Cassava brown streak virus evolves with a nucleotide-substitution rate that is typical for the family *Potyviridae*

Willard Mbewe [a,*], Settumba Mukasa [b], Mildred Ochwo-Ssemakula [b], Peter Sseruwagi [c], Fred Tairo [c], Joseph Ndunguru [c], Siobain Duffy [d,*]

[a] *Department of Biological Sciences, Malawi University of Science and Technology, P. O. Box 5196, Limbe, Malawi*
[b] *School of Agriculture and Environmental Science, Department of Agricultural Production, P. O. Box 7062, Makerere University, Kampala, Uganda*
[c] *Mikocheni Agricultural Research Institute, P.O. Box 6226, Dar es Slaam, Tanzania*
[d] *Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08901, United States*

ARTICLE INFO

ABSTRACT

The ipomoviruses (family *Potyviridae*) that cause cassava brown streak disease (cassava brown streak virus [CBSV] and Uganda cassava brown streak virus [UCBSV]) are damaging plant pathogens that affect the sustainability of cassava production in East and Central Africa. However, little is known about the rate at which the viruses evolve and when they emerged in Africa – which inform how easily these viruses can host shift and resist RNAi approaches for control. We present here the rates of evolution determined from the coat protein gene (CP) of CBSV (Temporal signal in a UCBSV dataset was not sufficient for comparable analysis). Our BEAST analysis estimated the CBSV CP evolves at a mean rate of 1.43×10^{-3} nucleotide substitutions per site per year, with the most recent common ancestor of sampled CBSV isolates existing in 1944 (95% HPD, between years 1922 – 1963). We compared the published measured and estimated rates of evolution of CPs from ten families of plant viruses and showed that CBSV is an average-evolving potyvirid, but that members of *Potyviridae* evolve more quickly than members of *Virgaviridae* and the single representatives of *Betaflexiviridae*, *Bunyaviridae*, *Caulimoviridae* and *Closteroviridae*.

1. Introduction

Cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV) are the two viruses implicated in etiology of cassava brown streak disease (CBSD) (Monger et al., 2001; Mbanzibwa et al., 2009; Winter et al., 2010), a threat to food security to more than 300 million people in Africa (Patil et al., 2015; Robson et al., 2023). These viruses belong to the genus *Ipomovirus* and family *Potyviridae* (Inoue-Nagata et al., 2022). These viruses are transmitted by whiteflies (most often *Bemisia tabaci* (Gennadius) (Maruthi et al., 2005) and disseminated by planting of infected cuttings.

CBSVs, like other members of the genus *Ipomovirus,* have a positive-sense single strand RNA (+ssRNA) genome that is translated into a polyprotein precursor that is cleaved by viral proteases to produce ten mature proteins denoted as P1, P3, 6K1, CI, 6K2, VPg, NIa-Pro, NIb, HAM1, and CP (Dombrovsky et al., 2014). An additional peptide, PIPO, is translated in the +2 reading frame relative to P3 via ribosomal frameshifting or transcriptional slippage at a high conserved motif at the 5′ end of PIPO (Chung et al., 2008).

The functions of CP gene of in members of *Potyviridae* (which we will informally refer to as 'potyviruses' – not just the members of genus *Potyvirus*) have been well documented (Atreya et al., 1990; Rybicki and Shukla, 1992; Varrelmann and Maiss, 2000; Dombrovsky et al., 2014). The CP is generally associated with the success of viral infection: it plays a role in vector transmission virion assembly, stability, and systematic infection, making it one of the most important genes for virus pathogenicity (Atreya et al., 1990; Varrelmann and Maiss, 2000). Thus, it is not surprising that the CP gene is frequently used to resolve phylogenetic relationships among viral isolates (Abdalla and Ali, 2021).

Although mechanisms driving faster evolutionary rates in CBSV than UCBSV were determined (Alicai et al., 2016), there is still limited knowledge on the rate of molecular evolution and the age of the sampled genetic diversity, reflected in the time to the most recent common ancestor (TMRCA). This information is critical to understanding the evolutionary fingerprint of any plant virus (Kosakovsky et al., 2010), and particularly whether they exhibit reduced rates of evolutionary

---

* Corresponding authors.
E-mail addresses: wmbewe@must.ac.mw (W. Mbewe), duffy@sebs.rutgers.edu (S. Duffy).

change, which in turn may have major implications on their ability to emerge in new host species or evade RNAi approaches for control (Prins et al., 2008; Patil et al., 2011). We undertook this study to estimate the nucleotide substitution rate of CBSV and UCBSV and calculate how recently these viruses may have emerged in cassava by the time to most recent common ancestor (TMRCA). We used CP sequences because there were more sequences of this genomic region, with reliable dates of isolation, over a wider date range, than of other genes or whole genomes in GenBank. The isolates used here in were collected in East and Central Africa from 1996 to 2019: the period from when CBSV was first demonstrated to cause CBSD (Monger et al., 2001) to date. We then compared the rate of evolution of CBSV to those previously reported for CPs of other plant viruses.

2. Materials and methods

2.1. Nucleotide sequences and alignment

Nucleotide sequences of full-length CP gene of CBSVs with known dates of isolation were downloaded from GenBank on 3rd September 2016. These sequences were from Kenya, Tanzania and Mozambique. Alignments were performed by ClustalW (Larkin et al., 2007) and manually edited using Se-Al (http://tree.bio.ed.ac.uk/software). Sequences were screened for recombination using six recombination detection programs within the RDP4 package (http://darwin.uvigo.es/rdp/rdp/html): RPD, GENCONV, MaxChi, Chimaera, Bootscan and 3Seq (Martin et al., 2015). The default detection thresholds were employed in all cases except that sequences were 'set linear'. The Bonferroni correction cutoff was set at highest acceptable P-value of 0.05. Sequences were considered recombinant when identified by at least three of the algorithms, and recombinant sequences were removed from the gene alignments. This ensured that the results have a high likelihood of being unaffected by recombination.

2.2. Molecular clock signal analysis

Maximum-Likelihood (ML) tree topologies were constructed from the alignments with the use of the best fitting nucleotide substitution model, the General Time Reversible with invariant sites and a gamma distribution (GTR+I+Γ_4, Tamura et al., 2021). We deployed an interactive regression approach to explore the association between CBSVs' CP genetic divergence through time and sampling dates by running TempEst (Rambaut et al., 2016). This allowed an evaluation of the temporal signal and assessment of how well a strict molecular clock might fit each of the data sets, as well allowed identification of potential outliers that deviated from the standard regression.

2.3. Estimation of TMRCA and epidemic reconstruction

The CBSV dataset was prepared for Bayesian MCMC analysis in BEAUti v 2.4.3, and then run in BEAST v1.10.4 (Suchard et al., 2018). We used marginal likelihoods via stepping stone sampling for three molecular clock models (strict, relaxed lognormal, and random local) and three demographic models (constant, exponential and Bayesian skyline) (Baele et al., 2016). After these 9 models were run for 100 million steps and evaluated, the best fitting model was run for 300 million steps (ensuring all parameters reached an ESS value ≥200). A control without sequence data was also run to ensure the priors did not independently determine the results of the analysis. Uncertainty in parameter estimates were reflected in the 95% highest probability density (HPD) values. Convergence in each BEAST run was assessed in Tracer v 1.6.0 (http://tree.bio.ed.ac.uk/software/tracer). The log and tree file for the relaxed lognormal clock with Bayesian skyline population priors were used for the reconstruction of the viral effective population size (Heled and Drummond, 2008, 2010; Parag et al., 2022). As further checks of the temporal signal in the dataset, we ran an isochronous version of the dataset (sequences identical, but all dated to the same year) with the best fitting model priors, and calculated the log Bayes Factor to determine if the tip-dated dataset was a better fit than the isochronous dataset (Duchene et al., 2020), and a Mantel test was run (using R scripts from Murray et al. 2016) to examine the degree of confounding between sampling date and phylogenetic clustering.

2.4. Selection analysis

Processes of natural selection over time mold every gene into a mosaic of sites that evolve rapidly or resist change, often captured in the site-specific ratio of non-synonymous to synonymous substitution rates (dN/dS, Kosakovsky et al., 2010). We deployed per-codon analyses using the following codon-based ML algorithms implemented in the web-server DATAMONKEY (http://www.datamonkey.org) (Kosakovsky and Frost, 2005): single-likelihood ancestor counting (SLAC) method, internal fixed method effects method (IFEL), Fixed Effects Likelihood (FEL) (Kosakovsky and Frost 2005; Massingham and Goldman, 2005) and the recently described Mixed Effects Model of Evolution (MEME) and Fast Unbiased Bayesian Approximation (FUBAR) (Murrell et al., 2013). All these methods incorporated the GTR model of nucleotide substitution, with phylogenetic trees inferred using the neighbor-joining method.

2.5. Comparison with other plant virus CP substitution rates

A survey of the literature revealed 42 published estimated substitution rates of plant virus coat/capsid proteins, mostly obtained with similar methods as in this study (i.e., BEAST analysis). All averages were calculated on log-transformed values. Statistical analyses and plotting were done in MS Excel.

3. Results

3.1. Temporal signal for UCBSV and CBSV

TempEst's molecular clock analysis of CBSV and UCBSV data sets placed the estimated TMRCA, based on crude root-to-tip regressions, at 1983 and 1575, respectively. The R^2 values were 0.23 for CBSV and 0.06 for UCBSV (Supplementary Figure 1a and 1b). Thus, the temporal signal of the UCBSV dataset was not well supported and consequently no further analyses were conducted with the UCBSV dataset. The final CBSV alignment, which did show temporal signal, is available as supplementary file S1.

3.2. Evolutionary history of CBSV

All 136 CP gene sequences of CBSV were included in the analyses as no recombination within the CP gene was detected. Bayesian coalescent methods were used to determine the rate of evolution and TMRCA were estimated using from this dataset of CBSV-CP gene sequences obtained over a period of 24 years (1996 to 2019). The TempEst results did not produce a very strong correlation between genetic divergence ($r = 0.48$) and time, and stepping stone sampling indicated that a relaxed molecular clock with a Bayesian Skyline demographic prior was the most appropriate model to use (MLE = −5605.34), which was a similar to that for the random local clock with Bayesian skyline demographic prior (MLE = −5605.86); other combinations of clock and demographic models were between −5615.06 and −5628.87. This tip-dated, heterochronous analysis was a much better fit to the data than a control analysis that examined the data as if all 136 sequences were isolated in the same year (log Bayes Factor = 349), indicating substantial temporal signal in our CBSV dataset. The mean rate of nucleotide substitution for relaxed log normal clock and Bayesian skyline prior is 1.43×10^{-3} nucleotide substitutions/site/year (n/s/y, 95% HPD, 1.4×10^{-4} – 8.86×10^{-3}). The TMRCA for the CBSV dataset was the year 1944 (95% HPD,

between years 1922 to 1963). The maximum clade credibility tree is shown in Fig. 1. The population change over time is visualized in a Bayesian skyline plot (Fig. 2). It shows that, despite epidemic spread in recent decades, the diversity of CBSV sequences declined between 2005 and 2015.

To further assess the temporal signal in our CBSV dataset, we conducted a Bayesian estimate of temporal signal, comparing our tip-dated estimate with one where the same sequences are considered to have been sampled in the same year. The comparison between the heterochronous and isochronous analyses strongly supported temporal signal in the data set (log Bayes Factor = 349). However, it appears that a CBSV isolate's location in the phylogenetic tree (Fig. 1) is correlated with its date of isolation, as is common in studies of viruses, especially during outbreaks (Murray et al. 2016). A Mantel test showed relatively a statistically significant, but fairly low correlation between genetic clustering and date of isolation in our analysis ($r = 0.25$, $p = 0.001$). Ten replicates with a clustered permutation test showed that the majority of replicates produced similar substitution rates and TMRCAs (supplemental file 2) – consistent with confounding between sampling date and genetic sequence. This reduces confidence in our estimated substitution rate.

3.2.1. Selection pressure

There was weak support for positive section acting upon the CP of CBSV using either SLAC, IFEL, FEL, MEME and FUBAR, with most of the gene under purifying selection. The 18th and 115th codons in the alignment (the 114th and 211st codon of the complete CP gene, relative to the CP of the reference sequence NC_012698.2) were detectably under positive selection by SLAC, FEL and IFEL, FUBAR (BF $\geq$0.6) and MEME (BF= 0.1).

3.2.2. Substitution rate

The estimated CBSV CP substitution rates are within the range of those previously reported in other RNA and ssDNA plant viruses (Duffy and Holmes, 2008, 2009; Gibbs et al., 2008; Pagán et al., 2010; Pagán and Holmes, 2010; Simmons et al., 2008), and CBSV is in the middle of the rates estimated for other potyviruses (Fig. 1, Table 1). Importantly, none of these other estimated substitution rates were conducted on datasets evaluated as rigorously for temporal signal (e.g., none had conducted a BETS analysis or a clustered permutation date randomization test), and so many may be similarly confounded by clustered isolates having been sampled in the same year. Based on our literature review, potyviruses (family *Potyviridae*) are the most well-studied in terms of evolutionary rates, with 11 different viruses having estimated substitution rates for their CP gene (means ranging from 7.2×10^{-3} to 1×10^{-4} n/s/y). Three other plant virus families have at least five published substitution rates for different viruses: geminivirus mean rates ranged from $3.8 \times 10{-3}$ to 1.3×10^{-4} n/s/y, sobemovirus (family *Sobemoviridae*) mean rates ranged from $3.5 \times 10{-2}$ to 6.2×10^{-4} n/s/y and tobamovirus (family *Virgaviridae*) mean rates were more narrowly distributed from 7.8×10^{-4} to 1.4×10^{-4} n/s/y.

When analyses were conducted with BEAST, the sampling time frame, molecular clock model and demographic prior are listed. Viruses in the table: African cassava mosaic virus (ACMV) arabis mosaic virus (ArMV), beet chlorosis virus (BChV), barley yellow dwarf virus (BYDV), cucurbit aphid-borne yellows virus (CABYV), cowpea aphid-borne mosaic virus (CABMV), cauliflower mosaic virus (CaMV), cassava brown streak virus (CBSV), citrus tristeza virus (CTV), cucumber green mottle mosaic virus (CGMMV), cyrtanthus elayus virus A (CyEVA), cereal yellow dwarf virus (CYDV), digitaria streak virus (DSV), East African cassava mosaic virus (EACMV), grapevine fanleaf virus (GFLV), grapevine red blotch virus (GRBV), maize streak virus (MSV), narcissus degeneration virus (NDV), narcissus latent virus (NLV), narcissus late season yellows virus (NLSYV), narcissus yellow stripe virus (NYSV), odontoglossum ringspot virus (ORSV), papaya leaf curl disease-causing viruses (PaLCD), pepper mild mottle virus (PMMoV), potato leaf roll

virus (PLRV), papaya ringspot virus (PRSV), potato virus S (PVS), potato virus Y (PVY), ribgrass mosaic virus (RMV), rice stunt virus (RSV), rice tungro spherical virus (RTSV), rice yellow mottle virus (RYMV), soybean dwarf virus (SbDV), sugarcane yellow leaf virus (ScYLV), tobacco mild green mosaic virus (TMGMV), tobacco mosaic virus (TMV), tomato mosaic virus (ToMV), tomato ringspot virus (ToRSV), tomato severe rugose virus (ToSRV), turnip mosaic virus (TuMV), turnip yellows virus (TuYV), tomato yellow leaf curl virus (TYLCV), zucchini yellow mosaic virus (ZYMV).

Some of the variation in substitution rates is due to the different lengths of time sequences were sampled in these different analyses. Shorter time spans are known to inflate estimated substitution rates (Harrison et al., 1997; Ho et al., 2005) and in fact some published rates of evolution for CP genes of plant viruses in the literature are based on datasets so short that they are acknowledged to be unreliable, and were therefore excluded from our metanalysis (e.g., blackcurrent reversion virus and strawberry mottle virus (Thompson et al., 2014)). We still see that signature of shorter sampling times affecting some of the rates included in our literature survey, with the noticeable outlier of a low rate for GRBV calculated from a dataset of only 8 years (Fig. 4). This unusually low rate is from a multiple-species analysis, looking at the diversification of grabloviruses over a much longer time frame than the other BEAST analyses included here, and therefore isn't perfectly comparable to studies looking at just the most recent common ancestor of a single species (Thompson 2022). The unequal sampling times also makes it more difficult to compare across studies. For instance, potyviruses in this dataset evolve more quickly than tobamoviruses (one-tailed *t*-test, $p = 0.0094$), but a greater proportion of the tobamovirus studies used older specimens than the potyvirus studies. While there is a clear trend in studies using similar time spans of samples that potyviruses have higher mean substitution rates, the conclusion based on all of the data is likely compromised by the sampling time effect. Among datasets with similar sampling times, potyviruses appear to evolve more quickly (mean rate 7.6×10^{-4} n/s/y) than the single representatives from four families: citrus tristeza virus, cauliflower mosaic virus, rice stunt virus and potato virus S. Potyviruses evolve at a similar rate to geminiviruses (two-tailed *t*-test, $p = 0.15$ excluding the anomalously low GRBV rate, $p = 0.08$ when including GRBV).

4. Discussion

Our initial screening showed CBSV had sufficient temporal signal to conduct this analysis, but UCBSV did not. CBSV has been detected longer than UCBSV (Monger et al., 2001), and previous research suggests UCBSV may be a complex of several species (Ndunguru et al., 2015) – the longer sampling times for CBSV and a more unambiguous monophyly may both contribute to CBSV being a more appropriate target for phylodynamic analysis. Clustered permutation date randomization tests conducted as a control after our analysis showed that the clustering of similar sequences sampled at the same time harmed our ability to confidently estimate the substation rate of CBSV's CP. This confounding can be due to sampling alone – it is not uncommon for plant diseases to be intensively sampled in outbreaks only, or intermittently as only a few students, separated in time, are supported to study the diversity of a pathogen (Hall et al. 2016). On the other hand, this confounding may reflect the true evolution of a pathogen, such as one that is prone to selective sweeps or the effects of genetic drift. As CBSV's diversity has declined in the observed time frame, the moderate correlation between genetic clustering and date of isolation may be either a sampling artefact or due to genuinely declining effective population size.

### 4.1. CBSV species could be a recent introduction

The coalescent analysis revealed that the most recent common ancestor of CBSV existed around 1944 (95% HPD, between years 1923–1963). Although UCBSV dataset was not included in the molecular
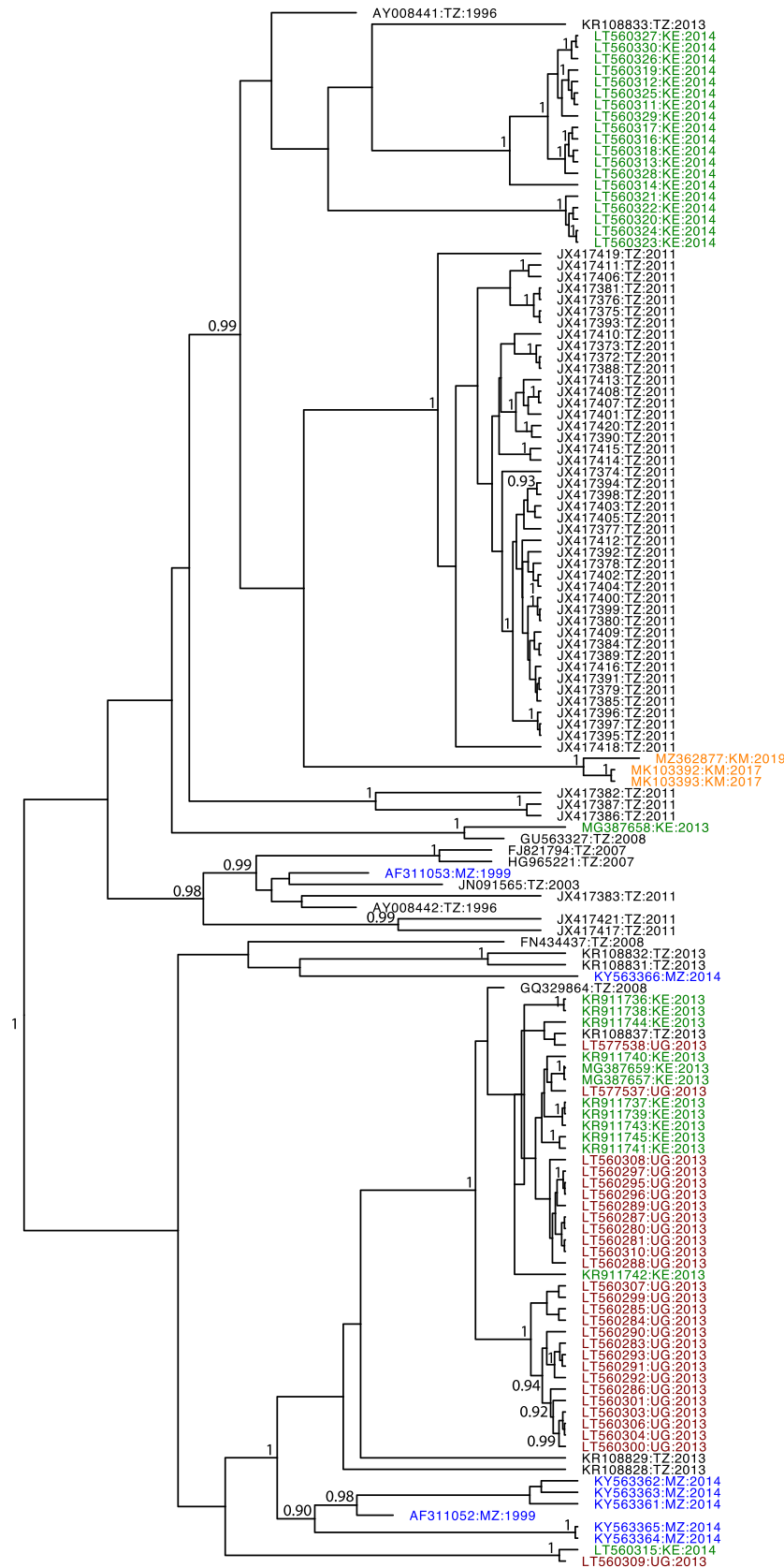
Fig. 1. The Maximum Clade Credibility (MCC) tree for the CBSV CP data set, based on a 300 million step run with a relaxed lognormal molecular clock and a Bayesian skyline demographic prior. Taxa are plotted by year of isolation. Posterior probability values above 0.9 are shown. Taxa are colored by country of isolation. The majority of the sequences are from Tanzania, and are shown in black, Kenyan isolates are shown in green, Ugandan isolates in brown, isolates from Mozambican isolates are shown in blue, and isolates from Comoros are shown in orange.
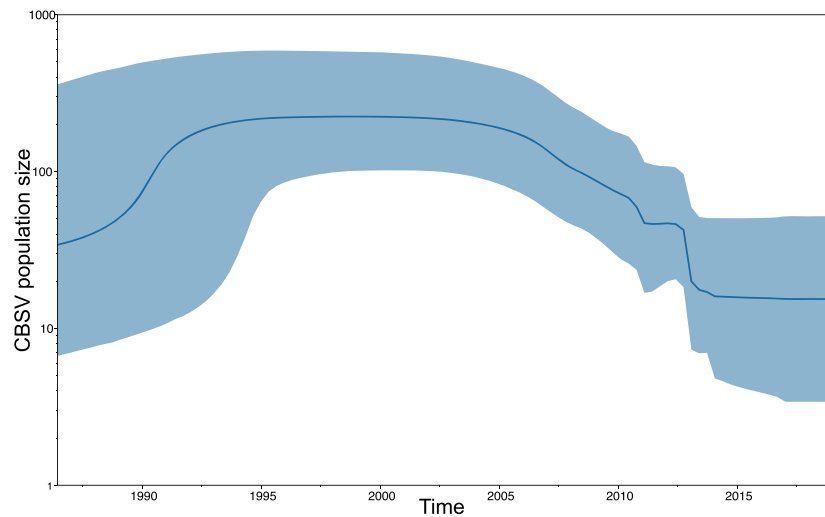
Fig. 2. Bayesian reconstruction of the effective population size of CBSV showing viral population dynamics over time. Sequences used in the tip-dated analysis were sampled from 1996 to 2019.
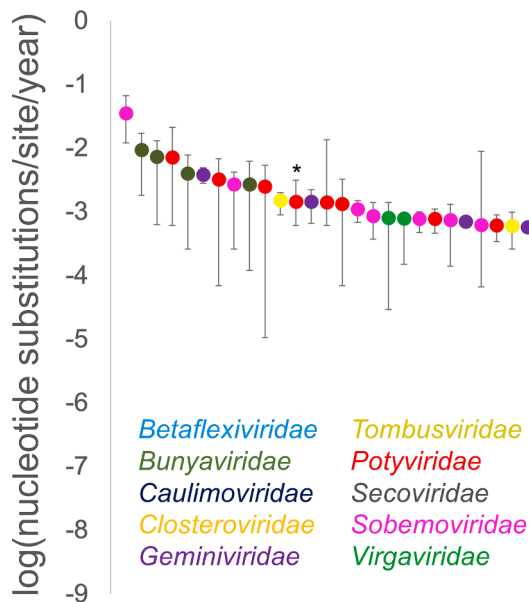
Fig. 3. Rates of evolution of plant virus CP genes. Error bars reflect 95% HPDs, and estimates without error bars were calculated without BEAST (Table 1). The rate for CBSV is indicated with an asterisk.

evolutionary analysis because of its lack of temporal signal, the crude root-to-tip regression in TempEst analysis estimated that its most common ancestor existed in 1575. This datum agrees with recent reports indicating a possibility of UCBSV being older that CBSV (Ateka et al., 2017), but our estimated age of CBSV is consistent with the timing of the first observation of brown streak symptoms in cassava. Cassava was introduced into Africa by the Portuguese in the late 16th century and CBSD was first reported around Lake Malawi region in the 1930s (Storey, 1936; Nichols, 1950; Tomlinson et al., 2018). Our results support the commonly held assumption that CBSV is likely the result of an indigenous African potyvirus host-shifting into cassava (Robson et al., 2023), and it is still an emerging cassava virus (Jacobson et al., 2018; Jones, 2021; Tugume et al., 2023). This has implication on disease management as virulence and spread of newly introduced pathogens is more detrimental to sustainable food security (Jones, 2021).

4.2. CBSV is going through small amounts of adaptive change

Our results are similar to previous reports of limited positive selection among the CP gene of several potyviruses, with the majority of sites under purifying selection (Cuevas et al., 2012; Gao et al., 2018; Nigam et al., 2019; Wokorach et al., 2020). The study has documented evidence for pervasive purifying selection events occurring in CP gene of CBSV. The significance of the two sites under positive selection in the CP gene of CBSV needs to be established, in a manner similar to studies CP gene of potato virus Y, which identified sites associated with significant modification of viral accumulation in different hosts, as well as its transmissibility by aphids (Moury et al., 2011). Positively selected sites in the sugarcane mosaic virus CP gene were reported to have an impact on host – virus – vector interactions (Li et al., 2012). Previous mutation studies in luteoviruses have indicated that sites under purifying selection on the CP surface provide critical knowledge on the biology of those viruses (Torres et al., 2005). For example, mutations in the acidic patch

Table 1

Details of evolutionary rates of coat protein gene of various plant viruses.

Virus	Genus (Family)	Date range	Clock model	Growth model	Mean substitution rate	95% HPD substitution rate	Reference
ACMV	*Begomovirus* (Geminiviridae)	1928 – 2014	Relaxed	Bayesian Skyline	1.3×10^{-4}	$8.0 \times 10^{-5} – 1.7 \times 10^{-4}$	Rieux et al., 2021
ArMV	*Nepovirus* (Secoviridae)	1991 – 2008	Relaxed	Constant	9.3×10^{-3}	$1.8 \times 10^{-3} – 1.7 \times 10^{-2}$	Thompson et al., 2014
BChV	*Polerovirus* (Solemoviridae)	1985 – 2005	Relaxed	Bayesian Skyline	2.7×10^{-3}	$2.6 \times 10^{-4} – 4.2 \times 10^{-3}$	Pagán and Holmes, 2010
BYDV	*Luteovirus* (Tombusviridae)	1917 – 2008	Relaxed	Bayesian Skyline	1.5×10^{-3}	$9.8 \times 10^{-4} – 2.0 \times 10^{-3}$	Pagán and Holmes, 2010
CABYV	*Polerovirus* (Solemoviridae)	2003 – 2008	Relaxed	Bayesian Skyline	3.5×10^{-2}	$1.2 \times 10^{-2} – 6.6 \times 10^{-2}$	Pagán and Holmes, 2010
CABMV	*Potyvirus* (Potyviridae)		Relaxed		1.4×10^{-4}		Gibbs et al., 2008
CaMV	*Caulimovirus* (Caulimoviridae)	1960 – 2010	Relaxed	Exponential	1.71×10^{-4}	$1.45 \times 10^{-5} – 3.87 \times 10^{-4}$	Yasaka et al., 2014
CBSV	*Ipomovirus* (Potyviridae)	1996 – 2019	Relaxed	Bayesian Skyline	1.43×10^{-3}	$1.4 \times 10^{-4} – 8.86 \times 10^{-3}$	This study
CGMMV	*Tobamovirus* (Virgaviridae)	1970 – 2008	Relaxed	Bayesian Skyline	8.0×10^{-4}	$2.9 \times 10^{-5} – 1.4 \times 10^{-3}$	Pagán et al., 2010
CTV	*Closterovirus* (Closteroviridae)	1990 – 2010	Strict	Bayesian Skyline	1.6×10^{-4}	$1.7 \times 10^{-5} – 3.2 \times 10^{-4}$	Silva et al., 2012
CYDV	*Polerovirus* (Solemoviridae)	1925 – 2005	Relaxed	Bayesian Skyline	7.4×10^{-4}	$1.4 \times 10^{-4} – 1.3 \times 10^{-3}$	Pagán and Holmes, 2010
CyEVA	*Potyvirus* (Potyviridae)	2008 – 2014	Relaxed	Constant	7.15×10^{-3}	$6.11 \times 10^{-4} – 2.12 \times 10^{-2}$	Ohshima et al., 2016
DSV	*Mastrevirus* (Geminiviridae)				1.27×10^{-4}		Ortega-del Campo et al., 2021
EACMV	*Begomovirus* (Geminiviridae)	1996 – 2009	Relaxed		3.83×10^{-3}	$2.81 \times 10^{-3} – 4.89 \times 10^{-3}$	De Bruyn et al., 2012
GFLV	*Nepovirus* (Secoviridae)	1991 – 2011	Relaxed	Exponential	7.3×10^{-3}	$6.3 \times 10^{-4} – 1.3 \times 10^{-2}$	Thompson et al., 2014
MSV	*Mastrevirus* (Geminiviridae)	1979 – 2006	Relaxed	Constant	7.0×10^{-4}	$6.2 \times 10^{-4} – 7.3 \times 10^{-4}$	Harkins et al., 2009
GRBV	*Grablovirus* (Geminiviridae)	2010–2017	Strict	Exponential	3.2×10^{-5}	$2.1 \times 10^{-5} – 4.3 \times 10^{-5}$	Thompson, 2022
NDV	*Potyvirus* (Potyviridae)	2000 – 2011	Relaxed	Exponential	6.16×10^{-3}	$1.31 \times 10^{-3} – 1.34 \times 10^{-2}$	Ohshima et al., 2016
NLSYV	*Potyvirus* (Potyviridae)	2000 – 2013	Relaxed	Constant	1.33×10^{-3}	$6.92 \times 10^{-5} – 3.23 \times 10^{-3}$	Ohshima et al., 2016
NLV	*Potyvirus* (Potyviridae)	2004 – 2012	Relaxed	Constant	2.48×10^{-3}	$1.06 \times 10^{-5} – 5.35 \times 10^{-3}$	Ohshima et al., 2016
NYSV	*Potyvirus* (Potyviridae)	2004 – 2013	Relaxed	Constant	$3.23 \times 10{-3}$	$1.35 \times 10^{-3} – 6.78 \times 10^{-3}$	Ohshima et al., 2016
ORSV	*Tobamovirus* (Virgaviridae)	1960 – 2004	Relaxed	Bayesian Skyline	7.8×10^{-4}	$1.5 \times 10^{-4} – 9.2 \times 10^{-4}$	Pagán et al., 2010
PaLCD	*Begomovirus* (Geminiviridae)	1997–2018	Strict	Constant	1.43×10^{-3}	$6.59 \times 10^{-4} – 2.22 \times 10^{-3}$	Srivastava et al., 2022
PLRV	*Polerovirus* (Solemoviridae)	1974 – 2008	Relaxed	Bayesian Skyline	6.2×10^{-4}	$6.6 \times 10^{-5} – 8.9 \times 10^{-3}$	Pagán and Holmes, 2010
PMMoV	*Tobamovirus* (Virgaviridae)	1972 – 2008	Relaxed	Bayesian Skyline	1.8×10^{-4}	$1.6 \times 10^{-5} – 3.9 \times 10^{-4}$	Pagán et al., 2010
PRSV	*Potyvirus* (Potyviridae)	1990 – 2016	Relaxed	Bayesian Skyline	7.7×10^{-4}	$4.7 \times 10^{-4} – 5.6 \times 10^{-4}$	Cabrera Mederos et al., 2019
PVS	*Carlavirus* (Betaflexiviridae)	1985 – 2014	Relaxed	Constant	3.3×10^{-4}	$1.3 \times 10^{-4} – 5.6 \times 10^{-4}$	Duan et al., 2018
PVY	*Potyvirus* (Potyviridae)	1938 – 2013	Relaxed	Constant	1.0×10^{-4}	$6.9 \times 10^{-5} – 1.3 \times 10^{-4}$	Adrian J. Gibbs et al., 2017
RMV	*Tobamovirus* (Virgaviridae)	1950 – 2001	Relaxed	Bayesian Skyline	1.4×10^{-4}	$1.3 \times 10^{-8} – 4.0 \times 10^{-4}$	Pagán et al., 2010
RSV	*Tenuivirus* (Bunyaviridae)	1997 – 2013	Strict	Constant	4.3×10^{-4}	$3.0 \times 10^{-4} – 5.7 \times 10^{-3}$	He et al., 2017
RTSV	*Waikavirus* (Secoviridae)	1995 – 2009	Relaxed	Exponential	4.0×10^{-3}	$2.6 \times 10^{-4} – 7.8 \times 10^{-3}$	Thompson et al., 2014
RYMV	*Sobemovirus* (Solemovirdae)	1975 – 2005	Relaxed	Bayesian Skyline	7.7×10^{-4}	$4.7 \times 10^{-4} – 1.0 \times 10^{-3}$	Fargette et al., 2008
SbDV	*Luteovirus* (Tombusviridae)	1990 – 2007	Relaxed	Bayesian Skyline	6.0×10^{-4}	$2.6 \times 10^{-4} – 9.9 \times 10^{-4}$	Pagán and Holmes, 2010
ScYLV	*Polerovirus* (Solemoviridae)	1989 – 2008	Relaxed	Bayesian Skyline	1.1×10^{-3}	$6.8 \times 10^{-4} – 1.5 \times 10^{-3}$	Pagán and Holmes, 2010
TMGMV	*Tobamovirus* (Virgaviridae)	1907 – 2008	Relaxed	Bayesian Skyline	1.7×10^{-4}	$1.7 \times 10^{-5} – 3.3 \times 10^{-4}$	Pagán et al., 2010
TMV	*Tobamovirus* (Virgaviridae)	1899 – 2008	Relaxed	Bayesian Skyline	1.6×10^{-4}	$5.2 \times 10^{-5} – 3.2 \times 10^{-3}$	Pagán et al., 2010
ToMV	*Tobamovirus* (Virgaviridae)	1975 – 2018	Strict	MASCOT	4.0×10^{-4}	$2.4 \times 10^{-4} – 5.6 \times 10^{-4}$	Xu et al., 2021
ToRSV	*Nepovirus* (Secoviridae)	1991 – 2007	Relaxed	Exponential	2.7×10^{-3}	$1.2 \times 10^{-4} – 6.2 \times 10^{-3}$	Thompson et al., 2014
ToSRV	*Begomovirus* (Geminiviridae)				5.70×10^{-4}		Pinto et al., 2021
TuMV	*Potyvirus* (Potyviridae)	1968 – 2007	Relaxed	Exponential	6.1×10^{-4}	$3.4 \times 10^{-4} – 8.9 \times 10^{-4}$	Nguyen et al., 2013

Table 1 (*continued*)

| Virus	Genus (Family)	Date range	Clock model	Growth model	Mean substitution rate	95% HPD substitution rate	Reference
TuYV	*Polerovirus (Solemoviridae)*	1980 – 2006	Relaxed	Bayesian Skyline	8.6×10^{-4}	$3.7 \times 10^{-4} - 1.4 \times 10^{-3}$	Pagán and Holmes, 2010
TYLCV	*Begomovirus (Geminiviridae)*	1988 – 2006	Relaxed	Exponential	4.6×10^{-4}	$6.6 \times 10^{-5} - 8.9 \times 10^{-4}$	Duffy and Holmes, 2008
ZYMV	*Potyvirus (Potyviridae)*	1984 – 2006	Relaxed	Exponential	$5.0 \times 10{-4}$	$1.8 \times 10{-4} - 8.8 \times 10{-4}$	Simmons et al., 2008
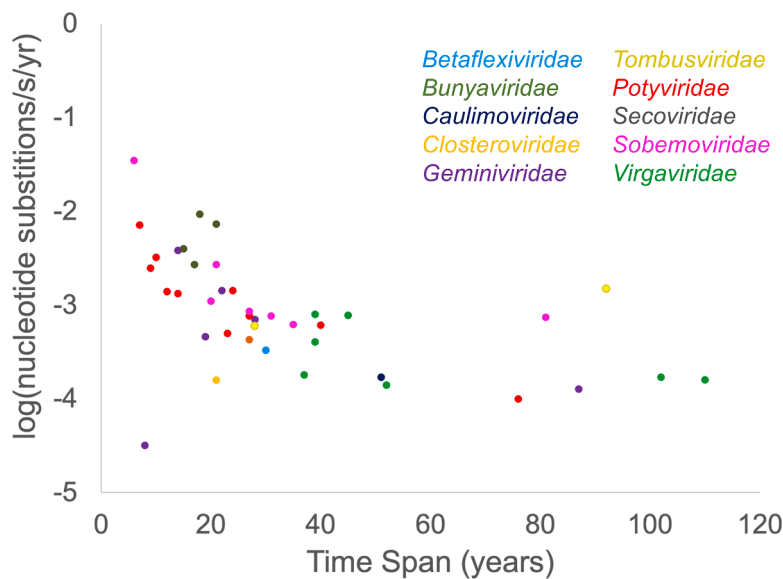
Fig. 4. The subset of CP substitutions rates calculated with BEAST are plotted against the time span of sampled viral sequences. The ten viral families are color-coded as in Fig. 3.

domain located in the surface loop of potato leafroll virus established the role of CP in virion assembly, systematic movement and aphid transmission (Lee et al., 2005). The CP gene of vectored plant viruses, which interact with both host and vector, are under more pressure to be unchanging, as evidenced by very low ratios of the rate of non-synonymous changes to the rate of synonymous changes (dN/dS) (Chare and Holmes, 2004; Jenkins et al., 2002).

4.3. CBSV CP is an average evolving potyvirus CP

The substitution rate is an important parameter in understanding evolution, as high mutation rates often lead to strong adaptability and high pathogenicity in viruses (Muniraju et al., 2014; Peck and Lauring, 2018). The study has documented nucleotide substitution rates in CP gene of CBSV that are about average for a potyvirus (CBSV: 1.43×10^{-3} n/s/y), just slightly higher than the average of the 10 other potyviruses: 9.0×10^{-4} n/s/y) – but it is still a fast-evolving rate befitting of an RNA virus (Koonin et al., 2022). CP genes are often not the fastest evolving genes in viruses (Pavesi, 2021), including in CBSVs (Alicai et al., 2016), so this rate may not reflect the evolvability of other parts of the CBSV genome to overcome host resistance (Elena et al., 2014; Lafforgue et al., 2011; Montarry et al., 2012) or adaptation to altered environmental conditions or novel crops (Duffy and Holmes, 2008). Comparing across other potyviruses, turnip mosaic virus' whole polyprotein has a higher substitution rate than that estimated for its CP (Kawakubo et al., 2021), and the P3 gene of potato virus Y has a higher substitution rate than its CP (Gao et al., 2020).

While the confounding between phylogenetic clustering and isolation date reduces the confidence we have in our substitution rate, our dataset performed robustly in other tests of temporal signal, similar to other published rates of CP evolution in plant viruses (all but eight of

which assessed temporal signal by randomized dates across their sequences without regard for clustered, similar sequences). Some studies were published prior to date randomization tests being piloted (Duffy and Holmes, 2009; Ramsden et al., 2009): CABMV, MSV, RYMV, TYLCV and ZYMV. Two additional rates were calculated without BEAST (meaning no date randomization test could be run: DSV and ToSRV) and one CP rate of evolution was published without any estimate of its temporal signal (RSV, He et al. 2017). However, the whole data set randomization test has been shown to yield false negative results – that is, data sets with significant flaws in their temporal signal, such as a high correlation between clustering genetically and dates of isolation, are given potentially false credibility as having temporal signal (Duchene et al. 2015, Murry et al. 2016). Therefore, we employed a more modern approach to verifying the temporal signal in our dataset and received strong support (Duchene et al. 2020). This support did not hold up to clustered permutation date randomization, implying that our calculated substitution rate should be interpreted with caution as it may be inaccurate. Arguably most of the collated plant virus rates have similarly confounded datasets, because of the common problem of uneven sampling practices (Hall et al. 2016; Murray et al. 2016). Therefore, our comparisons among rates are apples-to-apples – with potentially many of these rates needing revisiting in the future, when hopefully there will be more unbiased sampling of the viruses over time, and improvements in methods for pruning datasets to reduce the confounding effects of phylogenetic clustering and dates of isolation.

The differences of individual datasets, and the idiosyncratic ecological conditions of each virus, lead to variation in substitution rates even if the viruses within each family have similar mutation rates, complicating detection of overarching patterns (Hicks and Duffy, 2014). Nonetheless, potyviruses appear to evolve more quickly than several other plant virus families based on the limited data available.

Potyviruses are known to be capable of significant diversity (Nigam et al., 2019) and have high mutation frequencies (Khanal and Ali, 2021). Their more frequent emergence on novel hosts (Gibbs et al., 2020; Moury and Desbiez, 2020) may be due to higher rates of evolution than some other families of plant viruses.

5. Conclusion

CBSV, but not UCBSV, has been sufficiently sequenced in number and over sufficient time in East Africa to support a tip-dated coalescent analysis of its rate of evolution, which is just as expected for a member of *Potyviridae*.

Author statement

No AI technologies (large language models or otherwise) were used in the analysis or preparation of this manuscript.

CRediT authorship contribution statement

Willard Mbewe: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Settumba Mukasa:** Writing – review & editing, Supervision. **Mildred Ochwo-Ssemakula:** Writing – review & editing, Supervision, Resources. **Peter Sseruwagi:** Writing – review & editing, Supervision. **Fred Tairo:** Writing – review & editing, Supervision, Funding acquisition. **Joseph Ndunguru:** Writing – review & editing, Supervision, Funding acquisition. **Siobain Duffy:** Writing – review & editing, Visualization, Resources, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data used is publicly available on GenBank or in the published literature.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.virusres.2024.199397.

References

Abdalla, O.A., Ali, A., 2021. Genetic variability and evidence of a new subgroup in watermelon mosaic virus isolates. Pathogens. 10 (10), 1245. https://doi.org/10.3390/pathogens10101245.

Alicai, T., Ndunguru, J., Sseruwagi, P., Tairo, F., Okao-Okuja, G., Nanvubya, R., Kiiza, L., Kubatko, L., Kehoe, M.A., Boykin, L.M., 2016. Cassava brown streak virus has a

rapidly evolving genome: implications for virus speciation, variability, diagnosis and host resistance. Sci. Rep. 6, 36164. https://doi.org/10.1038/srep36164.

Ateka, E., Alicai, T., Ndunguru, J., Tairo, F., Sseruwagi, P., Kiarie, S., Makori, T., Kehoe, M.A., Boykin, L.M., 2017. Unusual occurrence of a DAG motif in the Ipomovirus Cassava brown streak virus and implications for its vector transmission. PLoS One 12 (11), e0187883. https://doi.org/10.1371/journal.pone.0187883.

Atreya, C.D., Raccah, B., Pirone, T.P., 1990. A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. Virology. 178 (1), 161–165. https://doi.org/10.1016/0042-6822(90)90389-9.

Baele, G., Lemey, P., Suchard, M.A., 2016. Genealogical working distributions for Bayesian model testing with phylogenetic uncertainty. Syst. Biol. 65 (2), 250–264.

Cabrera Mederos, D., Giolitti, F., Torres, C., Portal, O., 2019. Distribution and phylodynamics of papaya ringspot virus on Carica papaya in Cuba. Plant Pathol. 68 (2), 239–250. https://doi.org/10.1111/ppa.12942.

Chare, E.R., Holmes, E.C., 2004. Selectios pressures in the capsid genes of plant RNA viruses reflect mode of transmission. J. Gen. Virol. 85, 3149–3157.

Chung, B.Y.W., Miller, W.A., Atkins, J.F., Firth, A.E., 2008. An overlapping essential gene in the Potyviridae. Proceedings of the National Academy of Sciences 105 (15), 5897–5902. https://doi.org/10.1073/pnas.0800468105.

Cuevas, J.M., Delaunay, A., Visser, J.C., Bellstedt, D.U., Jacquot, E., Elena, S.F., 2012. Phylogeography and molecular evolution of potato virus Y. PLoS. One 7 (5), e37853. https://doi.org/10.1371/journal.pone.0037853.

De Bruyn, A., Villemot, J., Lefeuvre, P., Villar, E., Hoareau, M., Harimalala, M., Abdoul-Karime, A.L., Abdou-Chakour, C., Reynaud, B., Harkins, G.W., Varsani, A., Martin, D.P., Lett, J.M., 2012. East African cassava mosaic-like viruses from Africa to Indian ocean islands: moeclar diversity, evolutionary histor and geographical dissemination of a bipartitite begomovirus. BMC Evol. Biol. 12, 228.

Dombrovsky, A., Reingold, V., Antignus, Y., 2014. Ipomovirus - an atypical genus in the family Potyviridae transmitted by whiteflies. Pest. Manage Sci. 70 (10), 1553–1567. https://doi.org/10.1002/ps.3735.

Duan, G., Zhan, F., Du, Z., Ho, S.Y.W., Gao, F., 2018. Europe was a hub for the global spread of potato virus S in the 19th century. Virology. 525, 200–204. https://doi.org/10.1016/j.virol.2018.09.022.

Duchêne, S., Duchêne, D., Holmea, E.C., Ho, S.Y.W., 2015. The performance of date-randomization test in phylogenetic analyses of time-structured virus data. Molucular Biology and Evolution 32 (7), 1895–1906.

Duchene, S., Lemey, P., Stadler, ..T., Ho, S.Y.W., David A Duchene, D.A., Dhanasekaran, V., Baele, G., 2020. Bayesian Evaluation of Temporal Signal in Measurably Evolving Populations. Mol. Biol. Evol. 37 (11), 3363–3379.

Duffy, S., Holmes, E.C., 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. J. Virol. 82 (2), 957–965. https://doi.org/10.1128/JVI.01929-07.

Duffy, S., Holmes, E.C., 2009. Validation of high rates of nucleotide substitution in geminiviruses: Phylogenetic evidence from East African cassava mosaic viruses. Journal of General Virology 90 (6), 1539–1547. https://doi.org/10.1099/vir.0.009266-0.

Elena, S.F., Fraile, A., Garcia-Arenal, F., 2014. Chapter Three- Evolution and Emergence of Plant Viruses. Advances in Virus Research.

Fargette, D., Pinel, a, Rakotomalala, M., Sangu, E., Traoré, O., Sérémé, D., Sorho, F., Issaka, S., Hébrard, E., Séré, Y., Kanyeka, Z., Konaté, G, 2008. Rice yellow mottle virus, an RNA plant virus, evolves as rapidly as most RNA animal viruses. J. Virol. 82 (7), 3584–3589. https://doi.org/10.1128/JVI.02506-07.

Gao, F., Du, Z., Shen, J., Yang, H., Liao, F., 2018. Genetic diversity and molecular evolution of Ornithogalum mosaic virus based on the coat protein gene sequence. PeerJ. 6, e4550. https://doi.org/10.7717/peerj.4550.

Gibbs, A.J., MacKenzie, A.M., Wei, K.J., Gibbs, M.J., 2008. The potyviruses of Australia. Arch. Virol. 153 (8), 1411–1420. https://doi.org/10.1007/s00705-008-0134-6.

Gao, F., Kawakubo, S., Ho, S.Y.W., Ohshima, K., 2020. The evolutionary history and global spatio-temporal dynamics of potato virus Y. Virus Evol. 6, veaa056.

Gibbs, Adrian J., Hajizadeh, M., Ohshima, K., Jones, R.A.C., 2020. The potyviruses: An evolutionary synthesis is emerging. Viruses. 12 (2), 1–30. https://doi.org/10.3390/v12020132.

Gibbs, Adrian J., Ohshima, K., Yasaka, R., Mohammadi, M., Gibbs, M.J., Jones, R.A.C., 2017. The phylogenetics of the global population of potato virus y and its necrogenic recombinants. Virus. Evol. 3 (1) https://doi.org/10.1093/ve/vex002 vex002.

Hall, M.D., m Woolhouse, M.E.J., Rambaut, A., 2016. The effects of sampling strategy on the quality of reconstruction of viral population dynamics using Bayesian skyline family coalecent methods: A simulation study. Virus. Evol. 2 (1) vew003.

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., Varsani, A., 2009. Dating the origins of the maize-adapted strain of maize streak virus, MSV-A. Journal of General Virology 90 (12), 3066–3074. https://doi.org/10.1099/vir.0.015537-0.

Harrison, B.D., Zhou, X., Otim Nape, G.W., Liu, Y., Robinson, D.J., 1997. Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. Annals of Applied Biology 131 (3), 437–448. https://doi.org/10.1111/j.1744-7348.1997.tb05171.x.

He, M., Guan, S.Y., He, C.Q., 2017. Evolution of rice stripe virus. Mol. Phylogenet. Evol. 109, 343–350. https://doi.org/10.1016/j.ympev.2017.02.002.

Heled, J., Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. BMC Evol. Biol. 8, 289. https://doi.org/10.1186/1471-2148-8-289.

Heled, J., Drummond, A.J., 2010. Bayesian Inference of Species Trees from Multilocus Data. Mol. Biol. Evol. 27 (3), 570–580. https://doi.org/10.1093/molbev/msp274.

Hicks, A.L., Duffy, S., 2014. Cell Tropism Predicts Long-term Nucleotide Substitution Rates of Mammalian RNA Viruses. PLoS. Pathog. 10 (1), e1003838 https://doi.org/10.1371/journal.ppat.1003838.

Ho, S.Y.W., Phillips, M.J., Cooper, A., Drummond, A.J., 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Mol. Biol. Evol. 22 (7), 1561–1568. https://doi.org/10.1093/molbev/msi145.

Inoue-Nagata, A.K., Jordan, R., Kreuze, J., Li, F., López-Moya, J.J., Mäkinen, K., Ohshima, K., Wylie, S.J., Consortium, Ictv Report, 2022. ICTV Virus Taxonomy Profile: Potyviridae 2022. J. Gen. Virol. 103 (5), 001738 https://doi.org/10.1099/jgv.0.001738.

Jacobson, A.L., Duffy, S., Sseruwagi, P., 2018. Whitefly-transmitted viruses threatening cassava production in Africa. Curr. Opin. Virol. 33, 167–176. https://doi.org/10.1016/j.coviro.2018.08.016.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C., 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J. Mol. Evol. 54, 156–165.

Jones, R.A.C., 2021. Global plant virus disease pandemics and epidemics. Plants 10 (2), 233. https://doi.org/10.3390/plants10020233.

Kawakubo, S., Gao, F., Li, S., Tan, Z., Huang, Y.K., Adkar-Purushothama, C.R., Gurikar, C., Maneechoat, P., Chiemsombat, P., Aye, S.S., Furuya, N., Shevchenko, O., Špak, J., Škorić, D., Ho, S.Y.W., Ohshima, K., 2021. Genomic analysis of the brassica pathogen turnip mosaic potyvirus reveals its spread along the former trade routes of the silk road. Proc. Natl. Acad. Sci. U.S.A. 118, e2021221118.

Khanal, V., Ali, A., 2021. High mutation frequency and significant population differentiation in papaya ringspot virus-w isolates. Pathogens. 10 (10), 1278. https://doi.org/10.3390/pathogens10101278.

Koonin, E.V., Dolja, V.V., Krupovic, M., 2022. The logic of virus evolution. Cell Host and Microbe 30 (7), 917–929. https://doi.org/10.1016/j.chom.2022.06.008.

Kosakovsky Pond, S.L., Frost, S.D.W, 2005. Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics. 21 (10), 2531–2533. https://doi.org/10.1093/bioinformatics/bti320.

Kosakovsky Pond, S.L., Scheffler, K., Gravenor, M.B., Poon, A.F.Y., Frost, S.D.W, 2010. Evolutionary Fingerprinting of Genes. Mol. Biol. Evol. 27 (3), 520–536. https://doi.org/10.1093/molbev/msp260.

Lafforgue, G., Martínez, F., Sardanyés, J., de la Iglesia, F., Niu, Q.W., Lin, S.S., Solé, R.V., Chua, N.H., Darós, J.A., Elena, S.F., 2011. Tempo and Mode of Plant RNA Virus Escape from RNA Interference-Mediated Resistance. J. Virol. 85 (19), 9686–9695. https://doi.org/10.1128/jvi.05326-11.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., Mcgettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23 (21), 2947–2948. https://doi.org/10.1093/bioinformatics/btm404.

Lee, L., Kaplan, I.B., Ripoll, D.R., Liang, D., Palukaitis, P., Gray, S.M., 2005. A Surface Loop of the Potato Leafroll Virus Coat Protein Is Involved in Virion Assembly, Systemic Movement, and Aphid Transmission. J. Virol. 79 (2), 1207–1214. https://doi.org/10.1128/jvi.79.2.1207-1214.2005.

Li, Y., Jia, M., Jiang, Z., Zhou, T., Fan, Z., 2012. Molecular variation and recombination in RNA segment 10 of rice black-streaked dwarf virus isolated from China during 2007-2010. Arch. Virol. 157 (7), 1351–1356. https://doi.org/10.1007/s00705-012-1282-2.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus. Evol. 1 (1), 1–5. https://doi.org/10.1093/ve/vev003.

Maruthi, M.N., Hillocks, R.J., Mtunda, K., Raya, M.D., Muhanna, M., Kiozia, H., Rekha, A.R., Colvin, J., Thresh, J.M., 2005. Transmission of Cassava brown streak virus by Bemisia tabaci (Gennadius). Journal of Phytopathology 153 (5), 307–312.

Massingham, T., Goldman, N., 2005. Detecting amino acid sites under positive selection and purifying selection. Genetics 169 (3), 1753–1762. https://doi.org/10.1534/genetics.104.032144.

Mbanzibwa, D.R., Tian, Y., Tugume, A.K., Mukasa, S.B., Tairo, F., Kyamanywa, S., Kullaya, A., Valkonen, J.P.T., 2009. Genetically distinct strains of Cassava brown streak virus in the Lake Victoria basin and the Indian Ocean coastal area of East Africa. Arch. Virol. 154 (2), 353–359. https://doi.org/10.1007/s00705-008-0301-9.

Monger, W.A., Seal, S., Isaac, A.M., Foster, G.D., 2001. Molecular characterization of the Cassava brown streak virus coat protein. Plant Pathol. 50 (4), 527–534. https://doi.org/10.1046/j.1365-3059.2001.00589.x.

Montarry, J., Cartier, E., Jacquemond, M., Palloix, A., Moury, B., 2012. Virus adaptation to quantitative plant resistance: Erosion or breakdown? J. Evol. Biol. 25 (11), 2242–2252. https://doi.org/10.1111/j.1420-9101.2012.02600.x.

Moury, B., Caromel, B., Johansen, E., Simon, V., Chauvin, L., Jacquot, E., Kerlan, C., Lefebvre, V., 2011. The helper component proteinase cistron of Potato virus Y induces hypersensitivity and resistance in Potato genotypes carrying dominant resistance genes on chromosome IV. Molecular Plant-Microbe Interactions : MPMI 24 (7), 787–797. https://doi.org/10.1094/MPMI-10-10-0246.

Moury, B., Desbiez, C., 2020. Host range evolution of potyviruses: A global phylogenetic analysis. Viruses 12 (1), 111. https://doi.org/10.3390/v12010111.

Muniraju, M., Munir, M., Parthiban, A.R., Banyard, A.C., Bao, J., Wang, Z., Ayebazibwe, C., Ayelet, G., El Harrak, M., Mahapatra, M., Libeau, G., Batten, C., Parida, S., 2014. Molecular evolution of peste des petits ruminants virus. Emerg. Infect. Dis. 20 (12), 2023–2033. https://doi.org/10.3201/eid2012.140684.

Murray, G.G.R., Wang, F., Harrison, E.M., Paterson, G.K., Mather, A.E., Harris, S.R., Holmes, M.A., Rambaut, A., Welch, J.J., 2016. The effect of genetic structure on molecular dating and tests for temporal signal. Methods in Ecology & Evolution (7), 80–89. https://doi.org/10.1111/2041-210X.12466.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., Scheffler, K., 2013. FUBAR: A fast, unconstrained bayesian AppRoximation for inferring selection. Mol. Biol. Evol. 30 (5), 1196–1205. https://doi.org/10.1093/molbev/mst030.

Ndunguru, J., Sseruwagi, P., Tairo, F., Stomeo, F., Maina, S., Djinkeng, A., Kehoe, M., Boykin, L.M., Melcher, U., 2015. Analyses of twelve new whole genome sequences of cassava brown streak viruses and ugandan cassava brown streak viruses from East Africa: Diversity, supercomputing and evidence for further speciation. PLoS. One 10 (10), e0141939.

Nguyen, H.D., Tomitaka, Y., Ho, S.Y.W., Duchêne, S., Vetten, H.J., Lesemann, D., Walsh, J.A., Gibbs, A.J., Ohshima, K., 2013. Turnip Mosaic Potyvirus Probably First Spread to Eurasian Brassica Crops from Wild Orchids about 1000 Years Ago. PLoS. One 8 (2), e55336. https://doi.org/10.1371/journal.pone.0055336.

Nichols, R.F.W., 1950. The brown streak disease of cassava. East Afr. Agric. J. 15, 154–160.

Nigam, D., LaTourrette, K., Souza, P.F.N., Garcia-Ruiz, H., 2019. Genome-Wide Variation in Potyviruses. Front. Plant Sci. 10, 1–28. https://doi.org/10.3389/fpls.2019.01439.

Ohshima, K., Matsumoto, K., Yasaka, R., Nishiyama, M., Soejima, K., Korkmaz, S., Ho, S.Y.W., Gibbs, A.J., Takeshita, M., 2016. Temporal analysis of reassortment and molecular evolution of Cucumber mosaic virus: Extra clues from its segmented genome. Virology. 487, 188–197. https://doi.org/10.1016/j.virol.2015.09.024.

Ortega-del Campo, S., Grigoras, I., Timchenko, T., Gronenborn, B., 2021. Twenty years of evolution and diversification of digitaria streak virus in *Digitaria setigera*. Virus Evol. 7, 1–14.

Pagán, I., Firth, C., Holmes, E.C., 2010. Phylogenetic analysis reveals rapid evolutionary dynamics in the plant RNA virus genus tobamovirus. J. Mol. Evol. 71 (4), 298–307. https://doi.org/10.1007/s00239-010-9385-4.

Pagán, I., Holmes, E.C., 2010. Long-term evolution of the Luteoviridae: time scale and mode of virus speciation. J. Virol. 84 (12), 6177–6187. https://doi.org/10.1128/JVI.02160-09.

Parag, K.V., Pybus, O.G., Wu, C.H., 2022. Are Skyline Plot-Based Demographic Estimates Overly Dependent on Smoothing Prior Assumptions? Syst. Biol. 71 (1), 121–138. https://doi.org/10.1093/sysbio/syab037.

Patil, B.L., Legg, J.P., Kanju, E., Fauquet, C.M., 2015. Cassava brown streak disease: A threat to food security in Africa. Journal of General Virology 96 (5), 956–968.

Patil, B.L., Ogwok, E., Wagaba, H., Mohammed, I.U., Yadav, J.S., Bagewadi, B., Taylor, N.J., Kreuze, J.F., Maruthi, M.N., Alicai, T., Fauquet, C.M., 2011. RNAi-mediated resistance to diverse isolates belonging to two virus species involved in Cassava brown streak disease. Mol. Plant Pathol. 12 (1), 31–41.

Pavesi, A., 2021. Origin, evolution and stability of overlapping genes in viruses: A systematic review. Genes. (Basel) 12 (6), 809. https://doi.org/10.3390/genes12060809.

Peck, K.M., Lauring, A.S., 2018. Complexities of Viral Mutation Rates. J. Virol. 92 (14), e01031. https://doi.org/10.1128/jvi.01031-17, 17.

Pinto, V.B., Quadros, A.F.F., Godinho, M.T., Silva, J.C., Alfenas-Zerbini, P., Zerbini, F.M., 2021. Intra-host evolution of the ssDNA virus tomato severe rugose virus (ToSRV). Virus. Res. 292, 198234 https://doi.org/10.1016/j.virusres.2020.198234.

Prins, M., Laimer, M., Noris, E., Schubert, J., Wassenegger, M., Tepfer, M., 2008. Strategies for antiviral resistance in transgenic plants. Mol. Plant Pathol. 9 (1), 73–83. https://doi.org/10.1111/j.1364-3703.2007.00447.x.

Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus. Evol. 2 (1) https://doi.org/10.1093/ve/vew007 vew007.

Rieux, A., Campos, P., Duvermy, A., Scussel, S., Martin, D., Gaudeul, M., Lefeuvre, P., Becker, N., & Lett, J.M. (2021). Contribution of historical herbarium small RNAs to the reconstruction of a cassava mosaic geminivirus evolutionary history. Scientific Reports, 11(1), 21280. https://doi.org/10.1038/s41598-021-00518-w.

Ramsden, C., Holmes, E.C., Charleston, M.A., 2009. Hantavirus evolution in relation to its rodent and insectivore hosts: no evidence for codivergence. Mol. Biol. Evol. 26, 143–153.

Robson, F., Hird, D.L., Boa, E., 2023. Cassava brown streak: a deadly virus on the move. Plant Pathol. 73 (2), 221–241. https://doi.org/10.1111/ppa.13807.

Rybicki, E.P., Shukla, D.D., 1992. Coat protein phylogeny and systematics of potyviruses. Archives of Virology Supplementum 5, 139–170. http://www.ncbi.nlm.nih.gov/pubmed/1450733.

Silva, G., Marques, N., Nolasco, G., 2012. The evolutionary rate of citrus tristeza virus ranks among the rates of the slowest RNA viruses. Journal of General Virology 93 (2), 419–429. https://doi.org/10.1099/vir.0.036574-0.

Storey, H.H., 1936. Virus diseases of East African plants. East Afr. Agric. J. 1, 333–337.

Simmons, H.E., Holmes, E.C., Stephenson, A.G., 2008. Rapid evolutionary dynamics of zucchini yellow mosaic virus. Journal of General Virology 89 (4), 1081–1085. https://doi.org/10.1099/vir.0.83543-0.

Srivastava, A., Pandey, V., Sahu, A.K., Yadav, D., Al-Sadi, A.M., Shahid, M.S., Gaur, R.K., 2022. Evolutionary Dynamics of Begomoviruses and Its Satellites Infecting Papaya in India. Front. Microbiol. 13, 879413 https://doi.org/10.3389/fmicb.2022.879413.

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J. & Rambaut., A. (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution, 4, vey016. DOI:10.1093/ve/vey016.

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol. Biol. Evol. 38 (7), 3022–3027. https://doi.org/10.1093/molbev/msab120.

Thompson, J.R., 2022. Analysis of the genome of grapevine red blotch virus and related grabloviruses indicates diversification prior to the arrival of Vitis vinifera in North America. Journal of General Virology 103 (10), 001789. https://doi.org/10.1099/jgv.0.001789.

Thompson, J.R., Kamath, N., Perry, K.L., 2014. An evolutionary analysis of the Secoviridae family of viruses. PLoS. One 9 (9), e106305. https://doi.org/10.1371/journal.pone.0106305.

Tomlinson, K.R., Bailey, A.M., Alicai, T., Seal, S., Foster, G.D., 2018. Cassava brown streak disease: historical timeline, current knowledge and future prospects. Mol. Plant Pathol. 19 (5), 1282–1294. https://doi.org/10.1111/mpp.12613.

Torres, M.W., Corrêa, R.L., Schrago, C.G., 2005. Analysis of differential selective forces acting on the coat protein (P3) of the plant virus family Luteoviridae. Genetics and Molecular Research : GMR 4 (4), 790–802 doi.org/0157 [pii].

Tugume, A.K., Mbanzibwa, D.R., Alicai, T., Omongo, C.A., Maruthi, M.N., 2023. Endemism and Reemergence Potential of the Ipomovirus Sweet Potato Mild Mottle Virus (Family Potyviridae) in Eastern Africa: Half a Century of Mystery. Phytobiomes. J. 7 (1), 5–28. https://doi.org/10.1094/PBIOMES-05-22-0031-RVW.

Varrelmann, M., Maiss, E., 2000. Mutations in the coat protein gene of Plum pox virus suppress particle assembly, heterologous encapsidation and complementation in transgenic plants of Nicotiana benthamiana. Journal of General Virology 81 (3), 567–576. https://doi.org/10.1099/0022-1317-81-3-567.

Winter, S., Koerbler, M., Stein, B., Pietruszka, A., Paape, M., Butgereitt, A., 2010. Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa. Journal of General Virology 91 (5), 1365–1372.

Wokorach, G., Otim, G., Njuguna, J., Edema, H., Njung'e, V., Machuka, E.M., Yao, N., Stomeo, F., Echodu, R., 2020. Genomic analysis of Sweet potato feathery mottle virus from East Africa. Physiol. Mol. Plant Pathol. 110, 101473 https://doi.org/10.1016/j.pmpp.2020.101473.

Xu, Y., Zhang, S., Shen, J., Wu, Z., Du, Z., Gao, F., 2021. The phylogeographic history of tomato mosaic virus in Eurasia. Virology. 554, 42–47.

Yasaka, R., Nguyen, H.D., Ho, S.Y.W., Duchêne, S., Korkmaz, S., Katis, N., Takahashi, H., Gibbs, A.J., Ohshima, K., 2014. The temporal evolution and global spread of Cauliflower mosaic virus, a plant pararetrovirus. PLoS. One 9 (1), e85641. https://doi.org/10.1371/journal.pone.0085641.