Journal of Applied Horticulture, 28(1): 257-266, 2026
https://doi.org/10.37855/jah.2026.v28i01.41
Journal of Applied Horticulture
ISSN: 0972-1045

Phylogenetic and population analysis of *African Cassava mosaic* virus (ACMV) isolates from Cross River State based on the coat protein gene

E.E. Ekpiken[1]*, G.M. Ubi[2] and O.I. Eyong[1]

*[1]Department of Plant Science and Biotechnology, University of Cross River State, Nigeria. [2]Department of Genetics and Biotechnology, University of Calabar, Nigeria.*E-mail: etimekpiken@yahoo.com*

Abstract

African cassava mosaic virus (ACMV) is among the most economically significant begomoviruses affecting cassava (*Manihot esculenta* Crantz) production in sub-Saharan Africa, however its molecular population dynamics in Cross River State (CRS), Nigeria, have remained poorly characterized. This study presents the phylogenetic and population genetic analysis of cassava begomoviruses ACMV DNA-A isolates from CRS, based on coat protein (CP) gene sequences of 44 begomovirus isolates drawn from five geographic populations namely East/South Africa, Nigeria, Asia, West Africa, and CRS. Sequence alignment was performed using the MUSCLE algorithm in Bioedit, phylogenetic reconstruction was conducted in MEGA 12 under the General Time Reversible model with 1000 bootstrap replicates, and population genetic parameters were estimated using DnaSP v6.01. Phylogenetic analysis placed all three CRS isolates (CRS-P1, CRS-P2, CRS-P3) within the West African clade, with CRS-P1 and CRS-P3 clustering within a Nigerian–Ghanaian subgroup and CRS-P2 forming a distinct Nigerian–Togolese lineage, ruling out Asian or southern/central African origins. The CRS population exhibited the lowest nucleotide diversity ($\pi = 0.010$) and the strongest purifying selection (dN/dS = 0.03) across all populations, with only 9 segregating sites among three unique haplotypes. Due to a limited number of segregating sites, neutrality tests could not be performed for the CRS population. In contrast, the remaining populations exhibited significantly negative Tajima's D statistics, a pattern indicative of either a historical population expansion or the effects of purifying selection. Population differentiation analysis showed that the CRS sequences were significantly differentiated from Asian (Snn = 0.9000), East/South African (Snn = 0.7728), and Nigerian (Snn = 0.4792) populations, but not from West African populations. This suggests possible shared ancestry or ongoing regional gene flow between the CRS and West African populations.These findings indicate the CRS sequences harbors a genetically uniform and highly conserved ACMV lineage of West African origin, with important implications for disease surveillance and resistance breeding in Nigeria.

Key words: African cassava mosaic virus (ACMV), coat protein gene, phylogenetic analysis, population genetics, purifying selection, Cross River State, begomovirus, cassava mosaic disease, DNA-A, DnaSP

Introduction

Cassava (*Manihot esculenta* Crantz) is a vital food security crop for over 800 million people globally, particularly in sub-Saharan Africa where it serves as a primary staple food and source of income (Legg and Fauquet, 2004). Nigeria is the world's largest producer of cassava, contributing approximately 60 million tonnes annually, yet productivity remains constrained by various biotic stresses, most notably cassava mosaic disease (CMD) (FAO, 2018). CMD represents one of the most economically devastating plant diseases in Africa, caused by a complex of whitefly-transmitted begomoviruses in the family Geminiviridae that can result in yield losses of up to 80% in susceptible varieties (Patil and Fauquet, 2009).

African cassava mosaic virus (ACMV) is among the most widespread and economically important species within the cassava mosaic begomovirus (CMB) complex, alongside East African cassava mosaic virus (EACMV), East African cassava mosaic Cameroon virus (EACMCV), and several other related species (Legg *et al*., 2011). These viruses are characterized by circular, single-stranded DNA genomes organized into either monopartite or bipartite configurations, with DNA-A components encoding essential proteins including the coat protein (CP) gene that serves crucial functions in viral encapsidation, transmission, and host specificity (Fondong, 2013; Hanley-Bowdoin *et al*., 2013).

The coat protein gene has emerged as a critical molecular marker for phylogenetic and population genetic analyses of begomoviruses due to its relatively conserved nature and functional importance (Bull *et al*., 2006; Duffy and Holmes, 2009). Phylogenetic studies based on CP gene sequences have revealed significant genetic diversity within ACMV populations across different geographical regions, providing insights into viral evolution, migration patterns, and the emergence of new strains (Were *et al*., 2004). Moreover, population genetic analyses of the CP gene have demonstrated high rates of nucleotide substitution in cassava mosaic viruses, indicating rapid evolutionary dynamics that contribute to the adaptability and persistence of these pathogens (Duffy and Holmes, 2009). Recent studies of related begomoviruses, such as Tomato yellow leaf curl China virus, highlight that DNA viruses can display quasispecies-like population structures, with heterogeneous swarms of closely related variants driving rapid adaptation comparable to RNA

viruses (Ge *et al*., 2007). Such dynamics enhance the ability of begomoviruses to overcome resistance and persist under changing agro-ecological conditions.

Cross River State, located in the humid forest zone of south-southern Nigeria, represents a significant cassava production area with unique agro-ecological conditions that may influence viral population structure and diversity. The state borders Cameroon and lies within a region where multiple cassava mosaic virus species have been documented, creating opportunities for genetic recombination and the emergence of novel viral variants (Ekpiken *et al*., 2022). Comparative work on other plant viruses, such as Bean common mosaic virus (BCMV), demonstrates that recombination and purifying selection act as major forces shaping viral diversity, with phylogenetic structuring often linked to host range and geography (Moradi and Mehrvar 2019). These findings underscore the likelihood that similar evolutionary processes may drive cassava mosaic begomoviruses diversification in regions like Cross River State.

Understanding the genetic diversity and phylogenetic relationships of the virus isolates from Cross River State is essential for several reasons. First, it provides insights into the evolutionary history and dispersal patterns of the virus within West Africa, contributing to our broader understanding of begomovirus biogeography (Crespo-Bellido *et al*., 2024). Second, population genetic analyses can reveal patterns of gene flow, recombination, quasispecies dynamics, and selection pressures that shape viral evolution and adaptation to local conditions (Ge *et al*., 2007; Moradi and Mehrvar 2019). Third, such studies inform the development of effective disease management strategies, including the deployment of resistant cultivars and the prediction of viral epidemiological trends (Bull *et al*., 2006).

The objectives of this study were to conduct phylogenetic analysis of the isolates from Cross River State based on coat protein gene sequences to determine their evolutionary relationships with isolates from other regions, assess the population genetic structure and diversity of ACMV isolates within the state and identify potential recombination events and selection pressures acting on the coat protein gene.. This research contributes to the growing body of knowledge on cassava mosaic virus molecular epidemiology in West Africa and provides baseline data for future disease monitoring and management efforts in Nigeria.

Materials and methods

Sequence data collection and population assignment: This study analyzed 44 ACMV DNA-A cassava mosaic begomovirus sequences obtained from the National Centre for Biotechnology Information GenBank originating from diverse geographic regions with a view to investigate population genetic architecture and evolutionary dynamics (Table 1). Sequences were systematically stratified into five distinct geographic populations based on sampling locations and associated metadata: East/South Africa (n=24, 54.54%), Nigeria (n=5, 11.36%), Asia (n=7, 15.9%), West Africa (n=5, 11.36%), and Cross River State, Nigeria (n=3, 6.81%) from which three sequences in earlier studies were isolated (Ekpiken *et al*., 2022). This geographic distribution ensured adequate representation across major biogeographic regions while maintaining statistically robust sample sizes for most populations. All sequences underwent rigorous quality assessment protocols including verification of sequence authenticity, systematic removal of contaminating sequences, and confirmation of appropriate taxonomic assignment to ensure data integrity and analytical reliability. Due to the different length of the retrieved sequences, the aligned sequences were trimmed to 636 bp to remove the terminal gaps.

Table 1. List of the origins, hosts, and accession numbers of cassava mosaic begomovirus isolates/strains

Accession number	Location	Isolate/Clone Information	Host
MH251343	Nigeria	NI-94	*M. esculenta*
FM877473	Burkina Faso	Burkina Faso-Kamboinse-2008	*M. esculenta*
MZ507578	Nigeria	CRS-P1	*M. esculenta*
MZ507579	Nigeria	CRS-P2	*M. esculenta*
MG250147	Ghana	ACMV_GH:FM10A:13	*M. esculenta*
MG250085	Ghana	ACMV_GH:AK10A:13	*M. esculenta*
HG530111	Kenya	KE:mtw:CMD-MI79:12 (Clone MI79)	*M. esculenta*
MZ507580	Nigeria	CRS-P3	*M. esculenta*
EU155147	Togo	EAC05-50S	*M. esculenta*
AF259894	Cote d'Ivoire		*M. esculenta*
MG250100	Ghana	ACMV_GH:KW6A:13	*M. esculenta*
LC658964	Burkina Faso	ACMV-[BF:Nou:BFW670-15B:17]	*M. esculenta*
LC658957	Burkina Faso	ACMV-[BF:Sag:BFW213-3B:16]	*M. esculenta*
HE979766	Uganda	ACMV-[UG:Nam:CMD-MI31:12] (Clone MI31)	*M. esculenta*
HE979761	Uganda	ACMV-[UG:Nam:CMD-MI24:12] (Clone MI24)	*M. esculenta*
KR476371	Benin	BN:Con1017:14	*M. esculenta*
MG250126	Ghana	ACMV_GH:BW3A:13	*M. esculenta*
KM023678	Kenya	047_Kathiani (Strain)	*M. esculenta*
MK896240	Zambia	ZARI 1	*M. esculenta*
MK896247	Zambia	ZARI 9	*M. esculenta*
MT599665	Gabon	ODJ71	*M. esculenta*
KT869117	Zambia	ZM-LSK147	*M. esculenta*
KM023686	Kenya	33_Msambweni (Strain)	*M. esculenta*
AF011785		South African strain	*M. esculenta*
KM023674	Kenya	035_Siaya (Strain)	*M. esculenta*
MT861190	Ebonyi State, NG	AfikpoN1N11	*M. esculenta*
MT861210	Ebonyi State, NG	Obuba1Izzi1	*M. esculenta*
EU367500	Nigeria		*Soybean*
MN164632	Angola	954	*M. esculenta*
JQ804866	CAR	CF100	*M. esculenta*
FN435289	DR Congo	ACMV-Ybi8	*M. esculenta*
FN435259	DR Congo	ACMV-Ybi5	*M. esculenta*
GU580906	Angola	AOA	*M. esculenta*
MH251339	Nigeria	ON-16	*M. esculenta*
DQ490078	India	Nammakal	*M. esculenta*
KJ887643	Madagascar	MG:MG38A5:06	*M. esculenta*
PP346876	Vietnam	VN-HoaBinh-18	*M. esculenta*
AY211463	Yaoundé, CM	CM/YA	*M. esculenta*
AY211460	Cameroon	CM/AK2	*M. esculenta*
MN577576	Cambodia	Ome2	*M. esculenta*
KU308385	India	TVM4	*M. esculenta*
EU113300	India	Lucknow	*J. curcas*
AJ579307	India	Adivaram (Ker 20)	*M. esculenta*
AY738105	India	Thrissur	*M. esculenta*
MH251343	Nigeria	NI-94	*M. esculenta*
FM877473	Burkina Faso	Burkina Faso-Kamboinse-2008	*M. esculenta*

Central African Republic: CAR. Democratic Republic of the Congo: DR Congo.

Ebonyi, Nigeria = Ebonyi State, NG. Cameroon: Yaoundé = Yaoundé, CM

Sequence alignment and preprocessing: Multiple sequence alignment was performed using the MUSCLE algorithm implemented through the Bio.Align module of Biopython version 1.81 (Cock *et al*., 2009). To ensure alignment robustness, the sequences were evaluated using ClustalW algorithm in the Bioedit software version 7 (Hall 2004), with results systematically compared for consistency validation. Due to the different length of the sequences, the aligned sequences were trimmed to 636 bp to remove the terminal gaps on both ends.

Phylogenetic reconstruction and bootstrap validation: MEGA 12 software was used for inferring phylogenetic relationships. The best-fit nucleotide substitution model was selected using the maximum likelihood method and the tree reconstructed under the selected model. The specific nucleotide substitution model used in the analysis was the General Time Reversible (GTR) models.

Statistical support was assessed through bootstrap analysis using 1000 replicates for both methods. Bootstrap values provided confidence measures for individual nodes, identifying well-supported versus poorly-supported relationships. Trees were systematically compared with external phylogenetic data when available for independent validation of population genetic inferences.

Population genetic parameter estimation: Standard population genetic parameters were calculated for each geographic population using DnaSP version 6.01. The analysis encompassed fundamental diversity measures including the number of haplotypes (H), representing unique sequences within each population, and haplotype diversity (Hd), calculated as the probability that two randomly selected haplotypes differ according to Nei (1987). Additional parameters included the number of polymorphic sites (S), representing positions with nucleotide variation, and total mutations (η) inferred from phylogenetic relationships.

Genetic diversity was further quantified through calculation of the average number of nucleotide differences (k) between sequences within populations and nucleotide diversity (π), representing average nucleotide differences per site following Nei (1987). To assess potential selection pressures, synonymous sites (SS) and non-synonymous sites (NS) were identified through comprehensive codon-based analysis, enabling calculation of the dN/dS ratio representing the ratio of non-synonymous to synonymous substitution rates. This parameter suite provided detailed characterization of genetic diversity patterns and evolutionary processes operating within each population.

Neutrality testing: Assessment of neutral evolution employed three classical neutrality tests, each providing complementary insights into potential departures from neutral expectations. Tajima's D statistic (Tajima, 1989) was calculated using the standard formula $D = (\pi - \theta W) / \sqrt{Var}(\pi - \theta W)$, where π represents nucleotide diversity, θW is Watterson's theta estimator, and Var ($\pi - \theta W$) represents the variance of their difference. Statistical significance was rigorously evaluated using standard normal distribution with critical values established at $\alpha = 0.05$, following established protocols in population genetics literature.

Fu and Li's neutrality tests provided additional perspectives on potential selection through analysis of mutation frequency spectrum. The D* statistic compared singleton mutations to total mutations, while the F* statistic incorporated both singleton mutations and total mutations in different mathematical combinations, as described by Fu and Li (1993). Both statistics were appropriately standardized and assessed for significance using established critical values corresponding to respective sample sizes. This multi-test approach provided robust assessment of neutral evolution assumptions while accounting for different aspects of mutation frequency distribution.

Population differentiation analysis: Population genetic differentiation was assessed using a comprehensive suite of statistical measures extending beyond traditional FST analysis. Wright's FST (Wright, 1951) was calculated for all pairwise population comparisons using FST = (HT - HS) / HT, where HT represents total gene diversity and HS represents average gene diversity within populations.

Hudson's advanced differentiation statistics (Hudson *et al*., 1992) implemented genealogically-informed measures of population structure. The KS* and KST* statistics were calculated using formulations KS* = 1 - (KS / KT) and KST* = (KT - KS) / KT, where KS represents average differences within populations and KT represents average differences between populations. These measures account for genealogical relationships and provide enhanced sensitivity for detecting population differentiation.

The Z* statistic measured shared polymorphism between population pairs using Z* = (ZA + ZB) / (ZA + ZB + ZAB), where ZA and ZB represent differences within populations A and B, and ZAB represents differences between populations. The nearest neighbor statistic (Snn) was computed following Hudson (2000) using Snn = (1/n) Σ si, providing direct assessment of population clustering based on sequence similarity relationships.

Statistical validation: Statistical significance was assessed using appropriate methods tailored to each analysis type. Neutrality tests were evaluated against theoretical distributions or published critical values. Population differentiation statistics employed permutation tests and established significance thresholds. Phylogenetic support was quantified through 1000-replicate bootstrap analysis providing robust confidence measures.

Results and discussions

Phylogeny: The phylogenetic tree (Fig. 1) reveals that all three CRS sequences (CRS-P1, CRS-P2, and CRS-P3) belong to the same broad West African genetic lineage, though they occupy slightly different positions within the tree, suggesting some degree of intra-population genetic diversity while maintaining a clear regional identity.

CRS-P1 (MZ507578) and CRS-P3 (MZ507580) cluster together within the same sub-clade in the upper portion of the tree, grouping closely with sequences originating from Ghana (MG250147, MG250085), Benin (KR476371), Burkina Faso (LC658957), and Nigeria (MH251343). This grouping is supported by a bootstrap value of 61, indicating moderate phylogenetic confidence in their shared ancestry. The relatively short branch lengths associated with CRS-P1 and CRS-P3 further suggest that these two sequences are not markedly divergent from the reference strains within this cluster, pointing to a close genetic relationship with circulating West African strains.

CRS-P2 (MZ507579), while still firmly embedded within the West African lineage, occupies a slightly deeper and distinct position in the tree, clustering alongside sequences from Ghana (MG250100) and Burkina Faso (LC658964, MG250126). This placement indicates that although CRS-P2 shares a common regional ancestry with CRS-P1 and CRS-P3, it represents a somewhat divergent variant within the same broader genetic group, possibly reflecting localized evolutionary differentiation or a distinct transmission history.

Importantly, none of the three CRS sequences cluster with isolates from Asia, Southern Africa, or Central Africa, effectively ruling out those regions as sources of introduction. Their consistent placement within the West African clade strongly supports a local or regionally circulating origin, consistent with the epidemiological context of Nigeria and its neighboring countries. Taken together, the phylogenetic evidence suggests that CRS-P1, CRS-P2, and CRS-P3 are genetically related but not identical strains, all traceable to a common West African genetic background, with CRS-P2 exhibiting slightly greater divergence relative to CRS-P1 and CRS-P3.

The pairwise identity matrix (Fig. 2) reveals that CRS-P1, CRS-P2, and CRS-P3 share the highest nucleotide identity with each other and with sequences from Ghana, Burkina Faso, Benin, Ivory Coast, Togo, and Nigeria, confirming their close genetic relationship with West African strains. In contrast, all three CRS sequences show markedly lower identity when compared to sequences from India, Sri Lanka, Vietnam, Angola, Madagascar, and DRC, indicating considerable genetic divergence from these geographically distant lineages. This pattern is entirely consistent with the phylogenetic tree, reinforcing the conclusion that CRS-P1, CRS-P2, and CRS-P3 are of West African genetic origin, with no close molecular affinity to Asian or south/central African isolates.

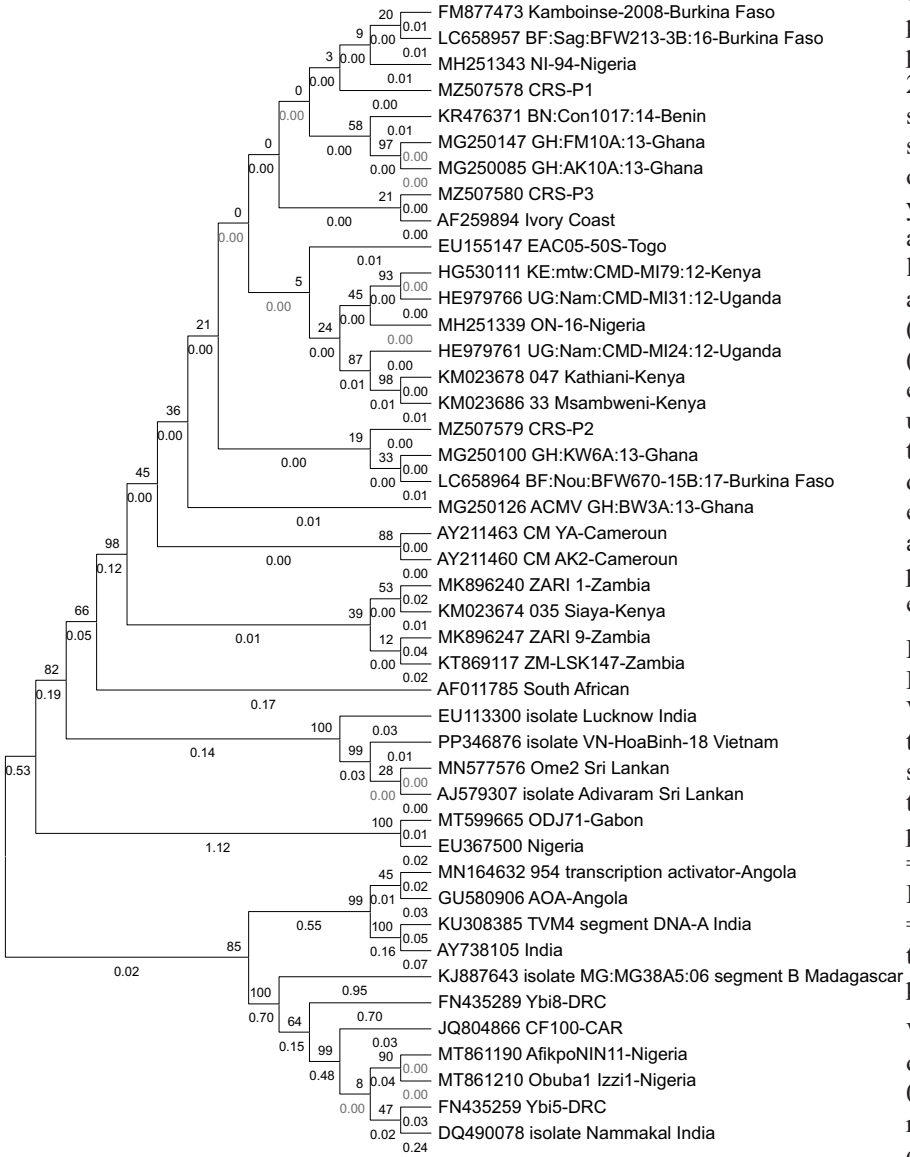

Fig. 1. Phylogenetic tree of ACMV coat protein gene isolates constructed using the Maximum Likelihood approach with 1000 bootstrap replicates. The tree shows the clustering of Nigerian isolates in relation to reference sequences from West, East, and Southern Africa, as well as Asia. Isolates CRS-P1 (MZ507578) and CRS-P3 (MZ507580) clustered within a Nigerian–Ghanaian subgroup, whereas CRS-P2 (MZ507579) grouped separately with a Togolese isolate, forming a Nigerian–Togolese subgroup. Bootstrap values >50% are shown at the nodes. The scale bar represents the number of nucleotide substitutions per site.

Population genetic diversity parameters: The genetic diversity analysis of ACMV sequences across five geographic populations revealed marked differences in evolutionary characteristics, with the Cross River State (CRS) population exhibiting a fundamentally distinct profile compared to all other regions (Table 2). The CRS population, represented by three sequences (MZ507578, MZ507579, MZ507580), showed minimal internal genetic variation, with only 9 segregating sites and 10 total mutations, yielding a mean pairwise difference (K) of 6.33 and a nucleotide diversity (π) of 0.010. This is the lowest nucleotide diversity recorded across all the analyzed populations. By comparison, Nigeria (π = 0.546), Asia (π = 0.453), East/South Africa (π = 0.359), and West Africa (π = 0.015) each exhibit substantially higher nucleotide diversity, underscoring the near-monomorphic character of the CRS sequences. It is suggested that the low diversity of the CRS population indicates it has experienced either severe genetic bottlenecking, a recent single-introduction event, or strong purifying selection that has systematically eliminated most variation (Table 2).

Haplotype diversity (Hd) was 1.000 for CRS, East/South Africa, and Asia, while Nigeria and West Africa registered Hd = 0.900. Although the Hd value for CRS appears high, all three sequences represent unique haplotypes despite their minimal genetic divergence. The mean pairwise differences were greatest for Asia (K = 188.57), followed by Nigeria (K = 164.20), East/South Africa (K = 47.04), West Africa (K = 9.00), and CRS (K = 6.33), further confirming the comparative genetic uniformity of the CRS population.

With respect to selection signatures, the dN/dS ratio for CRS was 0.03, derived from dN = 0.001 and dS = 0.048, it was the lowest ratio recorded across all populations and diagnostic of strong purifying (negative) selection where amino acid-altering substitutions are actively eliminated. This contrasts sharply with the ratios observed in Nigeria (dN/dS = 6.72), East/South Africa (dN/dS = 1.48), Asia (dN/dS = 4.85), and West Africa (dN/dS = 0.04). It is worthy of note that CRS population accumulated only

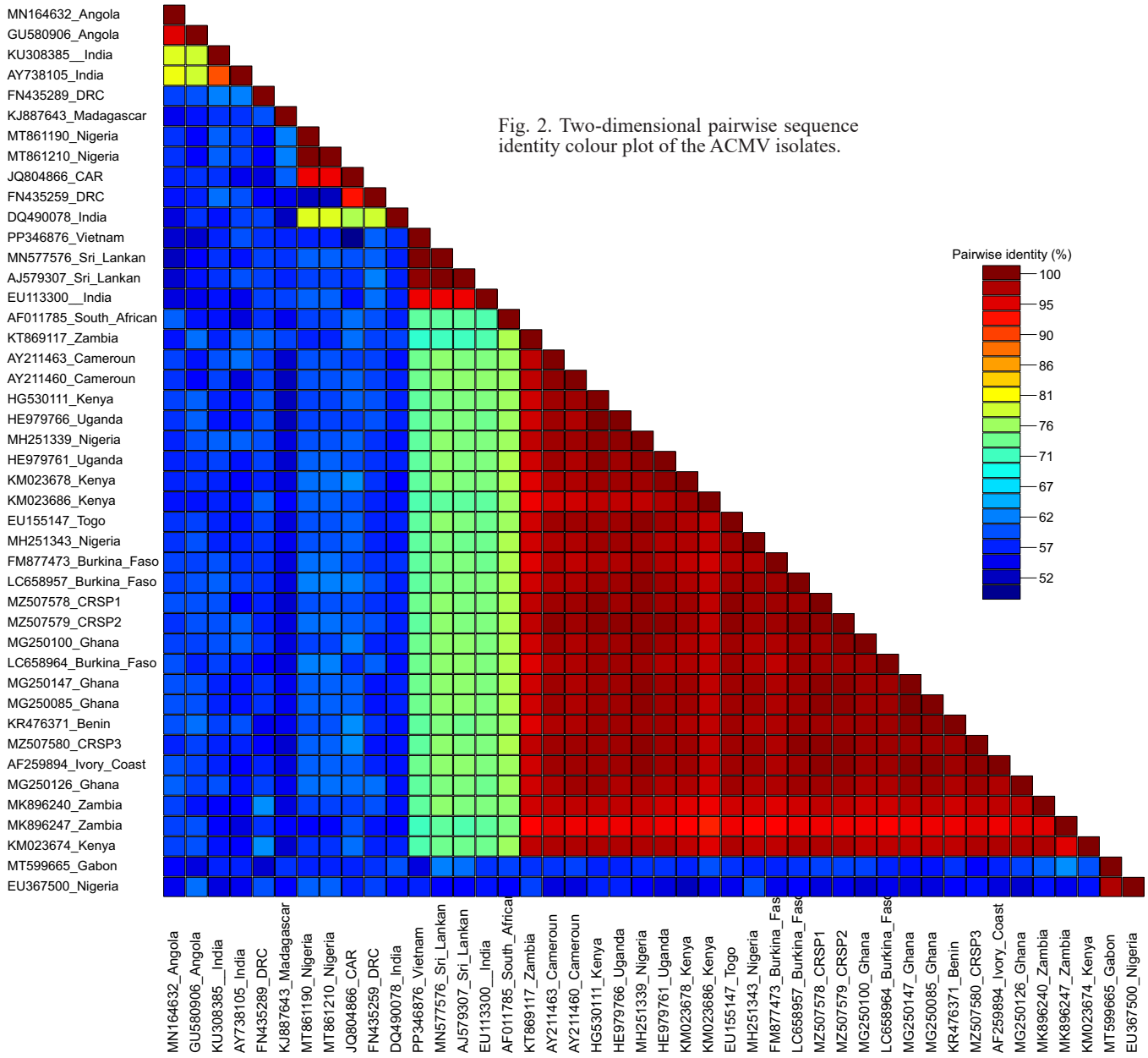

Fig. 2. Two-dimensional pairwise sequence identity colour plot of the ACMV isolates.

2 non-synonymous substitutions compared to 17 synonymous substitutions, consistent with the interpretation that synonymous changes don't alter the protein, so their accumulation relative to non-synonymous changes reflects selection preserving the protein sequence.

Neutrality tests and population demography: The neutrality test statistics for the CRS population could not be computed (Table 3). Tajima's D, Fu and Li's D*, and Fu and Li's F* all returned as not available (N/A), and Watterson's theta (θW) was also listed as N/A, reflecting limited polymorphic sites (S = 9, N = 3). This absence of computable diversity statistics among the populations examined strengthens the interpretation that CRS sequences are at or near fixation for a single genotype.

In contrast, all other populations yielded negative Tajima's D values, indicative of an excess of rare alleles typically associated with recent population expansion or purifying selection. East/South Africa showed the most significantly negative value (D = −2.9075, significant at α = 0.05), followed by Nigeria (D = −1.6598, not significant), Asia (D = −0.8878, not significant), and West Africa (D = −0.4620, not significant). Fu and Li's D* and F* statistics were largely non-significant across populations, with the exception of East/South Africa's Tajima's D reaching significance. Watterson's theta values were highest for Asia (θW = 222.04), followed by Nigeria (θW = 209.76), East/South Africa (θW = 167.10), and West Africa (θW = 9.60), again confirming the gradient of genetic diversity from the more diverse Asian and Nigerian populations toward the more constrained West African and CRS sequences (Table 3).

Population differentiation and gene flow patterns: Pairwise population differentiation statistics revealed complex patterns of genetic structure among ACMV populations (Table 4). The CRS versus Asia comparison showed FST = 0.0774, with highly significant KS* (0.4305), KST* (0.0377), and Snn (0.9000) values, indicating strong genetic differentiation. The high Snn value approaching 1 indicates that nearly every sequence's nearest genetic neighbor belongs to the same population, confirming clear population structure between CRS and Asian sequences. The CRS versus Nigeria comparison yielded FST = −0.0575, with significant KS* (0.4004), KST* (−0.0300), and Snn (0.4792) values, suggesting that despite originating from the same

geographic country, CRS sequences are genetically distinct from other Nigerian isolates, though less so than from Asian sequences.

The East/South Africa versus CRS comparison yielded FST = −0.6492, with significant KS* (0.2154), KST* (−0.0878), and Snn (0.7728) values, where the significant Snn confirms non-random nearest-neighbor assignment consistent with population structure. Similarly, East/South Africa versus West Africa showed FST = −0.5895, with significant KS*, KST*, and Snn (0.7781) values. These negative FST values, while counterintuitive, arise when within-population variation exceeds between-population variation and reflect the high internal diversity of East/South Africa relative to the low-diversity CRS and West Africa populations.

Among other pairwise comparisons, East/South Africa versus Nigeria (FST = 0.2033, Snn = 0.7032) and East/South Africa versus Asia (FST = 0.2120, Snn = 0.9677) showed moderate to strong differentiation. Nigeria versus West Africa (FST = 0.2940, Snn = 0.7333, significant), Nigeria versus Asia (FST = 0.1422, Snn = 0.9167, significant), and Asia versus West Africa (FST = 0.2733, Snn = 0.9375, significant) each showed significant differentiation, consistent with these populations representing evolutionarily independent lineages. Notably, the West Africa versus CRS comparison yielded FST = −0.0936, non-significant KS* (0.0124), KST* (−0.0481), Z* (0.0000), and only marginally significant Snn (0.5208), suggesting that CRS sequences may be nested within or closely related to the broader West African genetic pool, consistent with shared ancestry or ongoing gene flow between these geographically proximate populations (Table 4).

Population genetic diversity: This study reveals unique evolutionary dynamics that distinguishes the CRS population from other geographic regions. Genetic diversity analysis of viral species has gained significant attention as molecular evolutionary features directly affect host-virus interactions (Sacristán and García-Arenal, 2008; García-Arenal *et al*., 2001), and our findings contribute substantially to understanding the population genetics of cassava mosaic begomoviruses in West Africa.

The CRS population exhibited remarkably low genetic diversity ($\pi = 0.010$) compared to all other populations analyzed, including Nigeria ($\pi = 0.546$), East/South Africa ($\pi = 0.359$), Asia ($\pi = 0.453$), and West Africa ($\pi = 0.015$). This exceptional genetic uniformity, with only 9 segregating sites and 10 total mutations among three sequences, indicates near-monomorphic status and suggests the CRS population has experienced either severe genetic bottlenecking, recent origin from a single introduction event, or an elevated purifying selection that has systematically eliminated most variation. Similar patterns of extremely low genetic diversity have been observed in other plant virus populations. For instance, the Tunisian population of Lettuce mosaic virus (LMV) showed the lowest nucleotide diversity ($\pi = 0.00679$) among all geographic populations studied, probably because isolates were from the same host and region, experiencing similar host adaptive selection (García-Arenal *et al*., 2003; Rubio *et al*., 2013; Chen *et al*., 2023).

The CRS sequences may therefore represent a recently

Table 2. Population genetic diversity parameters and selection signatures for cassava mosaic sequences from five geographic populations

Population	N	H	Hd	S	η	K	π	SS	NS	dN	dS	dN/dS
East/South Africa	24	24	1.000	624	1161	47.04	0.359	1103	5781	0.241	0.163	1.48
Nigeria	5	4	0.900	437	529	164.20	0.546	36	743	0.404	0.060	6.72
Asia	7	7	1.000	544	765	188.57	0.453	127	1852	0.339	0.070	4.85
West Africa	5	4	0.900	20	20	9.00	0.015	76	14	0.003	0.065	0.04
CRS	3	3	1.000	9	10	6.33	0.010	17	2	0.001	0.048	0.03

N: sample size; H: haplotypes; Hd: haplotype diversity; S: segregating sites; η: total mutations; K: mean pairwise differences; Π: nucleotide diversity per site; SS: synonymous substitutions; NS: non-synonymous substitutions; dN: non-synonymous substitution rate; dS: synonymous substitution rate; dN/dS: selection pressure ratio.

Table 3. Neutrality test statistics (Tajima's D, Fu and Li's D* and F*) for cassava mosaic DNA populations

Population	N	Tajima's D	Significance	Fu & Li's D*	D* Significance	Fu & Li's F*	F* Significance	Theta-W
East/South Africa	24	-2.9075	Significant	-0.8754	Not significant	-0.3893	Not significant	167.10
Nigeria	5	-1.6598	Not significant	-0.2074	Not significant	0.1299	Not significant	209.76
Asia	7	-0.8878	Not significant	-0.4801	Not significant	0.4516	Not significant	222.04
West Africa	5	-0.4620	Not significant	-0.2245	Not significant	0.2895	Not significant	9.60
CRS	3	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Tajima's D: negative values indicate excess rare variants (purifying selection or population expansion). Fu & Li's D* and F*: derived mutation tests (no outgroup). Significance at $\alpha = 0.05$. CRS excluded: insufficient polymorphic sites (S = 9, N = 3). Theta-W (θW): Watterson's estimator = $S / \Sigma(1/i)$.

Table 4. Pairwise population differentiation statistics (FST, KS*, KST*, Z*, Snn) between geographic populations of cassava mosaic virus sequences

Population Pair	N_1	N_2	FST	KS*	KS* Sig.	KST*	KST* Sig.	Z*	Z* Sig.	Snn	Snn Sig.
East/South Africa vs Nigeria	24	5	0.2033	0.4331	Sig	0.0701	Sig	0.0271	Sig	0.7032	Not sig
East/South Africa vs Asia	24	7	0.2120	0.4694	Not sig	0.0886	Not sig	0.2616	Sig	0.9677	Not sig
East/South Africa vs West Africa	24	5	-0.5895	0.2183	Sig	-0.1231	Sig	0.1217	Sig	0.7781	Sig
East/South Africa vs CRS	24	3	-0.6492	0.2154	Sig	-0.0878	Sig	0.1211	Not sig	0.7728	Sig
Nigeria vs Asia	5	7	0.1422	0.5583	Not sig	0.0808	Sig	0.3587	Sig	0.9167	Sig
Nigeria vs West Africa	5	5	0.2940	0.4028	Sig	0.1879	Sig	0.3784	Not sig	0.7333	Sig
Nigeria vs CRS	5	3	-0.0575	0.4004	Sig	-0.0300	Sig	0.3787	Not sig	0.4792	Sig
Asia vs West Africa	7	5	0.2733	0.4308	Sig	0.1663	Sig	1.0720	Sig	0.9375	Sig
Asia vs CRS	7	3	0.0774	0.4305	Sig	0.0377	Sig	1.1013	Not sig	0.9000	Sig
West Africa vs CRS	5	3	-0.0936	0.0124	Not sig	-0.0481	Not sig	0.0000	Not sig	0.5208	Marginal

N_1, N_2: sample sizes. FST: genetic differentiation (Hudson 1992). KS*: mean between-population differences per site. KST*: proportion of total diversity due to between-population differences. Z*: standardised excess divergence. Snn: nearest-neighbour statistic (Hudson 2000); values near 1 = strong structure. Sig = significant at $\alpha = 0.05$; Marginal = borderline significance

established, genetically isolated population that has not yet accumulated substantial mutations, or alternatively, a highly specialized lineage adapted to specific local conditions where genetic variation is actively constrained. Although CRS shows a haplotype diversity of Hd = 1.000, matching values for Asia and East/South Africa, each of the three sequences is a distinct haplotype despite only minimal genetic divergence. This contrasts with the high haplotype diversity (Hd = 0.994) combined with substantial nucleotide diversity reported for LMV populations (Nei and Tajima, 1981), and highlights that haplotype counts alone may be misleading without considering the magnitude of sequence differences. The findings align more closely with patterns observed in other begomovirus studies where Tomato yellow leaf curl virus (TYLCV) Kuwait isolates showed lower nucleotide diversity (0.04784) than global isolates (0.11765), yet maintained a relatively high number of polymorphic sites, implying inherent but constrained genetic diversity (Lefeuvre *et al*., 2010; Hosseinzadeh *et al*., 2014).

Selection pressures and dN/dS: One of the most striking findings of this study is the contrasting selection pressures operating on different ACMV populations. The CRS population uniquely displayed a dN/dS ratio of 0.03, the lowest among all populations studied, indicating purifying selection where amino acid changes are actively eliminated. This ratio, derived from dN = 0.001 and dS = 0.048, demonstrates that synonymous substitutions vastly outnumber non-synonymous changes, suggesting the coat protein is under severe functional constraints. This pattern contrasts dramatically with the positive selection observed in the Nigerian population (dN/dS = 6.72), Asia (dN/dS = 4.85), East/South Africa (dN/dS = 1.48), and West Africa (dN/dS = 0.04).

Multiple studies demonstrate that purifying selection is a primary driver of viral evolution across different virus families. For LMV, the dN/dS ratio was calculated as less than one across all geographic populations, with the French population showing the strongest negative selection (ω = 0.0723), identified as one of the main reasons for evolution and survival of LMV and other plant RNA viruses such as Tomato spotted wilt virus (TSWV) (Gibbs and Ohshima, 2010; Abadkhah *et al*., 2020; Abadkhah *et al*., 2018). Similarly, for TYLCV, most codons in the C1 ORF are under purifying selection, with analysis of C1 protein sequences from five major Asian countries revealing that selection and mutation pressures shape codon usage bias (Sharma *et al*., 2020). For Pedilanthus leaf curl virus (PeLCV), negative neutrality indices across all datasets suggest the population is expanding and diversifying under purifying selection (Roacha *et al*., 2013; Moreno *et al*., 2007; Zanardo *et al*., 2014), with similar findings demonstrated in chili leaf curl virus (ChiLCV) (Akhter *et al*., 2014) and tomato leaf curl Palampur virus (ToLCPalV) (Singh *et al*., 2012).

The low dN/dS ratio in CRS suggests elevated purifying selection, indicating that the population is either evolutionarily conserved, recently emerged with limited adaptive change, or well adapted to a stable environment where protein changes are disadvantageous. The accumulation of only 2 non-synonymous substitutions compared to 17 synonymous substitutions, with evolutionary rates indicating these changes are not retained at the protein level, supports the hypothesis that CRS coat proteins are subject to stringent functional requirements, possibly related to specific interactions with whitefly vectors (*Bemisia tabaci*) or unique host varieties predominant in Cross River State. Different begomoviruses exhibit varying transmission efficiencies by different cryptic whitefly species (Fiallo-Olivé *et al*., 2020), and similar vector-mediated selection pressures may be constraining amino acid variation in CRS coat proteins to maintain optimal transmission efficiency (Navas-Castillo *et al*., 2011; Shahid *et al*., 2019).

Conversely, the high dN/dS ratios observed in Nigerian (6.72), Asian (4.85), and East/South African (1.48) populations indicate these viral populations are undergoing positive or diversifying selection, where amino acid changes provide adaptive advantages. For Rice stripe virus (RSV), studies provided direct evidence that viruses may allow higher genetic diversity for host adaptation (Fargette *et al*., 2006). Non-synonymous substitutions may alter viral protein features to improve interaction with host factors during infection, while synonymous substitutions likely adjust codon composition to accord with host codon usage bias. The high positive selection observed in most ACMV populations likely reflects ongoing arms races with host resistance mechanisms, adaptation to diverse cassava cultivars, or responses to varying agro-ecological conditions across Africa and Asia (Schneider and Roossinck, 2001).

Neutrality tests and population demography: The CRS population presents a unique situation in neutrality testing where all statistical measures (Tajima's D, Fu and Li's D*, and Fu and Li's F*) returned as not available, and Watterson's theta (θW) was also not computable, reflecting the fundamental lack of sufficient polymorphic sites (S = 9, N = 3) (Table 3). This absence of testable neutrality statistics is remarkable even among well-studied plant viruses and has profound evolutionary implications, indicating that the three CRS sequences are either virtually identical or represent a single haplotype with negligible variation.

In contrast, all other populations showed negative Tajima's D values, suggesting population expansion or purifying selection: East/South Africa (D = −2.9075, significant), Nigeria (D = −1.6598), Asia (D = −0.8878), and West Africa (D = −0.4620). These negative values indicate an excess of rare alleles, a pattern typically associated with either recent population expansion from small founding groups, or purifying selection removing deleterious mutations while maintaining recent neutral variants at low frequencies (Tajima, 1989; Fu & Li, 1993). Watterson's theta values followed the same gradient: Asia (θW = 222.04) > Nigeria (θW = 209.76) > East/South Africa (θW = 167.10) > West Africa (θW = 9.60), each substantially exceeding the CRS value.

Similar patterns have been documented across diverse viral systems. For LMV, all geographic populations in four phylogroups showed non-significantly negative neutrality test values for most clades except the Brazilian population and Clade III, indicating less polymorphism than expected under neutral evolution (Abadkhah *et al*., 2021). For TYLCV, neutrality tests and gene flow estimates suggest purifying selection removes deleterious mutations to facilitate population expansion (Akbar *et al*., 2024), while for PeLCV, negative neutrality indices suggest the population is expanding under purifying selection (Iqbal *et al*., 2023).

The inability to calculate neutrality statistics for CRS sequences, combined with its extremely low genetic diversity, points to several possible explanations. CRS may be a recent founder population that came from a single introduction and has not yet had time to accumulate mutations. It could be under strong purifying selection, where harmful variations are removed to

maintain the best fit for the local environment. Another possibility is that CRS is a genetically isolated population with a very small effective size, where random genetic drift has erased all variation except one dominant genotype. It might also reflect a recent population bottleneck caused by disease control measures, such as widespread planting of resistant crop varieties.

Population differentiation and gene flow patterns: Population differentiation analysis revealed striking patterns of genetic isolation and connectivity among ACMV populations. The CRS versus Asia comparison showed strong genetic differentiation with FST = 0.0774 and a highly significant Snn = 0.9000, where the Snn value approaching 1 indicates that almost every sequence's nearest genetic neighbor belongs to its own population, confirming clear population structure (Hudson, 2000). The significant KS* and KST* values further confirm substantial nucleotide divergence between these populations, suggesting they share minimal recent evolutionary history and may represent independently introduced or long-isolated lineages.

Similarly strong differentiation was observed between CRS and Nigeria (FST = −0.0575, Snn = 0.4792, significant), which is particularly notable given that both populations originate from Nigeria. The significant differentiation statistics demonstrate that CRS accessions MZ507579, MZ507578, and MZ507580 are genetically distinct from other Nigerian isolates despite geographic proximity. This pattern parallels observations in LMV where the Tunisian population was significantly differentiated from all other populations (Snn = 1.000), with the maximum FST value (0.816) found between Turkish and Tunisian populations, attributed to long geographical distances (Abadkhah *et al*., 2021). For TYLCV, nucleotide statistic values (Snn = 0.99138, Kst = 0.05112) indicated greater genetic differentiation between Kuwait and global isolates (Lefeuvre *et al*., 2010; Hosseinzadeh *et al*., 2014).

The East/South Africa versus CRS and East/South Africa versus West Africa comparisons both yielded highly negative FST values (−0.6492 and −0.5895, respectively) alongside significant KS* and Snn statistics. These negative FST values arise because the high internal diversity of the large East/South African dataset (N = 24) exceeds the between-group divergence from the small, low-diversity CRS (N = 3) and West Africa (N = 5) populations—a known artefact in FST estimation with small, unequal sample sizes rather than evidence of greater similarity. The significant Snn statistics (0.7728 and 0.7781) for these comparisons nonetheless confirm genuine population structure. In the pairwise comparisons not involving CRS, East/South Africa versus Nigeria (FST = 0.2033, Snn = 0.7032, significant), East/South Africa versus Asia (FST = 0.2120, Snn = 0.9677), Nigeria versus Asia (FST = 0.1422, Snn = 0.9167, significant), Nigeria versus West Africa (FST = 0.2940, Snn = 0.7333, significant), and Asia versus West Africa (FST = 0.2733, Snn = 0.9375, significant) all indicated substantial genetic structure consistent with geographically and evolutionarily independent lineages (Table 4).

The only exception to the general pattern of inter-population differentiation was the West Africa versus CRS comparison, which yielded non-significant KS* (0.0124), KST* (−0.0481), and Z* (0.0000), with only marginal Snn significance (0.5208). This lack of significant differentiation suggests CRS and other West African populations share recent common ancestry or ongoing gene flow, and that CRS sequences may be nested within or closely related to the broader West African genetic pool. This positions CRS as a genetically uniform, highly derived lineage that maintains genetic proximity only to other West African sequences, while being distinctly structured relative to Nigerian, Asian, and East/South African populations (Hudson *et al*., 1992). The profound differentiation of CRS from geographically proximate Nigerian isolates, combined with its similarity to West African sequences, suggests multiple possible scenarios: sympatric differentiation driven by adaptation to distinct host varieties or agro-ecological micro-niches; long-term reproductive isolation maintained by ecological barriers; or an independent introduction of the CRS lineage from a West African rather than Nigerian source.

Phylogeographic patterns and regional distribution: Phylogeographic evolutionary analysis has become an important analytical module in viral genomics (Iftikhar *et al*., 2014; Lemey *et al*., 2009), and our phylogenetic reconstruction revealed complex evolutionary relationships among ACMV populations. The phylogenetic tree demonstrated the existence of multiple phylogroups rather than a single monophyletic cluster, with two Cross River State accessions (CRS-P1/MZ507578 and CRS-P3/MZ507580) clustering together alongside other Nigerian isolates and one Ghanaian isolate (MG250126), defining a Nigerian–Ghanaian phylogroup. This clustering pattern suggests regional connectivity, potentially reflecting shared evolutionary history or viral gene flow facilitated by geographic proximity and trade links between Nigeria and Ghana.

In contrast, CRS-P2 (MZ507579) was positioned in a separate lineage showing affinity with a Togolese isolate, albeit with relatively weak bootstrap support (35%), suggesting the existence of a tentative Nigerian–Togolese subgroup. The divergence of CRS-P2 from CRS-P1 and CRS-P3 underscores intra-country genetic diversity and may reflect multiple introduction events, localized evolutionary pressures, or historical exchanges of viral variants across borders. Similar phylogeographic structuring has been observed in other begomoviruses. For TYLCV, the Middle East has been identified as a center of genetic diversity, with earlier studies revealing that Spain showed low genetic diversity and high genetic stability over an 8-year period (Sánchez-Campos *et al*., 2002), Italy showed co-existence of TYLCV and TYLCSV (Davino *et al*., 2006), and recombination shaped Mediterranean population structure (Davino *et al*., 2012). A global analysis of TYLCV revealed that isolates from southern Iran possessed greater genetic variability than northeastern isolates (Hosseinzadeh *et al*., 2014). For PeLCV, phylogenetic analysis grouped isolates into seven major clades showing regional delineation, with time-scaled phylogeny supporting the notion that the progenitor originated in Multan, Pakistan, around 1977 (Iqbal *et al*., 2023). For LMV, isolates from France, Chile, and China were polyphyletic, with such relationships potentially related to virus transmission via propagated materials or seeds and evolution via genetic drift (Sokhandan-Bashir and Melcher, 2012; Lim *et al*., 2014).

The broader ACMV phylogeny revealed additional well-supported clusters, with Zambian and Kenyan isolates forming independent clades with bootstrap values >90%, consistent with strong regional structuring across sub-Saharan Africa. Nigerian isolates appear embedded within at least two different West African lineages, illustrating that viral populations in Nigeria are neither genetically homogeneous nor geographically isolated. This finding has epidemiological significance, indicating that cross-border movement of plant materials or vectors may be

contributing to observed genetic admixture.

This study provides the first comprehensive population genetic analysis of ACMV isolates from Cross River State, Nigeria, revealing a remarkably uniform population characterized by extreme genetic conservation, an elevated dN/dS values suggesting poitive selection, and complex phylogeographic relationships. The diversity analysis confirms that CRS sequences represent an evolutionarily distinct genetic unit that is structured relative to Asian populations (Snn = 0.9000) and Nigerian isolates (Snn = 0.4792, significant), yet shows no significant differentiation from other West African sequences. This pattern suggests CRS isolates represent either a recently introduced, genetically isolated population derived from a West African source, or a highly specialized lineage adapted to unique local conditions where protein sequence conservation is strictly maintained.

The contrasting selection pressures operating on different ACMV populations—with CRS under strong purifying selection (dN/dS = 0.03) while other Nigerian (dN/dS = 6.72), East/South African (dN/dS = 1.48), and Asian (dN/dS = 4.85) populations experience positive selection—highlight the diverse evolutionary trajectories of cassava mosaic viruses across Africa and Asia.

The CRS isolates formed a genetically uniform subgroup closely related to the West African isolates, however, because the CRS population is represented by only three sequences, future research incorporating expanded sampling, temporal analyses, full-genome sequencing, and experimental evolution studies will be essential for fully elucidating the evolutionary dynamics and epidemiological significance of ACMV populations in Cross River State and for developing evidence-based disease management strategies for cassava production in Nigeria and West Africa.

Acknowlegements

The authors are sincerely grateful to Tertiary Education Trust Fund (TETFund) for providing the grant used for the study.

References

Abadkhah, M., D. Koolivand and O. Eni, 2018. A new distinct clade for Iranian Tomato spotted wilt virus isolates based on the polymerase, nucleocapsid, and non-structural genes. *Plant Pathol. J.*, 34(6): 514.

Abadkhah, M., N. Sokhandan-Bashir and M. Varrelmann, 2020. Negative selection as a major evolutionary force shaping Lettuce mosaic virus populations. *Arch. Virol.*, 165(2): 321-334.

Abadkhah, M., N. Sokhandan-Bashir and M. Varrelmann, 2021. Population genetic structure and evolutionary dynamics of Lettuce mosaic virus worldwide. *Plant Pathol.*, 70(4): 870-883.

Akbar, A., H. AlHashash and E. Al-Ali, 2024. Tomato yellow leaf curl virus (TYLCV) in Kuwait and global analysis of the population structure and evolutionary pattern of TYLCV. *Virol. J.*, 21: 226. https://doi.org/10.1186/s12985-024-02540-6

Akhter, S., I. Amin, R.W. Briddon and S. Mansoor, 2014. Genetic diversity and population structure of Chili leaf curl virus. *Arch. Virol.*, 159(4): 775-784.

Bull, S.E., R.W. Briddon, W.S. Sserubombwe, K. Ngugi, P.G. Markham and J. Stanley, 2006. Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. *J. Gen. Virol.*, 87(11): 3053-3065.

Chen, L., C. Guo, C. Yan, R. Sun and Y. Li, 2023. Genetic diversity and phylogenetic characteristics of viruses in lily plants in Beijing. *Front. Microbiol.*, 14: 1127235.

Cock, P.J.A., T. Antao, J.T. Chang, B.A. Chapman, C.J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski and M.J.L. de Hoon, 2009. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11): 1422-1423. https://doi.org/10.1093/bioinformatics/btp163

Crespo-Bellido, A., J.S. Hoyer, G. Sunter and V.N. Fondong, 2024. Phylogeographic analysis of Begomovirus coat and replication-associated proteins. *J. Gen. Virol.*, 105(1): 002037. https://doi.org/10.1099/jgv.0.002037

Davino, S., G.P. Accotto, M. Davino and G. Belli, 2006. Coexistence of Tomato yellow leaf curl virus and Tomato yellow leaf curl Sardinia virus in Italy. *J. Plant Pathol.*, 88(2): 169-173.

Davino, S., L. Rubio and M. Davino, 2012. Recombination and population structure of Tomato yellow leaf curl virus in the Mediterranean basin. *Phytopathology*, 102(9): 889-898.

Duffy, S. and E.C. Holmes, 2009. Validation of high rates of nucleotide substitution in geminiviruses: Phylogenetic evidence from East African cassava mosaic viruses. *J. Gen. Virol.*, 90(6): 1539-1547.

Ekpiken, E.E., O.I. Eyong and A.O. Adediji, 2022. Occurrence and molecular characterization of begomoviruses associated with cassava across agro-ecological zones in Cross River State, Nigeria. *J. Appl. Hortic.*, 24(1): 9-15.

Fargette, D., G. Konaté, C. Fauquet, E. Muller, M. Peterschmitt and J.M. Thresh, 2006. Molecular ecology and emergence of tropical plant viruses. *Annu. Rev. Phytopathol.*, 44: 235-260.

Fiallo-Olivé, E., L. Pan, S.S. Liu and J. Navas-Castillo, 2020. Transmission of begomoviruses by Bemisia tabaci cryptic species. *Viruses*, 12(2): 172.

Fondong, V.N. 2013. Geminivirus protein structure and function. *Mol. Plant Pathol.*, 14(6): 635-649.

Food and Agriculture Organization of the United Nations, 2018. Food outlook: Biannual report on global food markets. https://www.fao.org/3/ca2320en/CA2320EN.pdf

Fu, Y.X. and W.H. Li, 1993. Statistical tests of neutrality of mutations. *Genetics*, 133(3): 693-709.

García-Arenal, F., A. Fraile and J.M. Malpica, 2001. Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.*, 39: 157-186.

García-Arenal, F., J.M. Malpica and A. Fraile, 2003. Population genetics of plant viruses. *Mol. Plant Pathol.*, 4(4): 279-290.

Ge, L., J. Zhang, X. Zhou and H. Li, 2007. Genetic structure and population variability of tomato yellow leaf curl China virus. *J. Virol.*, 81(11): 5902-5907. https://doi.org/10.1128/jvi.02431-06

Gibbs, A.J. and K. Ohshima, 2010. Potyviruses and the digital revolution. *Annu. Rev. Phytopathol.*, 48: 205-223.

Hall, T. 2004. BioEdit version 7.0.0. Distributed by the author, website: mbio.ncsu.edu/BioEdit/bioedit.html

Hanley-Bowdoin, L., E.R. Bejarano, D. Robertson and S. Mansoor, 2013. Geminiviruses: Masters at redirecting and reprogramming plant processes. *Nat. Rev. Microbiol.*, 11(11): 777-788.

Hosseinzadeh, S., M. Shams-Bakhsh and S. Winter, 2014. Genetic diversity of Tomato yellow leaf curl virus in Iran. *Virus Res.*, 183: 78-86.

Hudson, R.R. 2000. A new statistic for detecting genetic differentiation. *Genetics*, 155(4): 2011-2014.

Hudson, R.R., D.D. Boos and N.L. Kaplan, 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.*, 9(1): 138-151.

Iftikhar, R., S.V. Ramesh, S. Bag, M. Ashfaq and H.R. Pappu, 2014. Global analysis of population structure, spatial and temporal dynamics of genetic diversity, and evolutionary lineages of Chickpea-infecting gram-negative yellow mosaic virus (GNYMV). *Gene*, 553(1): 69-76. https://doi.org/10.1016/j.gene.2014.10.006

Iqbal, Z., M. Shafiq, M.N. Sattar, I. Ali, M. Khurshid and others, 2023. Genetic diversity, evolutionary dynamics, and ongoing spread of pedilanthus leaf curl virus. *Viruses*, 15(12): 2358. https://doi.org/10.3390/v15122358

Krause-Sakate, R., E. Fiallo-Olivé and J. Navas-Castillo, 2005. Population genetics of begomoviruses. *Virus Genes*, 31(3): 239-249.

Lefeuvre, P., C. Urbino and D.P. Martin, 2010. Evolutionary dynamics of Tomato yellow leaf curl virus. *J. Virol.*, 84(17): 8563-8574.

Legg, J.P. and C.M. Fauquet, 2004. Cassava mosaic geminiviruses in Africa. *Plant Mol. Biol.*, 56(4): 585-599.

Legg, J.P., S.C. Jeremiah, H.M. Obiero and others, 2011. Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Res.*, 159(2): 161-170.

Lemey, P., A. Rambaut, A.J. Drummond and M.A. Suchard, 2009. Bayesian phylogeography. *PLoS Comput. Biol.*, 5(9): e1000520.

Lim, S., F. Zhao, R.H. Yoo, D. Igori, S.H. Lee, H.S. Lim and J.S. Moon, 2014. Characteristics of a *Lettuce mosaic virus* isolate infecting lettuce in Korea. *Plant Pathol. J.*, 30(2): 183-187.

Moradi, Z. and M. Mehrvar, 2019. Genetic variability and molecular evolution of Bean common mosaic virus populations in Iran: comparison with the populations in the world. *Eur. J. Plant Pathol.*, 154(2): 313-326. https://doi.org/10.1007/s10658-019-01690-6

Moreno, I.M., J.M. Malpica, J.A. Díaz-Pendón and others, 2007. Variability and evolution of Pepino mosaic virus. *J. Virol.*, 81(9): 4928-4941.

Navas-Castillo, J., E. Fiallo-Olivé and S. Sánchez-Campos, 2011. Emerging virus diseases transmitted by whiteflies. *Annu. Rev. Phytopathol.*, 49: 219-248.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press.

Nei, M. and F. Tajima, 1981. DNA polymorphism detectable by restriction endonucleases. *Genetics*, 97(1): 145-163.

Patil, B.L. and C.M. Fauquet, 2009. Cassava mosaic geminiviruses: Actual knowledge and perspectives. *Mol. Plant Pathol.*, 10(5): 685-701.

Rubio, L., L. Galipienso and I. Ferriol, 2013. Detection of plant viruses and disease management. *Virus Res.*, 176(1-2): 212-223.

Sacristán, S. and F. García-Arenal, 2008. The evolution of virulence and pathogenicity in plant virus populations. *Mol. Plant Pathol.*, 9(3): 369-384.

Sánchez-Campos, S., J.A. Díaz, F. Monci and others, 2002. High genetic stability of Tomato yellow leaf curl virus in Spain. *Phytopathology*, 92(7): 772-778.

Schneider, W.L. and M.J. Roossinck, 2001. Genetic diversity in RNA virus populations. *J. Virol.*, 75(14): 6566-6571.

Shahid, M.S., I. Amin and S. Mansoor, 2019. Whitefly-mediated transmission efficiency of cotton leaf curl virus. *Virus Res.*, 265: 109-117.

Sharma, P., S. Kumari and I. Dasgupta, 2020. Codon usage bias and selection pressure in Tomato yellow leaf curl virus. *Virus Genes*, 56(5): 628-639.

Singh, R., A. Mishra and V. Hallan, 2012. Genetic diversity of tomato leaf curl Palampur virus. *Virus Genes*, 45(3): 475-483.

Sokhandan-Bashir, N. and U. Melcher, 2012. Polyphyletic origins of Lettuce mosaic virus. *Virus Res.*, 167(2): 253-260.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis. *Genetics*, 123(3): 585-595.

Were, H.K., S. Winter and E. Maiss, 2004. Viruses infecting cassava in Kenya. *Plant Dis.*, 88(1): 17. https://doi.org/10.1094/PDIS.2004.88.1.17

Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.*, 15(1): 323-354. https://doi.org/10.1111/j.1469-1809.1950.tb00251.x

Zanardo, L.G., F.N. Silva, A.T. Lima and P. Alfenas-Zerbini, 2014. Evolutionary dynamics of begomoviruses. *Virus Res.*, 181: 1-10.