Makerere University Journal of Agricultural and Environmental Sciences
Vol. 12 (1). pp. 39 - 65, 2023

ISSN 1563-3721
eISSN: 2958-4795

Computational evidence that the *Ugandan Passiflora virus* likely evolved from the *Bean common mosaic necrosis virus* primarily through recombination

Ochwo-Ssemakula, M.[1], Mbewe, W.[2] and Badji, A.[1,3,4*]

[1]Department of Agricultural Production, College of Agricultural and Environmental Sciences, Makerere University, P. O. Box 7062, Kampala, Uganda.
[2]Department of Biological Sciences, Academy of Medical Sciences, Malawi University of Science and Technology, P. O. Box 5196, Limbe, Malawi
[3]Legumes Research Program, National Crops Resources Research Institute, National Agricultural Research Organization, P. O. Box 295, Entebbe, Uganda
[4]Makerere University Regional Centre for Crop Improvement, College of Agricultural and Environmental Sciences, Makerere University, P. O. Box 7062, Kampala, Uganda

*Corresponding author: dabgarf42@gmail.com

Abstract

The *Ugandan Passiflora virus* (UPV) is an economically devastating virus affecting passion fruit (*Passiflora edulis* Sims), and a close relative of the *Bean common mosaic necrosis virus* (BCMNV). Using publicly available sequences, we studied the evolutionary relationship between the two viral species to understand UPV's viral origin and evolutionary pressures associated with its emergence. Seven UPV virus sequences obtained from our previous study were used to identify 59 complete genome sequences with at least 70% similarity with our query and sequence coverage through batch-BLASTn on the NCBI website. Eleven recombination breakpoints (RBP) were revealed evidencing several recombination events within the viral sequences, including UPV as a recombinant of BCMNV (major parent) and a strain of the *Watermelon mosaic virus* (minor parent). Intensified selection and episodic selection, as well as the existence of positively selected sites and several selective sweep events were detected around the RBPs, indicating that the high recombination rates could have

maintained genetic diversity within the viruses in the presence of positive selection. Besides, some of the viruses, including UPV, harbored unique conserved protein domains, suggesting that their emergence could have been through function gains, allowing colonization of new environments and hosts. This is the first computational report that elucidates the viral evolutionary origin of UPV and provides hypothetical evidence that it likely evolved from BCMNV, mainly through recombination. The results we present, however, need to be further investigated with the discovery and inclusion of more UPV isolate sequences to confirm its evolutionary relationship with other potyviruses.

Key words: Host jump, plant virus, potyviruses, selection pressures, viral evolution

Introduction

Passion fruit (*Passiflora edulis* Sims) is a widely cultivated perennial vine native to South America, which thrives in tropical and subtropical regions worldwide. Its economic importance is evident in countries such as Australia, Hawaii, Malaysia, Sri Lanka, and Taiwan where the crop is commercially grown (Gioppato *et al.*, 2019; de Carvalho *et al*., 2021). In Africa, passion fruit plays a significant role in increasing the income of small-scale farmers, particularly in Uganda, where it serves as a crucial cash crop for disadvantaged farmers (Ochwo-Ssemakula *et al*., 2012; Atukunda *et al.*, 2018). However, viral diseases pose a significant threat to passion fruit production, with the woodiness disease being one of the most prevalent and destructive (Cerqueira-Silva *et al.*, 2014). The woodiness disease causes a range of symptoms, including foliar mosaic, fruit woodiness, and deformation, leading to substantial yield losses and a shortened lifespan for passion fruit orchards (Ochwo-Ssemakula *et al.*, 2012).

The causal agent of passion fruit woodiness disease (PWD) was identified as the *Passion fruit woodiness virus* (PWV) in Australia in 1973 (Cerqueira-Silva *et al*., 2014). The virus was later discovered in other regions, such as Brazil, which reported PWV as the primary pathogen responsible for PWD (Cerqueira-Silva *et al*., 2014). However, subsequent investigations in Brazil and South Africa revealed that the *Cowpea aphid-borne mosaic virus* (CABMV) was responsible for PWD in those regions (Mendes *et al*., 2022; Coutinho and Wingfield, 2022). In Japan, the *East Asian Passiflora virus* (EAPV) was identified as the causal agent (Iwai *et al*., 2006), while Uganda has the *Ugandan Passiflora virus* (UPV) associated with PWD (Ochwo-Ssemakula *et al*., 2012; Mbeyagala *et al*., 2019). UPV is an RNA potyvirus that shares evolutionary similarities with the *Bean common mosaic necrosis virus* (BCMNV), another potyvirus species found in common bean (Ochwo-Ssemakula *et al*., 2012). The availability of a near-complete genome sequence for UPV (accession number MK110656 in GenBank) has created opportunities for comparative evolutionary analyses, particularly in relation to BCMNV. These analyses

can provide insights into the origin and evolution of UPV, thereby informing future management strategies. Both UPV and BCMNV have single-stranded RNA genomes, approximately 9,500-10,000 and 9,200-9,800 nucleotides in length, respectively. Their genome organizations include untranslated regions (UTRs) at the 5' and 3' ends, containing regulatory elements for replication, translation, and packaging. Multiple open reading frames (ORFs) within the genome encode various proteins, with the coat protein (CP) gene specifically responsible for forming the viral capsid.

The evolution and adaptation of viruses, including plant viruses such as potyviruses, is driven by several important genetic phenomena, including recombination, selection pressure, selective sweeps, and genetic hitchhiking. These processes contribute to diversification and emergence of new viral variants, which can have implications for viral pathogenicity, host range, and transmission dynamics. Recombination is a significant phenomenon in the evolution of viruses and can play a crucial role in their emergence and adaptation to new hosts or environments. Recombination occurs when two different viral genomes co-infect the same host cell, leading to exchange of genetic material. This process can result in the generation of novel viral variants with altered characteristics, including enhanced virulence or altered host range (Gibbs and Weiller, 1999). Selection pressure is another important factor shaping the evolution of viruses. Viruses face constant selective pressures from the host immune system, antiviral treatments, and environmental factors. These pressures drive the selection of genetic variants that confer a fitness advantage, allowing the virus to better evade host defenses or adapt to new conditions (Holmes, 2009). Selective sweeps occur when a beneficial mutation or genetic variant rapidly spreads through a viral population due to strong positive selection. This process can lead to fixation of advantageous alleles and rapid evolution of viral populations (Smith and Haigh, 1974). Genetic hitchhiking refers to the phenomenon where a beneficial mutation or genetic variant spreads through a population due to its physical linkage to another beneficial allele. When one allele experiences positive selection, the genetic variants linked to it may also increase in frequency, even if they do not directly contribute to the fitness advantage (Smith and Haigh, 1974). This process can result in the spread of linked alleles that may have both positive and negative effects on viral fitness.

In this study, we conducted an in-depth analysis of the evolutionary relationship between UPV and other potyviruses, with a focus on providing computational evidence supporting the hypothesis that UPV likely evolved from BCMNV through recombination. By investigating the genetic variation and patterns of selection in the UPV genome, we aimed to identify signatures of recombination events and selective pressures that have shaped the evolution of this virus. Additionally, we examined the potential effects of selective sweeps and genetic hitchhiking on the spread of advantageous genetic variants in UPV populations.

By elucidating the genetic mechanisms underlying the emergence and spread of UPV, this research aims to contribute to improved prevention and management strategies for the woodiness disease in passion fruit. Understanding the evolutionary dynamics of UPV and its relationship to other potyviruses can inform the development of targeted approaches for disease prevention and management, ultimately safeguarding passion fruit production worldwide.

Materials and methods

Nucleotide sequence blasting and protein sequence retrieval

Seven UPV virus sequences (one complete genome [accession number MK110656.1] and six polyprotein sequences [accession numbers FJ896003.1, FJ896000.1, FJ896001.1, FJ896002.1, MK132862.1, and JF427623.1]), obtained from a previous study (Ochwo-Ssemakula *et al*., 2012), were batch-blasted (BLASTn) against the national NCBI databases. The blast parameters were: Standard Database = Nucleotide collection (nr/nt), Organism = Potyviridae, Optimized for = Somewhat similar sequences (blastn), Percent Identity = 70%, and Percent Coverage = 70%.

Pairwise similarities among the viral sequences were estimated using the Sequence Demarcation Tool (STD 1.2) with ClustalW as a multiple-alignment tool (Muhire, Varsani and Martin, 2014). Sequences were entered in the Open Reading Frame Finder (ORF Finder, https://www.ncbi.nlm.nih.gov/orffinder/) on the NCBI database to retrieve their corresponding protein sequences along with their untranslated nucleotide sequences (59-correspond). The "ATG only" option was selected to determine the start codon which corresponds with the selected genetic code "Standard." For each virus, the protein with the longest sequence was considered.

Analysis of viral recombination

Viral sequence recombination was assessed using the Recombination Detection Program (RDP) version 5 (Martin *et al*., 2021) with multiple methods: RDP, GENECONV, Bootscan, Maxchi, Chimaera, SiScan, PhylPro, LARD, and 3Seq. Recombination was considered if detected by at least three methods (Wainaina *et al*., 2019) at $P < 0.05$. The Datamonkey database (https://www.datamonkey.org, Weaver *et al*., 2018) was used with the GARD method (Kosakovsky Pond *et al*., 2006) to confirm recombination points within the 59-original codon-aligned sequences. Breakpoint-partitioned sequence data and corresponding phylogenetic trees for each partition were obtained as a nexus file for molecular evolution analyses.

Phylogenetic analyses and conserved protein domain (CPD) search

Original and partitioned sets of nucleotide sequences and protein sequences were multiple-aligned separately using ClustalX 2.1 (Thompson *et al*., 1997). The 59-correspond sequence alignment from ClustalX was additionally aligned using the JCoDA 1.4 program (Steinway *et al*., 2010) for downstream evolutionary analyses, resulting in an alignment of 3231 codons. A phylogenetic tree of the 59-original alignment was constructed in MEGA-X (Kumar *et al*., 2018) with the Sugarcane streak mosaic virus (SCSMV) complete genome sequence as an outgroup, using the maximum likelihood method with 1000 bootstrap repetitions. The aligned protein sequences underwent CPD batch-searches on the NCBI conserved domain search platform (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi, Marchler-Bauer *et al*., 2017; Lu *et al*., 2020) with default parameters. The complete analysis output was retrieved and locally analyzed to identify unique conserved domains for specific viral species.

Analysis of evolutionary processes

Identification of viral regions affected by selective sweeps

Selective sweeps were detected using two approaches: OmegaPlus 3.0.3 (Alachiotis *et al*., 2012), a linkage-disequilibrium-based method, and SweeD 3.0 (Pavlidis *et al*., 2013), a likelihood-based method. Both programs were implemented in the Ubuntu 20.04 Linux environment.

Pairwise nonsynonymous to synonymous rates among viral sequences

Selection pressures on viral sequences were determined by comparing the rate (K) of nonsynonymous (dN) to synonymous (dS) substitutions among the 59 codon-aligned viral sequences, with $K = dN/dS$. A $K < 1$ indicates purifying selection, $K > 1$ signifies positive selection, and $K = 1$ suggests neutral selection. For this analysis, K values between 0.98 and 1.02 were considered neutral selection. The online platform SNAP (Bromberg and Rost, 2007) was utilized, which employs a neural network-based method incorporating biophysical characteristics of nucleotide substitutions and evolutionary information to predict the likelihood of a mutation altering a protein (Bromberg and Rost, 2007).

Branch-level and codon-level selection pressures acting on viral sequences

Evolutionary processes were examined on the Datamonkey database to detect pervasive or episodic positive and negative selection at both branch and codon levels (Weaver *et al*., 2018; Spielman *et al*., 2019) within the 59 sequences. Branch-level methods were used for non-partitioned codon-aligned sequences, while codon-level methods were employed for partitioned sequences.

Investigation of branch-level selection pressures

Branch-level methods were employed to detect differential selection on specific branches compared to the rest of the phylogeny. These methods included RELAX (Wertheim *et al*., 2015), BUSTED (Branch-site Unrestricted Statistical Test for Episodic Diversification), and aBSREL (Smith *et al*., 2015).

RELAX determines whether a subset of branches underwent relaxed (or intensified) positive or negative selection compared to the reference branches. Statistical significance is determined through a Likelihood Ratio Test (Anisimova, 2019; Spielman *et al*., 2019). The RELAX analysis considered three clusters (18 in blue shade, 14 in pink shade, and 12 sequences in green shade) identified by the Maximum Likelihood Tree, as well as single-branch selections including recombinant and the minor and major parents of UPV (Fig. 1). All internal nodes linking selected sequences to their earliest common branch were also included in the test for each selection.

BUSTED examines whether branches in the phylogeny experience gene-wide positive pervasive or episodic selection. Each branch-site combination is assumed to evolve with specific ω values, and three-rate ω distributions are estimated from the entire alignment (Anisimova, 2019; Spielman *et al*., 2019). The same selection as in the RELAX method (Fig. 1) is applied to determine the foreground and reference lineages.

The aBSREL method identifies codons on specific branches subject to positive selection. The optimal number of evolutionary rate categories per branch is inferred using a small-sample Akaike Information Criterion correction (AICc). Statistical significance is assessed using a mixture of chi-square distributions, and the resulting P-values were corrected for multiple testing using a Bonferroni-Holms procedure to control false-positive rates (Anisimova, 2019; Spielman *et al*., 2019).

Investigation of codon-level selection pressures

The branch-level methods described above do not pinpoint the specific codons involved in positive selection on selected branches. However, the codon-level method, MEME (Mixed-Effect Model of Evolution, Murrell *et al*., 2012), offers the most powerful approach for detecting positively-selected codons and identifying the branches under selection on these codons (Spielman *et al*., 2019). MEME detects both pervasive and episodic selection, and positively-selected codons were considered significant at P-value thresholds of 0.05 and 0.1, with the latter being the default and less conservative threshold (Anisimova, 2019; Spielman *et al*., 2019). Branches were deemed under selection on positively-selected codons if their Empirical Bayes Factor (EBF) exceeded 100, the default threshold. To investigate pervasive diversifying selection, common in pathogens due to host immune escape and the evolutionary arms race, the FUBAR (Fast Unconstrained Bayesian AppRoximation)

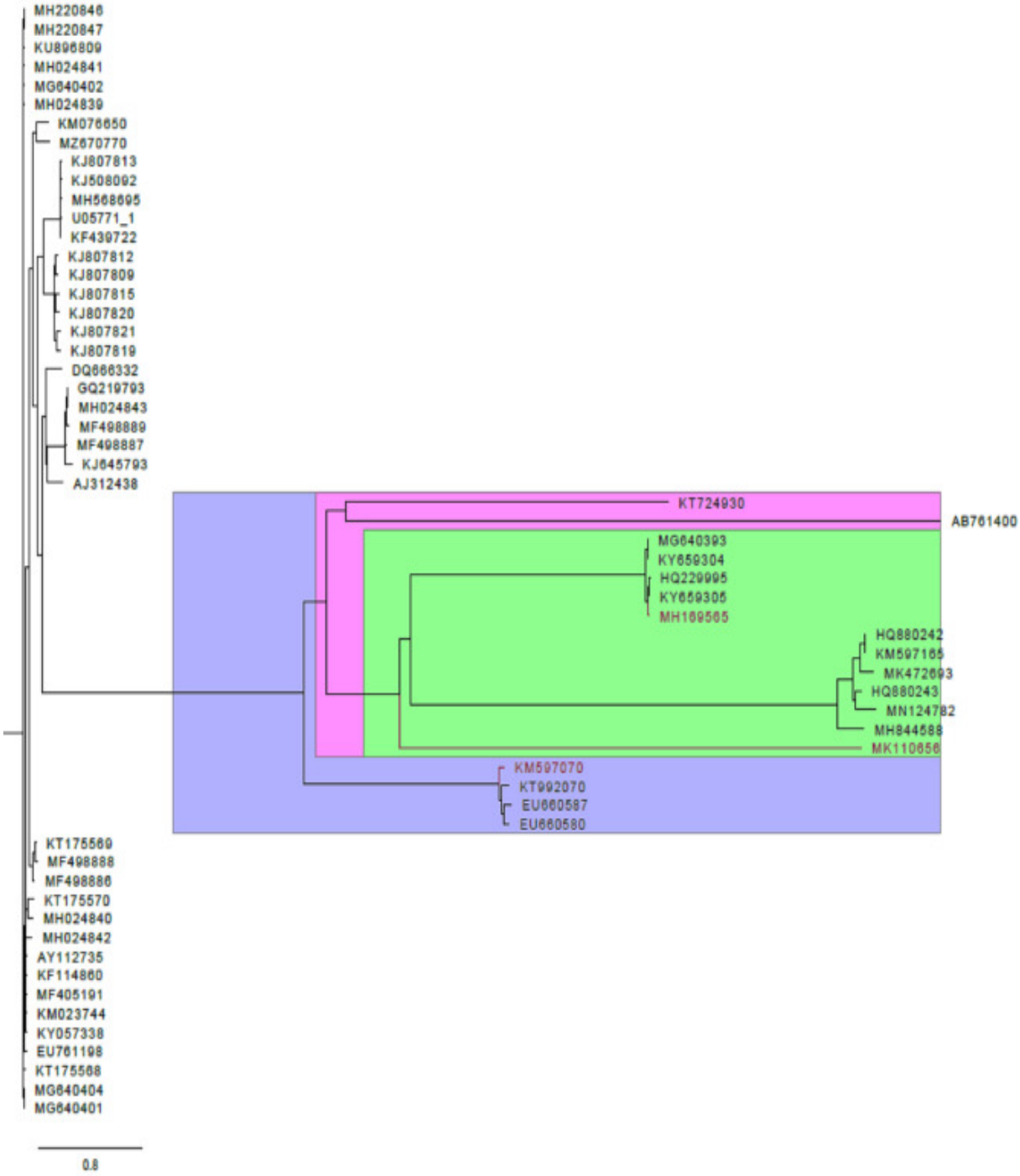


Figure 1. Phylogenetic tree showing the different sets of branches selected as a test in the RELAX and BUSTED analyses. The sequences shaded in green were first tested, then these plus the ones shaded in pink were tested, and finally, the sets shaded in green, pink, and purple were also tested together.

program implemented in the Datamonkey platform was employed. FUBAR identifies sites subject to strong selective pressures across the entire phylogeny, with a default posterior probability of 0.9 indicating pervasive positive (directional) selection (Anisimova, 2019; Spielman *et al*., 2019).

Results

Viral sequence identities

Blastn analysis yielded 100 hits, including 59 complete viral genomes (ranging from 9000nt to 11,000nt in length) and 40 polyproteins (S1[1]). The focus was on the 59 complete viral genomes, with percent identities ranging from 100% for the Ugandan Passiflora virus (UPV, accession number MK1106565.1) to 69.33% for Cowpea aphid-borne mosaic virus isolate CABMV/BJL1 (MN124782.1_CABMV_BJL1) (Fig. 2). The closest sequences to UPV in terms of percent identity were Bean common mosaic virus strain Peanut stripe virus isolate Laixi (KF439722.1_BCMV_psv_Laixi) (72.38%) and Bean common mosaic virus isolate LNSe (MH568695.1_BCMV_LNSe) (72.26%) (Fig. 3). Sequence lengths varied from 9608 bp for Bean common mosaic necrosis virus isolate N3 (BCMNV_N3, accession number MH169565.1) to 10151 bp for East African Passiflora virus isolate EAPV_IB_dpd (EAPV_IB_dpd, accession number KT724930.1).

Phylogenetic relationship, recombination, and selective sweeps within the viral sequences

The phylogenetic relationship among viral sequences revealed two main clusters: one consisting of 40 *Bean common mosaic virus* isolates and one *Peanut stripe virus* (PSV) isolate (Cluster 1), and another cluster comprising 18 viruses of different species, including UPV (Cluster 2) (Fig. 4). UPV closely clustered with five BCMNV isolates within the second cluster. This sub-cluster, along with six CABMV isolates, formed a larger group with a bootstrap value of 96%. This group of 12 viruses was further grouped with six isolates from three species: four *Watermelon mosaic virus* (WMV) isolates, one *Passion fruit woodiness virus* (PWV) isolate, and one EAPV isolate.

Recombination was detected within the viral sequences using both the GARD algorithm and the RDP5 software. The GARD algorithm identified 11 breakpoints in the viral sequence alignment, resulting in 12 partitions (1-536, 537-1256, 1257-1779, 1780-2677, 2678-3332, 3333-3993, 3994-6017, 6018-6752, 6753-7832, 7833-8667, 8668-9185) (Fig. 5A). The RDP5 software detected several triplets of recombinant

[1]a file with supplementary information for this paper was deposited with the journal and can be requested via mujaes.caes@mak.ac.ug

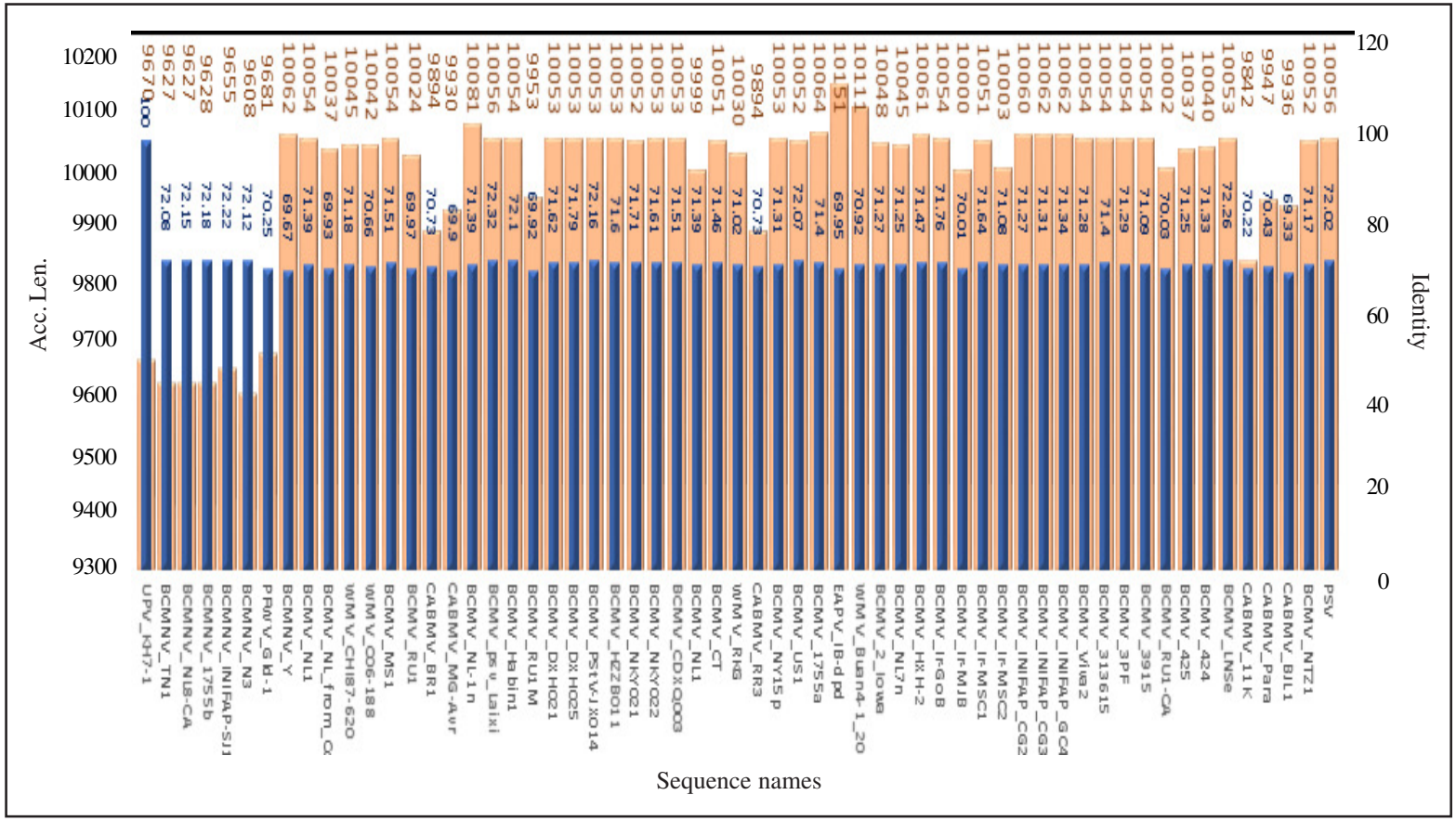


Figure 2. Query-to-hit sequence percent identity and sequence length of the 59 complete genome viral sequences resulting from the Blastn analysis.

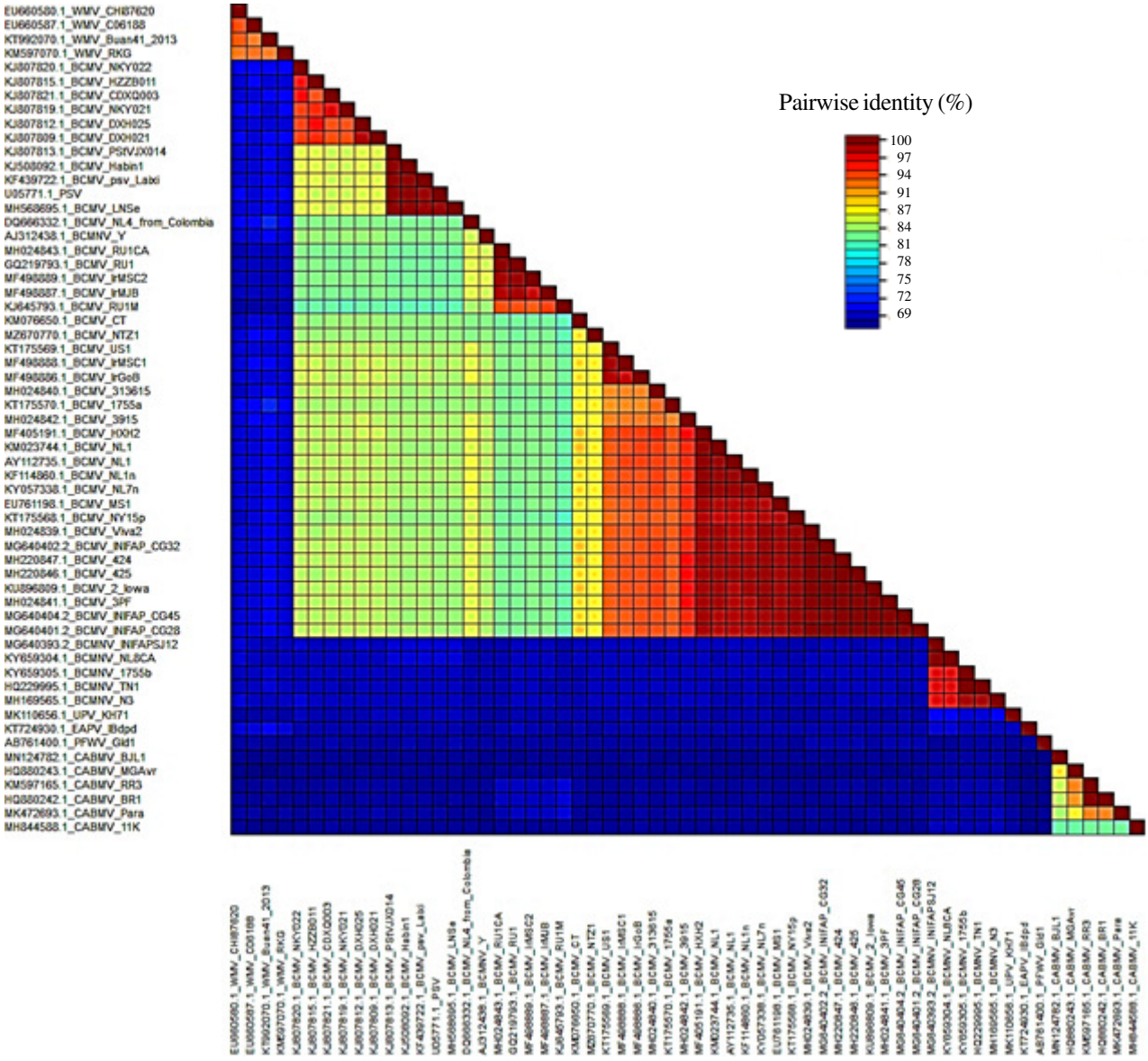


Figure 3. Pairwise percent identity among all the 59 viral sequences.

and parental sequences, forming recombination clusters (RC) where shared parental lines were present (Fig. 4). Fifty-six out of the 59 viruses showed evidence of involvement in viral recombination, with 35 as recombinants, 46 as minor parents, and 37 as major parents. UPV was identified as a recombinant, with BCMNV isolate N3 as the major parent and WMV isolate RKG as the minor parent (Fig. 4). However, this recombination event was flagged as potentially misidentified, although all three methods (Maxchi, Chimera, and 3seq) showed significance of at least 0.05.

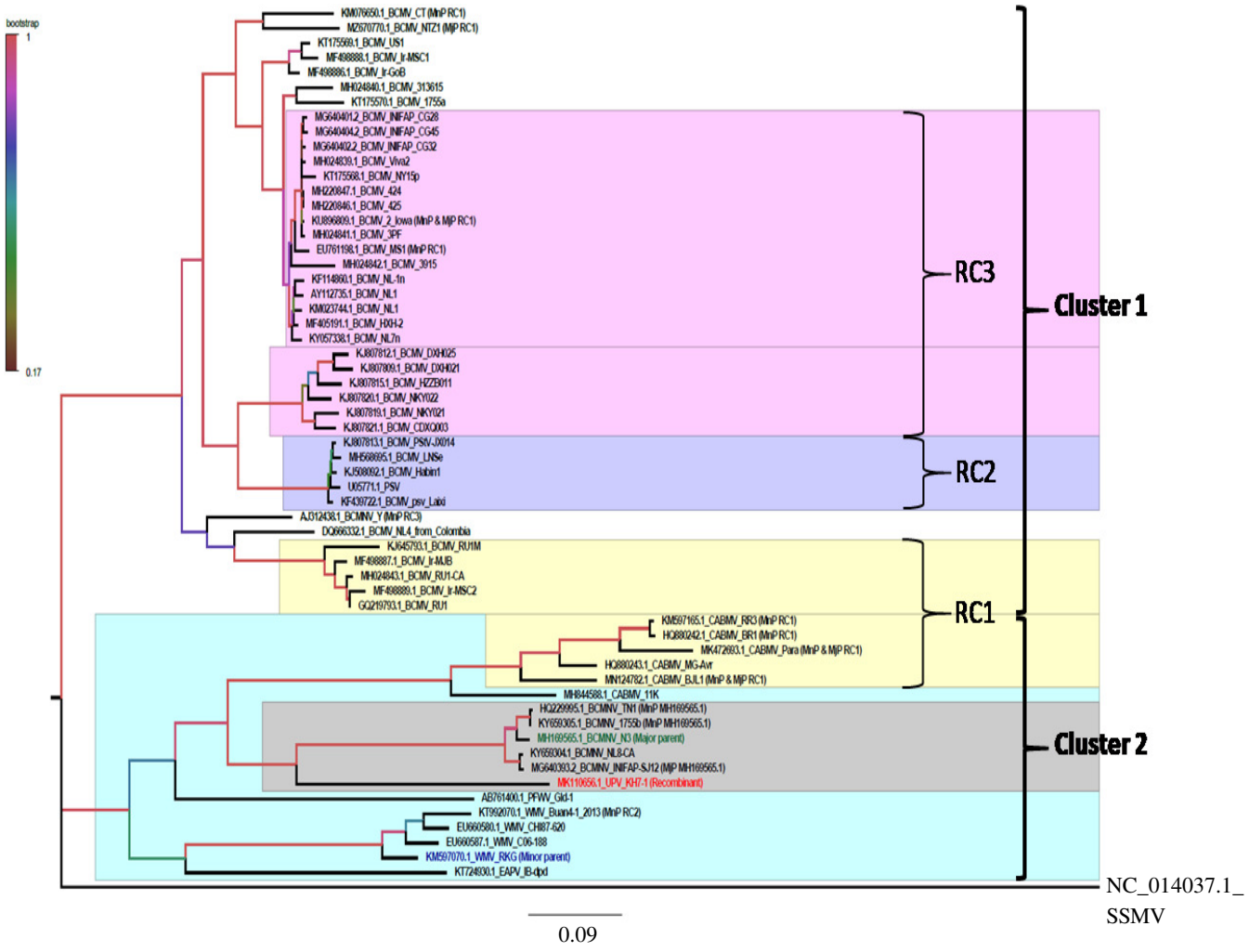


Figure 4. Maximum likelihood phylogenetic relationship amongst the viral sequences. Information is also provided for recombination clusters (RC) sharing major (MjP) and/ or minor (MnP) parents resulting from the recombination analysis. All viruses in a recombination cluster are detected as recombinants from parents either clustered with these or located in other clusters. The bootstrap color legend ranges for 1=100% to 0.17=17%.

The major parent of UPV, BCMNV isolate N3, appeared to be a product of another recombination event involving BCMNV isolate INIFAP SJ12 as the major parent and either or both BCMV isolate 1755b and BCMV strain TN1 as possible minor parents.

Recombination analysis revealed three additional recombination clusters (RCs) outside the one involving UPV and its parental sequences. RC1 consisted of isolates of

BCMNV and BCMV, with potential major or minor parents identified. RC2 and RC3 were uniquely represented by BCMV isolates, with specific minor parents but unknown major parents. Selective sweep analysis identified several genomic regions in the viral genome where selective sweeps occurred (Fig. 5C). These selective sweep regions were often found around recombination breakpoints, although results from the two methods did not match exactly. Similar trends were observed for positively selected sites detected by the MEME and FUBAR programs (Fig. 5D). Fourteen sites showed significant pervasive positive selection in specific partitions according to both methods (Fig. 5E).

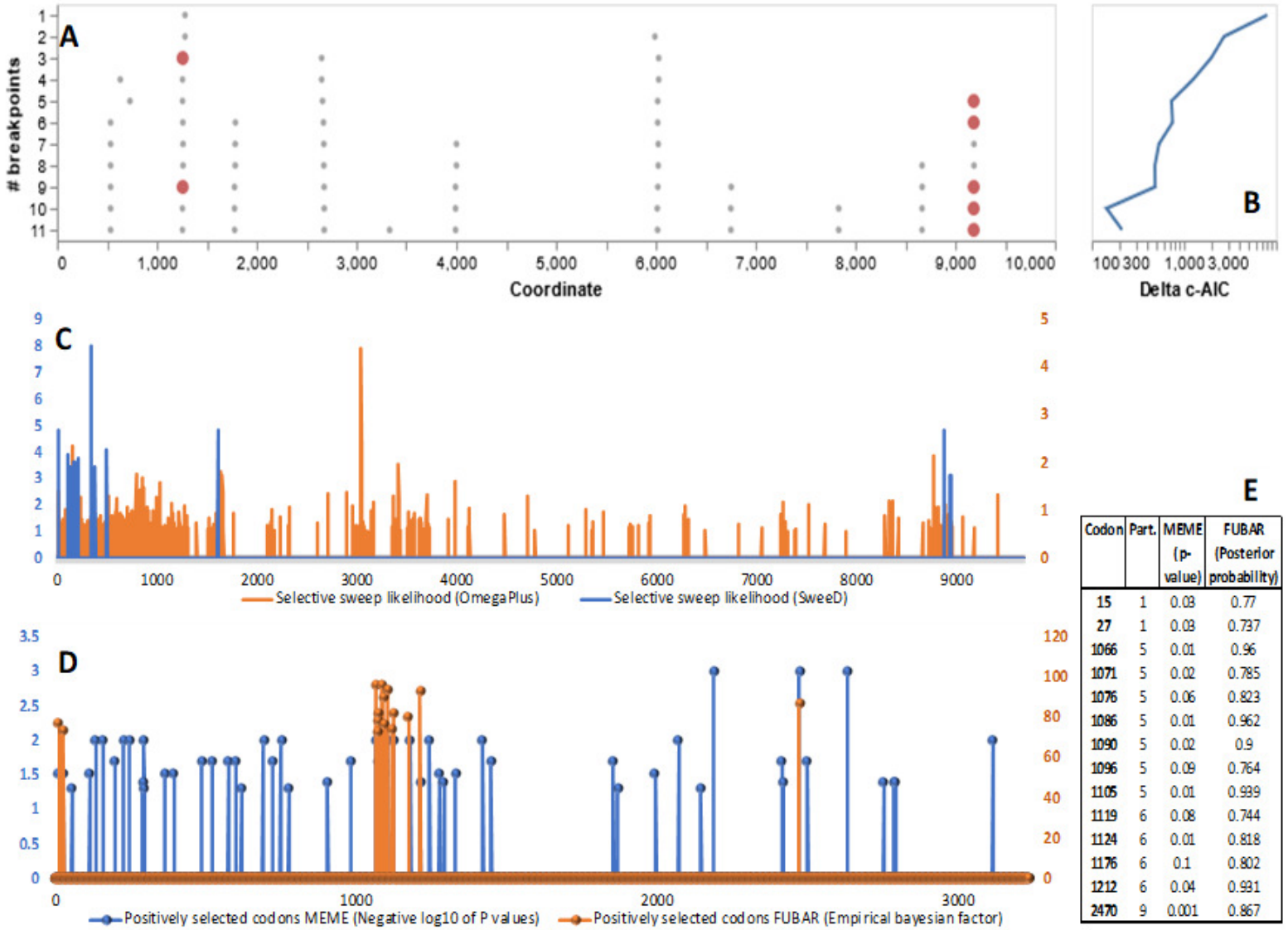
Pairwise viral sequence nonsynonymous to synonymous mutation rates

Most viral sequence pairs showed rates of nonsynonymous (dN) to synonymous (dS) mutation ratios below 1, indicating a prevalence of purifying selection among the studied viral genomes (Fig. 6). This negative selection pattern was observed in pairs involving the recombinant UPV, its potential major and minor parents, and all other sequences compared with UPV. Only a few viral sequence pairs exhibited a ratio (K) greater than 1. These pairs included specific combinations such as Bean common mosaic virus isolates DXH025 and DXH021, Watermelon mosaic virus isolates CHI87-620 and C06-188, Bean common mosaic necrosis virus isolates 1755b and N3, Bean common mosaic virus isolates MS1 and 424, and Bean common mosaic necrosis virus strain TN1 and isolate N3. Additionally, six viral sequence pairs showed K values close to 1, indicating neutral selection (Fig. 6).

Branch-level selection pressures on the viral genomes

With the RELAX algorithm, positive selection was observed in all tested sets of branches, except for the branch involving the UPV viral sequence (recombinant) and its minor parent KM597070.1_WMV_RKG (Table 1). BUSTED analysis also confirmed these results, except for the major parent of the UPV recombinant, MH169565.1_BCMNV_N3, which did not show significant evidence of gene-wide positive selection at a significance level of 0.05. Both analyses indicated that the UPV sequence experienced purifying selection rather than positive selection, consistent with the dN/dS ratios observed in pairwise comparisons of the virus.

To investigate individual branches with codons under positive selection, the aBSREL algorithm was used since RELAX and BUSTED do not provide such information. Among the 112 tested branches, the complexity model of 1 ω rate class described 62 branches, but none showed positive selection (Table 2). A 2 ω rate class model was best fit for 45 branches, covering 95% of the tree length, with 11 branches exhibiting positive selection at a p-value threshold of 0.05. The 3 ω rate class model described five branches (4.9% of the tree length), but no branch showed selection under this model.



E

Codon	Part.	MEME (p-value)	FUBAR (Posterior probability)
15	1	0.03	0.77
27	1	0.03	0.737
1066	5	0.01	0.96
1071	5	0.02	0.785
1076	5	0.06	0.823
1086	5	0.01	0.962
1090	5	0.02	0.9
1096	5	0.09	0.764
1105	5	0.01	0.939
1119	6	0.08	0.744
1124	6	0.01	0.818
1176	6	0.1	0.802
1212	6	0.04	0.931
2470	9	0.001	0.867

Figure 5. Structure with the viral sequences. A: The best placement of breakpoints inferred by the GARD algorithm for each number of breakpoints considered. B: The improvement in the corrected Akaike's Information Criterion (c-AIC) score between successive breakpoint numbers (log scale). C: Plot of the location of selective sweeps identified by SweeD (SD59) and OmegaPlus (OP59). D: Positively selected codons revealed by the MEME (blue) and FUBAR (brown) algorithms. E: Significantly positively selected codons under both MEME (P-value<0.1) and FUBAR models (Posterior probability: 0.7) Part. =Partition.

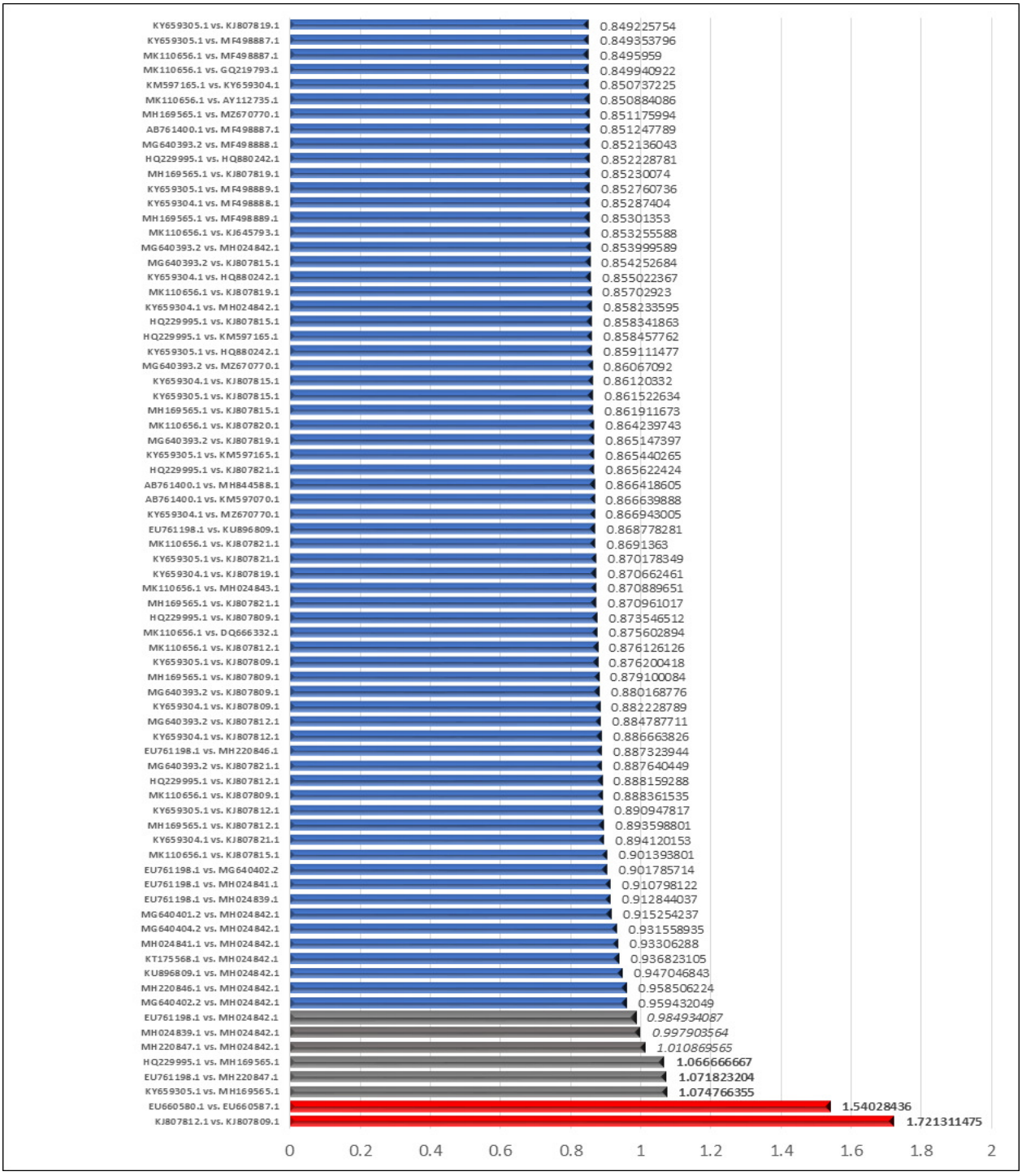


Figure 6. Pairwise ratios of nonsynonymous to synonymous substitutions among the viral sequences. In red are the sequence pairs that evidenced positive selection (diversifying selection) with a dN/dS>1, in grey pair evidencing neutral selection with dN/dS=1, and in blue are pairs that experienced negative selection (purifying).

Table 1. Evidence of intensified selection (a K > 1 with P-value <0.05 indicate intensified selection along the test branches) and episodic/pervasive selection along different test branch sets of the viral phylogenetic tree

Number of sequences selected test branches	K (RELAX)	P-value (RELAX)	Likelihood ratio (RELAX)	P-value BUSTED)
All 59 sequences	-	-	-	≤0.05
18	1.11	0.001	11.19	≤0.05
14	1.09	0.002	9.26	≤0.05
12	1.14	0	20.25	≤0.05
MK110656.1_UPV_KH71 (recombinant)	0.94	0.119^{NS}	2.43	0.500^{NS}
MH169565.1_BCMNV_N3 (major parent)	0.43	0.028	4.84	0.471^{NS}
KM597070.1_WMV_RKG (minor parent)	0.6	0.066^{NS}	3.39	0.155^{NS}

Table 2. Summary of the inferred aBSREL model complexity. Each row provides information about the branches that were best described by the given number of É rate categories

É rate classes	# of branches	% of branches	% of tree length	# under selection
1	62	55%	0.06%	0
2	45	40%	95%	11
3	5	4.50%	4.90%	0

The detailed results for the 11 branches that were statistically significant (P-value ≤ 0.05) for positive selection are not presented here (S2). Likelihood test ratios (LTR) ranged from 990.3798 to 13.1628. Notably, the Cowpea aphid-borne mosaic virus isolate MG-Avr (HQ880243.1) in Cluster 2 and RC1, the Bean common mosaic virus isolate RU1M (KJ645793.1) in Cluster 1 and RC1, and the Bean common mosaic virus isolate 1755a (KT175570.1) in Cluster 1 exhibited positive selection on specific branches. Nodes 59 and 63 were identified as statistically significant under positive selection by the aBSREL algorithm but did not show evidence of positive selection, except for the Cowpea aphid-borne mosaic virus isolate MG-Avr (HQ880243.1). These findings indicate that the UPV sequences did not experience branch-level or gene-wide positive selection, although certain codon sites within this sequence may have undergone positive selection.

Codon-level selection pressures on the viral genomes

Codon-level selection algorithms were employed to investigate individual codons undergoing positive selection on specific branches of the phylogeny. The MEME algorithm identified 104 codons under diversifying selection at a p-value of 0.1, and 56 codons at a p-value of 0.05 (S3). The number of branches exhibiting selection at these codons ranged from 6 to 0. There were 45 branches with an Empirical Bayes Factor (EBF) > 100, indicating evidence of positive selection at the codons identified by the MEME algorithm (S4). Out of these 45 branches, 43 were terminal branches representing viral sequences. The Bean common mosaic virus cowpea isolate Y (AJ312438) and the Bean common mosaic virus isolate NY15p (KT175568) had the highest number of codons under selection (6 codons). The UPV sequence (MK110656) showed positive selection at codons 2419 and 3230, while its major parent (MH169565) and minor parent (KM597070) exhibited positive selection at codons 2533 and 776, respectively.

The FUBAR algorithm identified five sites under pervasive positive/diversifying selection, while 2949 sites showed pervasive negative/purifying selection with a posterior probability threshold of 0.9. Similar to the observed selective sweeps, positively selected sites detected by the MEME and FUBAR algorithms tended to cluster around recombination breakpoints (Fig. 5D).

Conserved protein domains (CPD) enriched within the viral sequences

Conserved protein domains were found in all viral protein sequences, with some unique to specific sequences, particularly in Cluster 2 of the phylogenetic tree (Fig. 7). Examples include the BRR2 and DEXHc_HrpA domains found exclusively in UPV (MK1106565.1), and the comFa domain unique to the Watermelon mosaic virus isolate RKG (KM597070.1), which is a potential minor parent of UPV. The BRR2 domain is involved in replication, recombination, and repair, while DEXHc_HrpA is part of the HrpB-HrpA secretion system important for virulence-associated protein secretion. The comFa domain is associated with DNA uptake and also involved in replication, recombination, and repair.

The Bean common mosaic necrosis virus isolate N3, a potential major parent of UPV, and two other viruses in Cluster 2 (KY659305.1 and AB761400.1), however, did not have unique domains. The UPV minor parent, isolate KM597070.1, shared the DEXHc_DHX16 and DDXDc_reverse_gyrase domains with other WMV isolates (EU660587.1, EU660580.1, and KT992070.1). DEXHc_DHX16 is likely involved in pre-mRNA splicing, and reverse gyrase modifies DNA topology through ATP-dependent positive supercoiling. The Cosavirus_rdpr domain, with RNA-dependent RNA polymerase activity, was present in KM597070.1, UPV minor parent, and

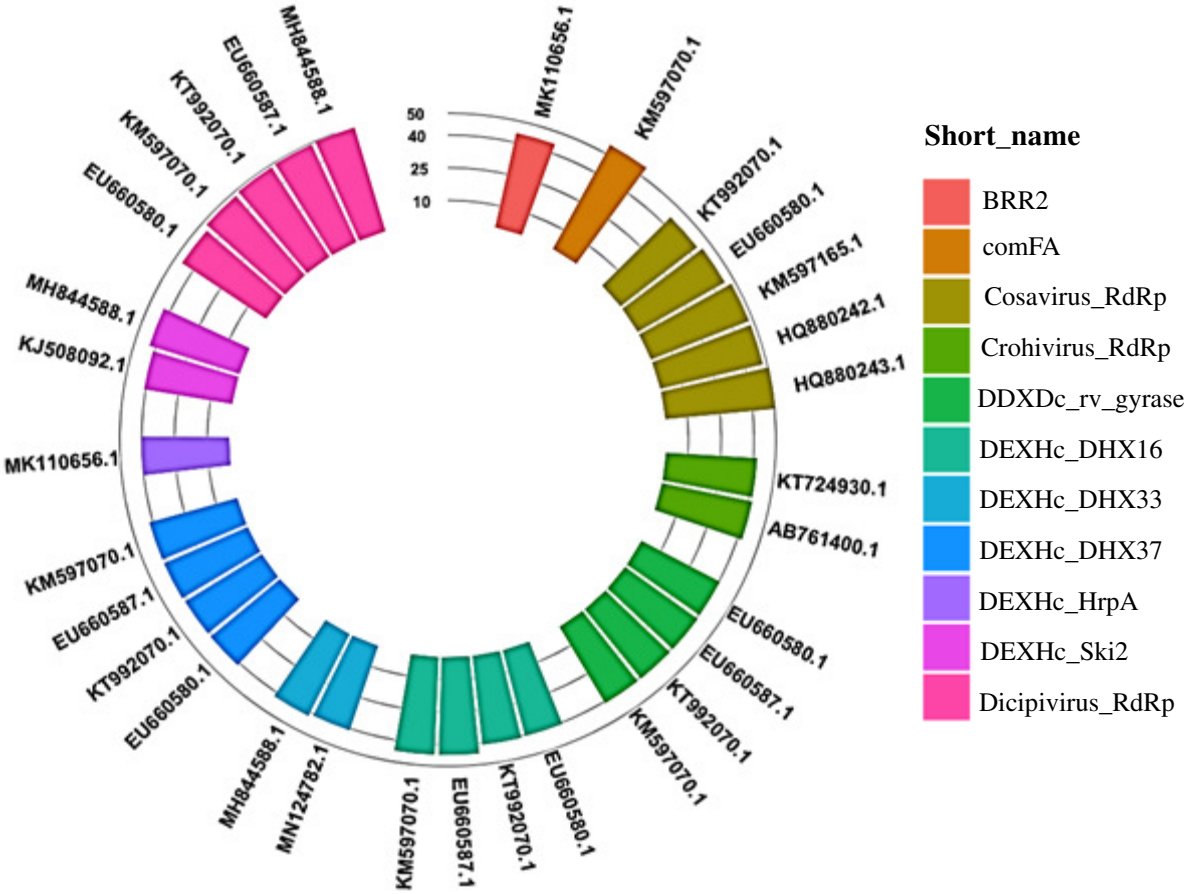


Figure 7. Sunburst of some conserved protein domains (inner circle bars) specific to few viral sequences. The height of each rectangle is proportional to the bitscore (ranging from 40.1313 to 54.3942) of the corresponding enriched conserved domain in the viral sequence.

other WMV isolates (KT992070.1 and EU660580.1) as well as CABMV isolates (HQ880242 and HQ880243).

In contrast, the DEXHc_Ski2 domain was exclusively found in the Bean common mosaic virus isolate Habin1 (KJ508092.1) and Cowpea aphid-borne mosaic virus isolate 11K (MH844588.1) protein sequences. The DEXHc_Ski2 domain belongs to the Ski2-like RNA helicases involved in RNA degradation, processing, and splicing. These helicases are part of the type II DEAD box helicase superfamily, which includes proteins involved in ATP-dependent RNA or DNA unwinding and contains the ATP-binding region.

Discussion

Analyzing 59 closely related full-length virus genomes has provided crucial insights into the evolutionary forces shaping UPV. Previously, studies had shed light on the recombination and evolutionary history of most plant potyviruses (Padhi and Ramu, 2011; Gell *et al*., 2015; Willemsen *et al*., 2018; Rao *et al*., 2020; Abdalla and Ali, 2021; Peng *et al*., 2021) but the evolutionary footprint of UPV, a recently characterized virus associated with passion fruit woodiness disease remained poorly understood. This information is vital for effectively managing the virus' impact on passion fruit productivity and commercialization (Ochwo-Ssemakula *et al*., 2012; Bancy *et al*., 2019; Mbeyagala *et al*., 2019).

UPV shares similarity with BCMNV, a potyvirus causing symptoms in common bean and other legume crops (Ochwo-Ssemakula *et al*., 2012; Wainaina *et al*., 2019; Worrall *et al*., 2019; Sengooba *et al*., 1997). BLASTN analysis confirmed the close relationship between UPV and BCMNV isolates (Ochwo-Ssemakula *et al*., 2012; Bancy *et al*., 2019; Mbeyagala *et al*., 2019). Recombination analysis revealed UPV as a recombinant with BCMNV as the major parent and WMV as the minor parent (Fred *et al*., 2017). Recombination events were also detected among other viral sequences, supported by recombination breakpoints (Petrzik, 2019; Nagy, 2008; Simon-Loriere and Holmes, 2011; Dolan *et al*., 2018). Previous studies on EAPV, another virus associated with passion fruit woodiness disease, have also shown evolutionary relationships with BCMV and BCMNV (Valli, López-Moya and García, 2007). Our findings emphasize the role of recombination in the emergence of new virus isolates such as UPV, which may acquire new biological properties (Sztuba-Soliñska *et al*., 2011; Green *et al*., 2017). The complex evolutionary journey of UPV warrants further investigation (Nagy, 2008; Sztuba-Soliñska *et al*., 2011; White *et al*., 2011; Miyashita, 2018; Bujarski and Kaesberg, 1986; Carpenter *et al*., 1995; Ali *et al*., 2006; Nuss, 2011). Our study reveals a distinct clade formed by UPV and BCMNV, highlighting the involvement of recombination in UPV's evolution (Sztuba-Soliñska *et al*., 2011).

Conserved protein domains, arising from mutations, recombination, and ligand interactions, are crucial for viral genomes (Li *et al*., 2019; Brito and Pinney, 2020; Aziz and Caetano-Anollés, 2021; Bordin *et al*., 2021). UPV possesses unique conserved domains, including BRR2 involved in replication, recombination, and repair (Rodamilans *et al*., 2018). The DEXHc_HrpA domain, shared with other potyviruses, facilitates the secretion of virulence-associated proteins (Marsh and Teichmann, 2010). Recombination and domain reshuffling contribute to protein evolution and complexity (Vogel, Teichmann and Pereira-Leal, 2005). Recombination events between

BCMNV and WMV, potentially through domain fusion, could have generated recombinant conserved protein domains in UPV (Basu *et al*., 2009). The acquisition of new domains and interactions may have enabled UPV to infect passion fruit (Vogel, Teichmann and Pereira-Leal, 2005; Li *et al*., 2019). These changes likely occurred in the polyprotein, which plays essential roles in transmission, host adaptation, and infectivity (Adams *et al*., 2005; Lopez-Moya and Garcia, 2008; Cui and Wang, 2016; Rodamilans *et al*., 2018).

Our study identified selective sweeps near recombination breakpoints, indicating recent beneficial mutations (Stephan, 2019). Selective sweeps occur when a beneficial mutation spreads throughout a population, driven by positive selection (Messer and Petrov, 2013). Positive selection, detected through codon-level and branch-level selection pressure analyses, is associated with selective sweeps (Stephan, 2019). Recombination interrupts genetic hitchhiking, allowing fixation of beneficial mutations without significant reduction in genetic variation (Stephan, 2019; Ebert and Fields, 2020).

Pairwise sequence dN/dS (K) comparisons showed that viral sequences, including UPV and its proposed parents, evolved under negative selection, except for two pairs with clear evidence of positive selection (Stephan, 2019). Potyviruses are generally subject to negative selection (Chiaki *et al*., 2016; Parizad *et al*., 2018; Nigam *et al*., 2019; Gibbs *et al*., 2020; Moradi and Mehrvar, 2021). Intensified selection was detected in some branches using the RELAX and BUSTED algorithms, but individual testing of UPV and its parental sequences showed non-significant results (Anisimova, 2019; Spielman *et al*., 2019). The aBSREL algorithm identified 11 branches with positive selection, including the major parent of UPV (MH169565.1_BCMNV_N3) (Anisimova, 2019). MEME and FUBAR methods detected codons under positive selection, suggesting their significance in host colonization by potyviruses (Nigam *et al*., 2019). Although UPV and its parental sequences did not explicitly show positive selection, these findings support the emergence of new potyvirus species through positive selection and recombination (Nigam *et al*., 2019; Forni *et al*., 2020). The role of positive selection, recombination, and conserved protein domain gain in shaping UPV's emergence warrants further investigation.

Understanding viral pathogen emergence is crucial for effective crop protection (Stephan, 2019). Recombination events and unique conserved protein domains in UPV indicate a potential host shift and increased virulence (Rodamilans *et al*., 2018). Monitoring viral populations, especially in intercropping systems, is vital due to mixed infections and recombination (Marsh and Teichmann, 2010). Early detection of

emerging viral pathogens like UPV facilitates deployment of appropriate management strategies, including crop rotation, breeding for resistance, and targeted antiviral use. Selective sweeps and positive selection in certain viral sequences imply ongoing adaptation (Stephan, 2019). Monitoring genetic diversity and evolutionary dynamics facilitates prediction of changes in virulence, host range, and resistance-breakdown (Nigam *et al*., 2019). This knowledge guides the development of durable resistance strategies and deployment of resistant cultivars, thus, safeguarding agricultural production.

Conclusion

Our study shows UPV's recombination from BCMNV, with minor input from WMV, contributing to unique protein domains and potentially enhancing virulence. Phylogenetic analysis supports UPV's close relation to BCMNV, indicating a shared evolutionary history. Selective sweeps and positive selection episodes suggest recent changes and highlight viral evolution's dynamic nature. These findings advance our understanding of plant viral pathogen emergence and evolution, impacting virus management and crop protection.

To advance our understanding of UPV and its evolutionary dynamics, the following research recommendations are suggested: Include additional UPV isolate sequences: Given that the study relied on only one UPV sequence, it is essential to prioritize the inclusion of additional UPV isolate sequences in future research. Expansion of the dataset will significantly enhance the statistical robustness of evolutionary inferences, enabling researchers to validate the findings and establish a more comprehensive evolutionary relationship between UPV and other potyviruses. Investigate functional implications and virus-host interactions: Further research should explore the functional implications of genetic changes observed in UPV and investigate the resulting virus-host interactions. Specifically, it is important to study the molecular interactions between UPV's adaptive factors and the defense mechanisms of passion fruit, as well as potential hosts in the Cucurbitaceae, Fabaceae, Solanaceae, and Chenopodiaceae families. These investigations will provide valuable insights into how UPV has adapted to different hosts and contribute to the development of more effective management strategies for UPV and woodiness disease in passion fruit. Addressing these recommendations will strengthen our understanding of UPV's evolution, its relationship with other potyviruses, and the mechanisms underlying its interactions with host plants. This knowledge will inform the development of targeted approaches for disease management control and prevention, ultimately safeguarding passion fruit production on a global scale.

Achnowledgements

A post-doctoral award at Makerere University and the National Crops Resources Research Instiute (NaCRRI) of Uganda by the Carnegie Corporation of New York (CCYN) through the Regional Universities Forum for Capacity Building in Agriculture (RUFORUM), Grant NumberRN/2020/RI/02.

References

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

Adams, M.J., Antoniw, J.F. and Beaudoin, F. 2005. Overview and analysis of the polyprotein cleavage sites in the family Potyviridae, *Molecular Plant Pathology*, 6(4):471-487. Available at: https://doi.org/10.1111/j.1364-3703. 2005.00296.x.

Alachiotis, N., Stamatakis, A. and Pavlidis, P. 2012. OmegaPlus: A scalable tool for rapid detection of selective sweeps in whole-genome datasets, *Bioinformatics*, 28(17):2274–2275. Available at: https://doi.org/10.1093/bioinformatics/bts419.

Ali, A., Li, H., Schneider, W. L., Sherman, D. J., Gray, S., Smith, D. and Roossinck, M. J. 2006. Analysis of genetic bottlenecks during horizontal transmission of Cucumber mosaic virus. *Journal of virology* 80(17):8345-8350.

Anisimova, M. 2019. Evolutionary Genomics Statistical and Computational Methods Second Edition Methods. Available at: http://www.springer.com/series/7651.

Atukunda, R., Sseruwagi, P., Karungi, J., Kyamanywa, S., Erbaugh, M. and Ochwo-Ssemakula, M. 2018. Farmers' knowledge of passion fruit virus diseases and their management in central Uganda. *African Journal of Horticultural Sciences* 13:53-64

Aziz, M.F. and Caetano-Anollés, G. 2021. Evolution of networks of protein domain organization, *Scientific reports*, 11(1), p. 12075. Available at: https://doi.org/10.1038/s41598-021-90498-8.

Bancy, W. W., Dora, C. K., Martina, K. and Mutuku, J. M. 2018. Molecular detection of Ugandan passiflora virus infecting passionfruit (*Passiflora edulis sims*) in Rwanda. *Annual Research & Review in Biology*. pp. 1-10.

Basu, M.K., Poliakov, E. and Rogozin, I.B. 2009. Domain mobility in proteins: Functional and evolutionary implications, *Briefings in Bioinformatics* 10(3): pp. 205–216. Available at: https://doi.org/10.1093/bib/bbn057.

Bordin, N., Sillitoe, I., Lees, J. G. and Orengo, C. 2021. Tracing evolution through protein structures: Nature captured in a few thousand folds. *Frontiers in Molecular Biosciences* 8:668184.

Brito, A.F. and Pinney, J.W. 2020. The evolution of protein domain repertoires: Shedding light on the origins of the Herpesviridae family, *Virus Evolution* 6(1): 1-12. Available at: https://doi.org/10.1093/ve/veaa001.

Bromberg, Y. and Rost, B. 2007. SNAP: Predict effect of non-synonymous polymorphisms on function, *Nucleic Acids Research* 35(11):3823=3835. Available at: https://doi.org/10.1093/nar/gkm238.

Bujarski, J.J. and Kaesberg, P. 1986. Genetic recombination between RNA components of a multipartite plant virus, *Nature* 321(6069). Available at: https://doi.org/10.1038/321528a0.

Carpenter, C. D., Oh, J. W., Zhang, C. and Simon, A. E. 1995. Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. *Journal of molecular biology* 245(5):608-622.

Carvalho, J. A. D., Jesus, J. G. D., Araujo, K. L., Serafim, M. E., Gilio, T. A. S. and Neves, L. G. 2021. Passion fruit (*Passiflora* spp.) species as sources of resistance to soil phytopathogens Fusarium solani and Fusarium oxysporum f. sp. passiflorae complex. *Revista Brasileira de Fruticultura* 43.

Cerqueira-Silva, C. B. M., Conceição, L. D. H. C. S., Souza, A. P. and Corrêa, R. X. 2014. A history of passion fruit woodiness disease with emphasis on the current situation in Brazil and prospects for Brazilian passion fruit cultivation. *European Journal of Plant Pathology* 139:261-270.

Chiaki, Y., Fukumoto, T., Nakamura, M. and Iwai, H. 2016. Population genetics analysis of East Asian Passiflora virus on Amami Oshima Island. *European Journal of Plant Pathology* 144:109-120.

Coutinho, T.A. and Wingfield, M.J. 2022. History of plant pathology in South Africa. BRIZA PUBLICATIONS. Edited by C. TERESA A. Pretoria: BRIZA PUBLICATIONS.

Cui, H. and Wang, A. 2016. Plum pox virus 6K1 protein is required for viral replication and targets the viral replication complex at the early stage of infection. *Journal of Virology* 90(10):5119–5131. Available at: https://doi.org/10.1128/jvi.00024-16.

Dolan, P.T., Whitfield, Z.J. and Andino, R. 2018. Mechanisms and concepts in RNA virus population dynamics and evolution, *Annual Review of Virology* 5:69–92. Available at: https://doi.org/10.1146/annurev-virology-101416-041718.

Ebert, D. and Fields, P.D. 2020. Host–parasite co-evolution and its genomic signature, *Nature Reviews Genetics* 21(12):754–768. Available at: https://doi.org/10.1038/s41576-020-0269-1.

Forni, D., Cagliani, R. and Sironi, M. 2020. Recombination and positive selection differentially shaped the diversity of betacoronavirus subgenera. *Viruses* 12(11). Available at: https://doi.org/10.3390/v12111313.

Gell, G., Sebestyén, E. and Balázs, E. 2015. Recombination analysis of maize dwarf mosaic virus (MDMV) in the Sugarcane mosaic virus (SCMV) subgroup of

potyviruses, *Virus Genes* 50(1). Available at: https://doi.org/10.1007/s11262-014-1142-0.

Gibbs, A. J., Hajizadeh, M., Ohshima, K. and Jones, R. A. 2020. The potyviruses: an evolutionary synthesis is emerging. *Viruses* 12(2):132.

Gibbs, M. J. and Weiller, G. F. 1999. Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. *Proceedings of the National Academy of Sciences* 96(14):8022-8027.

Gioppato, H. A., da Silva, M. B., Carrara, S., Palermo, B. R. Z., de Souza Moraes, T. and Dornelas, M. C. 2019. Genomic and transcriptomic approaches to understand Passiflora physiology and to contribute to passionfruit breeding. *Theoretical and Experimental Plant Physiology* 31:173-181.

Green, K. J., Brown, C. J., Gray, S. M. and Karasev, A. V. 2017. Phylogenetic study of recombinant strains of Potato virus Y. *Virology*, *507*, 40-52. Available at: https://doi.org/10.1016/j.virol.2017.03.018.

Holmes, E. C. 2009. The evolution and emergence of RNA viruses. Oxford University Press.

Iwai, H., Yamashita, Y., Nishi, N. and Nakamura, M. 2006. The potyvirus associated with the dappled fruit of *Passiflora edulis* in Kagoshima prefecture, Japan is the third strain of the proposed new species East Asian Passiflora virus (EAPV) phylogenetically distinguished from strains of Passion fruit woodiness virus. *Archives of Virology* 151:811-818. Available at: https://doi.org/10.1007/s00705-005-0659-x.

Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H. and Frost, S. D. 2006. GARD: A genetic algorithm for recombination detection. *Bioinformatics* 22(24):3096-3098. Available at: https://doi.org/10.1093/bioinformatics/btl474.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution 35*(6):1547. Available at: https://doi.org/10.1093/molbev/msy096.

Li, Y., Zhang, Y., Li, X., Yi, S. and Xu, J. 2019. Gain-of-function mutations: an emerging advantage for cancer biology. *Trends in Biochemical Sciences* 44(8): 659-674. Available at: https://doi.org/10.1016/j.tibs.2019.03.009.

Lopez-Moya, J.J. and Garcia, J.A. 2008. Potyviruses. pp. 313–322.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R. and Marchler-Bauer, A. 2020. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research* 48(D1):D265-D268. Available at: https://doi.org/10.1093/nar/gkz991.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S. and Bryant, S. H. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research* 45(D1):D200-D203. Available at: https://doi.org/10.1093/nar/gkw1129.

Marsh, J.A. and Teichmann, S.A. 2010. How do proteins gain new domains?, *Genome Biology* 11(7). Available at: https://doi.org/10.1186/gb-2010-11-7-126.

Martin, D. P., Varsani, A., Roumagnac, P., Botha, G., Maslamoney, S., Schwab, T., and Murrell, B. 2021. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evolution*, *7*(1), veaa087.. Available at: https://doi.org/10.1093/ve/veaa087.

Mbeyagala, E. K., Maina, S., Macharia, M. W., Mukasa, S. B. and Holton, T. 2019. Illumina sequencing reveals the first near-complete genome sequence of Ugandan Passiflora virus. *Microbiology Resource Announcements* 8(20): e00358-19. Available at: https://doi.org/10.1128/mra.00358-19.

Mendes, D. S., Viana, A. P., Santos, E. A., Cavalcante, N. R., Rodrigues, C. A., Lima, J. A. and Santos, H. A. 2022. Genetic gains in Passiflora for resistance to cowpea aphid-borne mosaic virus using recurrent selection. *Euphytica* 218(9): 132. Available at: https://doi.org/10.1007/s10681-022-03084-y.

Messer, P.W. and Petrov, D.A. 2013. Population genomics of rapid adaptation by soft selective sweeps, *Trends in Ecology and Evolution* 28(11):659–669. Available at: https://doi.org/10.1016/j.tree.2013.08.003.

Miyashita, S. 2018) Studies on replication and evolution mechanisms of plant RNA viruses, *Journal of General Plant Pathology*. Available at: https://doi.org/10.1007/s10327-018-0806-3.

Moradi, Z. and Mehrvar, M. 2021. Molecular characterization of two highly divergent Iranian johnsongrass mosaic virus isolates from Zea mays.*VirusDisease* 32(1): pp. 155–160. Available at: https://doi.org/10.1007/s13337-021-00671-w.

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(9). Available at: https://doi.org/10.1371/journal.pone.0108277.

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K. and Kosakovsky Pond, S. L. 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS genetics* 8(7):e1002764. Available at: https://doi.org/10.1371/journal.pgen.1002764.

Nagy, P.D. 2008. Recombination in plant RNA viruses. In: M.J. Roossinck (ed.) *Plant Virus Evolution*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 133–156. Available at: https://doi.org/10.1007/978-3-540-75763-4_8.

Nigam, D., LaTourrette, K., Souza, P. F. and Garcia-Ruiz, H. 2019. Genome-wide variation in potyviruses. *Frontiers in Plant Science* 10: 1439. Available at: https://doi.org/10.3389/fpls.2019.01439.

Nuss, D.L. 2011. Mycoviruses, RNA silencing, and viral RNA recombination. *Advances in Virus Research*. Available at: https://doi.org/10.1016/B978-0-12-385987-7.00002-6.

Ochwo-Ssemakula, M., Sengooba, T., Hakiza, J. J., Adipala, E., Edema, R., Redinbaugh, M. G. and Winter, S. 2012. Characterization and distribution of a

Potyvirus associated with passion fruit woodiness disease in Uganda. *Plant Disease* 96(5):659-665. Available at: https://doi.org/10.1094/PDIS-03-11-0263.

Padhi, A. and Ramu, K. 2011. Genomic evidence of intraspecific recombination in sugarcane mosaic virus, *Virus Genes* 42(2). Available at: https://doi.org/10.1007/s11262-010-0564-6.

Parizad, S., Dizadji, A., Koohi Habibi, M., Winter, S., Kalantari, S., Movi, S. and Ayllón, M. A. 2018. Description and genetic variation of a distinct species of Potyvirus infecting saffron (*Crocus sativus* L.) plants in major production regions in Iran. *Annals of Applied Biology* 173(3):233-242. Available at: https://doi.org/10.1111/aab.12456.

Pavlidis, P., Zivkoviæ, D., Stamatakis, A. and Alachiotis, N. 2013. SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. *Molecular Biology and Evolution* 30(9): 2224-2234. Available at: https://doi.org/10.1093/molbev/mst112.

Peng, B., Liu, L., Wu, H., Kang, B., Fei, Z. and Gu, Q. 2021. Interspecific Recombination Between Zucchini Tigre Mosaic Virus and Papaya Ringspot Virus Infecting Cucurbits in China. *Frontiers in Microbiology* 3387. Available at: https://doi.org/10.3389/fmicb.2021.773992.

Petrzik, K. 2019. Evolutionary forces at work in partitiviruses. *Virus Genes* 55(5): 563–573. Available at: https://doi.org/10.1007/s11262-019-01680-0.

Rao, S., Chen, X., Qiu, S., Peng, J., Zheng, H., Lu, Y. and Yan, F. 2020. Identification of two new isolates of chilli veinal mottle virus from different regions in China: Molecular diversity, phylogenetic and recombination analysis. *Frontiers in Microbiology* 11:616-171. Available at: https://doi.org/10.3389/fmicb.2020.616171.

Rodamilans, B., Shan, H., Pasin, F., and García, J. A. 2018. Plant viral proteases: Beyond the role of peptide cutters. *Frontiers in Plant Science* 9:666. Available at: https://doi.org/10.3389/fpls.2018.00666.

Sengooba, T. N., Spence, N. J., Walkey, D. G. A., Allen, D. J. and Femi Lana, A. 1997. The occurrence of bean common mosaic necrosis virus in wild and forage legumes in Uganda. *Plant Pathology* 46(1):95-103. Available at: https://doi.org/10.1046/j.1365-3059.1997.d01-12.x.

Simon-Loriere, E. and Holmes, E.C. 2011. Why do RNA viruses recombine? *Nature Reviews Microbiology* 9(8):617–626. Available at: https://doi.org/10.1038/nrmicro2614.

Smith, M. D., Wertheim, J. O., Weaver, S., Murrell, B., Scheffler, K. and Kosakovsky Pond, S. L. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular Biology and Evolution* 32(5):1342-1353. Available at: https://doi.org/10.1093/molbev/msv022.

Smith, J. M. and Haigh, J. 1974. The hitch-hiking effect of a favourable gene. *Genetics Research* 23(1):23-35.

Spielman, S. J., Weaver, S., Shank, S. D., Magalis, B. R., Li, M. and Kosakovsky Pond, S. L. 2019. Evolution of viral genomes: Interplay between selection, recombination, and other forces. *Evolutionary Genomics: Statistical and Computational Methods* 427-468. Available at: https://doi.org/10.1007/978-1-4939-9074-0_14.

Steinway, S. N., Dannenfelser, R., Laucius, C. D., Hayes, J. E. and Nayak, S. 2010. JCoDA: A tool for detecting evolutionary selection. *BMC Bioinformatics* 11:1-9.

Stephan, W. 2019. Selective sweeps. *Genetics* 211(1):5–13. Available at: https://doi.org/10.1534/genetics.118.301319.

Sztuba-Soliñska, J., Urbanowicz, A., Figlerowicz, M. and Bujarski, J. J. 2011. RNA-RNA recombination in plant virus replication and evolution. *Annual Review of Phytopathology* 49:415-443. Available at: https://doi.org/10.1146/annurev-phyto-072910-095351.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25(24):4876-4882. Available at: https://doi.org/10.1093/nar/25.24.4876.

Valli, A., López-Moya, J.J. and García, J.A. 2007. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. *Journal of General Virology* 88(3):1016–1028. Available at: https://doi.org/10.1099/vir.0.82402-0.

Vogel, C., Teichmann, S.A. and Pereira-Leal, J. 2005. The relationship between domain duplication and recombination. *Journal of Molecular Biology* 346(1): 355–365. Available at: https://doi.org/10.1016/j.jmb.2004.11.050.

Wainaina, J. M., Kubatko, L., Harvey, J., Ateka, E., Makori, T., Karanja, D. and Kehoe, M. A. 2019. Evolutionary insights of bean common mosaic necrosis virus and cowpea aphid-borne mosaic virus. *PeerJ 7*: e6297. Available at: https://doi.org/10.7717/peerj.6297.

Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V. and Kosakovsky Pond, S. L. 2018. Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution* 35(3):773-777. Available at: https://doi.org/10.1093/molbev/msx335.

Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. and Scheffler, K. 2015. RELAX: Detecting relaxed selection in a phylogenetic framework. *Molecular Biology and Evolution* 32(3):820-832. Available at: https://doi.org/10.1093/molbev/msu400.

White, K.A., Enjuanes, L. and Berkhout, B. 2011. RNA virus replication, transcription and recombination. *RNA Biology*. Available at: https://doi.org/10.4161/rna.8.2.15663.

Willemsen, A., Carrasco, J.L., Santiago, F. E. and Zwart, M.P. 2018. Going, going, gone: Predicting the fate of genomic insertions in plant RNA viruses. *Heredity* 121(5). Available at: https://doi.org/10.1038/s41437-018-0086-x.

Worrall, E. A., Hayward, A. C., Fletcher, S. J. and Mitter, N. 2019. Molecular characterization and analysis of conserved potyviral motifs in bean common mosaic virus (BCMV) for RNAi-mediated protection. *Archives of Virology* 164:181-194. Available at: https://doi.org/10.1007/s00705-018-4065-6.