., Tollenaere, C., Lebrun, M.-H., and Tharreau, D. 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. Front. Plant Sci. 13:1022348.

Konaté, G., and Traoré, O. 1992. Les hôtes réservoirs du virus de la striure du maïs (MSV) en zone soudano-sahélienne: Identification et distribution spatio-temporelle. Phytoprotection 73:111-117.

Koonin, E. V., and Dolja, V. V. 2018. Metaviromics: A tectonic shift in understanding virus evolution. Virus Res. 246:A1-A3.

Kraberger, S., Saumtally, S., Pande, D., Khoodoo, M. H. R., Dhayan, S., Dookun-Saumtally, A., Shepherd, D. N., Hartnady, P., Atkinson, R., Lakay, F. M., Hanson, B., Redhi, D., Monjane, A. L., Windram, O. P., Walters, M., Oluwafemi, S., Michel-Lett, J., Lefeuvre, P., Martin, D. P., and Varsani, A. 2017. Molecular diversity, geographic distribution and host range of monocot-infecting mastreviruses in Africa and surrounding islands. Virus Res. 238:171-178.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35:1547-1549.

Lefeuvre, P., Martin, D. P., Elena, S. F., Shepherd, D. N., Roumagnac, P., and Varsani, A. 2019. Evolution and ecology of plant viruses. Nat. Rev. Microbiol. 17:632-644.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 1:vev003.

Martin, D. P., and Shepherd, D. N. 2009. The epidemiology, economic impact and control of maize streak disease. Food Secur. 1:305-315.

Monjane, A. L., Dellicour, S., Hartnady, P., Oyeniran, K. A., Owor, B. E., Bezuidenhout, M., Linderme, D., Syed, R. A., Donaldson, L., Murray, S., Rybicki, E. P., Kvarnheden, A., Yazdkhasti, E., Lefeuvre, P., Froissart, R., Roumagnac, P., Shepherd, D. N., Harkins, G. W., Suchard, M. A., Lemey, P., Varsani, A., and Martin, D. P. 2020. Symptom evolution following the emergence of maize streak virus. eLife 9:e51984.

Monjane, A. L., Harkins, G. W., Martin, D. P., Lemey, P., Lefeuvre, P., Shepherd, D. N., Oluwafemi, S., Simuyandi, M., Zinga, I., Komba, E. K., Lakoutene, D. P., Mandakombo, N., Mboukoulida, J., Semballa, S., Tagne, A., Tiendrébéogo, F., Erdmann, J. B., van Antwerpen, T., Owor, B. E., Flett, B., Ramusi, M., Windram, O. P., Syed, R., Lett, J.-M., Briddon, R. W., Markham, P. G., Rybicki, E. P., and Varsani, A. 2011. Reconstructing the history of maize streak virus strain A dispersal to reveal diversification hot spots and its origin in Southern Africa. J. Virol. 85:9623-9636.

Moubset, O., François, S., Maclot, F., Palanga, E., Julian, C., Claude, L., Fernandez, E., Rott, P., Daugrois, J.-H., Antoine-Lorquin, A., Bernardo, P., Blouin, A. G., Temple, C., Kraberger, S., Fontenele, R. S., Harkins, G. W., Ma, Y., Marais, A., Candresse, T., Chéhida, S. B., Lefeuvre, P., Lett, J.-M., Varsani, A., Massart, S., Ogliastro, M., Martin, D. P., Filloux, D., and Roumagnac, P. 2022. Virion-associated nucleic acid-based metagenomics: A decade of advances in molecular characterization of plant viruses. Phytopathology 112:2253-2272.

Muhire, B. M., Varsani, A., and Martin, D. P. 2014. SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9:e108277.

Oyeniran, K. A., Hartnady, P., Claverie, S., Lefeuvre, P., Monjane, A. L., Donaldson, L., Lett, J.-M., Varsani, A., and Martin, D. P. 2021. How virulent are emerging maize-infecting mastreviruses? Arch. Virol. 166:955-959.

Peterschmitt, M., Granier, M., Frutos, R., and Reynaud, B. 1996. Infectivity and complete nucleotide sequence of the genome of a genetically distinct strain of maize streak virus from Reunion Island. Arch. Virol. 141:1637-1650.

Portères, R. 1970. Primary cradles of agriculture in the African continent. Pages 43-58 in: Papers in African Prehistory. Cambridge University Press, Cambridge, U.K.

Reynaud, B., Delatte, H., Peterschmitt, M., and Fargette, D. 2009. Effects of temperature increase on the epidemiology of three major vector-borne viruses. Eur. J. Plant Pathol. 123:269-280.

Rybicki, E. P. 2015. A top ten list for economically important plant viruses. Arch. Virol. 160:17-20.

Rybicki, E. P., and Pietersen, G. 1999. Plant virus disease problems in the developing world. Pages 127-175 in: Advances in Virus Research. Elsevier, Amsterdam, Netherlands.

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. 2019. The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 3:430-439.

Sereme, D., Neya, B. J., Bangratz, M., Brugidou, C., and Ouedraogo, I. 2014. First report of *Rice stripe necrosis virus* infecting rice in Burkina Faso. Plant Dis. 98:1451.

Shepherd, D. N., Martin, D. P., Van Der Walt, E., Dent, K., Varsani, A., and Rybicki, E. P. 2010. Maize streak virus: An old and complex 'emerging' pathogen. Mol. Plant Pathol. 11:1-12.

Soullier, G., Demont, M., Arouna, A., Lançon, F., and Mendez del Villar, P. 2020. The state of rice value chain upgrading in West Africa. Glob. Food Sec. 25:100365.

Tembo, M., Adediji, A. O., Bouvaine, S., Chikoti, P. C., Seal, S. E., and Silva, G. 2020. A quick and sensitive diagnostic tool for detection of *Maize streak virus*. Sci. Rep. 10:19633.

Tra Bi, S. C., Dje, K. T., Coulibaly, T., Soumahoro, B., and Tano, Y. 2020. Entomofaune du riz (*Oryza sativa* L.) en fonction des stades phénologiques dans un bas-fond, Daloa, Côte d'Ivoire. Afr. Sci. 16:98-113.

Urbino, C., Thébaud, G., Granier, M., Blanc, S., and Peterschmitt, M. 2008. A novel cloning strategy for isolating, genotyping and phenotyping genetic variants of geminiviruses. Virol. J. 5:135.

Varsani, A., Shepherd, D. N., Monjane, A. L., Owor, B. E., Erdmann, J. B., Rybicki, E. P., Peterschmitt, M., Briddon, R. W., Markham, P. G., Oluwafemi, S., Windram, O. P., Lefeuvre, P., Lett, J.-M., and Martin, D. P. 2008. Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. J. Gen. Virol. 89:2063-2074.

Wang, P., Liu, J., Lyu, Y., Huang, Z., Zhang, X., Sun, B., Li, P., Jing, X., Li, H., and Zhang, C. 2022. A review of vector-borne rice viruses. Viruses 14: 2258.