Phenotypic screening and molecular characterization of cassava mosaic disease resistance in Côte d'Ivoire cassava germplasm

William J-L. Amoakon[1,2]*, Mariam Combala[2,3], Justin S. Pita[2,3]*, J. Musembi Mutuku[2], Boni N'Zué[4], Daniel H. Otron[2,3], Edwige F. Yéo[2,3], Nazaire K. Kouassi[2,3] and Raoul Sié[1]

[1]UFR Sciences de la Nature, Université Nangui Abrogoua (UNA), Abidjan, Côte d'Ivoire, [2]The Central and West African Virus Epidemiology (WAVE) for Food Security Program, Pôle Scientifique et d'Innovation, Université Félix Houphouët-Boigny (UFHB), Abidjan, Côte d'Ivoire, [3]UFR Biosciences, Université Félix Houphouët-Boigny (UFHB), Abidjan, Côte d'Ivoire, [4]Centre National de Recherche Agronomique (CNRA), Bouaké, Côte d'Ivoire

Cassava is the staple food crop for hundreds of millions of people in Africa. In Côte d'Ivoire, it is a main source of calories for over 26 million people. However, cassava mosaic disease (CMD), caused by cassava mosaic geminiviruses (CMGs) threaten its production. The development, adoption, and use of CMD resistant varieties remain a key CMD management strategy. Therefore, 610 accessions from the Côte d'Ivoire national cassava germplasm maintained by Centre National de Recherche Agronomique (CNRA) research stations in the cities of Man and Bouaké were characterized to assess their resistance to CMD. We found 72 accessions which were symptomless at both sites. Thirty five out of these 72 accessions were infected by African cassava mosaic virus (ACMV) without any CMD symptoms and 37 were virus free. None of the 72 symptomless accessions were infected by East African cassava mosaic Cameroon virus (EACMCMV). The remaining 538 accessions were CMD infected and displayed clear CMD symptoms. The promising 72 accessions (10 locals and 62 improved) were then tested for the presence of resistance genes, CMD1, CMD2 or CMD3. Except for 2 accessions for which no resistance gene was found, the remaining 70 accessions contained one or more resistance genes. Genotyping of the 69 symptomless accessions revealed that 56 were unique genotypes and 13 were potential duplicates. Accessions were structured in three groups with the variabilities at the individual's level. These findings reveal the existence of potential tolerant/resistant cassava accessions in CNRA's germplasm, which can contribute to CMD control and the increase of cassava production in Côte d'Ivoire.

KEYWORDS

cassava germplasm, cassava mosaic disease, CMD resistance gene, genotyping, genetic diversity, SNP markers

1. Introduction

Cassava mosaic disease (CMD) is a major constraint to cassava (*Manihot esculenta* Crantz) production in Africa. Yield loss estimates in susceptible cultivars can be as high as 90 % (Vernier et al., 2018) and can be up to $2.7 billion USD (Patil and Fauquet, 2009). Therefore, losses due to CMD have an immediate impact on the food supply and threaten food security and the livelihoods of Africa's rapidly growing population (FAO, 2014). CMD is mainly spread by infected cuttings and transmitted by the whitefly *Bemisia tabaci* (Njoroge et al., 2017; Soro et al., 2021). CMD is caused by a complex of at least 11 Cassava mosaic begomoviruses (CMBs)

worldwide, of which nine occur in Africa and two are found in the Indian subcontinent (ICTV, 2019). They include members of the species *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), *East African cassava mosaic Cameroon virus* (EACMCMV), *East African cassava mosaic Malawi virus* (EACMMV), *East African cassava mosaic Zanzibar virus* (EACMZV), *East African cassava mosaic Kenya virus* (EACMKV), *Cassava mosaic Madagascar virus* (CMMGV), *African cassava mosaic Burkina Faso virus* (ACMBFV), *Indian cassava mosaic virus* (ICMV), *South African cassava mosaic virus* (SACMV), and *Sri Lankan cassava mosaic virus* (SLCMV) (ICTV, 2019).

Despite CMD and other cassava cultivation constraints, the crop has become the most important food crop in Africa because of its high productivity potential and its ability to grow on poor soils (Imakumbili et al., 2021). Given the rapid growth of the population and consequently, the increased demand of food, particularly in Africa urban areas, it is evident that the need for cassava will grow. In Côte d'Ivoire, cassava is the third most important source of calories after yam and rice (FAOSTAT, 2020), and with its ever-increasing demand, the production is no longer sufficient to meet the population's needs. This insufficient production (6.4 million tons per year) is due in part to CMD which considerably reduces cassava production (Dixon et al., 2000). The use of resistant varieties and the supply of healthy planting materials to farmers is an effective strategy to mitigate the impact of the disease. Resistant varieties can significantly reduce yields losses and the source of inoculum of the virus in the field (Rabbi et al., 2014). The first CMD-resistant varieties, developed by IITA in Nigeria, were hybrids obtained with *M. glaziovii,* found in Brazil, which conferred multigenic resistance (CMD1). Among these varieties, three have been largely popularized and have shown a very high CMD resistance with low or no yield losses. These are: TMS 60142, TMS 30337 and TMS 30572 (Vernier et al., 2018). In addition, the variety TME 419, originating from Togo, was improved by IITA-Nigeria and conferred with the CMD resistance. TME 419 is also a very productive and widely distributed variety. Currently, three CMD resistance genes, CMD1 (polygenetic recessive), CMD2 (monogenetic dominant) and CMD3 (quantitative trait loci, QTL, conferring resistance) were discovered and important molecular markers associated with CMD2 and CMD3 have been identified (Akano et al., 2002; Okogbenin and Fregene, 2002; Okogbenin et al., 2012). Our goal is to select the best accessions for CMD resistance from the CNRA germplasm to be used for breeding programs, and production and dissemination of clean seed materials throughout the country, with a view to reduce the impact of CMD to cassava production in Côte d'Iviore.

2. Material and methods

2.1. Description of the study sites and plant materials assessed

The field evaluation was conducted from September to October 2017 in two CNRA research stations located in the cities of Bouaké (7°40'N 5°05'W; Altitude: 399 m) and Man (7°20'N 7°36'W; Altitude: 1050 m). The most significant differences between the two locations were the altitude, the rainfall and the vegetation (Table 1). A total of 610 cassava accessions were assessed in this study, with 8 to 10 plants evaluated per accession. These accessions were originally from: Côte d'Ivoire (465 accessions), IITA-Nigeria (99 accessions), Central

African Republic (RCA; 18 accessions), Togo (7 accessions), Kenya (6 accessions), Madagascar (8 accessions), Congo (5 accessions) and Liberia (2) accessions. Independent from the origin, the germplasm included 429 local accessions and 181 improved ones, out of which 399 are local accessions from Côte d'Ivoire that are routinely used for local dishes while the remaining 67 Ivorian accessions are varieties improved mostly for yield, taste, and their aptitude for processing into attiéké (cassava couscous). The accessions from IITA-Nigeria, Kenya and Madagascar are all improved varieties presumed to be resistant to CMD. The remaining accessions from the other countries are mostly local accessions.

2.2. Classification of the accessions according to their level of resistance to cassava mosaic disease

Disease severity was scored based on visual assessment using a scale of 1–5 (Sseruwagi et al., 2004), where 1 represent absence of disease symptoms and 5 the most severe symptoms, including total leaf distortion and stunting of the entire plants (Figure 1). The mean severity (S_m) of each accession was calculated using the following formula:

$$S_m = \frac{\sum \text{scores of diseased plants}}{\text{Total number of diseased plants}}$$

The following scale was used to classify the accessions into different groups:

[1]: symptomless accessions (accessions with severity 1).
[2–3]: susceptible (S) accessions (accessions with severity from 2 to 3).
[3.1–5]: highly susceptible (HS) accessions (accessions with severity from 3.1 to 5).

2.3. Leaf sampling for molecular analyses

Samples were collected only from symptomless cassava accessions (accessions in which all plants showed no CMD symptoms). For each such accession, samples were collected from all plants evaluated. 8 to 10 samples were collected per cassava accession. A total of 2,025 samples were collected (709 samples from Bouaké, corresponding to 77 accessions and 1,316 samples from Man, corresponding to 154 accessions). The samples collected were placed in a zip lock bag and kept in a cool box for transportation to the laboratory where they were stored at −20°C.

2.4. Detection of cassava mosaic geminiviruses in the symptomless accessions

Total DNA extraction from cassava leaf samples was performed according to the protocol by Doyle and Doyle (1987). The concentration of each extracted DNA was determined using a spectrophotometer (Eppendorf) and adjusted to 50 ng for use in polymerase chain rection (PCR). Partial DNA was amplified by PCR using specific primer pairs for the detection of ACMV, EACMV

TABLE 1 Agro-ecological characteristics of cassava field assessment sites in Bouaké and Man, Côte d'Ivoire.

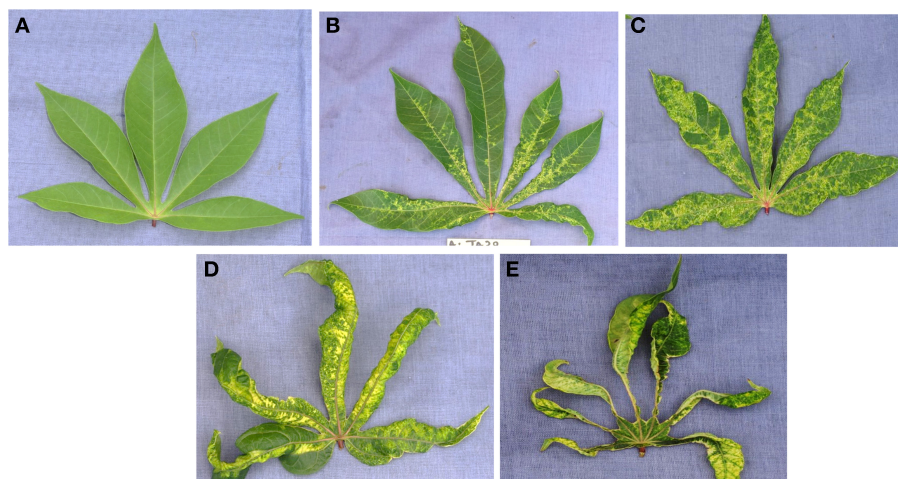| Location	Vegetation	Soil type	Rainfull (mm)	Relative humidity (%)	T° mean (°C)
Bouaké	Wooded savannah	Ferralitic	1,060	79	26.4
Man	Humid forest	Ferralitic and hydromorph	1,636.5	86	26.6

FIGURE 1
Symptoms of cassava mosaic disease observed on infected cassava leaves using a scale from 1 (no symptoms) to 5 (very severe symptoms). **(A)** = 1, **(B)** = 2, **(C)** = 3, **(D)** = 4, **(E)** = 5.

and EACMCMV (Table 2). The PCR reactions were carried out using GoTaq polymerase (Promega). The reaction mixture contained 1X Colorless GoTaq Reaction Buffer (Promega), 0.625 U GoTaq polymerase (Promega), 0.4 μM of each primer (synthesized by Eurogentec), 0.2 mM of dNTP (NEB), 1 mM MgCl$_2$ (Promega). The reaction consisted of an initial denaturation step at 94°C for 4 min, followed by 35 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min and a final extension of 72°C for 10 min. The PCR products (10 μl) were subjected to 1% agarose gel electrophoresis, stained with ethidium bromide, and visualized under UV light using a gel imager.

2.5. Classification of symptomless accessions into potentially resistant and tolerant accessions

The classification of potentially resistant and tolerant accessions to CMD was carried out based on the absence of symptoms on the accessions in both sites and the absence or presence of CMBs in the accession. Thus, an accession is tolerant to CMD when it is symptomless in both locations (Bouaké and Man) but is infected by CMB(s). An accession is potentially resistant to CMD when it is symptomless to CMD in both site and is virus free.

2.6. PCR detection of CMD resistance genes

Detection of CMD resistance genes was carried out on accessions which were symptomless in both locations. PCR reactions were performed with 4 pairs of primers (Table 3) to amplify CMD resistance genes (CMD1, CMD2 and CMD3). The reaction mixture

contained 1X Colorless GoTaq Reaction Buffer (Promega), 0.625 U GoTaq polymerase (Promega), 0.2 μM each primer (synthesized by Eurogentec), 0.2 mM of dNTP (NEB), 1 mM MgCl$_2$ (Promega), 100 ng of DNA in the mix final volume 25 μl. Previously characterized cassava accessions, TMS30572 and TMS98/0581, were used as positive controls for resistance to CMD2 while TMS97/2205 and TMS30572 were used as positive control for resistance to CMD3 and CMD1, respectively. Yacé kodjohon which is a local and highly susceptible accession to CMD, was used as a negative control (absence of resistance genes). The PCR amplification was run with the following conditions: 4 min at 94°C and 35 cycles of 30 s at 94°C, 1 min at 55°C and 1 min at 72°C followed by a final extension of 10 min at 72°C for the markers NS169, NS198, SSY28; and 4 min at 94°C, 35 cycles of 30 s at 94°C, 1 min at 50°C and 1 min at 72°C followed by a final extension at 72°C for 5 min for the marker RME1. The PCR products (10 μl) were analyzed by electrophoresis in a 3% agarose gel. After the migration, the gel was stained with ethidium bromide, and the electrophoretic profiles visualized under UV light using a gel imager.

2.7. Genotyping with SNPs markers

A sub-set of 36 single nucleotide polymorphism (SNP) markers were selected from Expressed Sequence Tag (EST) derived SNPs by Ferguson et al. (2012, 2019) and converted to Kompetitive allele specific PCR (KASP) primers (LGC Biosearch technologies, UK) as a cost-effective method to use in varietal identification and quality control. SNP markers were selected based on position and Polymorphic Information Content (PIC) value above 0.365 within the East African cassava germplasm to genotype the symptomless accessions.

TABLE 2 PCR primer pairs for ACMV, EACMV, and EACMCMV detection.

Virus	Primer	Primer sequence	DNA target	Size/amplified region
ACMV	JSP001	5′-ATG-TCG-AAG-CGA-CCA-GGA-GAT-3′	DNA-A	783 bp/CP
	JSP002	5′-TGT-TTA-TTA-ATT-GCC-AAT-ACT-3′		
	ACMVB1	5′-TCG-GGA-GTG-ATA-CAT-GCG-AAG-GC-3′	DNA-B	628 bp/BV1-BC1
	ACMVB2	5′-GGC-TAC-ACC-AGC-TAC-CTG-AAG-CT-3′		
EACMV	JSP001	5′-ATG-TCG-AAG-CGA-CCA-GGA-GAT-3′	DNA-A	780 bp/CP
	JSP003	5′-CCT-TTA-TTA-ATT-TGT-CAC-TGC-3′		
	EAB555F	5′-TAC-ATC-GGC-CTT-TGA-GTC-GCA-TGG-3′	DNA-B	544-560 bp/IR-BC1
	EAB555R	5′-CTT-ATT-AAC-GCC-TAT-ATA-AAC-ACC-3′		
EACMCMV	VNF031	5′ -GGA-TAC-AGA-TAG-GGT-TCC-CAC-3′	DNA-A	560 bp/AC2-AC3
	VNF032	5′-GAC-GAG-GAC-AAG-AAT-TCC-AAT-3′		

TABLE 3 PCR primers for the detection of CMD resistance markers.

Resistance genes	Primers	Marker systems	Sequences	Size (bp)
CMD1	SSRY40-F	SSR	5′-TGC-ATC-ATG-GTC-CAC-TCA-CT-3′	293
	SSRY40-R		5′-CAT-TCT-TTT-TCG-GCA-TTC-CAT-3′	
CMD2	NS169-F	SSR	5′-GTG-CGA-AAT-GGA-AAT-CAA-TG-3′	319
	NS169-R		5′-GCC-TTC-TCA-GCA-TAT-GGA-GC3′	
	RME1-F	SCAR	5′-ATG-TTA-ATG-TAA-TGA-AAG-AGC-3′	700
	RME1-R		5′-AGA-AGA-GGG-TAG-GAG-TTA-TGT-3′	
CMD3	NS198-F	SSR	5′-TGC-AGC-ATA-TCA-GGC-ATT-TC-3′	196
	NS198-R		5′-TGG-AAG-CAT-GCA-TCA-AAT-GT-3′	

SSR, simple sequences repeat; SCAR, Sequences characterized amplified region.

Two cuttings per accession were grown in a pot containing soil, kept for 3 months in a screenhouse to obtain fully expanded young leaves, then the BioArk collection kit was used to sample the leaves according to the LGC protocol.

Total genomic DNA was extracted from plant tissue using LGC's sbeadex TM DNA extraction. Sbeadex is a magnetic bead-based extraction technique that uses a two-step binding mechanism to allow for tight DNA binding, and a final wash with pure water to give a high level of quality and purity. Genotyping of the 36 markers was performed using the competitive allele-specific PCR (KASP™) system. KASP genotyping assays are based on competitive allele-specific PCR and enable bi-allelic scoring of SNPs and insertions/deletions at specific loci.

The raw data generated were analyzed and scored on a Cartesian plot, to assign a genotype to each DNA sample using LGC's proprietary Kraken software. Results of genotyping were presented as homozygotes (A:A, C:C, G:G and T:T) and heterozygotes (A:T, A:C, A:G, C:A, C:T, C:G, and G:T). One accession, Bocou1, was duplicated to serve as a control for identification of duplicate accessions.
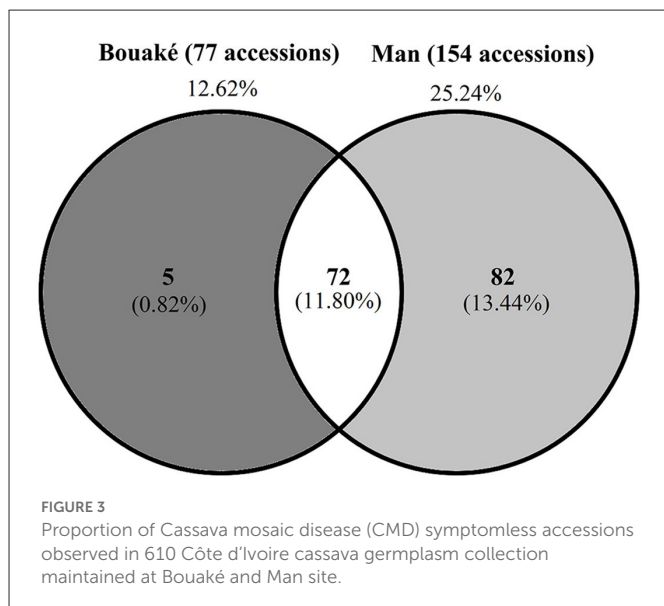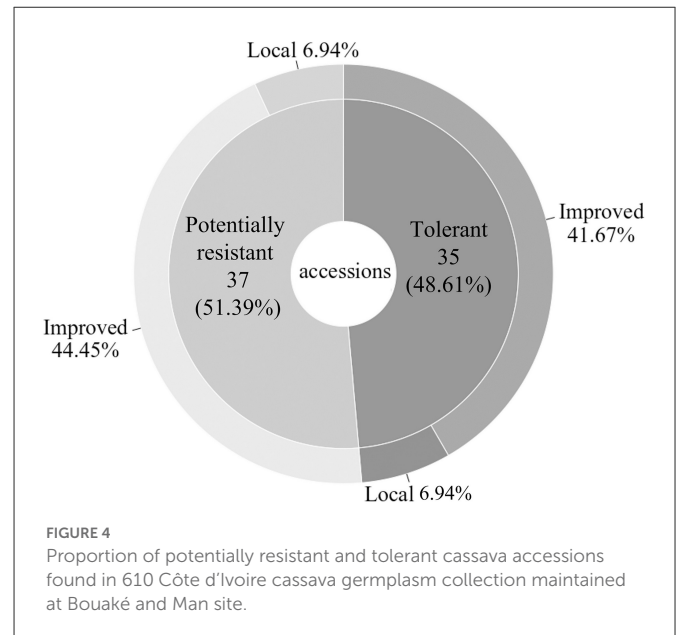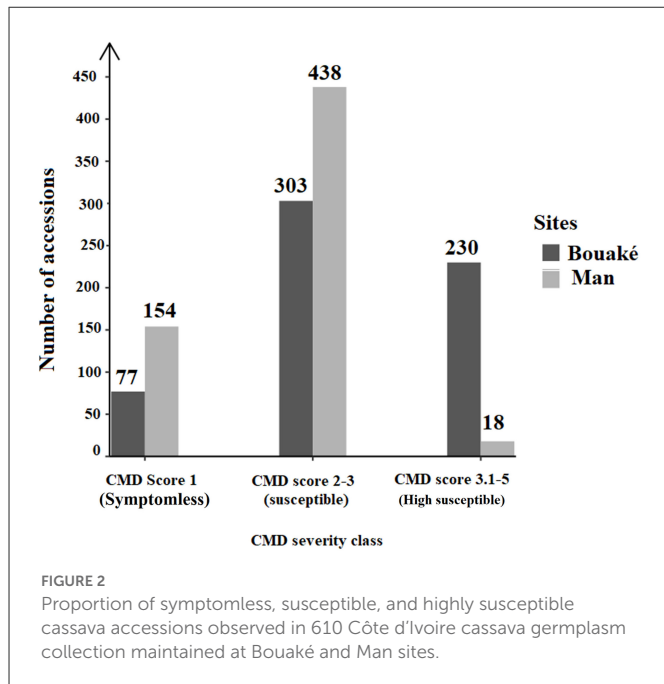
2.8. Statistical analyses

Missing data percentages for SNP markers and accessions were determined using the *missingno* function in the *poppr* package version 2.9.2 implemented in R software version 4.0.2. Markers and accessions with more than 6% missing data were removed from the initial matrix to avoid bias in the results. The genotype accumulation curve was performed using the genotype curve function from the *poppr* package implemented in R to ensure that the number of remaining markers was sufficient to discriminate cassava accessions.

All the genetic diversity parameters and F-statistics were calculated with the *HierFstat* package 0.04-22 version (De Meeûs and Goudet, 2008) implemented in R version 4.0.2, except for PIC which was calculated using *PICcalc* (Nagy et al., 2012). Hardy-Weinberg equilibrium (HWE) for each locus was checked by the Chi2 test at one degree of freedom (ddl = 1). Chi^2 values were calculated using the ADEGENET package (Jombart et al., 2010) in R software version 4.0.2.

Analysis of genetic structure was performed using Ascendant Hierarchical Clustering, Discriminant Analysis of Principal Components (DAPC) and molecular variance (AMOVA). A Ward's minimum variance hierarchical clustering dendrogram was constructed from the genetic distance using the *plot.phylog* algorithm from the *Ade4* package as implemented in R version 4.0.2. The critical distance threshold to declare whether two accessions (varieties) are identical or not was based on the genetic distance between two representatives of the same accessions (duplicated previously for genotyping). Any two accessions whose genetic distance was below 0.05 (dissimilarity coefficient, Ward's distance) were considered of the same genotype.

FIGURE 2
Proportion of symptomless, susceptible, and highly susceptible cassava accessions observed in 610 Côte d'Ivoire cassava germplasm collection maintained at Bouaké and Man sites.

FIGURE 4
Proportion of potentially resistant and tolerant cassava accessions found in 610 Côte d'Ivoire cassava germplasm collection maintained at Bouaké and Man site.

FIGURE 3
Proportion of Cassava mosaic disease (CMD) symptomless accessions observed in 610 Côte d'Ivoire cassava germplasm collection maintained at Bouaké and Man site.

DAPC was performed using the *Adegenet* package (Jombart et al., 2010) implemented in R version 4.0.2. The best number of clusters was determined by the find.cluster function. This function performs successive K-means using an increasing number of clusters (k), after transforming the data using a principal component analysis (PCA). For each model, a statistical measure of the quality of the Bayesian information criterion (BIC) was calculated, which allowed the selection of the best number of groups.

Analysis of molecular variance (AMOVA) was performed to assess the distribution of genetic variation among accessions using the *Poppr* package (Kamvar et al., 2021) implemented in R version 3.3.3. Plots were made with the ggplot2 package (Wickham, 2016).

3. Results

3.1. Classification of the accessions according to their level of resistance to cassava mosaic disease

Based on mean CMD severity calculations, we found that in Bouaké, 77 accessions (12.62%) were symptomless, 303 (49.67%) susceptible (S) and 230 (37.70%) highly susceptible (HS) accessions (Figure 2). In the Man site, 154 (25.25%) accessions were symptomless, 438 (71.80%) were susceptible (S) and 18 (28.51%) were highly susceptible (Figure 2). Compared to the Bouaké site, the Man site had the highest number of symptomless accessions. When we considered the symptomless accessions in Bouaké and Man, we found that 72 accessions are symptomless in both locations (Figure 3). We also checked the 72 accessions in the preservation plots planted 2 years before and found that they were indeed symptomless. These 72 symptomless accessions have been used for subsequent studies. Ten accessions out of the 72 symptomless accessions are local varieties and 62 are improved varieties (Supplementary Table 1). Fifty-seven of the 72 symptomless accessions are from Nigeria (IITA), 2 from Kenya and 13 from Côte d'Ivoire.

3.2. Detection of CMBs in the 72 symptomless accessions

Each of the 8 to 10 plants visually evaluated for the 72 symptomless accessions was tested by PCR using the primers for the detection of ACMV, EACMV and EACMCMV. The results showed that 35 (48.61%) symptomless accessions were infected by ACMV and are considered tolerant accessions. Thirty-seven accessions (51.39%) were virus free and are considered potentially resistant accessions (Figure 4). Surprisingly,

EACMCMV was not detected in any of the tolerant accessions (Supplementary Table 1). Out of the 37 accessions potentially resistant, five are local varieties (Bonoua34, Agba kangba3, AY14 (2), AY4 and N°35) and 32 are improved varieties. Regarding the 35 tolerant accessions, five are local varieties (Bahanin bou, Bou fouh4, Koko3, N°37 and 43) and 30 are improved varieties (Figure 4).

3.3. Detection of CMD resistance genes in the 72 symptomless accessions

Using four different markers, the 72 symptomless accessions were screened for CMD resistance genes. The markers NS169F/R and RME1F/R were used for the detection of the CMD2 resistance gene while the marker NS198F/R was used for the CMD3 resistance gene and the marker SSYR40F/R for the CMD1 resistance gene. PCR results showed that two accessions (2.78%) did not have

any CMD resistance gene, and 70 accessions (97.22%) had at least one CMD resistance gene (Table 4). The CMD2 resistance gene (RME1 and NS169), CMD1 (SSY40), and CMD3 (NS198) were all detected in 54 accessions (75%). Also, 13 accessions (18.05%) contain the three resistant genes, CMD2 (NS169), CMD1 (SSY40) and CMD3 (NS168). In addition, CMD1 (SSY40) and CMD3 (NS198) resistant gene were identified in 2 accessions (Bocou3 and TMS98/0002) and CMD2 (RME1 and NS169) in 1 accession (TMS99/0554) (Table 4). These three accessions belong to the group of tolerant accessions. Also, the two symptomless accessions without CMD resistance genes are local accessions [AY14(2) and AY4] and belong to the group of potentially resistant accessions. In conclusion, we observe that most of the symptomless accessions (potentially resistant and tolerant) have resistance genes for CMD. Data for individual accessions are reported in Supplementary Table 1. The genotypic coincidence, relative to the presence of markers linked to the CMD2 gene (NS169 × RME1), was higher (0.76).

TABLE 4 Combinations of CMD resistance markers found in 72 CMD symptomless cassava accessions in Côte d'Ivoire.

Markers	Number of accessions and detection percentage	
	Total	%
CMD2 (RME1+ NS169) + CMD1 (SSY40) + CMD3 (NS198)	54	75
CMD2 (NS169) + CMD1(SSY40) + CMD3 (NS168)	13	18.05
CMD1 (SSY40) + CMD3 (NS198)	2	2.78
CMD2 (NS169 + RME1)	1	1.39
No markers	2	2.78
Total	72	100%

FIGURE 5
Genotype accumulation curve showing the minimum number of SNPs (8) needed to differentiate all unique genotypes.
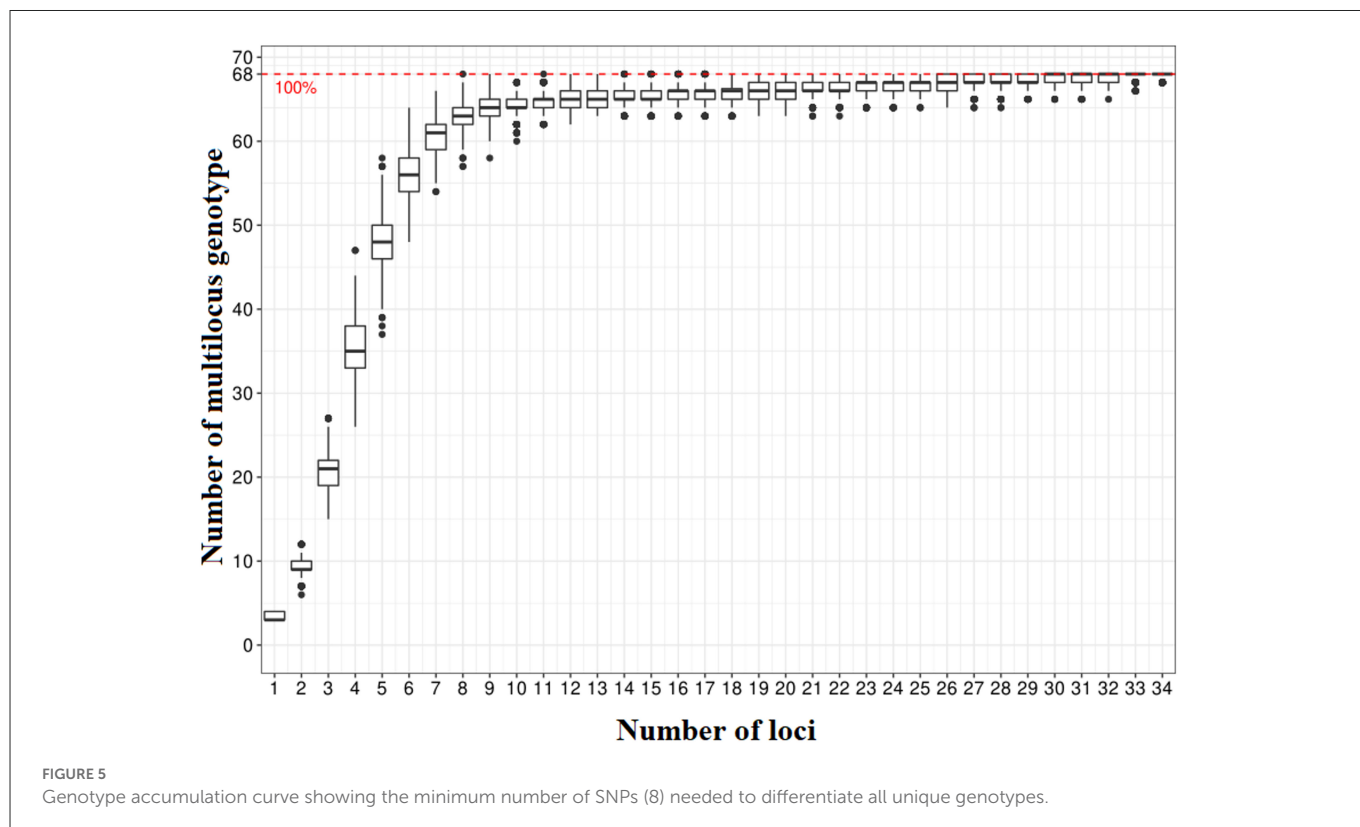
TABLE 5 Common genetic parameters and F-statistics for each locus of the 36 single nucleotide polymorphism (SNP) markers used for the genotyping selected Côte d'Ivoire cassava germplasm.

Markers	MaF	PIC	Ho	He	EHW test		Fis	Fst	Fit
					Chi^2	P-value			
Me_MEF_c_0556	0.28	0.32	0.41	0.4	0.02	0.89	−0.026	0.022	−0.004
Me_MEF_c_0566	0.3	0.33	0.41	0.42	0.12	0.73	−0.029	0.059	0.032
Me_MEF_c_0587	0.46	0.37	0.51	0.5	0.034	0.85	−0.068	0.047	−0.018
Me_MEF_c_0869	0.49	0.37	0.42	0.5	1.738	0.19	0.129	0.036	0.16
Me_MEF_c_0936	0.49	0.37	0.62	0.5	4.224	0.04*	−0.26	0.017	−0.239
Me_MEF_c_0979	0.48	0.37	0.55	0.5	0.74	0.39	−0.182	0.059	−0.112
Me_MEF_c_0981	0.29	0.33	0.46	0.41	1.105	0.29	−0.202	0.076	−0.111
Me_MEF_c_1018	0.46	0.37	0.46	0.5	0.18	0.67	0.053	0.024	0.076
Me_MEF_c_0363	0.46	0.37	0.52	0.5	0.306	0.58	−0.073	0.03	−0.04
Me_MEF_c_1074	0.3	0.33	0.26	0.42	10.171	0.00***	0.29	0.118	0.374
Me_MEF_c_1081	0.36	0.36	0.43	0.46	0.241	0.62	−0.038	0.096	0.061
Me_MEF_c_1179	0.46	0.37	0.65	0.5	6.816	0.01**	−0.321	0.012	−0.305
Me_MEF_c_1186	0.46	0.37	0.62	0.5	5.015	0.03*	−0.281	0.039	−0.231
Me_MEF_c_0153	0.49	0.37	0.49	0.5	0	0.10	−0.05	0.061	0.014
Me_MEF_c_1187	0.42	0.37	0.57	0.49	2.152	0.14	−0.254	0.091	−0.139
Me_MEF_c_3217	0.49	0.37	0.48	0.5	0.126	0.72	−0.024	0.071	0.049
Me_MEF_c_0262	0.38	0.36	0.48	0.47	0.008	0.93	−0.106	0.085	−0.012
Me_MEF_c_2368	0.4	0.36	0.46	0.48	0.02	0.89	0.018	0.032	0.05
Me_MEF_c_1361	0.44	0.37	0.64	0.49	6.631	0.01**	−0.314	0.011	−0.301
Me_MEF_c_1418	0.46	0.37	0.52	0.5	0.306	0.58	−0.222	0.151	−0.037
Me_MEF_c_2268	0.43	0.37	0.26	0.49	15.192	0.00***	0.162	0.357	0.461
Me_MEF_c_3025	0.39	0.36	0.49	0.48	0.082	0.77	−0.059	0.039	−0.018
Me_MEF_c_1568	0.4	0.37	0.54	0.48	1.141	0.29	−0.157	0.032	−0.119
Me_MEF_c_1585	0.15	0.22	0.29	0.25	2.062	0.15	−0.234	0.043	−0.181
Me_MEF_c_1671	0.39	0.36	0.49	0.48	0.082	0.77	−0.078	0.042	−0.033
Me_MEF_c_0227	0.36	0.35	0.51	0.46	0.798	0.37	−0.106	0.009	−0.096
Me_MEF_c_2177	0.32	0.34	0.52	0.43	2.792	0.10	−0.279	0.056	−0.208
Me_MEF_c_2297	0.26	0.31	0.39	0.38	0.102	0.75	−0.031	0.006	−0.024
Me_MEF_c_2515	0.42	0.37	0.49	0.49	0.009	0.93	−0.029	0.023	−0.005
Me_MEF_c_0284	0.49	0.37	0.61	0.5	4.323	0.04*	−0.314	0.081	−0.208
Me_MEF_c_2574	0.48	0.37	0.49	0.5	0.011	0.91	−0.271	0.231	0.022
Me_MEF_c_2644	0.38	0.36	0.58	0.47	3.789	0.05*	−0.256	0.021	−0.229
Me_MEF_c_2911	0.22	0.29	0.33	0.35	0.128	0.72	0.022	0.02	0.042
Me_MEF_c_3142	0.41	0.37	0.54	0.48	0.774	0.38	−0.147	0.034	−0.108
Me_MEF_c_0126	0.49	0.37	0.49	0.5	0	0.10	−0.077	0.089	0.019
Mean	**0.40**	**0.35*****	**0.49*****	**0.47*****	**-**	–	**−0.11*****	**0.06*****	**−0.04*****

MaF, major allele frequency; He, expected heterozygosity; Ho, observed heterozygosity; HWE, Hardy–Weinberg equilibrium.

Chi2 values of the test HWE; ddl = 1.

P-value: *P < 0.05; **P < 0.01; ***P < 0.001. F_{IT} = inbreeding coefficient of an individual into the whole population, F_{IS} = within-population inbreeding coefficient, FST = coefficient of differentiation, and PIC = polymorphic information content.

3.4. Analysis of the genetic diversity of the symptomless accessions using a set of 36 SNP markers

Markers and accessions with more than 6% missing data were removed from the initial matrix, leaving a final matrix consisting of 35 SNP markers and 69 accessions. The genotype accumulation curve based on multilocus genotype (MLGs) revealed that the minimum number of SNPs needed to differentiate the cassava accessions is eight (Figure 5). The major allele frequency (MaF) of each of the 35 loci selected for analyses was <0.95 (Table 5). All the 35 markers were therefore polymorphic in the study population. The common genetic parameters and genetic differentiation parameters estimated for each marker are reported in Table 5. The major allele frequency (MaF) value ranged from 0.22 to 0.49 with a mean of 0.4. The polymorphic information content (PIC) value ranged from 0.22 to 0.37 with a mean of 0.35. All markers had PIC $\geq$ 0.30, excluding Me.MEF. c.1585 with PIC = 0.23. The values of Ho and He ranged significantly from one marker to another. *He* varied from 0.25 to 0.50, with an average of 0.47. In contrast, *Ho* ranged from 0.26 to 0.58, with and an average of 0.49. The HWE analysis showed that, for 8 SNP markers, the rate of *Ho* was significantly different from that of *He* (Table 5). For two of them (Me_MEF.c.2268; Me_MEF_c.1074), this difference was highly significant ($P < 0.001$, Table 5). These two SNP markers showed a deficit of heterozygote (He>Ho), unlike the other six SNP markers where the heterozygote rate was high (Ho>He). The 69 symptomless accessions had a high heterozygote rate in the population (Fis = −0.11; $P = 0.001$). In addition, genetic differentiation between groups considering all 35 loci was low (Fst = 0.06; $P = 0.001$).

3.5. Population structure of the symptomless accessions by ascendant hierarchical clustering (AHC), discriminant analysis of principal components (DAPC), and identification of the potential duplicates

The ascendant hierarchical clustering was used to classify the accessions into three groups using a single dissimilarity coefficient of approximately 1.2 (Figure 6). Out of the 69 cassava accessions, 56 (81.16%) unique Multilocus Genotypes (MLGs) were identified. Based on their genetic distances of <0.05, the remaining 13 (18.84) accessions were potentially duplicates divided into one trio and five duos (Figure 6). The Bayesian Information Criterion (BIC) was used to distinguish an optimum number of three (03) groups in the cassava accessions (Figure 7A). The coefficient of membership of all accessions to each group varied from 80 to 100% (Figure 7B). The projection of the groups, on two axes, showed that they are all separated from each other (Figure 7C). Axis 1 expressed 64.2% of the total variability, while axis 2 expressed 35.80% of the total variability. The markers Me.MEF.c.1074 (0.075) and Me.MEF.c.2268 (0.15) contributed most to the structuring of the diversity of the cassava accessions along axis 1. For axis 2, the greatest contribution was due to the Me.MEF.c.1187 markers (0.06).

3.6. Analysis of molecular variance (AMOVA)

Analysis of the molecular variance of the 69 symptomless cassava accessions based on accession type (Improved or local), geographic origin, AHC groups and DAPC groups showed that the most significant differences in molecular variance of SNPs existed within individuals and not within the groups. This molecular variance between individuals ranged from 84.17 to 97.31% (Table 6). As for the molecular variance that existed between groups, it was very low and varied from 2.69% to 15.83% (Table 6).

4. Discussion

A plant is considered tolerant when it can endure an infection by a particular pathogen (in this case, a virus), without presenting symptoms and without showing serious disease (Cooper and Jones, 1984; Bos and Parlevliet, 1995). The tolerant plant will therefore be infected by the pathogen without showing any symptoms, sometimes due to a latent infection. According to Paudel and Sanfaçon (2018), tolerance can be described as a stable equilibrium between the virus and its host, an interaction in which each partner not only accommodate trade-offs for survival but also receive some benefits (e.g., protection of the plant against super-infection by virulent viruses; virus invasion of meristem tissues allowing vertical transmission). A plant is considered resistant when it prevents or limits the development or replication of the virus (Paudel and Sanfaçon, 2018). Thus, a diagnosis limited to the visual evaluation of symptoms on a plant is insufficient to determine its status (resistant, tolerant, or susceptible). This explains the approach used in our study which consisted in selecting symptomless accessions based on visual evaluation, then using molecular analysis we determined the real status of each evaluated accession. An assessment of CMD symptoms' severity was conducted on 610 cassava accessions planted in open fields in Bouaké and Man. We found that several accessions expressed severe and very severe CMD symptoms in both Bouaké and man and 72 accessions were symptomless in both locations. Based on this, we were able to classify the accessions into three groups, susceptible (S), Highly susceptible (HS), and symptomless accession respectively.

The symptomless accessions could be potentially CMD resistant accessions. Indeed, 86% (62) of the symptomless accessions are improved varieties with resistance to CMD and come from IITA-Nigeria, Kenya and Côte d'Ivoire. The remaining 10 potentially resistant accessions are local varieties from Côte d'Ivoire. This suggests that several local varieties with natural ability to resist CMD may exist in farmers' fields. This finding supports the report from Fauquet and Fargette (1990) that showed that some local varieties appear to be naturally resistant to CMD. However, molecular analyses detected ACMV in five of the local and 30 of the improved symptomless accessions. Consequently, the absence of symptoms on the leaves does not exclude the presence of virus infection. Indeed, as suggested by Soko et al. (2015) the expression of plant symptoms is a function of virus accumulation. Thus, the accessions infected with ACMV are healthy carriers (tolerant varieties) and could therefore constitute an important plant reservoir for ACMV, contributing to spreading the disease. Such tolerant varieties should not be used as a solution for the management and/or control CMD. For that reason, we suggest regular sanitation by tissue culture and routine
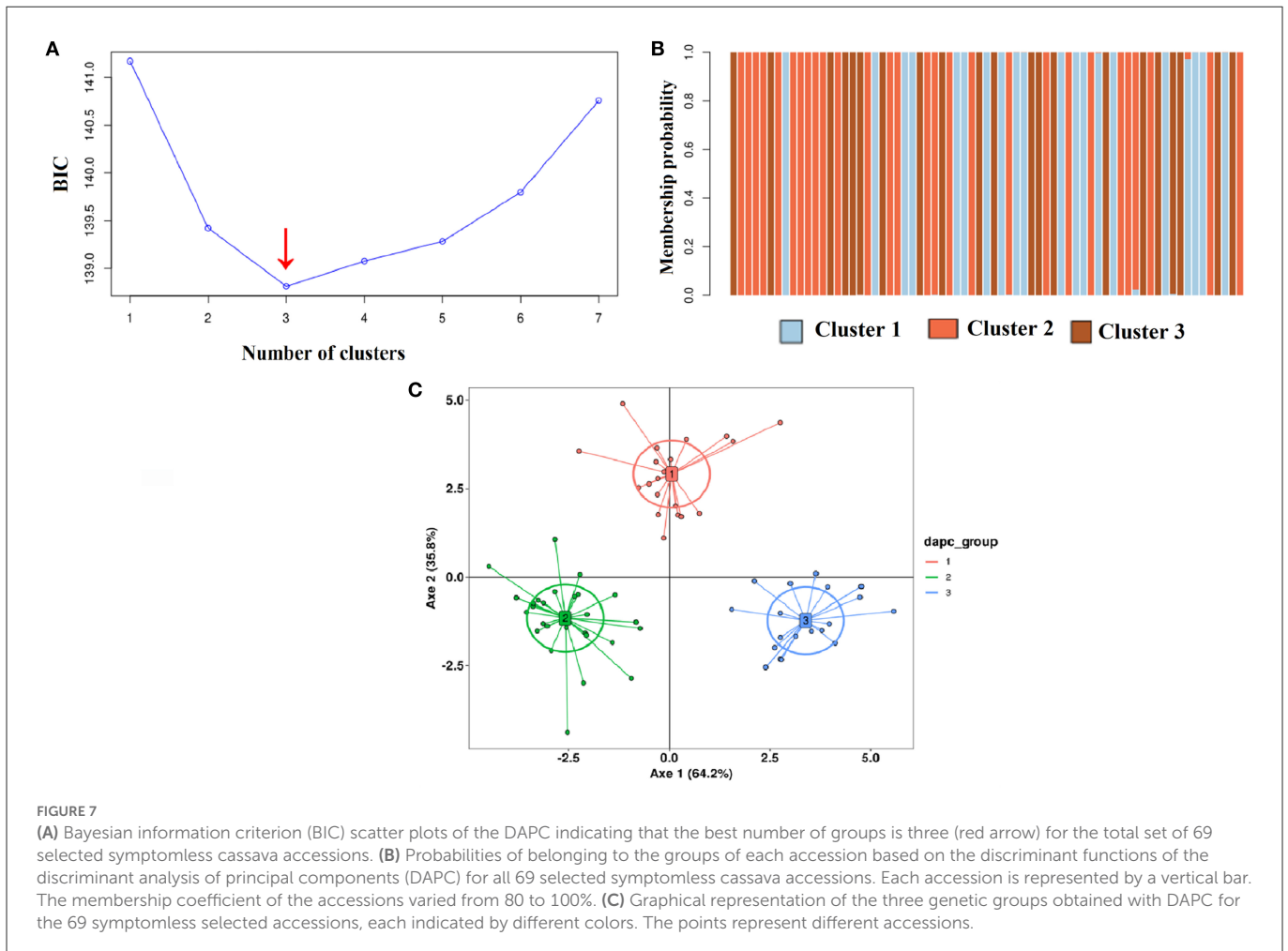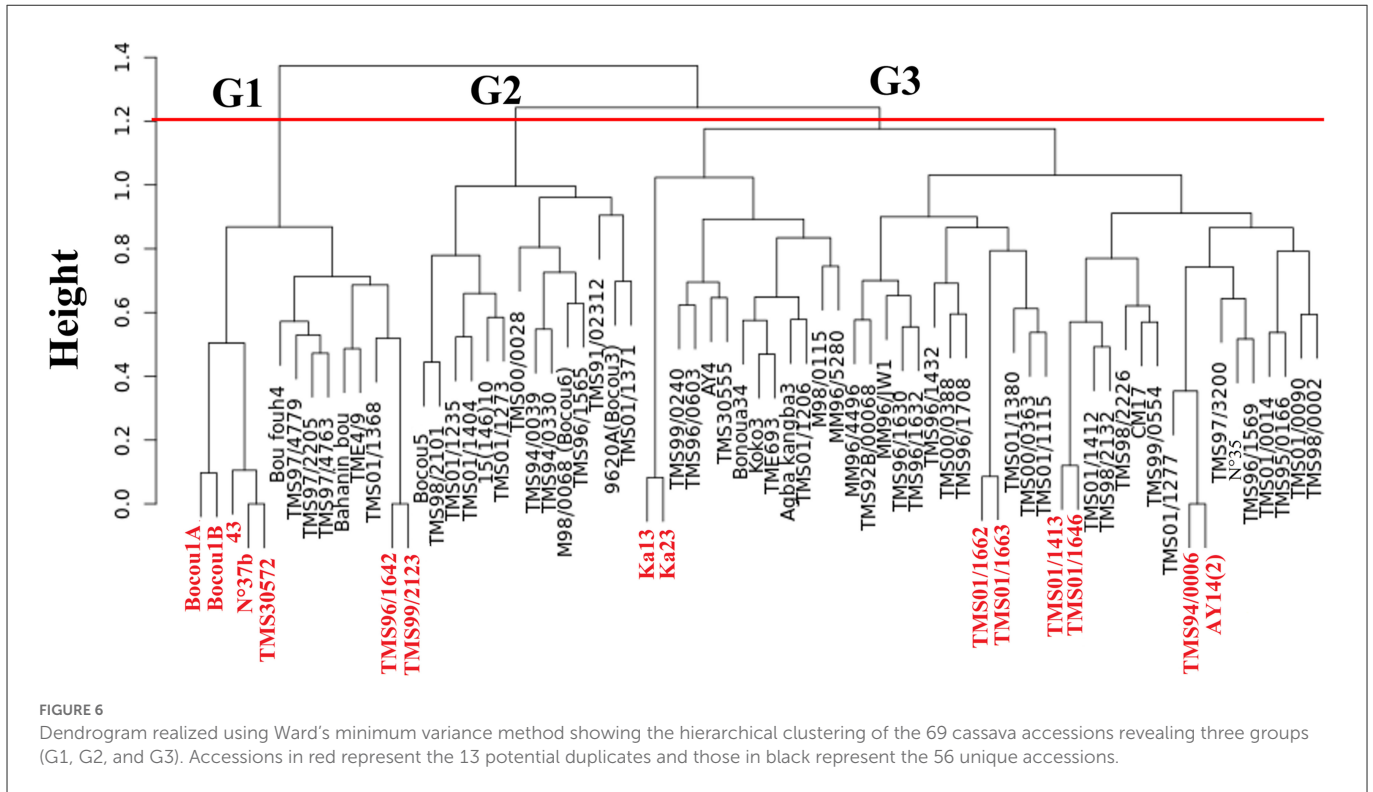
FIGURE 6
Dendrogram realized using Ward's minimum variance method showing the hierarchical clustering of the 69 cassava accessions revealing three groups (G1, G2, and G3). Accessions in red represent the 13 potential duplicates and those in black represent the 56 unique accessions.

FIGURE 7
(A) Bayesian information criterion (BIC) scatter plots of the DAPC indicating that the best number of groups is three (red arrow) for the total set of 69 selected symptomless cassava accessions. (B) Probabilities of belonging to the groups of each accession based on the discriminant functions of the discriminant analysis of principal components (DAPC) for all 69 selected symptomless cassava accessions. Each accession is represented by a vertical bar. The membership coefficient of the accessions varied from 80 to 100%. (C) Graphical representation of the three genetic groups obtained with DAPC for the 69 symptomless selected accessions, each indicated by different colors. The points represent different accessions.

TABLE 6 Analysis of molecular variance (AMOVA) for 69 individual cassava mosaic disease (CMD) tolerant or resistant Ivorian cassava accessions based on SNP markers.

Source of variation	Improved and local accessions		
	df	Mean Sq	% of variation
Between groups	1	11.29	2.69
Within individuals	67	7.88	97.31
Total	68	7.93	100.00

Source of variation	Geographical origins		
	df	Mean Sq	% of variation
Between groups	2	16.12	8.32
Within individuals	66	7.68	91.68
Total	68	7.93	100.00

Source of variation	HAC groups		
	df	Mean Sq	% of variation
Between groups	2	24.53	9.40
Within individuals	66	7.72	90.60
Total	68	8.21	100.00

Source of variation	DAPC groups		
	df	Mean Sq	% of variation
Between groups	2	37.15	15.83
Within individuals	66	7.04	84.17
Total	68	7.93	100.00

Two groups according to improved and local accessions, three groups according to geographic origins, three groups according to HAC and three groups according to DAPC.

virus indexing of these varieties if they are to be used for breeding for other important traits like high yield, processing suitability, etc. In addition, we found that the EACMV, which causes more severe symptoms than the ACMV, was not detected in the 72 symptomless accessions at either site (Bouaké or Man). These results suggest that these accessions could be resistant to the EACMV. In fact, we were able to identify the *East African cassava mosaic Cameroon virus (EACMCMV)* species in susceptible and highly susceptible accessions (unpublished data).

To establish a correlation between the symptomless accessions and the presence of resistance genes, we proceeded to the molecular detection of known CMD resistance genes in the 72 symptomless accessions. Specific markers to three resistance genes, CMD1 (polygenetic recessive), CMD2 (monogenetic dominant), and CMD3 (Quantitative trait loci) were used for molecular screening. Results showed that 97.22% accessions possess at least one CMD resistance genes. This high percentage (97.22%) of accessions carrying resistance genes can be explained by the fact that the most of these accessions have been improved for CMD resistance. This is evident in the fact that 79.16% (57) of these accessions came from IITA-Nigeria. Indeed, the first CMD resistant varieties developed by IITA in Nigeria are hybrids obtained from *M. glaziovii*, found in Brazil, which conferred multigenic resistance to the improved varieties developed. Some of these resistant varieties were introduced in Uganda in the 1980s and formed the basis of the cassava breeding program. Our results are similar to those of Kuria et al. (2017) who found resistance genes in cassava genotypes that did not show CMD symptoms. Curiously, the three resistance genes that were tested for were not detected in two local symptomless accessions [AY14(2) and AY4]. This suggests the existence of another source of resistance in these two accessions, which is to be determined.

To uncover the relationship between accessions AY14(2), AY4 and the other 70 symptomless accessions, we conducted a genetic diversity study on the 72 symptomless accessions, using a set of 36 SNPs markers. However, due to missing data, we were able to study only 69 accessions with 35 markers. Our results revealed a polymorphism rate of 100% for almost all the markers used, except for the Me_MEF_c_1585 marker with a PIC = 0.22, thus less informative according to the classification of Botstein et al. (1980). These results indicate that these markers are reliable for studying the diversity of cassava genetic resources. Indeed, these markers were validated in a study by Ferguson et al. (2012). Observed heterozygosity values ranged from 0.26 to 0.58 with an average of 0.49, while expected heterozygosity (He) values ranged from 0.25 to 0.50 with an average of 0.47. These results indicate high diversity within the 69 accessions studied. This was confirmed by the mean Fis value (-0.11) which indicated an excess of heterozygote in the subpopulation. In addition, the relatively low Fst value (0.05) indicated little genetic differentiation between subpopulations. Therefore, much of the genetic variability within accessions is explained by variation within individuals. Such observations could be partly related to the presence of improved varieties obtained from multiple crosses conducted by IITA and CNRA, and due to natural hybridizations, that occur in fields. Indeed, plants resulting from these natural hybridizations are often selected by farmers if they appear vigorous (Kizito, 2006). Through this action, they indirectly select genotypes that contribute to increase genetic variability in the field (Racchi et al., 2014). We identified 56 unique genotypes

and 13 potential duplicates in this study. This should be confirmed using a higher density genotyping. Our findings suggest the possible existence of the same cultivar under several different accession numbers (duplicates) within the Côte d'Ivoire cassava germplasm maintained at CNRA. The two methods (AHC and DAPC) of population structuring used allowed us to group the 69 accessions into three groups. The dendrogram allowed us to efficiently classify the accessions according to the genetic distance between them and to highlight the potentially duplicated accessions. Indeed, knowledge of genetic proximity is important for genetic crosses to maximize efficient hybridization (Sawler et al., 2013). The results of the dendogram also showed that the two accessions AY14 and AY14(2) are different accessions but belong to the same group (G3). However, the AY14(2) accession seems to be a duplicate of the TMS94/006 accession. But the number of markers used in this study is not sufficient to affirm this. We need to use more markers to know the real status of these two accessions.

Also, the contribution of alleles, to the structuring of accessions, identified by DAPC allows the identification of genomic regions responsible for genetic divergence between the three groups (De Meeûs and Goudet, 2008). However, AMOVA showed that the greatest molecular variation was present at the level of the individuals themselves and not the groups. These results show that the subpopulations were not clearly structured and did not vary from each other.

Finally, the resistant cultivars identified in this work can help farmers to select resistant cultivars and ultimately improve CMD management. These cultivars may also be prioritized for breeding programs to improve cassava against CMD.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ebi.ac.uk/biostudies/studies/S-BSST912?key=dbd75c6d-4732-4030-8755-44ddcbfac747.

Author contributions

JP, JM, NK, and RS initiated and designed the study. JP mobilize the fund for the research. MC, DO, EY, and WA collected samples and data and wrote the manuscript. WA, EY, and DO analyzed data.

JM, NK, RS, BN'Z, and JP reviewed the manuscript. All authors read, corrected, and approved the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fsufs.2022.1052437/full#supplementary-material

SUPPLEMENTARY TABLE 1
Molecular characterization of begomoviruses and detection of Casava mosaic disease (CMD) resistance markers in 72 CMD symptomless cassava accessions observed among 610 Côte d'Ivoire cassava germplasm collection maintained at Bouaké and Man site.

References

Akano, O. A., Dixon, A. G. O., Mba, C., Barrera, E., and Fregene, M. (2002). Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theor Appl Genet.* 105, 521–525. doi: 10.1007/s00122-002-0891-7

Bos, L., and Parlevliet, J. E. (1995). Concepts and terminology on plant/pest relationships: Toward consensus in plant pathology and crop protection. *Annu. Rev. Phytopathol.* 33, 69–102. doi: 10.1146/annurev.py.33.090195.000441

Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32, 314–331.

Cooper, J. I., and Jones, A. T. (1984). The use of terms for responses of plants to viruses: a reply to recent proposals. *Phytopathology.* 74, 379. doi: 10.1094/Phyto-74-379

De Meeûs, T., and Goudet, J. (2008). A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. *Infect Genet Evol.* 7, 731–735. doi: 10.1016/j.meegid.2007.07.005

Dixon, A., Hanna, R., Legg, J., Nzola-meso, M., and Toko, M. (2000). *Trip Report of an IITA multidisciplinary Team of Scientists on the Assessment of Cassava Pests and Diseases in Kinshasa and Bas Congo, Democratic Republic of Congo and Recommendations for further action (December 9-15, 2000). Reliefweb. 1-16.* Available online at: https://reliefweb.int/report/democratic-republic-congo/trip-report-iita-multidisciplinary-team-scientists-kinshasa-and-bas. (accessed August 06, 2022).

Doyle, J.eff J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19, 11–15.

FAO. (2014). *FAO Statistical Yearbook 2014 Africa Food and Agriculture*. Rome, Italy: FAO.

FAOSTAT. (2020). Available online at: http://www.fao.org/faostat/fr/#data/QC (accessed September 05, 2022).

Fauquet, C., and Fargette, D. (1990). African cassava mosaic virus: etiology, epidemiology and control. *Plant Dis*. 74, 404–411. doi: 10.1094/PD-74-0404

Ferguson, M. E., Hearne, S. J., Close, T. J., Wanamaker, S., William, A., Moskal., et al. (2012). Identification, validation and high-throughput genotyping of transcribed gene SNPs in cassava. *Theor. Appl. Genet*. 124, 685–695. doi: 10.1007/s00122-011-1739-9

Ferguson, M. E., Id, T. S., Kulakow, P., and Ceballos, H. (2019). A global overview of cassava genetic diversity. *PLoS ONE*. 4, 1–16. doi: 10.1371/journal.pone.0224763

ICTV. (2019). *Virus Taxonomy: 2019 Release*. Berlin, Germany: ICTV.

Imakumbili, M. L. E., Semu, E., Semoka, J. M. R., Abass, A., and Mkamilo, G. (2021). Heliyon Managing cassava growth on nutrient poor soils under different water stress conditions. *Heliyon*. 7, e07331. doi: 10.1016/j.heliyon.2021.e07331

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components : a new method for the analysis of genetically structured populations. *BMC Genetics*. 11, 94. doi: 10.1186/1471-2156-11-94

Kamvar, Z. N., Tabima, J. F., Everhart, S. E., Brooks, J. C., Krueger.-Hadfield, S. A., et al. (2021). *Genetic Analysis of Populations with Mixed Reproduction. CRAN.1-136*. https://cloud.r-project.org/package=poppr. (accessed August 03, 2022).

Kizito, E. B. (2006). Genetic and root growth studies in cassava (*Manihot esculenta* Crantz): implications for breeding. *Acta Univ. Agric. Sueciae*. 82, 1652–6880.

Kuria, P., Ilyas, M., Ateka, E., Miano, D., Onguso, J., Carrington, J. C., et al. (2017). Differential response of cassava genotypes to infection by cassava mosaic geminiviruses. *Virus Res*. 227, 69–81. doi: 10.1016/j.virusres.2016.09.022

Nagy, S., Poczai, P., Cernak, I., Gorji, A. M., Hegedus, G., and Taller, J. (2012). PICcalc : an online program to calculate polymorphic information content for molecular genetic studies. *Biochem Genet*. 50, 670–672. doi: 10.1007/s10528-012-9509-1

Njoroge, M. K., Mutisya, D. L., Miano, D. W., and Kilalo, D. C. (2017). Whitefly species efficiency in transmitting cassava mosaic and brown streak virus diseases. *Cogent Biol*. 3, 1311499. doi: 10.1080/23312025.2017.1311499

Okogbenin, E., Egesi, C. N., Olasanmi, B., Ogundapo, O., Kahya, S., Hurtado, P., et al. (2012). Molecular marker analysis and validation of resistance to cassava mosaic disease in elite cassava genotypes in Nigeria. *Crop Sci*. 52, 2576–2586. doi: 10.2135/cropsci2011.11.0586

Okogbenin, E., and Fregene, M. (2002). Genetic analysis and QTL mapping of early root bulking in an F 1 population of non-inbred parents in cassava (Manihot esculenta Crantz). *Theor Appl Genet*. 106,58–66. doi: 10.1007/s00122-002-1068-0

Patil, B. L., and Fauquet, C. M. (2009). Cassava mosaic geminiviruses : actual knowledge and perspectives. *Mol plant pathol*. 10, 685–701. doi: 10.1111/j.1364-3703.2009.00559.x

Paudel, D. B., and Sanfaçon, H. (2018). Exploring the diversity of mechanisms associated with plant tolerance to virus infection. *Front. Plant Sci*. 871, 1–20. doi: 10.3389/fpls.2018.01575

Rabbi, I. Y., Hamblin, M. T., Lava Kumar, P., Gedil, M. A., Ikpan, A. S., Jannink, J.-L., et al. (2014). High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Res*. 186, 87–96. doi: 10.1016/j.virusres.2013.12.028

Racchi, M. L., Bove, A., Turchi, A., Bashir, G., Battaglia, M., and Camussi, A. (2014). Genetic characterization of Libyan date palm resources by microsatellite markers. *Biotech*. 4, 21–32. doi: 10.1007/s13205-013-0116-6

Sawler, J., Reisch, B., Aradhya, M. K., Prins, B., Zhong, G., Simon, C., et al. (2013). Genomics Assisted Ancestry Deconvolution in Grape. *PLoS ONE*. 8, 1–8. doi: 10.1371/journal.pone.0080791

Soko, D. F., Ambroise, L., Siene, C., Kotchi, V., Gogbeu, J., Sere, Y., et al. (2015). Évaluation du niveau de sensibilité et de résistance des variétés de riz différentielles d' AfricaRice à huit isolats de la panachure jaune du riz (RYMV) de Gagnoa (Côte d' Ivoire). *J. Anim. Plant Sci*. 26, 4138–4149.

Soro, M., Tiendrébéogo, F., Pita, J. S., Traor,é, E. T., Som,é, K., Tibiri, E. B., et al. (2021). Epidemiological assessment of cassava mosaic disease in Burkina Faso. *Plant Pathol*. 70, 2207–2216. doi: 10.1111/ppa.13459

Sseruwagi, P., Sserubombwe, W. S., Legg, J. P., Ndunguru, J., and Thresh, J. M. (2004). Methods of surveying the incidence and severity of cassava mosaic disease and whitefly vector populations on cassava in Africa : a review. *Virus Res*. 100, 129–142. doi: 10.1016/j.virusres.2003.12.021

Vernier, P., N'Zué, B., and Zakhia-Rozis, N. (2018). Le manioc, entre culture alimentaire et filière agro-industrielle. *Quæ, CTA, Presses Agronomiques de Gembloux*. 1–208. doi: 10.35690/978-2-7592-2708-2

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York : Springer-Verlag New York. Available online at: http://ggplot2.org (accessed September 05, 2022).