International Journal of
Genetics and Molecular Biology

ACADEMIC JOURNALS
expand your knowledge

About IJGMB

The International Journal of Genetics and Molecular Biology (IJGMB) is a peer reviewed open access journal. The journal commenced publication in April 2009.

Indexing

AgBiotech News and Information, AgBiotechNet, Agroforestry Abstracts, Animal Breeding Abstracts, Animal Production Database, Animal Science, Biofuels Abstracts, Botanical Pesticides, CAB Abstracts, CABI's Global Health Database, Chemical Abstracts (CAS Source Index), CNKI Scholar, Crop Physiology Abstracts, Crop Science Database, Dairy Science Abstracts, Dimensions Database, Environmental Impact, Environmental Science Database, Field Crop Abstracts, Forest Science, Horticultural Science, Horticultural Science Abstracts, Information Matrix for the Analysis of Journals (MIAR), Nutrition Abstracts and Reviews Series B: Livestock Feeds and Feeding, Nutrition and Food Sciences, Parasitology Database, Plant Breeding Abstracts, Plant Genetic Resources Abstracts, Plant Genetics and Breeding Database, Plant Growth Regulator Abstracts, Plant Protection Database, Protozoological Abstracts, ResearchGate Journal Impact, Review of Agricultural Entomology, Review of Aromatic and Medicinal Plants, Review of Plant Pathology, Rice Abstracts, Seed Abstracts, Soil Science Database, Soils and Fertilizers, Abstracts, TROPAG & RURAL, Tropical Diseases Bulletin, Veterinary Bulletin, Veterinary Science Database, VetMed Resource, Weed Abstracts, WorldCat

Open Access Policy

Open Access is a publication model that enables the dissemination of research articles to the global community without any form of restriction. All articles published under open access can be accessed by anyone with internet connection.

The International Journal of Genetics and Molecular Biology is an Open Access journal. Abstracts and full texts of all articles published in this journal are freely accessible to everyone immediately after publication without any form of restriction.

Please see list of all published articles.

Article License

All articles published by the International Journal of Genetics and Molecular Biology are licensed under the Creative Commons Attribution 4.0 International License. This permits anyone to copy, redistribute, remix, transmit and adapt the work provided the original work and source is appropriately cited. Citation should include the article's DOI. The article license is displayed both on the abstract page and the full-text PDF of each article.

Please see more details about Creative Commons Attribution License 4.0

Article Copyright

When an article is published in the journal, the author(s) of the article retain the copyright. Author(s) may republish the article as part of a book or other materials.

A copyright statement is displayed both on the abstract page and the full-text PDF of each article.

Example: Copyright ©2016 Author(s) retains the copyright of this article.

Self-Archiving Policy

The International Journal of Genetics and Molecular Biology is a RoMEO green journal. This permits authors to archive any version of their article they find most suitable, including the published version on their institutional repository and any other suitable website.

Digital Archiving Policy

The International Journal of Genetics and Molecular Biology is committed to the long-term preservation of its content. All articles published by the journal are preserved by Portico. In addition, the journal encourages authors to archive the published version of their articles on their institutional repositories and as well as other appropriate websites

Metadata Harvesting

The International Journal of Genetics and Molecular Biology encourages metadata harvesting of all its content. The journal fully supports the Open Archives Initiative. See **Harvesting Parameter**

Memberships and Standards

OPEN ACCESS

Academic Journals strongly supports the Open Access initiative. Abstracts and full texts of all articles published by Academic Journals are freely accessible to everyone immediately after publication.

creative commons

All articles published by Academic Journals are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). This permits anyone to copy, redistribute, remix, transmit and adapt the work provided the original work and source is appropriately cited.

Crossref

Crossref is an association of scholarly publishers that developed Digital Object Identification (DOI) system for the unique identification published materials. Academic Journals is a member of Crossref and uses the DOI system. All articles published by Academic Journals are issued DOI.

Similarity Check powered by iThenticate is an initiative started by CrossRef to help its members actively engage in efforts to prevent scholarly and professional plagiarism. Academic Journals is a member of Similarity Check.

CrossRef Cited-by Linking (formerly Forward Linking) is a service that allows you to discover how your publications are being cited and to incorporate that information into your online publication platform. Academic Journals is a member of CrossRef Cited-by.

<idpf>

Academic Journals is a member of the International Digital Publishing Forum (IDPF). The IDPF is the global trade and standards organization dedicated to the development and promotion of electronic publishing and content consumption.

Contact

Table of Content

International Journal of Genetics and Molecular Biology

Full Length Research Paper

Morphological diversity patterns among selected elite Shea trees (*Vitellaria paradoxa* C.F. Gaertn.) from Tchologo and Bagoué districts in Northern Côte d'Ivoire

YAO Saraka Didier Martial[1*], DIARRASSOUBA Nafan[1], ATTIKORA Affi Jean Paul[2], FOFANA Inza Jesus[1], DAGO Dougba Noel[1] and SILUE Souleymane[1]

[1]Département Biochimie-Génétique, University of Peleforo Gon Coulibaly (UPGC), UFR Sciences Biologiques, Unité Pédagogique et de Recherche (UPR) de Génétique, BP 1328 Korhogo, Côte d'Ivoire.
[2]University of Liège, Gembloux Agro-Bio-Tech, Passage des Déportés, 2 B-5030 Gembloux, Belgium.

Agromorphological diversity structure of the elite shea trees identified in village lands and conserved *in situ* in the districts of Bagoué and Tchologo by the shea breeding program of the University of Peleforo Gon Coulibaly (UPGC, Côte d'Ivoire), are not known. In the present study, we characterized the agromorphological parameters of 220 elite shea trees using a set of 12 quantitative traits. The results showed that elite shea trees population has been structured into three morphological clusters or genetic pools that do not overlap with the original geographic areas. Morphological Cluster I contain elite shea trees with small trunk diameters carrying large leaves and producing fewer fruits per tree. Morphological Cluster II consisted of elite shea trees with stronger trunks bearing small leaves and producing a high number of fruits per tree. Morphological Cluster III regrouped elite shea trees of medium trunk diameters carrying medium sized leaves; fruit production level is intermediate compare to preceding groups. The elite shea trees of morphological Clusters II, which are more interesting from an agronomic point of view, can be used as grafting trees for the production of high-yielding grafted plants for farmers in Côte d'Ivoire.

Key words: elite shea trees, genetic improvement, phenotypic variability, Northern Côte d'Ivoire.

INTRODUCTION

Vitellaria paradoxa C.F. Gaertn., commonly known as 'shea butter tree' or 'shea tree' in English and '*karité*' in French, is a plant species of the *Sapotaceae* family that grows naturally in Sudano-Sahelian belt of Africa (Hall et al., 1996; Diarrassouba et al., 2007). The geographical distribution of shea tree extends from Senegal to Uganda at latitudes between 2° and 8° North in East Africa, 7° and 12° North in Central Africa and 9° and 14° North in

*Corresponding author. E-mail: didierys@yahoo.fr. Tel: +225 04737926.

West Africa (Naughton et al., 2015). The large distribution of shea doubled by practicing allogamy as a mode of reproduction suggests a high intra-specific diversity (Diarrassouba et al., 2007). It is already reported the existence of two subspecies within the species *V. paradoxa* that are *paradoxa* and *nilotica* (Gwali et al., 2014). Butter of the western subspecies *paradoxa*, present in Côte d'Ivoire, is rich in stearic acid and gives a solid oil calling butter at ambient temperature, while the oil from the eastern subspecies *nilotica* is especially rich in oleic acid with liquid oil at ambient temperature. The species is widely known for its oil from nut which is used in cooking, cosmetics and traditional medicine (Diarrassouba et al., 2009a; Soro et al., 2011). The marketing of almonds and butter on local markets provides substantial income for women who are involved in the sector (Diarrassouba et al., 2008).

Despite the economic importance of shea butter as a multipurpose product, the erosion of genetic diversity and density of shea trees in natural agroforestry park is increasing over time. The cumulative effects of violent wind uprooting trees, the cutting downs for household needs, plowing for the installation of fields and the systematic collecting of fruits in natural parks by rural populations limit natural regeneration of shea trees (Boussim, 1991; Senou, 2000; Diarrassouba et al., 2009a). To overcome these problems, progressive domestication and genetic improvement of the species are being considered. Previous research has focused on fruit production (Lamien, 2006; Aleza et al., 2018), vegetative propagation (Bonkoungou et al., 1988), molecular diversity of populations (Bouvet et al., 2004; Fontaine et al., 2004; Sanou et al., 2005; Gwali et al., 2014 ; Abdulai et al., 2016), the spatial structure of populations (Kelly et al., 2004; Sanou et al., 2006, Aleza et al., 2015), the parasitism of the shea trees (Bayala et al., 2009; Samaké et al., 2011), the variability of fruit yields (Lamien et al., 2007; Bondé et al., 2019), etc. In Côte d'Ivoire especially, researches on V. paradoxa by Salé et al. (1991) reported that some characters such as fruit number, leaf and fruit size, leaf density, flowering and fruit ripening times are related to nut and butter yields. Also, in an agroforestry park of shea trees only 26% of trees are good producers and 15% of them have stable production (Salé et al., 1991). Similarly, quantitative and qualitative traits have proved highly relevant for probing morphological diversity in park of shea trees in Tengrela in Northern Côte d'Ivoire (Diarrassouba et al., 2007; Diarrassouba et al., 2009b). Only from qualitative morphological traits related to fruits and leaves, five morphological varieties of shea tree have been identified in the shea tree park at Tengrela (Diarrassouba et al., 2009b). The research activities carried out in the context of vegetative propagation enable breeding program of the shea tree of Côte d'Ivoire to multiply grafted plant materials and disseminate selected elite shea trees in village lands (Yao et al., 2019).

However, in the light of previous researches, there are still questions about the best ways and means of conserving shea genetic resources and improving the productivity of this species. Until today, the agromorphological characteristics of elite shea trees, identified in village terroirs and preserved *in situ* in the Bagoué and Tchologo districts, are not known in Côte d'Ivoire. Such knowledge would make help to set up breeding program of shea tree in Côte d'Ivoire and to judiciously involve the identified elite individuals in the subsequent hybridization plans for the creation of improved plant material. This study aims to know phenotypic diversity patterns among selected elite plant from Tchologo and Bagoué districts constituting the *in situ* collection of shea trees in Northern Côte d'Ivoire, using quantitative agromorphological traits.

MATERIALS AND METHODS

Study area

The study was conducted in the administrative districts of Bagoué and Tchologo (Figure 1). The districts of Tchologo and Bagoué are located in Northern Côte d'Ivoire between 9° 31 'and 9° 35' North latitude and 5° 11 'and 6° 29' West longitude.

Ecological zones of districts and departments selected for the study

The Northern region of Côte d'Ivoire (Figure 1) has a Sudano Guinean climate characterized by two major seasons. The dry season runs from November to April and the rainy season covers the period from May to October. The rainy season has an annual rainfall of around 1, 200 mm per year (Brou, 2005). The vegetation is of Sudano Guinean type with gallery forests along the rivers and a predominance of wooded and grassy savannas (N'Guessan et al., 2015). The geological formations of the zone consist of a succession of bands of schistose rocks, magmatic rocks and plutonic rocks from which several types of soils are derived, namely Ferralsols, Cambisols, Fluvisols and Luvisols (Yace, 2002; Kone et al., 2009).

Plant material

The study was conducted on 220 elite shea trees constituting the *in situ* collection of the University of Peleforo Gon Coulibaly (UPGC) (Figure 1). The elite shea tree is a natural tree presenting good agronomic characteristics in village lands and recognized by the famers themselves basing on theirs knowledges about shea tree. Like that, the elite shea trees were identified according participatory method by the farmers from a survey based on criteria such as the high fruit yield of tree, the sweet taste of the fruit pulp, the high size of the fruit, the early flowering every year and the regularity of production of the tree. The survey was conducted by the shea breeding program of UPGC, Côte d'Ivoire in collaboration with *Agence Nationale d'Appui au Développement Rural* (ANADER, Côte d'Ivoire). The elite shea trees were selected in four localities that are Tengrela (73 trees), Kouto (50 trees), Boundiali (23 trees), Ferkéssédougou (21 trees), Ouangolodougou (12 trees) and Kong (41 trees).

Morphological parameter measurements

The data was collected on elite shea trees during the period of
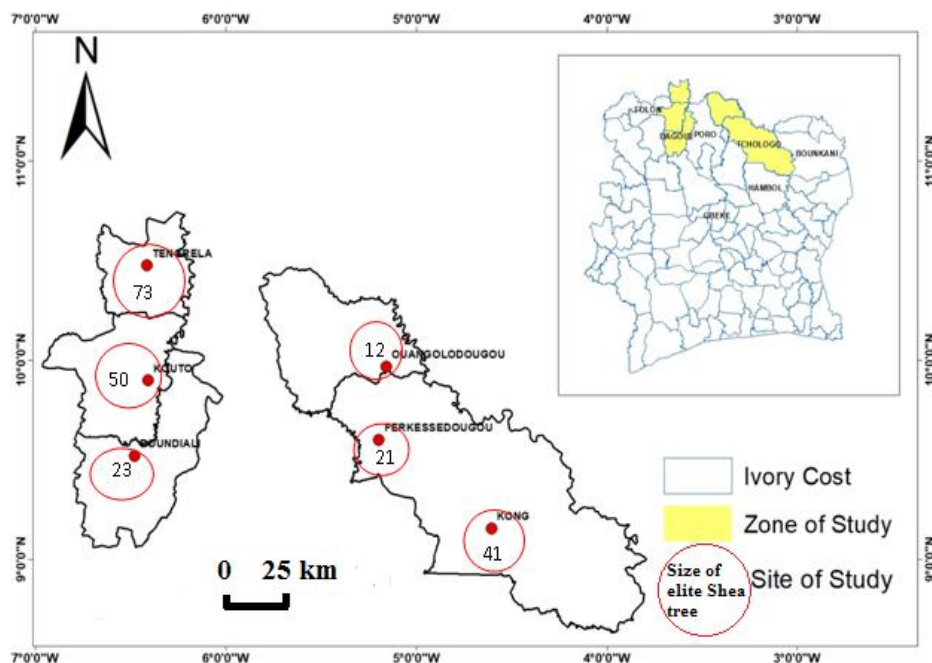
Figure 1. Study site maps showing Bagoué and Tchologo districts and sample size of observed elite shea trees.
Source: Shea breeding program of UPGC, Côte d'Ivoire, December (2019).

Table 1. Measured morphological traits, codes and unity.

Traits (SI Unity)	Code	Unity (SI)
Girth of trunk	GT	cm
Petiole length	PL	cm
Limb length	LL	cm
Limb width	LW	cm
Limb length/Limb width ratio	LL/LW	-
Fruit per tree	FT	Unit
Nut per fruit	NF	Unit
Nut length	NL	cm
Nut width	NW,	cm
Nut lenght/Nut width ratio	NL/NW	-
Nut volume	NV	cm^3
Nut weight	NWG	g

fruiting (May to July 2017), to characterize both vegetative aspects (trunk and leaf) and some descriptors related to the fruits. A total of 12 agromorphological quantitative traits were evaluated (Table 1, Figure 2). Quantitative traits such as girth of trunk, petiole length, limb length, limb width, fruit per tree, nut per fruit, nut length, nut width and nut weight were measured directly on the tree. For nut per fruit, nut length, nut width and nut weight evaluations ten nuts were considered per tree and the average values were retained. Thus the nuts extracted per well-developed fruit (when they fall down) were count. The length and width of each extracted nuts were measured using a sliding caliper type tool. The weight of each extracted nuts was assessed using an electronic scale. The parameters such as Limb length/Limb width ratio, Nut length/Nut width ratio and nut volume were calculated. The shape of the shea nut has been assimilated to a cylinder and its volume (NV) has been determined according to the mathematical expression:

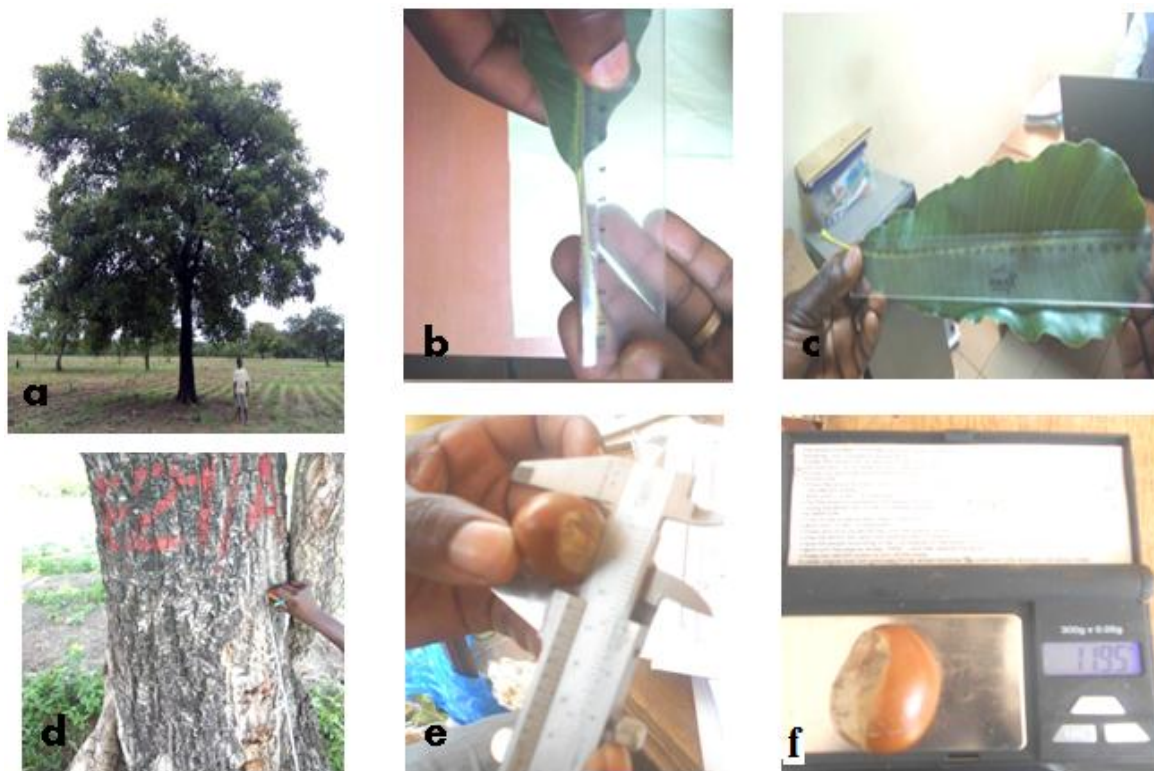$$NV = \pi \times r^2 \times h$$

where $r^2 = NW/2$, h = NL and $\pi = 3.14$.

Data analysis

First, data collected on elite shea trees have been submitted to a descriptive statistic analysis. The minimum, maximum and mean

Figure 2. Measures of morphological traits on the trunk, leaf and the seed of elite shea trees in the Bagoué and Tchologo districts, Northern Côte d'Ivoire. (a) Elite shea tree at the field, (b and c) Leaf size measures, (d) Measure of girth of trunk at 130 cm above the soil level (e and f). Measures of nut size and nut weight.

values, standard deviations and coefficients of variation have been determined for all quantitative traits. Then, a multiple analysis of variance (MANOVA) has been realized. This analysis was performed from all studied agromorphological traits and has been done in order to test the significance of all studied traits for elite shea trees discrimination when "district" or "department" was fixed as factor. Once the significance of the effect of the "district or department" factor is verified, the specific traits that contributed to the difference between localities (district or department) were identified from t-test when two districts (Bagoué vs. Tchologo) are compared and analysis of variance (ANOVA) when comparisons were done between six departments (Tengrela, Kouto, Boundiali, Ferkessedougou, Ouangolodougou and Kong) at the risk threshold of 5%. Any significant ANOVA (p <0.05) was followed a post-ANOVA test like Student Newman Keuls (SNK) test. Before MANOVA and ANOVA tests, the normality of each quantitative trait was verified from Shapiro-Wilk test at 5% of probability. Finally, the structure of the morphological diversity of elite shea trees was done from Principal Component Analysis (PCA), Hierarchical Classification Analysis (HCA) and Discriminant Analysis (DA). For PCA, the first four principal components were retained. Cluster analysis (Unweighted Pair Group Method Analysis, UPGMA) was performed using the matrix of average population values based on the elite shea tree's matrix of means. The clusters were then represented in a dendrogram. Also, the Pearson correlations between the morphological characters were revealed from Heat map using a quick and semi-automatic computational bio-statistical pipeline developed by Dago et al. (2019) in a simple programming language with R software. The other statistical analyzes were carried out using STATISTICA version 7.1 (StatSoft Inc., France) and SPSS version 20 (IBM Corp., USA) softwares.

RESULTS

Morphological variability of elite shea trees

The multiple analysis of variance (MANOVA) showed that all 12 variables made distinctions between the elite shea trees of two districts at the scale of the districts (Wilks'Lambda test; F = 4.100; p <0.001) (Table 2). Likewise, at the scale the department, differences between six departments were observed with all studied traits (Wilks'Lambda test; F = 3.100; p <0.001) (Table 2). Among the agromorphological traits that contributed to the structure of elite shea trees per district or department, the highly significant differences were observed in the expression of trunk girth (61 to 287 cm), the fruit number per tree (241 to 3903 fruits.tree^{-1}) and the nut weight (1.05 to 17.92 g) (Table 2).

Principal Component Analysis (PCA) identified four factors accounting for 75.95% of the variability observed among elite shea trees. The first identified factor accounting for 30.48% of the total variability reflects the nut size. The second identified factor accumulates 18.02% of the variability and defines the trunk robustness and fruit yield per tree. The third factor captured 15.01% of the variability and essentially explains the nut shape. The fourth factor identified accounted for 12.38% of the

Table 2. Variations in the trait expression in elite Shea trees and associated MANOVA, Student and ANOVA tests from district and department factors in Northern Côte d'Ivoire.

Traits (SI Unity)	Mean	Range	t-test with factor district		ANOVA test with factor department	
			F	p-value	F	p-value
Girth of trunk (cm)	149.25	50.00-287.00	18.63	<0.001	18.69	<0.001
Petiole length (cm)	8.50	5.18-29.42	4.92	0.027	1.71	0.131
Limb length (cm)	14.84	9.34-21.58	4.28	0.039	2.77	0.018
Limb width (cm)	4.58	2.82-9.94	5.87	0.016	1.71	0.133
Limb length/Limb width ratio	3.28	1.10-4.83	1.57	0.211	1.76	0.120
Fruit per tree	1774.51	241.00-3903.00	18.63	<0.001	18.69	<0.001
Nut per fruit	1.02	1.00 – 2.00	3.51	0.062	1.89	0.097
Nut length (cm)	3.18	2.26-6.60	5.29	0.022	3.83	<0.01
Nut width (cm)	2.38	1.21-3.15	3.85	0.051	2.71	0.021
Nut length/Nut width ratio	1.34	1.06-2.94	1.26	0.260	1.15	0.335
Nut volume (cm^3)	14.44	2.63-26.34	8.51	0.033	4.73	<0.01
Nut weight (g)	8.96	1.05-17.92	14.73	<0.001	6.17	<0.001
Wilks'Lambda test	-	-	4.10	<0.001	3.10	<0.001

Table 3. Factor loadings in the first four factor components.

Factor components	First	Second	Third	Fourth
Eigenvalues	3.65	2.17	1.80	1.48
Variance (%)	30.48	18.08	15.01	12.38
Cumulative variance (%)	30.48	48.56	63.57	75.95
Girth of trunk	-0.38	0.74	0.33	0.00
Petiole lengh	0.39	-0.36	0.13	0.31
Limb lengh	0.42	-0.52	0.26	0.28
Limb width	0.31	-0.47	0.37	-0.57
Limb lenght/Limb width ratio	0.05	0.04	-0.13	0.96
Fruit per tree	-0.38	0.75	0.33	0.00
Nut per fruit	0.38	0.03	0.48	-0.08
Nut lenght	0.70	0.32	-0.57	-0.14
Nut width	0.85	0.31	0.27	0.04
Nut lenght/Nut width ratio	0.07	0.07	-0.84	-0.19
Nut volume	0.92	0.34	-0.06	-0.04
Nut weight	0.86	0.30	0.09	0.04

total variability and reflected leaf shape (Table 3).

The analysis of the heat map revealed four categories of studied traits: (i) the trunk girth and the fruit number per tree, (ii) the nut length / nut width and limb length / limb width ratio, (iii) nut size (NV, NWG, NW and NL) and (iv) leaf size (LL, LW and LP) and nut number per fruit (see color key in Figure 3). Positive correlations were revealed within the descriptors of each identified category. In these categories, the strongest positive correlations were recorded in categories (i) and (iii). The highest correlation (r = 0.98) was obtained between the trunk girth at 130 cm (GT 130) and the fruit number per tree (fruit / tree). Highly significant values (p < 0.001) of

positive correlations were also observed between nut weight and nut width (r = 0.80), between nut weight and nut length (r = 0.58) and between nut weight and nut volume (r = 0.84). For foliar characteristics, the longer leaf is wider (r = 0.56, p = 0.01) also (see color key in Figure 3).

Morphological clusters of elite shea trees

The Hierarchical Classification Analysis (HCA) of the 220 elite of shea trees performed according to the Unweighted Pair Group Method Analysis (UPGMA)
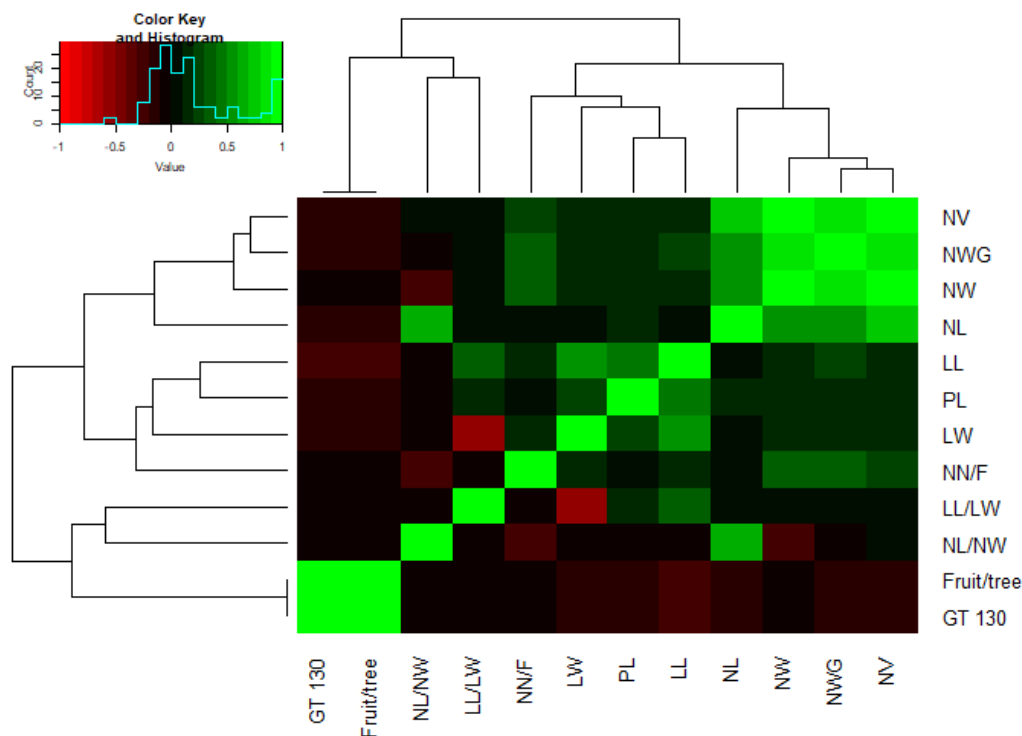
Figure 3. Pearson correlation heat map showing relationships between morphological traits measured on elite shea trees in Bagoué and Tchologo districts in Northern Côte d'Ivoire. GT 130: Girth of trunk at 130 cm above the soil; PL: Petiole lengh; LL: Limb lengh; LW: Limb width; LL/LW: Limb lenght/Limb width ratio; FT: Fruit per tree; NF: Nut per fruit; NL: Nut length; NW: Nut width; NL/NW: Nut lenght/Nut width ratio; NV: Nut volume; NWG: Nut weight.

aggregation criterion from the Euclidean distances of the quantitative traits revealed three morphological clusters (Figure 4). The multiple analysis of variance (MANOVA) performed on these clusters from the set of 12 studied morphological traits showed a significant difference (Wilks'Lambda test; F = 20.22; p <0.001) (Table 4). Morphological cluster I consisted of 139 elite shea trees, of which 83 were from the Bagoué district and 56 elite shea trees coming from Tchologo district. Elite shea trees of morphological cluster I have small trunk diameters (121.21± 26.73 cm) with large leaves (15.21± 2.45 cm limb length and 4.69± 0.73 cm limb width) and producing less fresh fruits per tree (1341± 413 fruits per tree) (Table 4, and Figure 5). Morphological cluster II consisted of 16 elite shea trees that represent 7.27% of studied sample: 10 trees from Bagoué and 6 trees from Tchologo. Trees of morphological cluster II express high values of trunk diameter (244.56 ± 20.98 cm) carrying small leaves (13.66 ±1.68 cm limb length and 4.17± 0.84 cm limb width) and producing a high number of fresh fruits per tree (3247±324.26 fruits per tree) (Table 4 and Figure 5). Morphological cluster III contains 65 elite shea trees or 29.54% of all studied elite shea trees with 53 and 12 from Bagoué and Tchologo districts respectively. Elite shea trees from morphological cluster III express medium values of trunk diameter (185.78 ± 14.56 cm) with

medium-sized leaves (14.34 ± 2.15 cm limb length and 4.45 ± 1.04 cm limb width) and an intermediate fruit production (2339 ± 225 fresh fruits per tree) in relation to the two preceding morphological clusters (Table 4 and Figure 5).

DISCUSSION

The agromorphological diversity study is important approach for the management of plant genetic resources. Likewise, Djekota (2014) reported that agromorphological approach constitutes the first step to start shea tree selection. For a long time, shea resources management has been mainly based on the classification of farmers for the conservation, domestication and selection (Lovett and Haq, 2000). However, these local knowledges retained for shea elite trees must necessarily be refined by scientifically proven knowledge in order to optimize improvement strategies for the species (Masters, 2002).

The agromorphological traits measured on the trunk, leaves and fruits showed significant variations between districts or even departments in Northern Côte d'Ivoire. These results suggest the significant influence of the environment on the expression of morphological characters in shea tree. These results can be explained
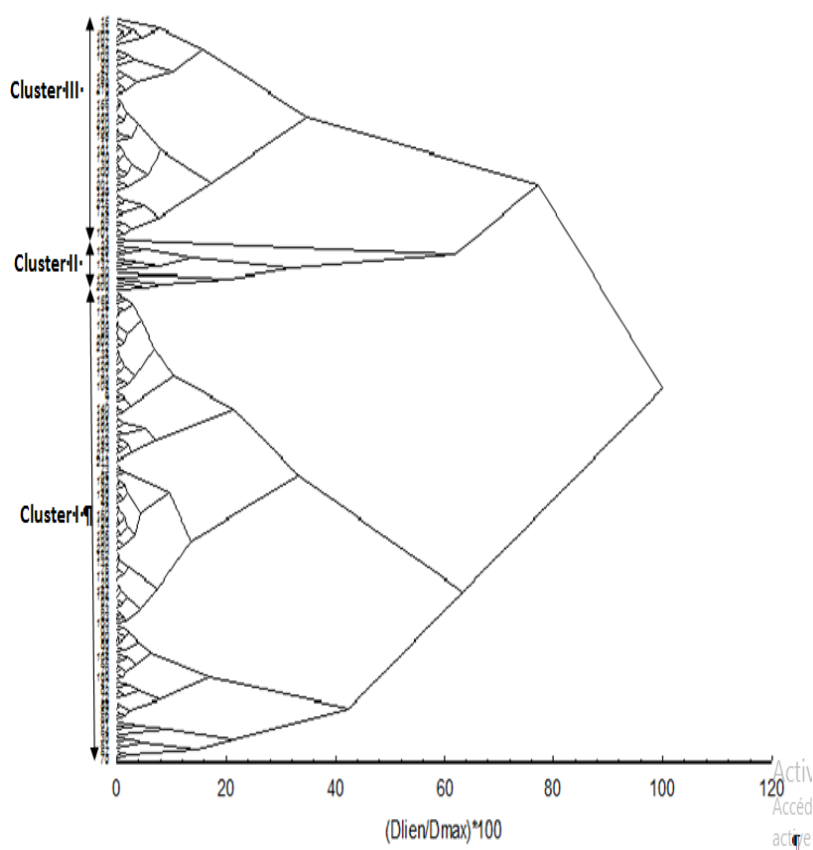
Figure 4. Dendrogram UPGMA of 220 elites shea trees in Bagoué and Tchologo districts, Northern Côte d'Ivoire.

Table 4. Characteristics of three morphological clusters identified from Hierarchical Cluster Analysis within 220 elite shea trees in Northern Côte d'Ivoire.

Traits (SI Unity)	Means ± standart deviation			F	p-value
	Cluster I (N=139)	Cluster II (N=16)	Cluster III (N=65)		
Girth of trunk (cm)	121.21 ± 26.73^c	244.56 ± 20.98^a	185.78 ± 14.56^b	311.91	<0.001
Petiole lengh (cm)	8.73 ±2.34^a	7.74 ± 1.08^a	8.21 ± 1.52^a	2.59	0.077
Limb lengh (cm)	15.21 ± 2.45^a	13.66 ±1.68^b	14.34 ± 2.15^{ab}	5.41	0.005
Limb width (cm)	4.69 ±0.73^a	4.17 ±0.84^b	4.45 ± 1.04^{ab}	3.84	0.022
Limb lenght/Limb width ratio	3.26 ±0.41^a	3.34 ±0.44^a	3.30 ±0.54^a	0.31	0.726
Fruit per tree	1341.00±413.00^c	3247±324.26^a	2339.00±225.00^b	311.91	<0.001
Nut per fruit	1.02±0.75^a	1.02 ± 0.42^a	1.02 ± 0.78^a	0.12	0.887
Nut lenght (cm)	3.22 ±0.44 ^{a}	3.19 ±0.22^a	3.10 ±0.31^a	1.95	0.144
Nut width (cm)	2.39 ± 0.24^a	2.39 ± 0.14^a	2.36 ±0.20^a	0.41	0.667
Nut lenght/Nut width ratio	1.35 ±0.18^a	1.33 ±0.06^a	1.31 ±0.10^a	1.32	0.266
Nut volume (cm^3)	14.74 ±4.04^a	14.47 ±2.43^a	13.80 ±3.20^a	1.39	0.249
Nut weight (g)	9.11 ± 2.81^a	9.08 ± 1.74^a	8.62 ±2.11^a	0.84	0.432
Wilks'Lambda test	-	-	-	19.22	<0.001

by a savannah gradient observed in the localities of study. In fact, on the ecological level, the Sudanese and Sudano-Guinean savannah are encountered in the study area (Bagoué and Tchologo). The climatic characteristics of these two savannah types being not similar would be at the origin of the environmental effects observed in the
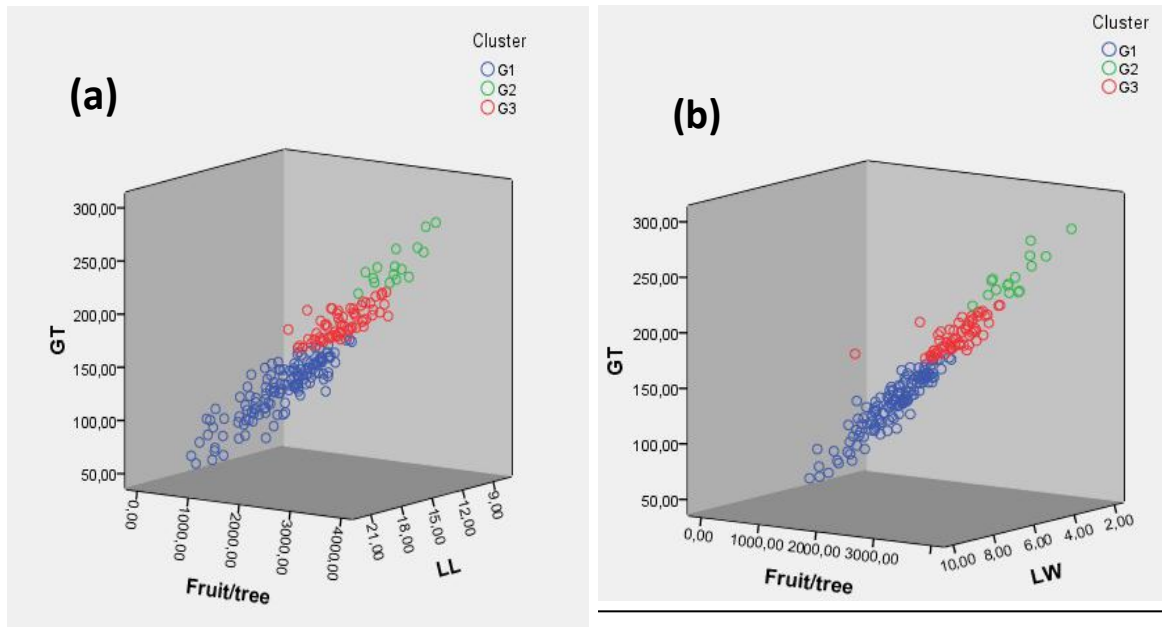
Figure 5. 3-D Scatter showing morphological diversity structure of the three clusters in the space formed from axes (X) Fruit per tree, (Y) Girth of trunk and (Z) Limb length or Limb width. GT: Girth of trunk at 130 cm; LL: Limb lengh; LW: Limb width; FT: Fruit per tree.

variation of the morphological characters expression as reported by McGowen et al. (2010). The impacts of climate on the expression of morphological characters have also been reported in Northern Equator (Maranz and Wiesman, 2003), Mali (Sanou et al., 2006; Tchabi and Adechi, 2014), Benin (Kougblénou et al., 2012; Kafilatou et al., 2015) and Uganda (Gwali et al., 2012). Moreover, in area of study the mineral composition of the soil, varying from one ecological zone to another, would have also influenced the expression of morphological characters as reported by Sanou et al. (2006), Moore (2008) and Bondé et al. (2019) respectively in studies on the vegetative characteristics of shea tree parks in Mali, Eastern Ghana and West Africa. The environmental effect significantly influencing the expression of agromorphological characters in shea tree can lead to a possible structuring of elite shea trees according to districts or departments. But, analysis of the diversity of all the elite shea trees without prior fixing of the district or department factors gave three morphological clusters. These morphological clusters do not correspond to the geographical structuring of the identified elite shea trees. These results would indicate that the geographic zone factor alone cannot constitute the basic factor in the structuring of identified elite shea trees. Indeed, the natural and human selections could be at the origin of these groupings. The selection of elite trees based on farmers' preferences would have led to the selection of trees with identical performances from one locality to another (Gwali et al., 2012; Karambiri et al., 2016). Similarly, the allogamous nature of shea tree has been

found to cause important gene flow between populations (Vaughan et al., 2007; Abasse et al., 2011). This would have led to similar morphological characteristics of shea trees from one locality to another.

The results show that the variability factors reflecting the dimensions of the nut, although more important in explaining the variability in the shea tree, did not make it possible to differentiate the three morphological clusters of elite trees identified in the Bagoué and Tchologo districts. The three morphological clusters differed mainly in the girth of trunk and the number of fruits per tree and secondarily in the size of the leaves. These results demonstrate that girth of trunk could be a good indicator for selection of producing trees, especially since a higher correlation (r = 0.99) was obtained between trunk girth and fruit number per tree. In the study areas, elite shea trees are spared and maintained by the farmers during agricultural practices according to preference criteria such as the taste of the pulp, the fruit size and the oil content. This way of phenotypic selection made by farmers in village lands would have led to retaining elite trees or semi-domesticated trees with similar performances depending on the characteristics of the nut.

The results reveal four categories of studied traits presenting significant positive interrelationships between them: (i) the trunk girth and the fruit number per tree, (ii) the nut length / nut width and limb length / limb width ratio, (iii) nut size (NV, NWG, NW and NL) and (iv) leaf size (LL, LW and LP) and nut number per fruit. This result indicates that the use of a single pomological descriptor such as the nut width is sufficient to characterize the

shea nut. For example there are 80% odds to make a good estimate of the nut weight by referring to its width. Also, a single foliar descriptor such as the leaf length can be retained to discriminate populations of shea trees. Similarly, there is a significant correlation between trunk girth and fruit number per tree. This positive correlation indicates that fruit yield increases with the tree robustness. Meanwhile the increase of the fruit production is not linear with the age of the shea tree. Indeed, Nouvellet et al. (2006) showed that old shea tree individuals with big diameter are less or not productive compare to young ones with medium diameter in shea trees park of Ténéfina, Mali. Thus, the correlations found in the categories of characters highlighted in the present study in shea tree have already been mostly reported in the earlier works of Diarrassouba et al. (2007). Taking correlations into account may reduce the number of descriptors as the biological information provided by two positively correlated descriptors is similar as reported by Yao et al. (2015) concerning the coconut trees.

Conclusion

This study was conduct with a view to characterize the morphological diversity of selected elite shea trees identified in the Bagoué and Tchologo districts of Côte d'Ivoire. The results revealed three morphological clusters of elite shea trees from Bagoué and Tchologo districts in Northern Côte d'Ivoire. While waiting to use these three gene pools for hybrid creation, individuals of morphological cluster II showing attractive agronomic performances, can serve as graft-producing trees for the production of high-yielding grafted plants in Côte d'Ivoire.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Abasse T, Weber J, Katkore B, Boureima M, Larwanou M, Kalinganire A (2011). Morphological variation in *Balanites aegyptiaca* fruits and seeds within and among parkland agroforests in eastern Niger. Agroforestry Systems 81(1):57-66.

Abdulai I, Krutovsky KV, Finkeldey R (2016). Morphological and genetic diversity of shea tree (*Vitellaria paradoxa*) in the savannah regions of Ghana. Genetic Resources and Crop Evolution 64(6):1253-1268.

Aleza K, Wala K, Bayala J, Villamor GB, Dourma M, Atakpama W, Akpagana K (2015). Population structure and regeneration status of *Vitellaria Paradoxa* C. F. Gaertn under different land management regimes in Atacora department, Benin. Agroforestry Systems 89(3):511-523.

Aleza K, Villamor GB, Nyarko BK, Wala K, Akpagana K (2018). Shea (*Vitellaria paradoxa* C. F. Gaertn) fruit yield assessment and management by farm households in the Atacora district of Benin. *PLoS ONE* 13(1).

Bayala J, Ky-Dembélé C, Boussim IJ (2009). Technique de Taille Sylvicole pour la Lutte Contre les Parasites Loranthaceae et l'Amélioration de la Production Fruitière du Karité. Fiche Technique, INERA, Burkina Faso.

Bondé L, Ouédraogo O, Ouédraogo I, Thiombiano A, Boussim I (2019). Variability and estimating in fruiting of shea tree (*Vitellaria paradoxa* C.F. Gaertn.) associated to climatic conditions in West Africa: implications for sustainable management and development. Plant Production Science 22(2):143-158.

Bonkoungou E, Zerbo J, Billan A (1988). Multiplication végétative du karité (*Butyrospermum paradoxum*) au Burkina Faso. In: IRBRT-ISNIIDR, 1988. Séminaire national sur la valorisation du karité pour le développement national. Bilan et perspectives, pp 45-54.

Boussim I (1991). Contribution à l'étude des *Tapinanthus* Parasites du karité au Burkina Faso. Thèse de Doctorat 3ème Cycle, Université de Ouagadougou, Ouagadougou, Burkina Faso, 131p.

Brou YT (2005). Climat, mutations socio-économiques et paysages en Côte d'Ivoire. Mémoire de synthèse des activités Scientifiques présenté en vue de l'obtention de l'habilitation à diriger des recherches. Université des Sciences et Technologies de LILLE, France, p. 212.

Bouvet JM, Fontaine C, Sanou H, Cardi C (2004). An analysis of the pattern of genetic variation in *Vitellaria paradoxa* C. F. Gaertn. Using RAPD markers. Agroforestry Systems 60(1):61-69.

Dago DN, Fofana IJ, Diarrassouba N, Barro ML, Moroh JL, Dagnogo O, Loukou NE, Yao SDM, Silué S, Giovanni M (2019). A Quick Computational Statistical Pipeline Developed in R Programing Environment for Agronomic Metric Data Analysis. American Journal of Bioinformatics Research 9(1):22-44.

Diarrassouba N, N'guessan A, Koffi E, Sangaré A (2007). Evaluation des performances de quelques descripteurs quantitatifs et leur utilisation dans la structuration de la population d'un parc naturel de karité en Côte d'Ivoire. Plant Genetic Resource Newsletter 152:65-72.

Diarrassouba N, Koffi KE, N'Guessan KA, Patrick V, Sangare A (2008). Connaissances locales et leur utilisation dans la gestion des parcs à karité en Côte d'Ivoire. Afrika Focus 21(1):77-96.

Diarrassouba N, Bup ND, Fofana IJ, Sangare A (2009a). Varieties fixing at shea trees (*Vitellaria paradoxa* C.F. Gaertn.) Using morphological traits in Côte d'Ivoire. Geneconserve 8(33):752-780.

Diarrassouba N, Fofana JI, Bakayoko A, Nguessan AK, Sangare A (2009b). Influence des systèmes agraires sur la dynamique de régénération naturelle du karité : *Vitellaria paradoxa* C.F. Gaertn. (*Sapotaceae*) en Côte d'Ivoire. Agronomie Africaine 21(1):49-58.

Djekota C (2014). Etude de la variabilité morphologique et taxonomique des morphotypes chez le karité (*Vitellaria paradoxa* C.F. Gaertn. sous-espèce *paradoxa. Sapotaceae*) dans la région du Mandoul au Tchad. Thèse de Doctorat, Université Cheikh Anta Diop de Dakar, Sénégal, p. 151.

Fontaine C, Lovett PN, Sanou H, Maley J, Bouvet JM (2004). Genetic diversity of the shea tree (*Vitellaria paradoxa* C.F. Gaertn) detected by RAPD and chloroplast microsatellite markers. Heredity 93(6):639–648.

Gwali S, Nakabonge G, Okullo JBL, Eilu G, Nyeko P, Vuzi P (2012). Morphological variation among shea tree (*Vitellaria paradoxa* subsp. *nilotica*) 'ethnovarieties' in Uganda. Genetic Resources and Crop Evolution 59(8):1883-1898.

Gwali S, Vaillant CA, Nakabongea G, Okulloa J, Eilua G, Muchugid A, Bouvet C J (2014). Genetic diversity in shea tree (*Vitellaria paradoxa* subspecies *nilotica*) ethno-varieties in Uganda assessed with microsatellite markers. Forests Trees and Livelihoods 24(3):163-175.

Hall JB, Aebischer DP, Tomlison HF, Osei-Amaning E, Hindle JR (1996). *Vitellaria paradoxa*: A monograph Vitellaria paradoxa: a

monograph (Publication No. 8). University of Wales. Bangor, UK: School of Agricultural and Forest Sciences.

Kafilatou ST, Léonard AE, Vincent E, Eliassou SH (2015). Agro-morphological variability of shea populations (*Vitellaria paradoxa* CF Gaertn) in the Township of Bassila. Benin Republic. Journal of Plant Breeding and Crop Science 7(2):28-37.

Karambiri M, Elias M, Vinceti B, Grosse A (2016). Exploring local knowledge and preferences for shea (*Vitellaria paradoxa*) ethnovarieties in Southwest Burkina Faso through a gender and ethnic lens. Forests, Trees and Livelihoods 26(1):13-28.

Kelly BA, Bouvet JM, Picard N (2004). Size class distribution and spatial pattern of *Vitellaria paradoxa* in relation to farmers' practices in Mali. Agroforestry System 60(1):3-11.

Kone B, Diatta S, Sylvester O, Yoro G, Mameri C, Desiré DD, Ayemou A (2009). Estimation de la fertilité potentielle des ferralsols par la couleur. Canadian Journal of Soil Science 89(3):331-342.

Kougblénou N, Ahouansou RH, Aïssi MV, Houssou P, Padonou W, Fandohan P, Mensah GA, Soumanou MM (2012). Caractérisation physique du fruit et valeur nutritionnelle de la pulpe de karité (*Vitellaria paradoxa*) collecté dans différents parcs au Bénin. Bulletin de la Recherche Agronomique du Bénin, pp. 1840-7099.

Lamien N (2006). Fructification du karité *(Vitellaria paradoxa* Gaertn.F.. Sapotaceae): Facteurs de déperdition. Amélioration et Prévision des rendements à Bondoukuy. Ouest du Burkina Faso. Thèse de Doctorat, Université de Ouagadougou, Burkina Faso, p. 108.

Lamien N, Tigabu M, Guinko S, Oden PC (2007). Variations in dendrometric and fruiting characters of *Vitellaria paradoxa* populations and multivariate models for estimation of fruit yield. Agroforestry System 69(1):1-11.

Lovett PN, Haq N (2000). Diversity of the Shea nut tree (*Vitellaria paradoxa* C.F. Gaertn.) in Ghana. Genetic Resources and Crop Evolution 47(3):293–304.

Maranz S, Wiesman Z (2003). Evidence for indigenous selection and distribution of the shea tree, *Vitellaria paradoxa*, and its potential significance to prevailing parkland savanna tree patterns in sub-Saharan Africa north of the equator. Journal of Biogeographie 30(10):1505-1516.

Masters E (2002). La ressource en karité: vue d'ensemble de la recherche et du développement en Afrique. Présentation à atelier international sur le traitement, la valorisation et le commerce du karité en Afrique. Centre de Suivi Ecologique Dakar, Sénégal du 4 au 6 mars 2002.

McGowen MH, Vaillancourt RE, Pilbeam DJ, Potts BM (2010). Sources of variation in self incompatibility in the Australian forest tree, *Eucalyptus globulus*. Annals of Botany 105 (5):737-745.

Moore S (2008). The role of *Vitellaria Paradoxa* in poverty reduction and food security in the Upper East region of Ghana. Earth and Environment 3:209-245.

Naughton CC, Lovett PN, Mihelcic JR (2015) Land suitability modeling of shea (*Vitellaria paradoxa*) distribution across sub-Saharan Africa. Applied Geography 58:217-22.

N'Guessan A, Diarrassouba N, Konan A, Nangha KY, Inza JF, Kouamé AY (2015). Indicateurs de dégradation physique des sols dans le nord de la Côte d'Ivoire : cas de Boundiali et Ferkéssédougou. Afrique Science 11:3.

Nouvellet Y, Kassamba A, Besse F (2006). Le parc à karité au Mali: inventaire, volume, houppiers et production fruitière. Bois et Forêts des Tropiques 287:6-20.

Salé G, Boussim J, Raynal-Roques A, Brunck F (1991). Le Karité une richesse potentielle. Perspectives de recherche pour améliorer sa production. Revue Bois et forets des Tropiques 28(2):11-23.

Samaké O, Dakouo JM, Kalinganire A, Bayala J, Bréhima K (2011). Techniques de déparasitage et gestion du karité. ICRAF Technical Manual No. 15. Nairobi: World Agroforestry Centre.

Sanou H, Lovett PN, Bouvet JM (2005). Comparison of quantitative and molecular variation in agroforestry populations of the shea tree (*Vitellaria paradoxa* C.F. Gaertn.) in Mali. Molecular Ecology 14(8):2601-2610.

Sanou H, Picard N, Lovett PN, Dembele M, Korbo A, Diarisso D, Bouvet JM (2006). Phenotypic variation of agromorphological traits of the shea tree, *Vitellaria paradoxa* C.F.Gaertn, in Mali. Genetic Resources and Crop Evolution 53(1):145-161.

Senou O (2000). Les peuplements de karité *(Vitellaria paradoxa* Gaertn C.F.) dans le cercle de Koutiala au Sud de Mali: répartition, structure et parasitisme par les *Tapinanthus*. Mémoire de DEA, Université de Ouagadougou, Burkina Faso, p. 84.

Soro D, Traore K, Kassi NJ (2011). Variabilité des caractères morphologiques chez le karité (*Vitellaria paradoxa*) dans le Nord de la Côte d'Ivoire. International Journal of Biological and Chemical Sciences 5(3):1201-1214.

Tchabi V, Adechi KO (2014). Occupation des terres, typologie et structure des faciès de végétation à Vitellaria paradoxa (Gaertn F.) de la Commune de Toukountouna au Nord-Bénin. International Journal of Biological and Chemical Sciences 8(4): 1684-1696.

Vaughan DA, Balazs E, Heslop-Harrison JS (2007). From crop domestication to super-domestication. Annals of Botany 100(5):893-901.

Yace I (2002). Initiation à la géologie. L"exemple de la Côte d"Ivoire et de l"Afrique de l"Ouest. Pétrologie, Géologie régionale. Ed. CEDA, SODEMI p 183.

Yao SDM, Konan JL, Sie R, Diarrassouba N, Lekadou TT, Koffi EBZ, Yoboue K, Bourdeix R, Issali AE, Doh F, Allou K, Zoro Bi IA (2015). Fiabilité d'une liste minimale de descripteurs agromorphologiques recommandée par le COGENT dans l'étude de la diversité génétique du cocotier (*Cocos nucifera* L.). Journal of Animal and Plant Sciences 26 (1):4006-4022.

Yao SDM, Alui KA, Kouame NMT, Ble PA, Kone B, Diarrassouba N (2019). Réussir le « greffage en fente simple » et le « greffage de côté dans l'aubier » du karité. Journal of Applied Biosciences 137(1):13961-13972.

International Journal of Genetics and Molecular Biology

AJ ACADEMIC JOURNALS
expand your knowledge

Full Length Research Paper

Mapping of fourteen high-risk human papillomavirus genotypes by molecular detection in sexually active women in the West African sub-region

Théodora Mahoukèdè Zohoncon[1,2,3]*, **Wendkuuni Florencia Djigma**[1,2,4], **Abdoul Karim Ouattara**[1,2], **Ina M.A Traore**[1,2], **R. Alice Ouedraogo**[1,2], **Esther M.A Traoré**[1,2], **Prosper Bado**[1,2], **T. Clarisse Ouedraogo**[1,2], **Birama Diarra**[1,2], **Maimouna Ilboudo**[1,2], **Callinice D. Capo-chichi**[5], **Dorcas Obiri-Yeboah**[6], **Simplice Damintoti Karou**[2,7], **Mady Nayama**[8], **Apollinaire Horo**[9], **Kouame Privat Kouakou**[10], **Moutawakilou Gomina**[11], **Souleymane Ouattara**[12], **Ambaliou Sanni**[5], **Akpona Simon**[11], **Charlemagne Ouedraogo**[2,3] and **Jacques Simpore**[1,2,3,4]

[1]Laboratory of Molecular Biology and Genetics (LABIOGENE), University of Ouagadougou, 03 BP 7021 Ouagadougou 03, Burkina Faso.
[2]Pietro Annigoni Biomolecular Research Centre (CERBA), 01 BP 364 Ouagadougou 01, Burkina Faso.
[3]University Saint Thomas d'Aquin, Faculty of Medicine, 06 BP 10212 Ouagadougou 06, Burkina Faso.
[4]University of Ouagadougou, UFR/SDS, 03 BP 7021 Ouagadougou 03, Burkina Faso.
[5]University of Abomey-Calavi, Benin, 01 BP 526 Cotonou, Benin.
[6]University of Cape Coast, School of Medical Sciences, Department of Microbiology and Immunology, University Post Office, Tel. +233 332 132 440; Fax: +233 332 132 484, Ghana.
[7]University of Lomé, Togo, 01BP1515, Lomé, Togo.
[8]University Abdoul Moumouni, Niamey, BP 237 – Niamey, Niger.
[9]University Félix Houphouët Boigny, UFR Des Sciences Médicales D'Abidjan,01 BPV 34 Abidjan, Ivory Coast.
[10]University of Bouaké, Côte d'Ivoire, CHU de Bouaké, BP V 18 01, Bouaké Ivory Coast.
[11]University of Parakou, Benin, BP 123 Parakou, Benin.
[12]University Polytechnique of Bobo, 01 BP 1091, Bobo-Dioulasso - Burkina Faso.

The aim of this study was to determine the distribution of high-risk human papillomavirus genotypes (HR-HPV) in women from the general population of five West African countries. This was a cross-sectional descriptive study, involving 2133 women from nine cities of five West African countries: Benin, Burkina Faso, Côte d'Ivoire, Niger and Togo. Women were screened for precancerous cervical lesions and HR-HPV infection. The detection of HR-HPV was done by a multiplex real-time PCR on extracted viral DNA. The average age of the women in this study was 35.06 ± 10.00 years with a range of 15 to 65 years. The overall prevalence of high-risk HPV infection among general population sample of women in five West African countries was 33.61% (717/2133). The prevalence of dysplasia was 8.81%. In decreasing order of frequency, the genotypes found were: HPV 52 followed by HPV 31, 59, 51, 66, 45, 68, 56, 56, 58, 35, 39, 18, 33 and 16. The prevalence of HPV16/18 (bivalent vaccine types) was 7.02%. This study reveals a high prevalence of HPV 52 in West Africa. The extent and diversity of HR-HPV genotypes in these West African countries deserve special attention for prevention.

Key words: High-risk HPV, real time PCR, genotypes, women, epidemiology, West Africa.

INTRODUCTION

Human papillomavirus (HPV) infection and cervical cancer remain a major concern worldwide, especially in sub-Saharan Africa where cervical cancer, induced by high-risk HPV (HR-HPV), is the leading cause of cancer death in women. In addition, the slow and insidious evolution of this condition as well as the absence of systematic screening would explain why it is most often diagnosed at a late stage. HPV infection is the most common sexually transmitted infection (STI) in the world, with 660 million people infected according to the World Health Organization (WHO). The WHO estimates that the annual incidence of cervical cancer was 500,000 with more than 90% of cases in developing countries. In sub-Saharan Africa, invasive cervical cancer is the most common cancer in women with more than 75,000 new cases and more than 50,000 deaths per year (Ferlay, et al., 2010). In Africa, the prevalence of HPV infection reaches 21.3% with significant regional variations: 33.6% in East Africa, 21.5% in West Africa and 21% in Southern Africa (Ferlay et al., 2010) countries, both nationally and internationally, especially in developing countries, cancer has a negative impact on the general health of the family and results in a loss of income and huge health expenditures, as it mainly affects the economically productive age group. In Burkina Faso, annual number of cervical cancer cases is estimated at 2.517 and cervical cancer deaths are found to be 2.081 per year (ICO/IARC, 2018).

   The best means of control and prevention through the use of prophylactic vaccination against HPV, is not available to all populations, both urban and rural. In addition, 12 years after the release of the first two HPV vaccines (2006), despite GAVI's efforts, they remain expensive and are not yet accessible to the entire population of the West African sub-region. Some pharmaceutical companies have made efforts to reduce the cost of the vaccine in order to expand HPV vaccination campaigns for girls in some African countries. However, the HPV vaccines available on the market only cover two HR-HPV genotypes, HPV16 and HPV18. HPV16 and 18 genotypes are believed to be the most prevalent in Europe and the rest of the world, while preliminary studies by Djigma et al. (2011); Ouédraogo et al. (2011); Zohoncon et al. (2013) and Ouedraogo *et al.* (2015) and Rahimy et al. 2015) in Burkina Faso have rather shown a high prevalence of the HPV 30 and 50 family. A deep knowledge of circulating HPV genotypes is of a high interest in specific populations for the development of effective HPV vaccine covering the predominant genotypes in these populations. A large

sample study is therefore crucial to determine the circulating genotypes in the general population on the one hand and in cervical cancer cases on the other hand. For effective control of cervical cancer, a preventable malignant tumor through prophylactic vaccination is done. This study aims to describe the molecular epidemiology of high-risk HPV genotypes in women without cervical lesions in nine cities of five West African countries.

MATERIALS AND METHODS

Study type and population

This was a cross-sectional, descriptive study that collected 2133 endocervical samples from the cervix of women in the general population without cervical lesions. The samples came from five West African countries: Benin, Burkina Faso, Côte d'Ivoire, Niger and Togo.

Inclusion criteria

Included were all non-pregnant women and girls who freely consented after receiving information on the study.

Criteria for non-inclusion

Not included in the study were women or girls who were virgins or pregnant or who had a total hysterectomy.

Collection of samples

After sensitization on HPV infection prevention and cervical cancer risk, and after obtaining free and informed consent of women, a questionnaire was administered to women to collect socio-demographic, behavioural and clinical information, and an endocervical swab was performed at the cervix of women; followed by screening for precancerous cervical lesions by IVA/VILI. The samples collected were sent to the CERBA/LABIOGENE molecular biology and genetics laboratory, University Joseph Ki-Zerbo, Burkina Faso, for molecular analyses.

Extraction of viral DNA from HR-HPV

The DNA extraction was done using DNA-Sorb-A kit (Sacace Biotechnologies, Como, Italy) by following the protocol supplied by the manufacturer.

Real-time HR-HPV detection by multiplex PCR

Detection of high-risk HPV genotypes was made by real-time PCR using "HPV Genotypes 14 Real-TM Quant" kit (Sacace Biotechnologies, Como, Italy) and Sacycler-96 Real time PCR v.7.3 (SACACE Biotechnologies, Como, Italy). This genotyping is based

*Corresponding author. E-mail: zohoncont@gmail.com. Tel: +226 25361232, Fax: +226 25363242.

on multiplex real time PCR amplification for each sample and the β-globin gene was used as internal control. The "HPV Genotypes 14 Real-TM Quant" kit allowed to detect the following 14 high-risk HPV genotypes such as HPV16, HPV18, HPV31, HPV33, HPV35, HPV39, HPV45, HPV51, HPV52, HPV56, HPV58, HPV59, HPV66 and HPV68. Each sample was subjected to multiplex amplification in 4 tubes and each tube contained primers of the target regions (L1 gene and oncoproteins E6 and E7) of three or four types of HPV-HR and of the human beta-globin gene as control internal. For each sample we had respectively for the 4 tubes: PCR-mix-1 16, 18, 31, IC; PCR-mix-1 39, 45, 59, IC; PCR-mix-1 33, 35, 56, 68; PCR-mix-1 51, 52, 58, 66. The pre-PCR steps consisted in: preparing the Mix solution (PCR-buffer-FRT + Hot Start DNA Polymerase) and the Reaction Mix solution (Mix solution + each PCR- mix-1).

For each sample, 15 µL of the Reaction Mix solution was introduced, into the 4 tubes and add 10 µL of the extracted DNA. The total volume of the reaction was 25 µL. This PCR reaction mixture contained in sterile 0.2 mL microtubes was introduced onto the plate of the SaCycler-96 Real Time PCR v.7.3 (Sacace Biotechnologie, Italy) for amplification. The PCR program used was as follows: 1 cycle of 95°C for 15 min; 5 cycles of 95°C for 05 s, 60°C for 20 s, 72°C for 15 s; 40 cycles of 95°C for 05 s, 60°C for 30 s and 72°C for 15 s.

Ethics approval and consent to participate

This study was approved by the Health Research Ethics Committee of Burkina Faso with the reference number 2016-02-0012 on 03/02/2016. All study participants gave their free written and informed consent according to the Helsinki Declarations.

Data analysis

Data were entered and analyzed using the IBM SPSS software in its 21 version and Epi Info 6. The Chi-square test was used for comparisons with a significant difference for $p < 0.05$.

RESULTS

Sociodemographic, behavioral and clinical characteristics of the study population

The average age of the women in this study was 35.06 ± 10.0 years a range of 15 to 65 years. The median age was 34 years. The 25 - 34 years age group was in the majority with 39.05% (833/2133) of women in the study population. Married women accounted for 73.18% (1561/2133) of the population; and secondary education was the majority at 38.44%. The average age at first intercourse was 18.5 ± 3.28 years with extremes of 6 to 30 years. Women reported having only one sexual partner in 83.87% (1789/2133) of cases. The frequency of sexual intercourse was on average twice a week in 51.90% of cases. Among the women in the study, 67.83% (1183/1744) did not use condoms; 60.57% (1292/2133) did not use a contraceptive method; 31.65% (675/2133) had a history of sexually transmitted infection (STI) and 2.30% (49/2133) reported being HIV positive. Screening for precancerous and cancerous cervical lesions by visual inspection with acetic acid (VIA) and

visual inspection with lugol (VILI) among the women in the study had a dysplasia prevalence of 8.81% or 188 positive VIA/VILI. Table 1 shows the characteristics of the study population.

Prevalence of high-risk HPV infection among women in the general population

The overall prevalence of high-risk HPV infection among women in the general population of the five West African countries was 33.61% (717/2133). By country, the prevalence of high-risk HPV infection was 34.78% (160/460) in Benin; 37.09% (171/461) in Burkina Faso; 39.67% (192/484) in Côte d'Ivoire; 12% (30/250) in Niger and 34.31% (164/478) in Togo. Figure 1 shows the prevalence of HR-HPV infection by city, with nine cities in the five West African countries.

Frequency of HR-HPV genotypes in women in the general population of the five West African countries

Cumulative total number of genotypes identified in HPV-infected women was 1068 genotypes. HPV52 was the most common genotype (Table 2). Figure 2 shows the frequencies of the 14 high-risk HPV genotypes detected in our study.

The presence of the different high-risk oncogenic HPV genotypes in the women in our study is shown in Figure 3. Figure 3 shows the mapping of high-risk HPV genotypes in the five countries of our study: Benin, Burkina Faso, Cote d'Ivoire, Niger and Togo. Among women in our study, without cervical lesions and infected with HR-HPV, the prevalence of HPV16/18 (bivalent vaccine types) was 7.02% (Figure 4).

Multiple and isolated infections

Of the 717 women infected with HPV, 250 or 34.87% had a high-risk multiple HPV infection and 467 women (65.13%) had an isolated infection. The number of HR-HPV genotypes per woman ranged from 1 to 6. Multiple infections with 2 and 3 genotypes were in the majority, with 161/250 (64.40%) and 70/250 (28.00%) of cases respectively. Among the two-genotype combinations, one woman was infected with HPV16/18; while among the three-genotype co-infections, two women were infected with HPV16/18/31. Table 3 presents the different combinations of multiple infection with 4, 5 and 6 high-risk HPV genotypes obtained in this study.

HPV and risk factors

Age, marital status, education level, occupation, number of sexual partners, frequency of sexual intercourse, condom non-use and STI history, contraceptive use, HIV

Table 1. Socio-demographic, behavioural and clinical characteristics of the study population.

Characteristics of the study population	Number of employees	%
Age groups (years)		
15-24	262	12.28
25-34	833	39.05
35-44	686	32.17
45-54	256	12.00
55-65	96	4.50
Marital status		
Married	1561	73.18
Single	497	23.30
Widow	65	3.05
Divorced	10	0.47
Level of study		
Uneducated	536	25.13
Primary school	394	18.47
Secondary school	820	38.44
Academic	383	17.96
Profession		
Student	176	8.25
Housewife	820	38.44
Employee	487	22.83
Informal sector	650	30.48
1st sexual intercourse		
≤18ans	1239	58.09
>18 years old	894	41.91
Number of sexual partners		
1	1789	83.87
2	218	10.22
3	101	4.74
4	1	0.05
Frequency of sexual intercourse		
≤2 times/week	947	55.87
>2 times/week	748	44.13
Contraception		
Yes	841	39.43
No	1292	60.57
Condom use		
Yes	561	32.17
No	1183	67.83
HIV serology		
Positive	49	2.30
Negative	574	26.91
Unknown	1510	70.79

Table 1. Contd.

IST ATCD		
Yes	675	31.65
No	601	28.18
Unknown	857	40.17
VIA/VILI		
Positive	188	8.83
Negative	1945	91.17
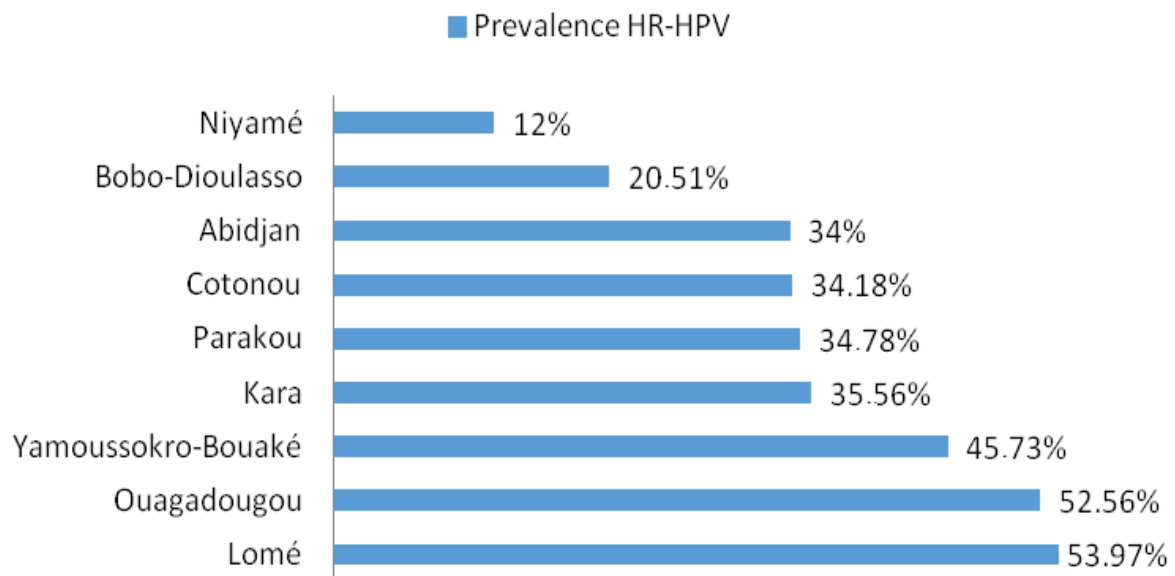
Figure 1. Prevalence of HR-HPV infection in nine cities in five West African countries.

Table 2. Frequency of single genotypes infections in 5 West African countries.

Genotypes	Countries N = 2133				
	Benin % (N = 460)	Burkina Faso % (N = 461)	Ivory Coast % (N = 484)	Niger % (N = 250)	Togo % (N = 478)
HPV 16	1.88	1.17	2.60	0.4	4.81
HPV18	6.25	6.43	5.21	0.8	3.37
HPV 31	6.25	4.09	11.98	1.6	18.75
HPV 33	2.50	0	1.6	0.4	3.37
HPV 35	13.13	4.09	7.30	1.6	11.54
HPV 39	9.38	9.94	4.17	1.2	5.28
HPV 45	16.88	7.60	8.85	0.8	7.21
HPV 51	12.50	8.77	4.69	0.8	13.94
HPV 52	16.88	11.11	14.58	2.8	16.83
HPV 56	6.88	8.77	10.42	0.4	16.34
HPV 58	12.50	7.60	5.21	0.4	9.13
HPV 59	7.5	15.78	2.60	1.2	6.73
HPV 66	7.5	15.78	5.21	2	9.62
HPV 68	8.75	8.77	16.14	1.6	10.58
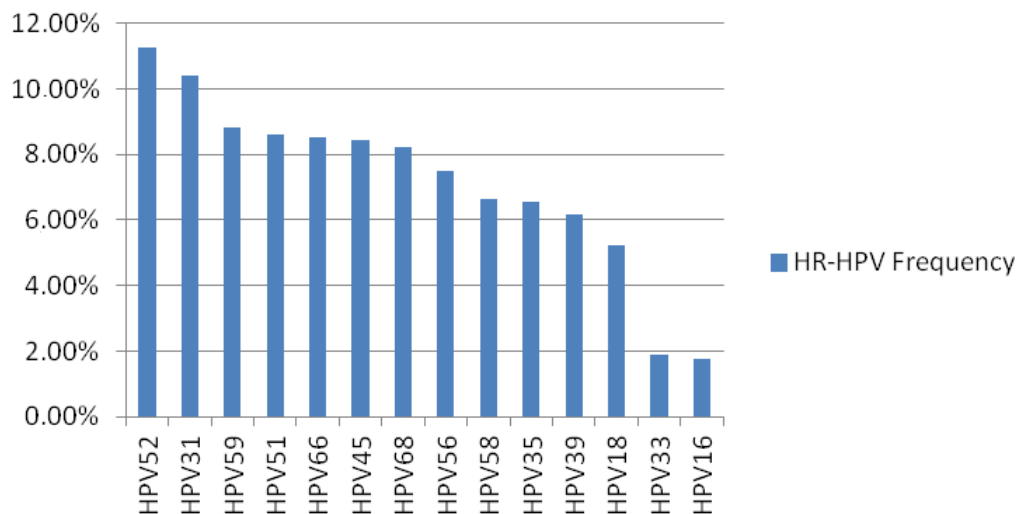
Figure 2. Frequency of high-risk HPV genotypes in women in the general population, sexually active without cervical lesions, infected with HR-HPV in five countries in the West African sub-region.

Source: Dr A. K. OUATTARA
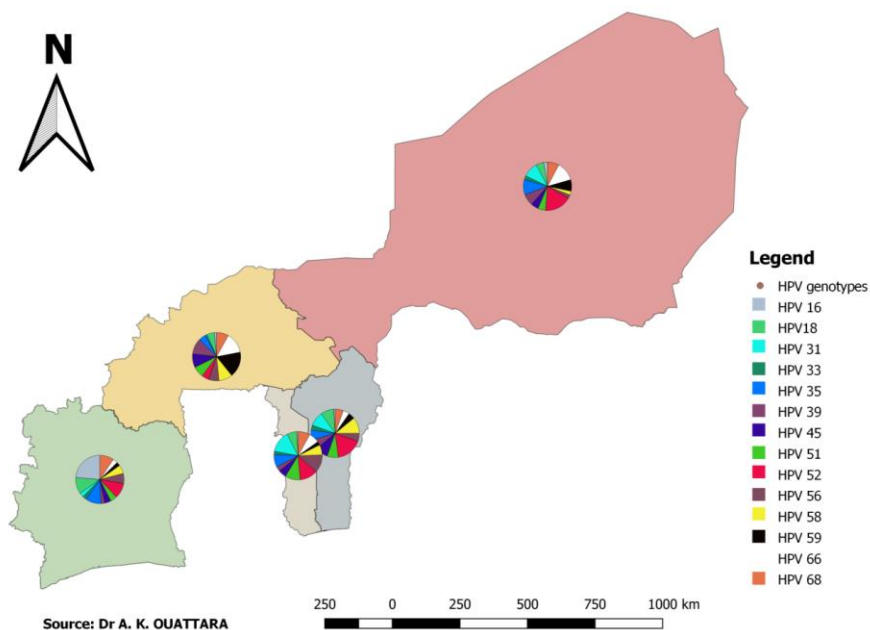
Figure 3. Mapping of fourteen HR-HPV genotypes in five West African countries: Benin, Burkina Faso, Cote d'Ivoire, Niger and Togo.

infection and presence of VIA/VILI dysplasia positive, were significantly associated with HR-HPV infection. Table 4 presents the risk factors for HPV infection in women in our study.

DISCUSSION

In our study, the overall prevalence of high-risk HPV infection among women in nine (09) cities in five (05) West African countries was 33.61% (717/2133). The prevalence of high-risk HPV infection in each country varied. According to the literature, the prevalence of HPV infection varies according to region; in West Africa this prevalence is estimated at 21.5% (WHO/ICO, 2009), which is lower than those in our study: either the general prevalence or those found in each of the five countries in our study except the prevalence found in Niger. It should
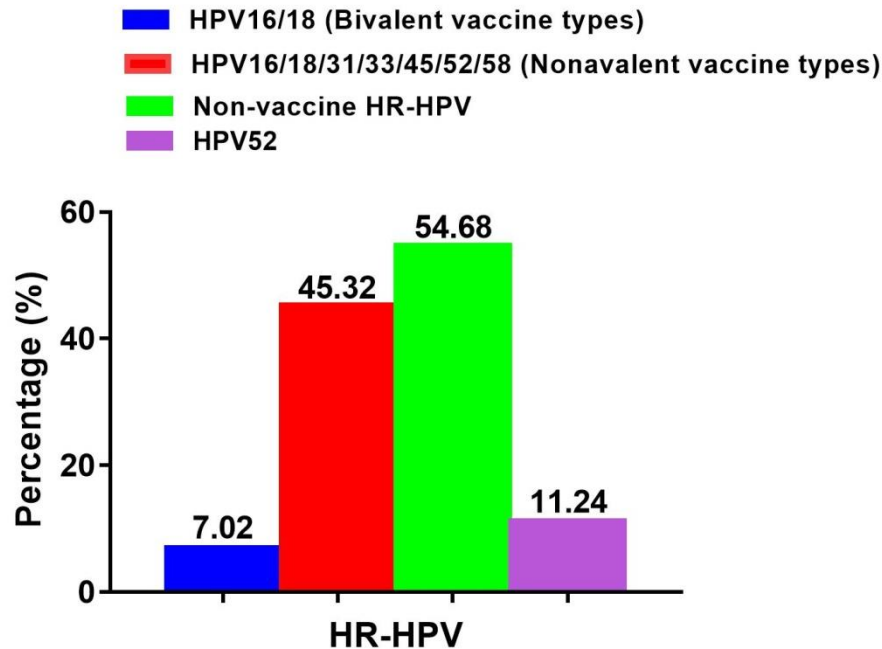
Figure 4. Frequency of HR-HPV genotypes: bivalent vaccine types, nonavalent vaccine types, non-vaccine HR-HPV and HPV52 in women of the general population in West Africa (N = 2133).

Table 3. Multiple infection with 4, 5 and 6 high-risk HPV genotypes.

Multiple infections with 4, 5 and 6 genotypes	Number of employees
HPV16/31/31/39/59/59/58/68	1
HPV16/51/51/52/56/66	1
HPV31/35/35/45/52/68	2
HPV45/52/52/58/59/66	1
HPV58/45/52/66/59	2
HPV16/39/45/59	1
HPV16/18/45/52	1
HPV18/45/52/58	4
HPV31/51/51/56/59	1
HPV31/33/33/35/52	4
HPV31/51/51/56/59	2
HPV35/51/52/58	1
HPV35/58/66/68	1
HPV33/51/51/56/68	1
HPV58/59/59/66/68	1
HPV51/52/52/56/68	1
Total	25

be noted, however, that the data in this study are 9 years old and that the prevalence of HPV has certainly changed positively since 2010. This is worrying in the sense that the prevalence of cervical cancer could also change in parallel.

Our prevalence are consistent with those of other African authors such as: 42.6% in Ghana (Obiri-Yeboah et al. 2017), 54% in South Africa (Adler et al., 2013), 76% in Tanzania (Watson-Jones et al. 2013). In contrast, Wang et al. (2018) reported a general prevalence of HPV

Table 4. Risk factors for HPV infection among women in the general population in nine cities in five West African countries.

Factors	HPV+ N = 717 (%)	HPV- N = 1416	OR (CI 95%)	P value
Age groups (n = 2133)				
15-24 years, n = 262	118 (45.04)	144		
25-34 years, n = 833	332 (39.86)	501	0.6 (0.38 - 0.95)	0.03
35-44 years, n = 686	229 (33.38)	457	0.5 (0.28 - 0.72)	< 0.001
45-54 years, n = 256	83 (32.42)	173	0.4 (0.26 - 0.73)	0.001
55-65 years, n = 96	50 (52.08)	46	Ref.	Ref.
Civil status (n = 2133)				
Unmarried, n = 497	271 (54.53)	226	2.5 (2.00 - 3.11)	< 0.001
Divorced, n = 10	7 (70.00)	3	4.1 (1.03 - 16.7)	0.02
Widow, n = 65	27 (41.53)	38	-	-
Married, n = 1561	752 (48.17)	809	Ref.	Ref.
Education (n = 2133)				
No, n = 536	178 (33.21)	358	0.5 (0.34 - 0.60)	< 0.001
Primary, n = 394	125 (31.73)	269	0.4 (0.31 - 0.58)	< 0.001
Secondary, n = 820	308 (37.56)	512	0.5 (0.42 – 0.71)	< 0.001
University, n = 383	200 (52.23)	183	Ref.	Ref.
Occupation (n = 2133)				
Purple/student, n = 176	87 (49.43)	89		
Housewives, n = 820	223 (27.20)	597	0.5 (0.38 - 0.64)	< 0.001
Informal sector, n = 650	249 (38.31)	401		
Salaried, n = 487	210 (43.12)	277	Ref.	Ref.
Sexual partners (n = 2133)				
0, n = 24	9 (37.50)	15		
1, n = 1789	622 (34.77)	1167		
2, n = 218	152 (69.72)	66	4.6 (2.72 – 7.86)	< 0.001
3, n = 101	37 (36.63)	64	Ref.	Ref.
4, n = 1	0 (0.0)	1		
Contraception (n = 2133)				
Yes, n = 841	295 (35.08)	546	0.8 (0.67 – 1.0)	0.03
No, n = 1292	483 (37.41)	809	Ref.	Ref.
Sexual intercourse (n = 2133)				
Thrice a week, n = 541	221 (40.85)	320	Ref.	Ref.
Twice a week, n = 947	276 (29.14)	671	0.6 (0.48 – 0.74)	< 0.001
Once a month, n = 207	106 (51.21)	101	1.5 (1.10 – 2.10)	0.01
Condom (n = 1744)				
No, n = 1183	380 (32.12)	803	Ref.	Ref.
Often, n = 482	195 (40.46)	287	1.4 (1.15 – 1.79)	0.001
Sometimes, n = 79	40 (50.63)	39	2.2 (1.37 - 3.42)	0.001
STD history, n = 1276				
Yes, n = 675	327 (48.44)	348	1.8 (1.47 – 2.31)	< 0.001
No, n = 601	203 (33.77)	398	Ref.	Ref.

Table 4. Contd.

HIV history, n = 623				
Yes, n = 49	23 (46.94)	26	1.8 (1.02 – 3.29)	0.04
No, n = 574	187 (32.57)	387	Ref.	Ref.
VIA (n = 2133)				
Positive, n = 188	128 (68.9)	60	4.0 (2.80 – 5.61)	< 0.001
Negative, n = 1945	684 (35.2)	1261	Ref.	Ref.
VILI (n = 2133)				
Positive, n = 235	194 (82.55)	41	10.0 (6.86 -14.49)	< 0.001
Negative, n = 1898	615 (32.40)	1283	Ref.	Ref.

Legend, Ref. chosen as reference.

infection of 14.5% in China, which is lower than in our study. In our study, the 25 to 34 year of age group was 39.05% or 813/2133 and was the age group of women most affected by HPV infection. These infected young women are the most vulnerable layer, and therefore at risk of developing cervical cancer later on. According to some authors, with regard to age-specific prevalence, young women in the 20-25 age group have the highest prevalence (>20%). Prevalence then declines rapidly with age, reflecting the most often transient nature of HPV infection. According to these authors, this decrease is much more pronounced in countries with high socio-economic levels and in these countries, prevalence is less than 10% beyond the 30 to 35 age group. In addition, they report that a re-augmentation is generally observed in women of menopausal age, without the causes of this increase being clearly established (De Sanjosé et al., 2007; Louie et al., 2008).

Another study on the carrying of HPV infection in women in the general population reported extremes of age from 17 to 68 years in China, which is similar to ours, that is 15 to 65 years (Wang et al., 2018). The fourteen high-risk HPV genotypes investigated in our study were all identified. The cumulative number of genotypes cytology had reported the presence of HPV genotypes 18, 31, 39, 45, 16, 35, 52 and 58 (Zohoncon et al., 2016a; Zohoncon et al., 2016b). Chen et al. (2018) had identified HPV16, 18, 58, 52, 33, 31, 68, 45, 66 and HPV 39 in invasive cervical cancers and the most predominant were HPV16, 18, 58 and 52. Other studies have reported the presence of HPV genotypes 35, 52, 52, 31, 58, 58, 59, 39, 39, 51, 51, 56, 16, 18, 33, 45, 66 in women in the general population (Zohoncon et al., 2013; Ouedraogo et al., 2015; Traore et al., 2016; Obiri-Yeboah et al., 2017, Ouedraogo et al., 2018). In women with normal cytology in Pakistan, Aziz et al. (2018) reported the presence of high-risk HPV genotypes such as HPV45 (12.5%), HPV33 (8.33%), HPV18 (6.25%) and HPV16 (4.16%). The prevalence of HPV16/18 (bivalent vaccine types) in our study was 7.02% in women of the general population in West Africa. ICO/IARC reported that prevalence of

identified in high-risk HPV-infected women was 1068 genotypes in total. HPV52 was the most common genotype followed by HPV31, HPV59, HPV51, HPV66, HPV45, HPV68, HPV56, HPV58, HPV35, HPV39, HPV18, HPV33 and HPV16. All these genotypes are at high oncogenic risk. Wang et al. (2018) reported the presence of HPV16, HPV58, and HPV52 genotypes as the most common genotypes among women in China. Other authors such as Yuan et al. (2019) reported that HPV52 and HPV58 infection are as common as HPV16 infection.

The importance of knowing the high-risk HPV genotypes circulating in our countries lies in the fact that these genotypes are oncogenic and their involvement in cervical cancer is well established. The presence of these high-risk HPVs in the West African population in our study merits preventive action. The presence of high-risk HPV genotypes in cervical cancers historically confirmed in the West African region must also be taken into account, but genotypes present in the general population should not be overlooked. However, some studies in Benin and Burkina Faso on high-risk HPV genotypes involved in cervical cancer and histologically confirmed precancerous lesions in anatomy and pathological HPV 16 and/or HPV 18 among women with normal cytology in subregion Western Africa was 4.3% (ICO/IARC, 2018). These studies report a relatively low frequency of HPV16 and 18 and a predominance of other high-risk HPV genotypes.

When considering the two types of female population such as women in the general population or women without cervical lesions and women with cervical cancer, the finding seems to be that the accumulation of other high-risk HPV genotypes is higher compared to the accumulation of HPV16 and 18 except in Europe or in some countries of the world where HPV16 and 18 are found in 70% of cervical cancer cases (ICO/IARC, 2018). Cervical cancer is one of the few cancers that can be prevented by controlling HPV infection through prophylactic vaccination and screening / early diagnosis / treatment. The currently available vaccines such as

Cervarix (bivalent HPV 16 and 18), Gardasil 4 (quadrivalent HPV16, 18, 6 and 11) and Gardasil 9 (nonavalent HPV 16, 18, 31, 33, 45, 52, 58, 6 and 11) are the prophylactic vaccines that different countries use. HPV vaccines have the potential to reduce the incidence of cervical and other anogenital cancers. But the choice of vaccine should be directed towards effective control of this scourge.

In addition, in our study, 34.87% of women infected with HPV had multiple infections and the number of high-risk HPV genotypes per woman ranged from 1 to 6. Multiple infections with 2 and 3 genotypes were in the majority, with 161/250 (64.40%) and 70/250 (28.00%) respectively. The frequency of high-risk multiple HPV infection in our study is higher than those reported by other authors: 14.9% in China (Wang et al., 2018), 19.8% in the United States (Monsonego et al., 2015), 24.3% in Italy (Panatto et al., 2013) and lower than the 48.1% reported by Kavanagh et al. (2013). These differences can be explained by the size of the study populations, the type of population, the number of genotypes sought and the risk factors. Several multiple infections in this study (for example HPV16/31/39/59/58/68; HPV16/51/52/56/66; HPV31/35/45/52/68; HPV16/39/45/59; HPV16/18/45/52; HPV18/45/52/58) are combinations of HPV genotypes found in invasive cervical cancer, hence the importance of focusing on genotypes found in the general population.

In this study, risk factors were significantly associated with HPV infection among women. Other studies have also noted these risk factors that influence or increase the risk (Monsonego et al., 2015; Aziz, et al., 2018; Ouedraogo et al., 2018). This mapping of high-risk HPV genotypes circulating in women in the general population shows diversity in the distribution of genotypes and raises questions about effective prophylactic actions to control HPV. However, mapping high-risk HPV genotypes in invasive cervical cancer in West Africa would strengthen this control.

Conclusion

This study provides us with a mapping of high-risk HPV among sexually active women in the general population in nine cities in five West African countries. It shows a predominance of the HPV 52 genotype followed by HPV 31, 59, 51, 66, 45, 68, 56, 58, 35, 39, 18, 33 and 16 and a prevalence of high-risk HPV infection ranging from 12 to 50%. The HPV genotypes predominant in the general population in West Africa are not HPV16 and 18. Is it due to viral clearance or genetic mechanisms? As cervical cancer is one of the few preventable cancers, it is crucial to place emphasis on prophylactic vaccination against broad spectrum HPV, adapted to the African context.

ACKNOWLEDGEMENTS

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Adler D, Laher F, Wallace M, Grzesik K, Jaspan H, Bekker LG, Gray G, Valley-Omar Z, Allan B, Williamson AL (2013). "High rate of multiple concurrent Human Papillomavirus infections among HIV-uninfected South African Adolescents." Journal of Immunological Techniques in Infectious Diseases 2(1):1000106.

Aziz H, Iqbal H, Mahmood H, Fatima S, Faheem M, Sattar AA, Tabassum S, Napper S, Batool S, Rasheed N (2018). "Human papillomavirus infection in females with normal cervical cytology: Genotyping and phylogenetic analysis among women in Punjab, Pakistan." International Journal of Infectious Diseases 66:83-89.

Chen Z, Zhou J, Chen Y, Zhu J (2018). "Distribution of human papillomavirus genotypes and its relationship to clinicopathology in invasive cervical carcinoma in Zhejiang Province, China." Journal of Cancer Research and Therapeutics 14(4):780-784.

De Sanjosé S, Diaz M, Castellsagué X, Clifford G, Bruni L, Muñoz N, Bosch FX (2007). "Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis." The Lancet infectious diseases 7(7):453-459.

Djigma FW, Ouedraogo C, Karou DS, Sagna T, Bisseye C, Zeba M, Ouermi D, Gnoula C, Pietra V, Ghilat-Avoid-Belem NW, Sanogo K, Sempore J, Pignatelli S, Ferri AM, Nikiema JB, Simpore J (2011). "Prevalence and genotype characterization of human papillomaviruses among HIV-seropositive in Ouagadougou, Burkina Faso." Acta Tropica 117(3):202-206.

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010). "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008." International journal of cancer 127(12):2893-2917.

ICO/IARC (Initial Coin Offering /International Agency for Research on Cancer) (2018). Human Papillomavirus and Related Diseases in Burkina Faso, ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre): 70.

Kavanagh K, Sinka K, Cuschieri K, Love J, Potts A, Pollock KG, Cubie H, Donaghy M, Robertson C (2013). "Estimation of HPV prevalence in young women in Scotland; monitoring of future vaccine impact." BMC Infectious Diseases 13(1):519.

Louie K, Didelot MN, Damay A, Nagot N, Mayaud P, Segondy M (2008). "Papillomavirus humains (HPV) et cancers associés: aspects épidémiologiques." Revue francophone des laboratoires 405:27-34.

Monsonego J, Cox JT, Behrens C, Sandri M, Franco EL, Yap PS, Huh W (2015). "Prevalence of high-risk human papilloma virus genotypes and associated risk of cervical precancerous lesions in a large US screening population: data from the ATHENA trial". Gynecologic Oncology 137(1):47-54.

Obiri-Yeboah D, Akakpo PK, Mutocheluh M, Adjei-Danso E, Allornuvor G, Amoako-Sakyi D, Adu-Sarkodie Y, Mayaud P (2017). "Epidemiology of cervical human papillomavirus (HPV) infection and squamous intraepithelial lesions (SIL) among a cohort of HIV-infected and uninfected Ghanaian women." BMC Cancer 17(1):688.

Ouedraogo CM, Djigma FW, Bisseye C, Sagna T, Zeba M, Ouermi D, Karou SD, Pietra V, Buelli F, Ghilat-Avoid-Belem NW, Sanogo

K, Sempore J, Moret R, Pignatelli S, Nikiema JB, Simpore J (2011). "[Epidemiology, characterization of genotypes of human papillomavirus in a population of women in Ouagadougou]." Journal de gynécologie obstétrique et biologie de la reproduction (Paris) 40(7):633-638.

Ouedraogo CM, Rahimy RM, Zohoncon TM, Djigma FW, Yonli AT, Ouermi D, Sanni A, Lankoande J, Simpore J (2015). "[Epidemiology and characterization of high-risk genotypes of human Papillomavirus in a population of sexually active adolescents in Ouagadougou]." Journal de gynécologie obstétrique et biologie de la reproduction (Paris) 44(8):715-722.

Ouedraogo RA, Zohoncon TM, Guigma SP, Angele Traore IM, Ouattara AK, Ouedraogo M, Djigma FW, Obiri-Yeboah D, Ouedraogo C, Simpore J (2018). "Oncogenic human papillomavirus infection and genotypes characterization among sexually active women in Tenkodogo at Burkina Faso, West Africa." Papillomavirus Research 6:22-26.

Panatto D, Amicizia D, Tanzi E, Bianchi S, Frati ER, Zotti CM, Lai PL, Bechini A, Rossi S, Gasparini R (2013). "Prevalence of human papillomavirus in young Italian women with normal cytology: how should we adapt the national vaccination policy?" BMC Infectious Diseases 13(1):575.

Traore IM, Zohoncon TM, Dembele A, Djigma FW, Obiri-Yeboah D, Traore G, Bambara M, Ouedraogo C, Traore Y, Simpore J (2016). "Molecular Characterization of High-Risk Human Papillomavirus in Women in Bobo-Dioulasso, Burkina Faso." BioMed Research International, 7092583.

Traore IMA, Zohoncon TM, Ndo O, Djigma FW, Obiri-Yeboah D, Compaore TR, Guigma SP, Yonli AT, Traore G, Ouedraogo P, Ouedraogo CMR, Traore Y, Simpore J (2016). "Oncogenic Human Papillomavirus Infection and Genotype Characterization among Women in Orodara, Western Burkina Faso." Pakistan Journal of Biological Sciences 19(7):306-311.

Wang X, Ji Y, Li J, Dong H, Zhu B, Zhou Y, Wang J, Zhou X, Wang Y, Peppelenbosch MP, Pan Q, Ji X, Liu D (2018). "Prevalence of human papillomavirus infection in women in the Autonomous Region of Inner Mongolia: A population-based study of a Chinese ethnic minority." Journal of Medical Virology 90(1):148-156.

Watson-Jones D, Baisley K, Brown J, Kavishe B, Andreasen A, Changalucha J, Mayaud P, Kapiga S, Gumodoka B, Hayes RJ, de Sanjose S (2013). "High prevalence and incidence of human papillomavirus in a cohort of healthy young African female subjects." Sexually Transmitted Infections 89(5):358-365.

WHO/ICO World Health Organization/Initial Coin Offering (2009). Human papilloma virus and related cancers in Ethiopia.

Yuan XW, Li YJ, Qiu Q, Luo ZY, Zhao XF(2019). "Prevalence and genotype distribution of human papillomavirus among 9945 women from the Nanhai area of Foshan". BMC infectious diseases 19(1):1-6.

Zohoncon TM, Bado P, Ouermi D, Traoré E, Djigma F, Traore IMA, Yonli AT, Ouédraogo C, Akpona SA, Lompo O (2016a). "Human papillomavirus genotypes involved in invasive cervical cancer from formalin-fixed, paraffin embedded tissues in Ouagadougou, Burkina Faso." International Journal of Current Research 8(9):39314-39318.

Zohoncon TM, Bisseye C, Djigma FW, Yonli AT, Compaore TR, Sagna T, Ouermi D, Ouedraogo CM, Pietra V, Nikiema JB, Akpona SA, Simpore J (2013). "Prevalence of HPV High-Risk Genotypes in Three Cohorts of Women in Ouagadougou (Burkina Faso)" Mediterranean Journal of Hematology and Infectious Diseases 5(1):e2013059.

Zohoncon TM, Ouedraogo TC, Brun LVC, Obiri-Yeboah D, Djigma WF, Kabibou S, Ouattara S, Gomina M, Yonli AT, Bazie V, Ouedraogo C, Lompo O, Akpona SA, Simpore J (2016b). "Molecular Epidemiology of High-Risk Human Papillomavirus in High-Grade Cervical Intraepithelial Neoplasia and in Cervical Cancer in Parakou, Republic of Benin." Pakistan Journal of Biological Sciences 19(2):49-56.

International Journal of Genetics and Molecular Biology

Full Length Research Paper

Distribution of groundnut rosette disease and sequence diversity of groundnut rosette virus associated satellite RNA (Sat-RNA) in Western Kenya

Benard Mukoye[1,2*], Anthony Simiyu Mabele[1], Millicent Florence Owuor Ndonga[1], Bonphace Collins Mangeni[1] and Hassan Karakacha Were[1]

[1]Masinde Muliro University of Science and Technology (MMUST), Kakamega, Kenya.
[2]Kenya Plant Health Inspectorate Service (KEPHIS), Nairobi, Kenya.

Production of groundnuts (*Arachis hypogaea* L.) in Western Kenya is mainly constrained by groundnut rosette disease (GRD) which cause up to 100% yield loss. This disease expresses different symptoms as a result of variations in the groundnut rosette virus (GRV) associated satellite-ribonucleic acid (GRV Sat-RNA). Over the past 20 years, no work had been done to document the status of the disease in Kenya. Additionally, no sequences of any of the GRD associated viruses were available in the GeneBank from Kenya. This study determined the distribution of GRD and the genetic diversity of GRV Sat-RNA. Sampling was done in main groundnut growing areas of Western Kenya during the long and short rain seasons in 2016/2017. Total RNA was extracted from the leafy samples collected using RNeasy Mini Kit (Qiagen) according to the manufacturers' protocol and used for double stranded cDNA synthesis using the SuperScript II kit. DNA libraries were sequenced on the Illumina MiSeq platform (Illumina). Reads were used for *de novo* assembly and contigs aligned to the viral genomes database using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database. Average GRD incidence was 53 and 41% in the short and long rain seasons, respectively. Chlorotic rosette was the dominant symptom followed by green rosette and mosaic. Nucleotide sequences of Sat-RNA revealed identities of 88 to 100% between the Kenyan isolates and those from Malawi, Nigeria and Ghana. All Kenya isolates clustered closest with green rosette variants of Malawi except one which clustered with chlorotic/yellow blotch variants. Rosette is widely distributed in Western Kenya and occurs wherever groundnuts are grown. The variations of GRD symptoms in Western Kenya could be due to the existence of different variants of Sat-RNA or other agents.

Key words: Groundnuts, satellite-ribonucleic acid (Sat-RNA), diversity, Western Kenya.

INTRODUCTION

Groundnut (*Arachis hypogaea L.)* is the fifth most important annual oilseed and food legume crop. It is grown in diverse environments throughout the semi-arid and sub-tropical regions, in nearly 100 countries, in the six continents of the world (Kumar and Waliyar, 2007).

Groundnut production is of great value in terms of income and nutrition for smallholder farmers in East Africa (Kidula et al., 2010; Okello et al., 2010). Resource poor smallholder farmers grow nearly 75 to 80% of the world's groundnuts in developing countries obtaining yields of

500 to 800 kg/ha, as opposed to the potential yield of >2.5 t/ha (Kayondo et al., 2014). In Western Kenya, an average of 600 to 700 kg/ha is achieved which is less than 30 to 50% of the potential yield (Kidula et al., 2010). Low yields are mainly attributed to poor quality seeds, drought, poor agronomic practices, numerous pests and diseases caused by numerous pathogenic viruses, fungi, bacteria and nematodes (Mutegi, 2010; Okello et al., 2010). Among the viral diseases, groundnut rosette disease (GRD) is the most devastating in sub-Saharan Africa (SSA) causing an estimated annual loss of US$156 million every year (Waliyar et al., 2007). The disease is caused by association between groundnut rosette assistor virus (GRAV), groundnut rosette virus (GRV) and a Satellite-RNA (Sat-RNA) of GRV (Taliansky and Robinson, 2003). Variations in Sat-RNA have been shown to be responsible for different rosette symptoms (chlorotic, green and mosaic rosette) (Taliansky and Robinson, 1997; Olorunju et al., 2001; Kayondo et al., 2014). Both chlorotic and green rosette symptoms occur throughout the SSA, and sometimes occur in the same field (Mugisa et al., 2016). A less common third symptom variant, called mosaic rosette, resulting from mixed infection by the Sat-RNA causing chlorotic and green mottled variant, has been reported from East Africa (Scott et al., 1996; Waliyar et al., 2007). Infected groundnut leaves may also show symptoms other than the typical chlorotic or green rosette (Naidu et al., 1999).

In Eastern Uganda, green rosette symptoms predominate (Okello et al., 2014). This is in contrast with the findings of Wangai et al. (2001), who reported that chlorotic rosette symptoms of GRD have been the predominant form throughout SSA and Western Kenya. The dynamics of the GRD virus symptomatology, therefore, needs constant monitoring. For example, in Nigeria, a there was shift from green to chlorotic rosette over a period of about 20 years. The shift could be due to changes in the genome sequences of GRD associated agents or other factors (Okello et al., 2014).

Survey done by Wangai et al. (2001) showed that GRD incidence ranged between 40% in areas of Western Kenya surveyed in the groundnut growing seasons of 1997-1998 and Sat-RNA shared 89 to 95% nucleotide identity with those from Malawi and Nigeria. Since then, no other survey has been conducted to ascertain the current status of GRD in the region. In addition, no genomic sequences of any of the GRD associated viruses existed in the GeneBank from Western Kenya. With the dynamics of the disease, this hinders proper diagnosis of GRD and development of management strategies. This study determined the distribution of GRD and assessed the sequence diversity of Sat-RNA of isolates from Western Kenya.

MATERIALS AND METHODS

Field survey

The GRD diagnostic sampling was conducted in all the major groundnut growing areas in Western Kenya during the short rains (October to December 2016) and long rains season (May to July, 2017). The following counties were covered: Bungoma, Busia, Homabay, Kakamega, Siaya and Vihiga. Sampling of groundnut farms was done by stopping at regular predetermined intervals, of 3 to 8 km along motorable roads that traverses each sampling area. The survey was conducted, by walking through groundnut field and visually inspecting groundnut crops for symptomatic leaves. Disease incidence was calculated according to Reddy (1991), as the percentage of plants showing GRD virus symptoms, to the total number of plants observed in the field. GRD viral incidence was scored using a rating scale according to Reddy (1991) where: low incidence = 1-20%; moderate incidence = 21-49%, and high incidence = 50-100%. The types of GRD symptoms observed were recorded. The collected data on GRD virus incidence and severity, was subjected to analysis of variance (ANOVA), using Statistical Analysis System (SAS) program version 9.3.1 software (SAS Institute, 2013). Pairwise comparisons of means were done using Least Significance Differences (LSD) for multiple-means comparison method at $P \leq 0.05$ confidence level.

Symptomatic and asymptomatic leaves were collected in 10 ml falcon tubes containing RNA*later*® RNA Stabilization Solution and put in a cool box. The samples were kept in the fridge and used for molecular studies. Geographical Positioning System (GPS) (entrex venture HC GARMIN^{TM}), was used to record the latitude, longitude and altitude of the sampled farms.

RNA extraction, sequencing and sequence analyses

Total RNA was extracted from the leaf samples using RNeasy Mini Kit **(**Qiagen**)** according to the manufacturers' protocol and used for double-stranded cDNA synthesis using the SuperScript II (Thermo Fisher Scientific, Waltham, USA) kit. The cDNA was column-purified with the DNA Clean and Concentrator^{TM}-5-DNA kit (Zymo Research, Irvine, USA). The samples were then processed with the transposon-based chemistry library preparation kit (Nextera XT, Illumina) following manufacturer's instructions. The fragment sizes structure of the DNA libraries was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA). The indexed denatured DNA libraries were sequenced (200-bp paired-end sequencing) on the Illumina MiSeq platform (Illumina).

Reads quality check was done using FastQC (version 0.11.5). Reads were then trimmed to remove poor quality sequences using Trimmomatic (V 0.36) software (Bolger et al., 2014). Trimmed reads (Haas et al., 2013) were used for *de novo* assembly and contigs aligned to the viral genomes database (ftp://ftp. ncbi.nih.gov/genomes/Viruses/all.fna.tar.gz/, downloaded on October 2017) using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database (Altschul et al., 1990). Complete and partial GRV Sat-RNA sequences used for comparison and phylogenetic analyses were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/). Phylogenetic analyses and comparisons were performed using the MEGA X (Kumar et al., 2018). Tobacco bushy top virus - KU997687.1 TBTV was used as an outgroup.

*Corresponding author. E-mail: btemukoye@gmail.com.

Table 1. Mean GRD incidence (%) per County.

County	Season	N	*Mean (%)	Std. error of mean (+/-)
Bungoma	Long rain	45	30.89[b]	4.534
	Short rain	47	66.51[b]	4.295
Busia	Long rain	74	43.36[b]	3.526
	Short rain	108	46.56[b]	2.728
Homabay	Long rain	73	48.60[b]	3.919
	Short rain	55	48.22[b]	4.025
Kakamega	Long rain	30	43.47[b]	5.283
	Short rain	17	94.12[a]	4.779
Siaya	Long rain	31	33.94[b]	4.820
	Short rain	26	43.23[b]	6.645
Vihiga	Short rain	20	47.50[b]	6.412
Total	Long rain	253	41.51	1.962
	Short rain	273	53.04	1.909

*Means with the same letter within the column are not significantly different.

Table 2: Percentage frequency of the occurrence of the three main GRD symptoms in western Kenya.

Symptom	Percent
Chlorotic rosette (CR)	58.6
Green rosette (GR)	27.4
Mosaic (MS)	14.1

RESULTS

A total of 526 farms were sampled in six (6) counties (253 in long rain and 273 in short rain). Generally, GRD incidence was high during the short rain season (53%) than the long rain season (41%) in all counties. High mean GRD incidence was recorded in Kakamega in the short rain season (68.92%) while moderate incidence was in Bungoma (30.89%) during the long rain season. There was a significant difference in GRD incidence among the counties (p=0.011, df=521, F=3.322). Siaya County had the overall lowest incidence which was significantly different from that of Kakamega but did not vary significantly from that of Bungoma, Busia, Vihiga and Homabay counties (Table 1).

Generally, GRD infected plants were dwarf with increased tillering although some were tall but expressed other major symptoms associated with GRD. The main symptoms observed in all counties in order of abundance, starting from the most common, were chlorotic rosette, green rosette and mosaic. Chlorotic rosette was recorded in 58.6% of the surveyed farms, green rosette in 27.4% while mosaic was observed 14.1% of farms (Table 2). Other symptoms observed were leaf rolling, upward leaf curling and severe leaf bunching (Figure 1). The distribution of the major symptoms is shown in Figure 2.

Diversity of GRV Sat-RNA

Six complete genomes of GRV Sat-RNA were assembled. The sequences varied slightly in the number of nucleotides (nt) ranging between 896 and 901 nt (Table 3).

The six Sat-RNAs from Kenya were then compared with those from the GeneBank. In the phylogenetic tree, all Kenyan isolates formed two distinct clusters together with Malawian isolates. Isolates E7 and E8 clustered with M11S, isolates BUG1-21, BG3-18 and KG8-1 clustered together with M16S while isolate EG16-5 is grouped with

Figure 1. Some of the virus-like symptoms observed in the surveyed fields. A: Dwarfed plant with green rosette; B: Severe chlorosis (yellow) on young leaves and dwarfing; C: Severe young leaf rolling and bunching on a dwarfed plant; D: Mosaic mostly on young leaves.

M24S. All Nigerian isolates clustered together is similar to Ghanaian isolates. Sequence identities between 88 and 100% of the Kenyan isolates and those from Malawi, Nigeria and Ghana were revealed. Very close identities between 92 and 100% were observed between the Kenyan isolates and those from Malawi, followed by Nigerian isolates (90-93%) and least with Ghanaian isolates (86-89%). Isolate BUG1-21 had 100, 99 and 98% identities with M16S, M12S, and M11S, respectively, which are all green rosette variants and 94% with M24S (chlorotic variant). While the other Western Kenya isolates (KG8-1, BUG1-21, BG3-18, E7 and E8) had 92 to 95% identity with Malawian isolate M24S (chlorotic rosette variant), isolate EG16-5 (Kakamega) showed the closest identity (97%) with this isolate. The same isolate EG16-5 was the only that clustered together with M24S, all chlorotic isolates (Z29702.1, Z29703.1) and yellow blotch (Z29710.1, Z29711.1). Isolates E7 and E8 were closest to Malawian isolate M11S with 97 and 99% identity, respectively. Isolates BG3-18 and KG8-1 were closest to Malawian isolates M16S displaying 97% identity (Figure 3).

DISCUSSION

Groundnut rosette disease is the most prevalent disease of groundnuts in Western Kenya. The disease was recorded in every county that was surveyed with incidences of up to 100%. The short rain season recorded higher incidence (53%) than the long rains (41%). This can be attributed to the high vector pressure during the short rains as compared to the long rains season when the aphid pressure is low as a result of heavy rains that wash the insects away. This concurs with the findings by Mugisa et al. (2016) that periods of long rains negatively affected GRD progression when aphid pressure is low.

All major GRD symptoms were observed in the surveyed region with chlorotic rosette being the most prevalent followed by green rosette. This concurs with Wangai et al. (2001) who reported chlorotic rosette to be the most prevalent GRD symptom in the region. The high prevalence of the chlorotic rosette could also be attributed to its higher transmission efficiency compared to green rosette. In a study, Misari et al. (1988a), reported
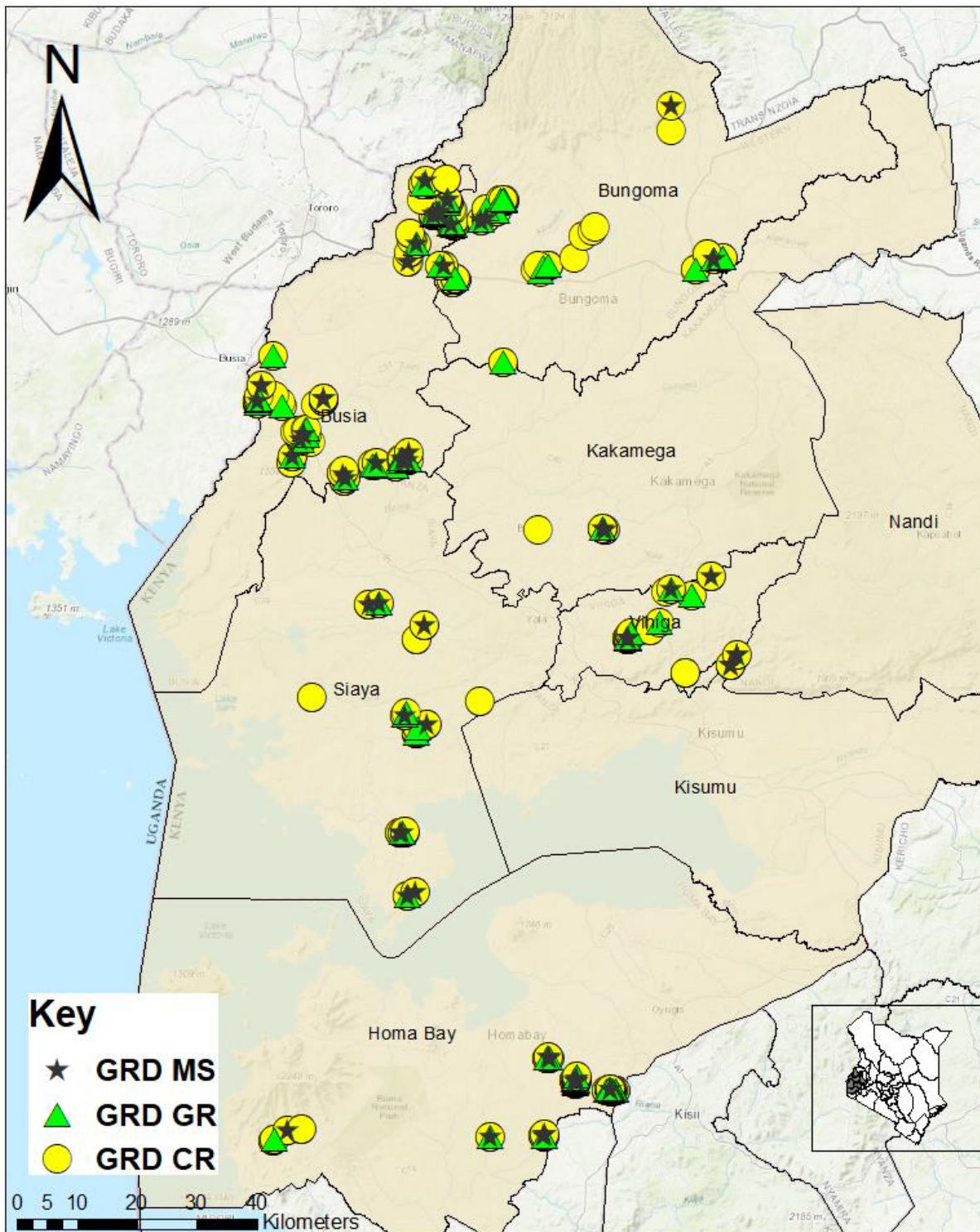
Figure 2. A map of Western Kenya showing the distribution of GRD symptoms in the surveyed counties. **MS**-Mosaic, **GR**-Green Rosette, **CR**-Chlorotic Rosette.

minimum acquisition feeding periods of 4 and 8 h for chlorotic and green rosette, respectively and the median latent periods of 26.4 and 38.4 h, respectively, for chlorotic and green rosette. In this study, a new symptom, the mosaic, which had not been previously reported in Western Kenya, was observed in most of the

Table 3. Description of the Sat-RNA sequences assembled.

Sample ID	Sat-RNA ID	Sequence length (nt)	County of origin
EG16	EG16-5	901	Kakamega
E7	E7	896	Siaya
E8	E8	897	Busia
BUG1	BUG1-21	901	Busia
KG8	KG8-1	898	Kakamega
BG3	BG3-18	901	Bungoma
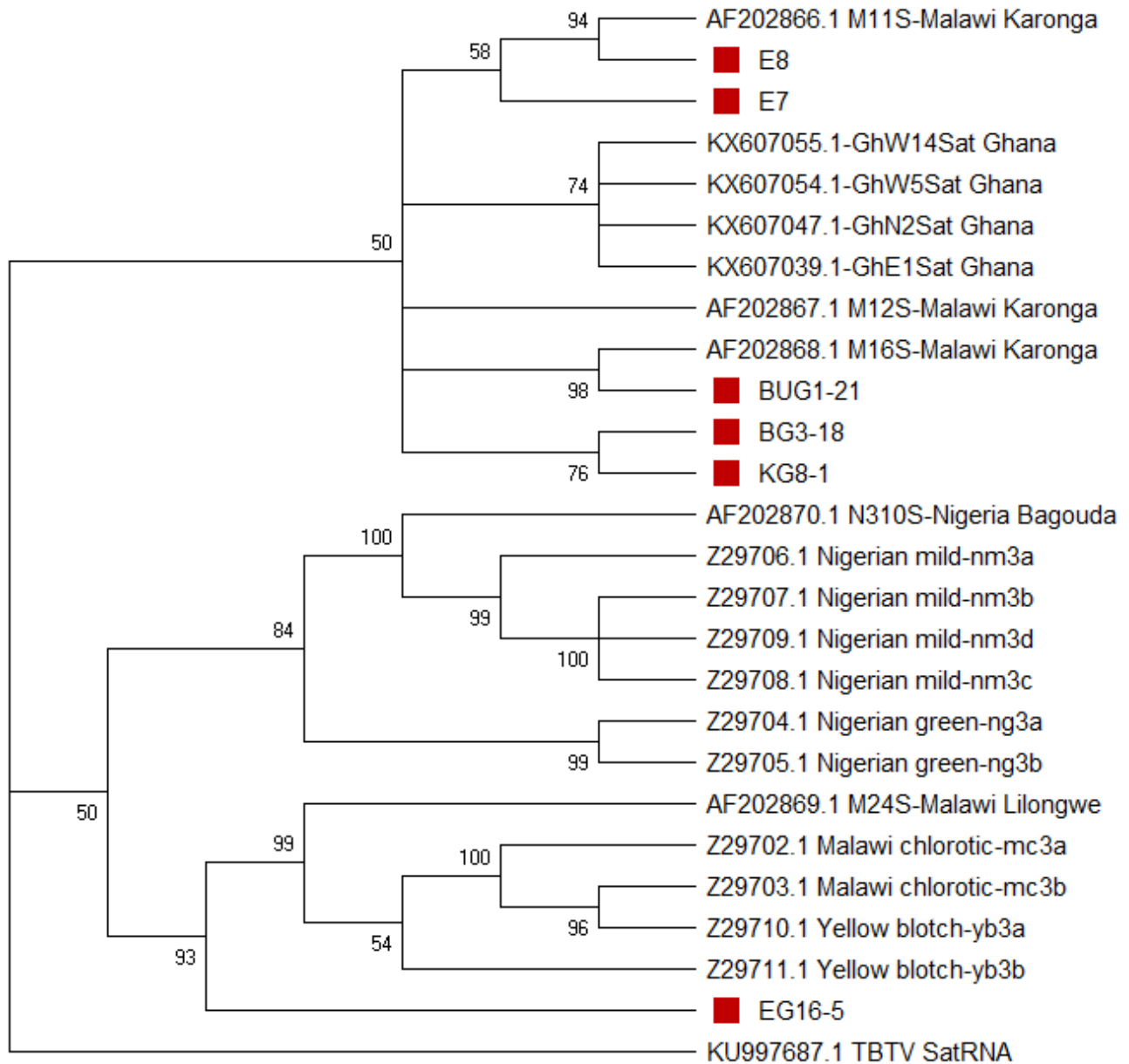
Figure 3. Phylogenetic tree of Western Kenya Sat-RNA and GeneBank isolates. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). The tree is rooted on Sat-RNA of a distantly related *Umbravirus* (Tobacco bushy top virus - KU997687.1 TBTV). Bootstrap confidence values (500 replications) are shown.

surveyed counties. This suggests that there is evolution of new variants of Sat-RNA that might be causing these new symptoms. A total of 10 variants of Sat-RNA have been reported as being associated with the various GRD symptoms (Blok et al., 1994). A mixture of either variants, especially the chlorotic and green rosette and/or the mild ones, are likely to induce the mosaic symptoms (Naidu et al., 1998). It is therefore possible that the variants of sat-RNA reported in this study occur in Western Kenya in mixed infections, thus causing the mosaic observed. It is worth noting that from the Next Generation Sequences (NGS) used in this study, order than GRV Sat-RNA, other viruses were detected (data not shown) and could be the reason for some of the new symptoms observed on groundnuts (Mukoye et al., 2018).

The Western Kenya Sat-RNAs sequences showed close identity (92-100%) to Malawian isolates than those from Ghana and Nigeria (88-93%). This implies that the genetic diversity of the Sat-RNA become more varied with wide geographical distance. Kenya and Malawi are located in Eastern Africa while Ghana and Nigeria are in West Africa thus having a wider geographical separation than Malawi. This finding is in line with Wangai et al. (2001) who observed a closer sequence relationship between Kenyan Sat-RNA isolates and those from Malawi. However, this study has reported sequence identity of up to 100% with Malawian isolates as opposed to 95% reported by Wangai et al. (2001). This suggests that more variants of Sat-RNA exist in Western Kenya that are contributing to the diverse symptoms expressed by GRD. Since this study used NGS which has been demonstrated to be more reliable in detection of new or poorly characterized viruses (Rott et al., 2017), it has revealed new variants of Sat-RNA in Western Kenya. Besides, there were variations among the Western Kenya Sat-RNA isolates similar to Malawian isolates where they formed distinct clusters in the phylogenetic tree. The isolate EG16-5 was the most distinct and clustered together with chlorotic and yellow blotch Sat-RNA variants. This suggests that this isolate is related to the chlorotic rosette symptom that was prevalent in the surveyed areas.

This study concludes that GRD is still the major viral disease of groundnuts in Western Kenya and occurs wherever groundnuts are grown in the region. The disease expresses varied symptoms with chlorotic rosette being the most prevalent form. The observed variations in the symptoms were due to the presence of diverse variants of the symptom inducing agent, Sat-RNA. The use of NGS has revealed that new variants of Sat-RNA exist in Western Kenya.

The six Kenyan Sat-RNAs have been deposited in the GeneBank with accession numbers LC469779, LC472299, LC472300, LC472301, LC472302 and LC472303.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. Journal of Molecular Biology 215:403-410.

Blok VC, Ziegler A, Robinson DJ, Murant AF (1994). Sequences of 10 variants of the satelitte- like RNA -3 of groundnut rosette virus. Virology 202:25-32.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8:1494-1512.

Kayondo SI, Rubaihayo PR, Ntare BR, Gibson PI, Edema R, Ozimati A, Okello DK (2014). Genetics of resistance to groundnut rosette virus disease: African Crop Science Journal 22:21-29.

Kidula N, Okoko N, Bravo-Ureta BE, Thuo M, Wasilwa L (2010). A preliminary analysis of yield differences in groundnuts between research and non-research farmers in Kenya. In paper presented at the 12th KARI biennial scientific conference, 8-12 November 2010, Naiobi Kenya.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35:1547-1549.

Kumar PL, Waliyar F (2007). Diagnosis and detection of viruses infecting ICRISAT mandate crops: Methods Manual. Patancheru 502 324, Andhra Pradesh, India; International Crops Research Institute for the Semi-Arid Tropics.133 p.

Misari SM, Abraham JM, Demski JW, Ansa OA, Kuhn CW, Casper R, Breyel E (1988a). Aphid transmission of the viruses causing chlorotic and green rosette diseases of peanut in Nigeria. Plant Disease 72:250-253.

Mugisa IO, Karungi J, Akello B, Ochwo-Ssemakula MKN, Biruma M, Okello DK, Otim G (2016). Determinants of groundnut rosette virus disease occurrence in Uganda. Elseviercropprotectionjournal.http://dx.doi.org/10.1016/j.cropro.2015.10.019.

Mukoye B, Mangeni BC, Sue J, Ndong'a MFO, Were HK (2018). Next generation sequencing as a tool in modern pest risk analysis: a case study of groundnuts (Arachis hypogaea) as a potential host of new viruses in Western Kenya. Kenya Plant Health Inspectorate Service 2nd phytosanitary conference, 4th – 8th June, 2018, Nairobi – Kenya.

Mutegi CK (2010). The extend of aflatoxin and aspergillus section flavi, penicillium spp. and Rhizopus spp. contamination of peanuts from households in Western Kenya and the causative factors of contamination. PhD dissertation, University of Kwazulu-Natal,Pietermaritzburg . South Africa.

Naidu RA, Bottenberg H, Subrahmanyam P, Kimminns FM, Robinson DJ, Thresh JM (1998). Epidemiology of groundnut rosette virus disease: Current status and future research needs. Annals of Applied Biology 132:525-548.

Naidu RA, Kimmins FM, Deom CM, Subrahimanyam P, Chiyeubekeza AJ, van der Merwe PJA (1999). Groundnut rosette: Avirus disease affecting groundnut production in Sub-Saharan Africa. Plant Disease 83:700-709.

Okello DK, Birima M, Deom CM (2010). Overview of groundnut

research in Uganda: Post, present and future. African Journal of Biotechnology 9:6448-6459.

Okello DV, Akello LB, Tukamuhabwa P, Odongo TL, Ochwo-Ssemakula M, Adriko J, Deom CM (2014). Groundnut rosette disease symptom types, distribution and management of the disease in Uganda. African Journal of Plant Science 8:153-163.

Olorunju PE, Ntare BR, Pande S, Reddy SV (2001). Additional sources of resistance to groundnut rosette disease in groundnut germplasm and breeding lines. Annals of Applied Biology 159:259-268.

Reddy DVR (1991). Groundnut viruses and virus diseases; Distribution, identification and control. Review of Plant Pathology 70:665-678.

Rott M, Xiang Y, Boyes I, Belton M, Saeed H, Kesanakurti P, Hayes S, Lawrence T, Birch C, Bhagwat B, Rast H (2017). Application of Next Generation Sequencing for diagnostic testing of tree fruit viruses and viroids. Plant Disease 101:1489-1499.

SAS Institute (2013). SAS/STAT 9.3.1: User guide. SAS Publishers, India.

Scott KP, Farmer MJ, Robinson DJ, Torrence L, Murant AF (1996). Comparison of the coat protein of groundnut rosette assistor virus with those of other *luteovirus*. Annals of Applied Biology 128:77-83.

Taliansky ME, Robinson DJ (1997). Trans-acting untranslated elements of groundnut rosette virus satelitte RNA are involved in symptom production. Journal of General Virology 78:1277-1285.

Taliansky ME, Robinson DJ (2003). Molecular Biology of umbraviruses: Phantom warriors. Journal of General Virology 84:1951-1960.

Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10:512-526.

Waliyar F, Kumar PL, Ntare BR, Monyo E, Nigam SN, Reddy AS, Osiru M, Diallo AT (2007). A Century of Research on Groundnut Rosette Disease and its Management. Information Bulletin no.75.Patancheru 502 324, Andhra Pradesh, India. International Crops Research Institute for the Semi-Arid Tropics 40 p. ISBN 978-92-9066-501-4.

Wangai AW, Pappu SS, Pappu HR, Okoko N, Deom CM, Naidu RA (2001). Distribution and characteristics of groundnut rosette disease in Kenya. Plant Disease 85:470-474.

(AJ) **ACADEMIC JOURNALS** expand your knowledge

International Journal of Genetics and Molecular Biology

Full Length Research Paper

First detection of the hemizygote frameshift variant of the DMD gene in a 13-year-old patient affected by dystrophinopathy (Duchenne DMD and Becker BMD) at Saint Camille Hospital of Ouagadougou, Burkina Faso: Case report

Théodora Mahoukèdè Zohoncon[1,2,3,4*], Paul Ouedraogo[1,4], Marius Belemgnegre[1], Abdoul Karim Ouattara[2,3], L. Bienvenue Tingueri[1] and Jacques Simpore[1,2,3,4]

[1]Saint Camille Hospital of Ouagadougou (HOSCO), 09 BP 444 Ouagadougou 09, Burkina Faso.
[2]Laboratory of Molecular Biology and Genetics (LABIOGENE), University Joseph KI-ZERBO, 03 BP 7021 Ouagadougou 03, Burkina Faso.
[3]Pietro Annigoni Biomolecular Research Center (CERBA), 01 BP 364 Ouagadougou 01, Burkina Faso.
[4]Faculty of Medicine, University Saint Thomas d'Aquin (USTA), 06 BP 10212 Ouagadougou 06, Burkina Faso.

Genetic diseases are poorly reported in sub-Saharan Africa, especially in Burkina Faso. The reasons for this reality are multifactorial including, the difficulty of diagnostic confirmation, financial accessibility and even the difficulty with referral of cases by medical staff. Genetic diseases, although relatively rare, exist in families and deserve special attention in our paraclinical assessments. Few cases can be diagnosed in sub-Saharan Africa. The clinical case we are reporting here is a dystrophynopathy by mutation of the Duchenne muscular dystrophy (DMD) gene. This is the first confirmed detection of a frameshift mutation in the DMD gene in a boy received at Saint Camille Hospital in Ouagadougou, Burkina Faso. This is a 13-year-old MR boy, from a family of four siblings, all male, from the eastern region of Burkina Faso. The boy had a clinical picture of myopathy with difficulty in walking, frequent falls, myogenic syndrome with stool sign, Gowers sign and scapula alata, all leading to a suspicion of dystrophinopathy with a request for genetic analysis. The DMD gene responsible for the disease is located on the X chromosome (Xp21.2-p21.1). The study of the dystrophin gene (DMD) was done using three methods, namely MLPA, high throughput sequencing and Sanger sequencing. The results led to the identification of a frameshift mutation of exon 71 in the DMD gene: it is a hemizygotic variant with ribosomal shift of the DMD gene NM_004006.2 (DMD): c.10258del p.(Ser3420Leufs*25). This clinical case led for the first time in Burkina Faso to the confirmed diagnosis of hereditary muscular dystrophy resulting from a mutation with a frameshift in exon 71 of the DMD gene in the hemizygotic state in a 13-year-old boy, a student and the eldest sibling of 4 boys, three of whom have myopathy.

Key words: Mutation, Duchenne muscular dystrophy (DMD) gene, dystrophinopathy, duchenne versus becker, case report.

INTRODUCTION

The frame shift mutation of the DMD gene is rarely reported in the literature in patients with dystrophinopathy. This is the first detection of a mutation with ribosomal shift in the DMD gene in a patient from Burkina Faso. This clinical case aims to report the existence of a rare mutation of the DMD gene in sub-Saharan Africa.

Duchenne and Becker muscular dystrophies (DMD and BMD, respectively) result from mutations in the DMD gene (MIM #300377) encoding a membrane cytoskeleton protein, dystrophin. These are genetic diseases linked to sex chromosomes X and variable clinical expression in women (usually healthy carriers) due to random inactivation of the X chromosome while male hemizygotes manifest the disease. It is transmitted recessively in women and dominant in men.

The alteration on the dystrophin sequence resulting in a total absence or partial deficiency (truncated protein and/or in decreased quantity) in dystrophin has consequences that are expressed in several ways. In skeletal muscles, there is a loss of membrane integrity of the muscle fibers responsible for muscular dystrophy (Muntoni et al., 2003; Straathof et al., 2015) while in cardiac muscle abnormal expression of dystrophin is responsible for isolated cardiomyopathy or most often associated with muscular dystrophy (Deburgrave et al., 2007; Ortez et al., 2019). On the other hand, cardiomyopathies linked to the absence of dystrophin can be noted both in DMD and/or BMD patients and in women who are said to transmit these pathologies (Bushby et al., 1993; Mori-Yoshimura et al., 2018).

DMD is clinically characterized by progressive muscle weakness, with an incidence of 1 in 3500 to 6000 men (Gospe et al., 1989). Similarly, BMD is characterized by high phenotypic variability, ranging from severe muscle weakness with cardiomyopathy and early death to exercise-induced myalgia and/or muscle cramps associated with an increase in serum creatine kinase (CK) activity (Bakker et al., 1997; Melacini et al., 1996; Rahimov and Kunkel, 2013; Kwiatkowska et al., 2020). However, cardiomyopathy has been found in BMD patients without skeletal muscle weakness (Serrano and Munoz-Canoves, 2017). These two genetic diseases (BMD and DMD) are caused by the mutation of the DMD gene, which codes for a protein called dystrophin, which is essential for structural muscle stability. Genetic mutations in the DMD gene are very heterogeneous and can be deletions (65%), duplications (5.10%) and point mutations (10 - 15%). The DMD gene is subject to high proportion of neo mutations. The DMD gene, located on the short arm of the X chromosome (Xp21.2-p21.1), is the largest human gene known to date. With a size of 2.2 million base pairs, it occupies 0.1% of the human genome and 1.5% of the X chromosome sequence. The 14 kilobase (kb) messenger RNA is composed of 79 exons that constitute only 0.6% of the overall gene sequence (Nakamura et al., 2015).

Major genomic rearrangements such as deletions or duplications of one or more exons are in the majority (70% of cases), and point mutations are found in 20 to 30% of cases (Gospe et al., 1989; Cripe and Tobias, 2013; Buddhe et al., 2018). The severity of the phenotype (DMD versus BMD) can be correlated with the impact of the mutation on dystrophin expression in more than 90% of cases. Exceptions to this so-called reading framework rule result from various mechanisms and are observed in less than 10% of cases. In addition to the benefit of having a confirmed diagnosis to the patient, identification of the mutation allows genetic counselling adapted to the family, and a potential inclusion of the patient in therapeutic trials based on the genotype (Gospe et al., 1989; Straathof et al., 2015).

At present, there is no curative treatment. Corticosteroids, currently indicated in the early stages of the disease, are the only ones that have shown a change in the natural history of the disease, independent of the genetic mutation cases (Cripe and Tobias, 2013).

CASE PRESENTATION

Demographic details

Our patient is a 13-year-old boy from a family of four children, all male, from the eastern region of Burkina Faso. The patient is a student in the 5th grade.

Medical history

The patient had a history of normal birth and psychomotor development until the age of 11 years when the first symptoms appeared and motivated a medical consultation. The first signs were of difficulty walking, and the physical examination had not noted any visible deformation. The hypothesis of sickle cell disease was raised and hemoglobin electrophoresis was requested resulting in the detection of AC genotype.

He had a second medical consultation two months after the first one with the reason being difficulty walking and fatigability for 1 month. The physical examination noted:

*Corresponding author. E-mail: zohoncont1@yahoo.fr. Tel: +226 25361232. Fax: +226 25363242.

Table 1. Biological and cardiac profiles of the MR patient and family.

Persons	Age (years old)	CK (U/L)	LDH (U/L)	AST (U/L)	ALT (U/L)	Echocardiography
RM Patient	13	8171		109.3	130	Normal
Patient's mother	33	295	759	43	47	Normal
Patient's father	35	125	313	34	39	Normal
1st Brother of the patient	9	6184	489	77	192	Normal
2nd Brother of the patient	5	20167	907	276	248	3 mm perimembrane IVC with a left-right micro-shunt
3rd Brother of the patient	1	120	243	10	32	Normal

CK, serum creatine kinase; LDH, lactate dehydrogenase ; AST, Aspartate aminotransferase ; ALT, alanine aminotransferase ; IVC, interventricular communication.

Good general condition, normal colored and anicteric mucous membranes, decreased motor strength in the lower limbs with motor strength (FM) = 3/5 in both limbs and predominant in the thigh muscles. No sensory disturbances. Osteotendinous reflexes (ROT) were reduced. The diagnostic differential of myopathy and neuropathy was made. The patient was referred to neurology department.

About ten days later, a medical consultation was conducted for progressive weakness of the lower limbs associated with frequent falls. The examination notes a good general condition, a myogenic syndrome with stool sign, Gouers sign, scapula alata. The diagnosis of limb-girdle muscular dystrophy was hypothesized. Paraclinical examinations were requested: CK = 6351 U/L with standards between 80 and 200 U/L (i.e. more than 30 times normal), ALT = 167 U/L, AST = 102 U/L, creatinine = 40.84 micromol/L,hemogram shows a white blood cell count at 6090 cells/µL; hemoglobin level = 12.4 g/dL; platelet count = 441000/µL. The electrocardiogram, echocardiogram, electroneuromyogram, and blood ionogram were without particular abnormality. The medical prescription was Défal cp 30 mg ½ cp per day in the morning; Kaleoride cp 600 mg ½ cp at noon; Albendazole cp 1cp/jr for 3 days; Prazol capsules 20 mg 1 capsule in the morning and a low sodium diet. There was no secondary effect related to drug treatment.

Under this treatment, the child during neurological follow-up and CK monitoring had elevated CK levels at 8171 U/L; AST at 109.3 U/L; ALT at 130 U/L. This high CK level motivated the consultation, which found a myogenic syndrome with stool sign, Gowers sign and scapula alata with the notion of difficulty walking, frequent falling and fatigability. The cardiac Doppler echo was within normal limits. We therefore continued with a study of the DMD gene in search of a mutation. The request for the study of the DMD gene was preceded by a family interview looking for events related to dystrophinopathy in the maternal family. The child was put under kinesitherapy sessions.
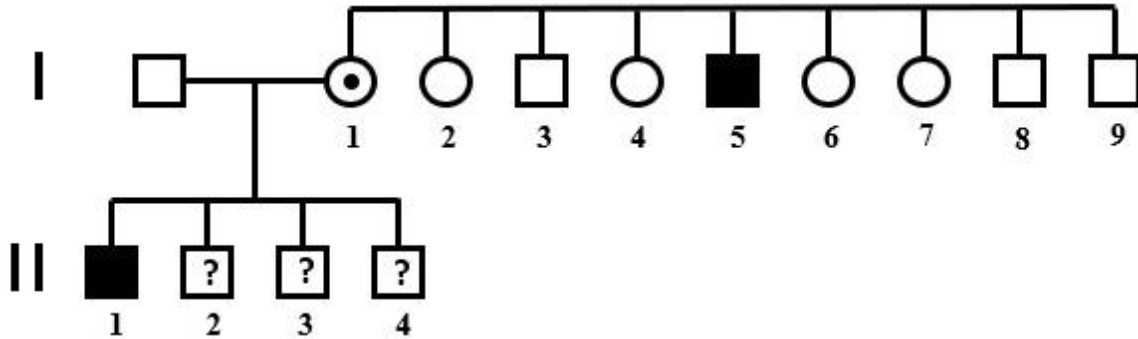
Family history

There is a notion of similar undocumented disorders in our patient's maternal uncle (one of the brothers of the MR patient's mother's mother). This uncle was said to have had difficulty walking, frequent falls with an inability to recover resulting in disability at the age of 30 years. Indeed, the mother of the MR patient comes from a sibling of 9 children, including 4 boys (including the uncle who has difficulty getting up) and 5 girls (including the patient's mother). The MR patient's mother's biological test scores found CK at 295 U/L and LDH at 759 U/L. While the CK of the MR patient's father had 125 U/L and LDH = 313 U/L. Of the three brothers of the MR patient, respectively aged 9, 5 and 1 year old, two had high CK values (Table 1). The 5-year-old boy with high CK (20167 U/L) also had a small 3 mm perimembrane interventricular communication (IVC) on the cardiac Doppler with a left-right micro-shunt without impact on the cardiac chambers (Table 1). The latter with a very high CK value was placed on the same treatment as the MR patient.

The pedigree of the parents and the MR patient is shown in Figure 1.

Genetic analysis of the DMD gene in the MR patient

Genomic DNA was extracted from peripheral blood leucocytes according to standard protocols.

The identification of the mutation of the hemizygous frame shift variant of the DMD gene NM_004006.2(DMD): c.10258del p. (Ser3420Leufs*25) in our MR patient was performed by three molecular biology and genetics techniques such as Multiplex Ligation-dependant Probe Amplification (MLPA), High throughput sequencing using next-generation sequencing (NGS) and confirmation of the DMD variant by targeted Sanger sequencing; at the Biochemistry and Molecular Biology Laboratory Grand Est of the University Hospital of Lyon in France. The first

Figure 1. DMD patient's family history and pedigree, □, Male ; ○Female ; ⊙Heterozygote female ; ■ Affected male ; ⍰,Probably affected male.

step consisted of the search for the most frequent lesions: deletions and duplications of exons using MLPA.

Multiplex ligation-dependant probe amplification:

This technique allows the number of copies for each of the 79 exons to be estimated using standardized probes that highlight both deletions and duplications. This fast and reliable technique using SALSA KIT P034/P035 MRC-Holland, is one of the most commonly used in routine, both for index cases and for the status search of women at risk: it solves the majority of quantitative anomalies on the DMD gene. DNA quality can sometimes be a limiting criterion for good interpretation and the anomaly of an isolated exon must always be controlled by another technique (Muntoni et al., 2003).

The MLPA analysis initially performed in our MR patient did not detect any duplication or deletion of the DMD gene. Then high throughput sequencing was performed in search of mutations.

High throughput sequencing

This technique was performed using preparation of the Kappa Nimblegen library and Illumina NextSeq™500 Sequencing System (at a reading length of 2 × 150 bp). Considering the family medical history of MR, a panel of genes was compiled, which have been reported in literature to be linked to muscular dystrophy: Limb-girdle muscular dystrophy (LGMD) is a genetically and clinically heterogeneous group of rare muscular dystrophies. The list of genes in the "LGMD main genes" panel considered was: ANOS, CAPN3, CAV3, DMD, DYSF, FKRF, SGCA, SGCG.

Computer analysis Home-made Pipeline "Papillyon": produced sequences aligned and compared to the

human reference genome (GRCh37/hg19). Targeted analysis of the DMD gene (NM_004006.2) was done. The intronic exons and borders of the 79 exons of the DMD gene are covered with a depth greater than 30 X. This high throughput sequencing analysis allowed us to detect in our MR patient, a hemizygotic frameshift variant of the DMD gene NM_004006.2 (DMD): c.10258del p. (Ser3420Leufs*25) located on exon 71. This variant is absent from LOVD but has been reported in the literature in a patient with severe Becker phenotype (Straathof et al., 2015; Mornet and Rivier, 2017).

The confirmation of the DMD variant was done by targeted Sanger sequencing. The pathogenic variant was confirmed by conventional dideoxy sequencing using Big-Dye Terminators (Life Technologies) after specific amplification. Sanger sequencing is a long and costly but effective technique, serving as a reference technique to validate results provided by new high throughput sequencing techniques. Sanger sequencing: Amplification by the "classical" PCR technique then bidirectional capillary sequencing on ABI 3500xL. The sequences produced were compared with the human reference genome (GRCh37/hg19) with SeqScape v3, Life Technologies software. The search for mutation of the DMD gene in the MR patient was positive with the identification of a hemizygous frameshift variant of the DMD gene NM_004006.2(DMD): c.10258del p. (Ser3420Leufs*25) located on exon 71. The discovery of this variant made it possible to diagnose dystrophynopathy in the MR patient (Table 2). A muscle biopsy was not obtained to determine the exact phenotype (DMD or BMD) by immunohistochemistry and Western blot. No Adverse and unanticipated events were noted.

DISCUSSION

Duchenne muscular dystrophy (DMD) and Becker

Table 2. Point mutation identified in DMD gene in RM patient.

Position change	Type of mutation	Nucleotide change	Amino acid change
Exon 71	Deletion : Frameshift	c.10258del	p.Ser3420Leufs*25

muscular dystrophy (BMD) are monogenic X-linked recessive disorder and are both located on Xp21.2-p21.1, the short (p) arm of the X chromosome between positions 21.2 and 21.1 (OMIM*300377).

Cases of genetic diseases are underestimated in sub-Saharan Africa due to a lack of access to diagnostic means of confirmation. The case of our patient, who is a young boy, a student (with normal cognitive development), clinically presenting a myopathy with a suspicion of Duchenne or Becker dystrophynopathy, could have gone unnoticed for the confirmation of the effective existence of a mutation of the DMD gene. In addition, a rare mutation has been identified in our patient. The MLPA analysis had not noted any deletion or duplication; and it was the high throughput sequencing that made it possible to detect this rare mutation, a frameshift variant of the DMD gene in the hemizygous state located on exon 71, considered pathogenic and part of the new mutations of the DMD gene (Mornet and Rivier, 2017).

This variant is absent in the Leiden Open Variation database (LOVD) but has been reported in the literature in a patient with a Severe Becker phenotype and normal cognitive development, but a muscle biopsy could not be performed for the RNA study (Mornet and Rivier, 2017). In addition, Straathof et al. found this pathogenic novel mutation: a frameshift mutation in exon 71 (c.10258del, p.ser3420-Leufs25*) in one of their patients who also had a severe BMD phenotype with normal cognitive development (Straathof et al., 2015).

This genotype/phenotype discrepancy could be explained by a catch-up of the exon 71 skipping reading frame on some of the transcripts. This mechanism has been reported for some frameshift variants in exon 71. These are frameshift mutations c.193delG and c.4922insT in two patients with BMD in whom muscle biopsy followed by Western blot shows a slightly reduced size of the dystrophin protein cases (Cripe and Tobias, 2013).

The deletion is frame shift, so we expect to have a DMD profile with the absence of the dysthrophin protein. But with regard to the clinic, particularly the age at which symptoms begin, such as difficulty walking (after the age of 10 years), the phenotype would be that of a BMD with partial production of dystrophin. The consequences and life expectancy are not the same for these two forms, hence the indication of a muscle biopsy is necessary to affirm DMD or BMD. In our MR patient we did not obtain a muscle biopsy to determine the exact phenotype (DMD or BMD) by immunohistochemistry and Western blot. The MR child was put on kinesitherapy and cardiac, respiratory and Achilles tendon monitoring is required. The mother of the MR patient is probably a carrier and the other children will have the same follow-up.

Dystrophin is one of the largest proteins in the body, located on the inner surface of the skeletal muscle fibre, where it provides a structural link between the cytoskeleton and the sarcolemma, via a set of proteins called the "dystrophin associated protein complex" (CPAD) (Péréon et al. 2015).

The clinical signs of myopathy noted in our patient are explained by the involvement of the dystrophy gene. The mechanical contraction action of muscle fibers generates micro lesions of the sarcolemma, in the absence of the link between the muscle fiber membrane and myofibrils represented by dystrophin and CPAD (Campbell, 1995; Petrof et al., 1993). However, we did not observe any cardiac muscle abnormalities or cognitive impairment in the MR patient.

It is established that in Duchenne muscular dystrophy (DMD), not only is there a lack of dystrophin, but the expression of the proteins that make up the CPAD is also much reduced. The excessive fragility of the membrane with respect to mechanical stress and the increase in its permeability, the dysregulation of calcium homeostasis, the disturbances of NO synthase, oxidative stress, all of which result from the loss of dystrophin, are at the origin of muscle necrosis, followed by an initially effective regeneration. Over time, it becomes exhausted, inflammation and endomysial fibrosis set in, with the progressive replacement of muscle fibres by fibroadipular tissue and the loss of muscle function (Péréon et al., 2015).

Genetic counselling was offered and carried out to the family of the MR patient with multidisciplinary follow-up. This clinical case reports dystrophinopathy (Duchenne DMD versus Becker BMD) for which the confirmatory genetic diagnosis is not available in Burkina Faso. This is the first detection of a rare mutation of the DMD gene to Frameshift type mutation in a 13-year-old boy (from a sibling of four boys, three of whom have biological signs of myopathy) from Burkina Faso, at Saint Camille Hospital in Ouagadougou.

This potentially fatal mutation was detected following high throughput sequencing and confirmed by Sanger sequencing. This genetic disease requires a multidisciplinary team and the appropriate treatment is

still being investigated.

ACKNOWLEDGEMENTS

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ABBREVIATIONS

MR, Malo RABO (pseudonym of the patient); **DMD,** Duchenne muscular dystrophy; **BMD,** Becker muscular dystrophy; **MLPA,** Multiplex-ligation dependent probe amplification; **CPAD,** Complex of dystrophin-associated proteins; **CK,** Creatine kinase; **AST,** Aspartate aminotransferase; **ALT,** Alanine aminotransferase; **IVC,** interventricular communication; **CERBA,** Pietro Annigoni Biomolecular Research Center; **LABIOGENE,** Laboratory of Molecular Biology and Genetics; **HOSCO,** Saint Camille Hospital of Ouagadougou; **USTA,** University Saint Thomas d'Aquin.

REFERENCES

Bakker E, Jennekens FG, De Visser M, Wintzen AR (1997). Duchenne and Becker muscular dystrophies In: Diagnostic criteria for neuromuscular disorders, pp. 1-4.

Buddhe S, Cripe L, Friedland-Little J, Kertesz N, Eghtesady P, Finder J, Hor K, Judge DP, Kinnett K, McNally EM, Raman S, Thompson WR, Wagner KR, Olson AK (2018). Cardiac management of the patient with duchenne muscular dystrophy. Pediatrics 142(2):S72-S81.

Bushby KM, Gardner-Medwin D, Nicholson LV, Johnson MA, Haggerty ID, Cleghorn NJ, Harris JB, Bhattacharya SS (1993). The clinical, genetic and dystrophin characteristics of Becker muscular dystrophy. II. Correlation of phenotype with genetic and protein abnormalities. Journal of Neurology 240(2):105-112.

Campbell KP (1995). Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. Cell 80(5):675-679.

Cripe LH, Tobias JD (2013). Cardiac considerations in the operative management of the patient with Duchenne or Becker muscular dystrophy. Pediatric Anesthesia 23(9):777-784.

Deburgrave N, Daoud F, Llense S, Barbot JC, Recan D, Peccate C, Burghes AH, Beroud C, Garcia L, Kaplan JC, Chelly J, Leturcq F (2007). Protein- and mRNA-based phenotype-genotype correlations in DMD/BMD with point mutations and molecular basis for BMD with nonsense and frameshift mutations in the DMD gene. Human Mutation 28(2):183-195.

Gospe SM, Lazaro RP, Lava NS, Grootscholten PM, Scott MO, Fischbeck KH (1989). Familial X-linked myalgia and cramps: a nonprogressive myopathy associated with a deletion in the dystrophin gene. Neurology 39(10):1277-1280.

Kwiatkowska J, Meyer-Szary J, Bazgier M, Fijałkowska J, Wierzba J, Glińska A, Dorniak K (2020). Left ventricular volumes and function affected by myocardial fibrosis in patients with Duchenne and Becker muscular dystrophies: a preliminary magnetic resonance study. Kardiologia Polska, pp. 1-10.

Melacini P, Fanin M, Danieli GA, Villanova C, Martinello F, Miorin M, Freda MP, Miorelli M, Mostacciuolo ML, Fasoli G, Angelini C, Dalla Volta S (1996). Myocardial involvement is very frequent among patients affected with subclinical Becker's muscular dystrophy. Circulation 94(12):3168-3175.

Mori-Yoshimura M, Mitsuhashi S, Nakamura H, Komaki H, Goto K, Yonemoto N, Takeuchi F, Hayashi YK, Murata M, Takahashi Y, Nishino I, Takeda S, Kimura E (2018). Characteristics of Japanese Patients with Becker Muscular Dystrophy and Intermediate Muscular Dystrophy in a Japanese National Registry of Muscular Dystrophy (Remudy): Heterogeneity and Clinical Variation. Journal of Neuromuscular Diseases 5(2):193-203.

Mornet D, Rivier F (2017). What future for dystrophin? *Les cahiers de myologie* (15):17-21.

Muntoni F, Torelli S, Ferlini A (2003). Dystrophin and mutations: one gene, several proteins, multiple phenotypes. The Lancet Neurology 2(12):731-740.

Nakamura A (2015). X-Linked Dilated Cardiomyopathy: A Cardiospecific Phenotype of Dystrophinopathy. Pharmaceuticals (Basel) 8(2):303-320.

Ortez C, Natera de Benito D, Carrera Garcia L, Exposito J, Nolasco G, Nascimento A (2019). Advances in the treatment of Duchenne muscular dystrophy. Medicina 79(3):77-81.

Péréon Y, Mercier S, Magot A (2015). Duchenne muscular dystrophy pathophysiology. Archives de pédiatrie. 22(12Suppl1):12S18-23

Petrof BJ, Shrager JB, Stedman HH, Kelly AM, Sweeney HL (1993). Dystrophin protects the sarcolemma from stresses developed during muscle contraction. Academy of Sciences of the United States of America 90(8):3710-3714.

Rahimov F, Kunkel LM (2013). The cell biology of disease: cellular and molecular mechanisms underlying muscular dystrophy. Journal of Cell Biology 201(4):499-510.

Serrano AL, Munoz-Canoves P (2017). Fibrosis development in early-onset muscular dystrophies: Mechanisms and translational implications. Seminars in cell and developmental biology 64:181-190.

Straathof CS, Van Heusden D, Ippel PF, Post JG, Voermans NC, De Visser M, Brusse E, Van Den Bergen JC, Van Der Kooi AJ, Verschuuren JJ, Ginjaar HB (2015). Diagnosis of becker muscular dystrophy: Results of Re-analysis of DNA samples. Muscle and Nerve 53(1):44-48.

International Journal of Genetics and Molecular Biology

Full Length Research Paper

Verification of single nucleotide polymorphism (SNP) markers associated with maize (*Zea mays. L*) streak virus resistance in early generation maize lines

Solomon Shibeshi[1]*, V.O. Adetimirin[1], Lava Kumar[2], Menkir Abebe[2] and Melaku Gedil[2]

[1]Pan African university, Life and Earth Science Institute (Including Health and Agriculture), Department of Agronomy, University of Ibadan, Ibadan, Oyo State, Nigeria.
[2]International Institute of Tropical Agriculture, Department of virology, Ibadan, Oyo State, Nigeria.

Maize streak virus (MSV) is a devastating viral disease of maize in sub-Saharan Africa. The disease may cause up to 100% yield loss in susceptible crops. The use of molecular markers can facilitate the development of varieties resistant to the virus. The objective of this study is to assess the usefulness of Single Nucleotide Polymorphism (SNP) markers linked to MSV resistance in selecting for resistance at early generations of inbred line development in maize. A total of 160 maize lines were genotyped with three SNP markers that are linked to MSV resistance. These lines were tested for their reaction to MSV through artificial inoculation using viruliferous *Cicadulina triangular* at the three leaf stage; maize streak virus symptom was scored from 7 days after inoculation for six weeks at weekly interval on a scale of 1 to 5. MSV titer on the upper and lower leaves was determined using Direct Antigen Coating Enzyme linked Immunosorbent Assay (DAC-ELISA). One hundred and forty-two (142) of the 160 maize lines had the favourable marker allele for MSV resistance while 18 maize lines did not have the allele. Differences among the 160 maize lines for MSV symptom on upper leaves at six week after inoculation were significant *(P< 0.01)*. Favourable allele of the SNP markers was significantly associated with MSV symptom score at 6 week after inoculation and MSV titer status. The percentage of maize lines with desirable marker allele with resistance based on symptoms score and ELISA were 97.9 and 93%, respectively. The three SNP markers showed high efficiency in the identification of MSV resistant maize lines and therefore have potential for use in marker-assisted selection. The SNP markers were not effective in detecting MSV resistance in few genotypes, indicating a need to develop other markers for resistance.

Key words: ELISA, genotyping, inoculation, maize streak virus, SNPs.

INTRODUCTION

Maize (*Zea mays. L*) is among the most important cereals globally, along with wheat and rice, providing basic diet to millions of people in sub-Saharan Africa (SSA) (Gebrekidan et al*.,* 1992). Despite its importance, the

production and productivity are hampered by many biotic and abiotic stresses. Abiotic constraints include drought, low soil fertility, soil acidity, and expensive farm inputs, among others. The major biotic stresses include weeds such as striga, pests such as stem borers and weevils, and diseases caused by fungi and viruses, all of which contribute to substantial yield losses (Cairns et al., 2012). Viral diseases have been a major threat among which "streak disease" or maize streak virus disease (MSVD) is by far the most important; a widespread virus responsible for high grain yield loss in SSA (Guthrie, 1978). Maize Streak Virus Disease (MSVD) is considered to be the third most important disease of maize in the world, after gray leaf spot and northern corn leaf blight (Pratt and Gordon, 2006).

MSVD, caused by the maize streak virus (MSV, genus: *Masterevirus*, family: *Geminiviridae*) is transmitted by leafhoppers of the genus *Cicadulina.* The disease is characterized by yellow streaks which run parallel to the leaf veins. In susceptible maize lines severe infection may result in stunting, inter-veinal necrosis, chlorosis, and death of affected individuals. Streaks develop along the veins on most of the leaf lamina and, as the virus is systemic, symptoms appear only on the infected and subsequent leaves (Thottappilly et al., 1993).

Similar to many insect-transmitted virus diseases, the annual incidence of MSVD in farmers' fields varies from insignificant in some years to epidemic proportions in others (Efron et al., 1989). Outbreaks of the disease are often associated with drought conditions or irregular rains such as those that occurred in West Africa in 1983 and 1984 (Rossel and Thottappilly, 1985). MSVD can result in 100% yield loss in susceptible maize lines (Magenya et al., 2008). Yield losses vary with the time of infection and varietal resistance (Guthrie, 1978). Field trials relying on natural infection in East Africa detected yield losses between 33 and 56% (Guthrie, 1978), whereas losses of 100% were reported in many countries in West Africa (Fajemisin et al., 1986). Trials conducted between 1983 and 1985, presented by Fajemisin et al. (1986), reported a yield reduction of 71 to 93% in maize due to MSV.

It is difficult to manage MSVD due to its variability and unpredictable vector migratory patterns (Vivek et al., 2010). Furthermore, there exist grasses which are host reservoirs for both the insect vector and the virus (Vivek et al., 2010). Traditionally, the disease can be controlled through cultural, chemical and physical measures (Wambugu and Wafula, 2000). However, chemical sprays can only kill the insect vector found within a maize field at a given time, but since the leafhoppers are migratory insects and can travel several miles, the use of chemicals can be very ineffective especially because not all fields are sprayed at the same time. In addition, use of chemicals has been regarded as environmentally unsafe and requires continues monitoring as more insects migrate back after the chemical loses potency (Njuguna et al., 1990). On the other hand, use of cultural measures such as crop rotation, early planting, intercropping is not efficient enough to control this insect transmitted virus.

The development and deployment of resistant varieties is a more appropriate and cost- effective approach to controlling MSVD. Significant progress has been made in breeding maize for resistance to MSVD through conventional methods. These methods involve crossing of the best plants possessing the most desirable traits such as high yield, disease resistance, or any other character that is preferred by farmers. With this method, it may take over eight years to produce a variety that has acceptable levels of resistance, and there is usually no guarantee that the resistance will hold for different virus strains. In most cases, resistance has been reported to break down in different environments, partly due to difference in the strains of the pathogen and mode of inheritance of the resistance, and partly due to maize line-by environment interactions (Njuguna et al., 1990).

Development and application of marker assisted selection (MAS) in crop improvement has become a useful technique for breeders. Since MSVD resistance trait has a high heritability and is controlled by a few genes (Welz et al., 1998), the application of markers in transferring gene for resistance is quite possible and quicker to assay than in conventional breeding. Molecular markers can help to select individuals carrying target genes in a segregating population based on patterns of tightly linked markers rather than on their phenotypes.

The Quantitative Trait Loci (QTLs) for resistance to MSVD have been identified and mapped in maize. Welz et al. (1998) detected a major QTL on chromosome 1 and minor QTL on chromosome 2, 3 and 4. The presence of a major QTL for resistance to MSVD on the short arm of chromosome 1 has been replicated by other authors (Lagat et al., 2008); this QTL explains 50 to 60% of phenotypic variation, although other minor QTLs have also been implicated. Alleles at this locus were additive or partially dominant depending on the resistance source (Redinbaugh et al., 2004). Due to the availability of high-throughput single nucleotide polymorphism (SNP) detection and validation technologies, the development of SNP markers is becoming a routine process, especially in crops with reference genome.

Identification of candidate genes linked markers can be used for forward breeding for MSV resistance in maize. Since the germplasms developed at IITA (International Institute of Tropical Agriculture) have recovery type of resistance, that is, plants which upon virus infection develop severity symptoms but the symptom severity is reduced in the leaves that develop subsequent to virus infection, thereby allowing plant to recover from infection. However, other QTLs may confer resistance. Consequently, there is a need to investigate the effectiveness of the known SNP markers in maize germplasm having recovery type of resistance for MSVD resistance. The objective of this study was to assess the usefulness of SNP markers in selecting MSV resistance

in early generation maize line development.

MATERIALS AND METHODS

Maize germplasm and insect vector sources

The experiment was carried out in the laboratories and screen house at the International Institute of Tropical Agriculture (IITA), Ibadan (190 m; Latitude 7° 29'11.99"N, longitude 3° 54'2.88"E), Nigeria. The location has mean and minimum temperatures of 26.5 and 21.4°C, respectively, a mean relative humidity of 74.6%. The maize lines used for the experiment were obtained from IITA Maize Breeding Unit. A total of 160 maize lines with known pedigree were used in the experiment. Of the 160, 151 lines were selected from 3003 S_1 lines developed from nine different bi-parental crosses; 9 inbred lines which were resistant to MSV were included in the experiment. A highly MSV-sensitive maize Line (Pool 16) was used as MSV susceptible check. All the maize lines were genotyped using three SNP markers (snpZM0020-PZE101093951, snpZM0021-PZE0186065237 and snpZM0022-PZE0186365075) linked to the gene for MSV resistance. The markers which were developed at CIMMYT (International Maize and Wheat Improvement Center, 1985) detect the presence or absence of favorable allele for MSV resistance gene on Chromosome 1 (Nair et al., 2015). Leafhopper cultures and MSV isolates used in this study were obtained from the IITA Virology and Molecular Diagnostic Unit at Ibadan. The insects were bred artificially in cages on pearl millet *(Pennisetum typhoides)* seedlings, and all the young adults used were able to transmit MSV following their feeding on infected host plants.

Experimental design

The 160 maize lines were arranged in a completely randomized design (CRD) with two replications. Each maize line was considered a treatment. Eight seeds of each maize line and four seeds of the susceptible maize line (control) were sown per pot on March, 2018. After germination the control line was thinned to one, each pot had 1 plant of pool 16 as a susceptible check.

MSV inoculation and disease assessment

Plants of the test entries and susceptible controls were inoculated under screen house condition using viruliferous leafhoppers. Leafhopper colonies reared on pear millet seedlings were transferred onto MSV-infected maize maintained in cages for a virus acquisition access period (AAP) of 48 h. Viruliferous leafhoppers from maize were subsequently transferred onto one week old seedlings of test lines maintained in cages and allowed 48 h inoculation access period (IAP) to facilitate virus inoculation of test plants. This procedure allowed uniform inoculation of all the maize lines. The inoculated plants were removed from the cage followed by insecticide spray to kill insect vectors. The inoculated plants were transferred to an insect-proof screen house for observations. MSV infection and symptom severity score was assessed based on visual observation of inoculated plants for six weeks at weekly intervals, beginning from seventh day after inoculation. A 1 to 5 rating scale was used as described by Beyene et al. (2012) where 1 = < 10% of the leaf area covered with streaks, 2 = 11-25% of the leaf area covered with streaks, 3 = 26-50% of the leaf area covered with streaks, 4 = 51-75% of the leaf area covered with streaks and 5 = >75% of the leaf area covered with streaks. The scores were used to define resistance categories as follows: 0=Immune, 1.0-2.0= resistant, 2.1-3.0 = moderately resistant, 3.1-

4.0 = susceptible and 4.1-5.0 =highly susceptible.

Sample collection and MSV detection by DAC-ELISA

Leaf samples were collected 42-days after inoculation from the oldest leaves available at the lower portion of the plants for leaf symptoms and virus titer in early formed leaves; and the youngest leaves at the upper portion of the plants were sampled for symptoms and virus titer in newly formed leaf, which were used for virus detection using enzyme-linked immunosorbent assay (ELISA). The samples collected were immediately wrapped in aluminum foils to prevent dehydration and placed in labelled transparent polythene zip lock bags. Thereafter, the samples were transported to laboratory on ice. Samples were tested by Direct Antigen Coating (DAC)-ELISA as described by Peterschmitt et al. (1992).

The DAC-ELISA was the serological virus detection method used to determine the relative titer of MSV in each tested maize lines. DAC-ELISA, also referred as Antigen Coated Plate (ACP)-ELISA involved coating of viral antigen to the ELISA plate surface and the antigen was used for virus detection with primary antibody (anti-MSV polyclonal antibody, sourced from IITA); the antigen-antibody complex was detected by the enzyme-labeled secondary antibody (anti-rabbit antibody) tagged with an enzyme (alkaline phosphate). The reaction of antigen and antibody was detected using chromogenic substrate p-nitrophenoyl phosphate and the color intensity was measured at an absorbance of 405 nm using a 96-well spectrophotometer. About 100 mg of maize leaf sample from the lower and upper leaves obtained from the maize lines were ground with 0.5ml of coating buffer (1.59 g of Na_2CO_3, 2.93 g of $NaHCO_3$, 1 L of distilled water pH 9.6). 100 µl of the ground sample was dispensed into each well of a new ELISA plate and kept in a refrigerator for overnight incubation at 4°C. The plate was washed three times with PBS-Tween [2.38 g of Na_2HPO_4, 0.4 g of KH_2PO_4, 0.4 g of KCl, 16.0 g of NaCl, 2 L of Distilled water and Tween -20 to 0.2% (v/v) pH 7.4] allowing 3 min interval between washes. The MSV polyclonal antibody cross adsorbed with healthy maize extract was prepared and used at a final dilution of 1/5000 (v/v) and 100 µl was dispensed into wells of the ELISA plate. The plate was covered and placed in a humid chamber and incubated at 37 °C for 1 h.

After incubation, the plate was washed three time with-PBS-Tween. 1 µl of anti-rabbit enzyme with 15 ml conjugate buffer (PBS-TPO) at a ratio of 1/15000 (v/v) was prepared and 100 µl of the solution dispensed into each well of the ELISA plate and incubated at 37°c for 1 h. The plate was washed again three times with PBS-T and 100 µl of substrate was dispensed to each well of the plate to detect the positive reactions. Washing with PBS was done at each stage of the ELISA procedure after incubation for 1 h at 37 °c allowing three minutes interval between two washes to remove non-binded materials and to avoid contamination of sample. MSV-diseased, healthy and buffer samples were placed in each ELISA plate as control at the middle and edge wells of the plate. Each sample was tested in duplicate wells, and each plate included positive (sap extracted from known MSV infected tissue) and negative (healthy plant) controls from stock maintained at IITA.

Absorbance at 405 nm was measured after 30, 60, 90 and 120 min with microplate reader after the addition of substrate. An overnight reading was also done after keeping the microplate in refrigerator (4°C) overnight. Absorbance values were considered positive when the optical density readings were at least twice that of the mean of the negative control.

Disease severity assessment by image analysis

For accurate quantification of streak symptoms induced by MSV, leaf image captured by a digital camera was assessed using Leaf Doctor Software as described by Martin and Rybicki (1998). For

Table 1. Analysis of variance for the mean symptom severity score for tested maize lines over the five weeks of assessment after inoculation.

Source of variation	DF	Sum of squares	Mean square	F Value	Pr>F
Line	159	569.35	3.58	116.28**	<0.01
Week	4	598.32	149.58	4857.17**	<0.01
Line * Week	636	215.81	0.34	11.02**	<0.01
Error	800	24.64	0.03		
LSD$_{(0.05)}$	0.19				
CV	5.4				

DF= degree of freedom, CV= coefficient of variation.

each maize line, about 12 cm long portion at the middle part of the 5th leaf was captured. Scaled black cardboard background was used to maintain the size and quality of the picture. The Leaf Doctor software provided percentage of the diseased (streak area) and healthy portion (asymptomatic area) of the camera-captured portion of the leaf.

Data analysis

Analysis of variance was carried out on data collected symptoms severity on the leaves of inoculated plants, percentage of diseased portion of the leaf from leaf image analysis and ELISA test result using (PROC GLM in SAS program). Correlation analysis was carried out using PROC CORR in SAS. The area under disease progress curve (AUDPC) values was calculated for the disease severity scores using the formula of Wilcoxson et al*., (1975) as follows:

$$AUDPC = \Sigma \left(\frac{x_i + x_{i+1}}{2} \right) \cdot (t_{i+1} - t_i)$$

Where, x_i is the disease rating on date i and t_i is the time (in calendar days) on which x_i recorded; i= 1, 2…5

RESULTS

Genotyping for maize streak virus resistance with SNP markers

The maize lines showed variation in marker allele. Majority of the tested maize lines had favorable allele for MSV resistance. Of the total 160 maize lines tested, 142 had favorable allele for MSV resistance while 18 lines did not have the favorable allele after SNP genotyping; the 18 lines originated from SW5-S-C6-18-2-1-B-TZISTR1248, SW5-S-C6-18-3-1-B-TZISTR1248 and KS27-S-C3-2-6-2-B-1-TZISTR1262 bi-parental crosses.

MSV symptom scores of 160 maize lines and relationship with alleles of SNP markers

Artificial inoculation of MSV successfully induced disease development in the maize lines. Continuous, narrow chlorotic streaks appeared on secondary and tertiary leaf vein of the plants within 3 to 5 days after inoculation in all inoculated maize lines as well as Pool 16, the control. The mean symptom severity scores at weeks 2, 3, 4, 5 and 6 weeks after inoculations were 4, 3.8, 3.2, 2.8 and 2.3, respectively. Thus, there was a reduction in disease symptoms over time. In general, the disease symptoms reduced from the lower leaves to the upper leaves. Significant differences (*P<0.01*) in MSV symptoms severity were obtained among the 160 maize lines at six week after inoculation (Table 1).

On the basis of mean symptom score assessed at 42 days after inoculation, 61 of the lines had symptoms scores between 1.0 -2.0 (resistant), 88 had scores between 2.1-3.0 (moderate resistant), 8 were rated 3.1-4.0 (susceptible) and 3 had scores of 4.1-5.0 (highly susceptible) (Table 2). Of the 142 maize lines that had the favorable allele for MSV resistance, 61 lines had MSV symptoms scores in the range 1.0 - 2.0, and 78 had scores in the range 2.1-3.0 while 3 had scores in the range 3.1-4.0 (Table 2). None of the 18 maize lines without favorable allele for resistance to MSV had symptoms scores in the range 1.0- 2.0. Ten of the maize lines had MSV symptoms scores between 2.1 and 3.0; 5 had scores in the range 3.1-4.0, and 3 had scores in the range 4.1-5.0 (Table 2). It was interesting to observe that 10 of the 18 lines that did not have the favorable allele for resistance to MSV were moderately resistant. None of the maize lines were immune to MSV. For the susceptible control, 93.1 % of the plants tested in this study had symptoms score of 4.1-5.0 (highly susceptible) while 6.9 % had scores of 3.1-4.0 (susceptible). This distribution observed for Pool-16 was within expectation and indicates the effectiveness of the inoculation procedure, using *C.triangular*. The very low scores observed for some of the lines indicate the availability of promising germplasms that can be used to improve resistance to MSV in breeding programmes in Africa.

A Chi-square value of 52.25 obtained from the test of association between allele of the SNP markers and MSV symptoms was significant *(P<0.01)* indicating that the marker allele was associated with resistance to MSV in

Table 2. Distribution of maize lines reaction to maize streak virus based on MSV score at 6 week after inoculation and desirable alleles of SNP markers.

| Parameter | Total no. of maize lines | | Classification based on SNP markers | | | |
| | | | No favorable allele | | Favorable allele | |
MSV severity score class	Frequency	%	Frequency	%	Frequency	%
1.0 - 2.0	61	38.1	0	0.0	61	43.0
2.1 -3.0	88	55.0	10	55.5	78	54.9
3.1 - 4.0	8	5.0	5	27.8	3	2.1
4.1 -5.0	3	1.9	3	16.7	0	0.0
Total	160	100	18	100	142	100

Table 3. Distribution of percentage recovery of 160 maize lines tested under artificial inoculation of MSV.

| Symptom severity score | % Recovery | | | |
	No. of lines	Range	Mean	%
1.0 - 2.0	61	33.1-77.2	59.4	38.1
2.1 -3.0	88	0.0-52.9	33.6	55
3.1 - 4.0	8	0.0-36.7	22.2	5
4.1 -5.0	3	0.0-10.0	5.0	1.9
Total	**160**			**100**

the population studied. Majority of the maize lines tested were moderately resistant to MSV following successful inoculation. Of the 142, maize lines with the favorable allele for resistance, 139 were resistant, a success rate of 79.9%. However, the success rate in identifying susceptible maize lines was 44.4% (8 out of 18). These results indicate that the SNP markers were useful in identifying majority of resistant maize lines; however, they were not able to identify all maize lines resistant to MSV. Ten maize lines that did not have the favorable allele for resistance and which on the basis of the marker allele would have been regarded as susceptible, had some moderate level of resistance.

**Percentage recovery from disease and area under disease progress curve**

Some lines had high percentage recovery while recovery for other lines was low. Percent recovery varied from 0 to 77.2% (Table 3). Amongst the 160 tested lines, 58 lines had above 50% recovery percentage whereas 102 lines had lower than 50% recovery percentage. Of the 58 lines with high percentage recovery only one line did not carry the favorable allele for MSV resistance markers. Among the maize lines with the favorable marker allele for resistance, 43% showed high recovery percentage (33.1-77.2), 54.9% showed moderate recovery (23.9-52.9) while 2.1% did not show appreciable recovery (4.5-23.3).

There were significant differences ($P<0.01$) among the maize lines for AUDPC value. Low AUDPC value indicates high resistance while high AUDPC value infers the susceptibility. Entries 152,153,158, 17 and 145 had the lowest AUDPC values of 43.8, 50.1, 53.2, 54.7 and 55.2 respectively. These lines showed high resistance to MSV, with a considerable decline in symptoms severity scores (Figure 1). Entries 115, 105, 112,140 and 110 had the highest AUDPC values of 139.1, 137.1 135.6, 126.7 and 122.5 respectively, and were therefore regarded as susceptible (Figure 2). In effect, the trends in symptoms severity scores of highly resistant lines were different from the trends observed for the highly susceptible lines. This difference provides unique information on the nature of resistance in the lines studied, a type of resistance in which host plants limit the spread and development of symptoms with growth.

The MSV damage of lines from the bi-parental crosses also showed considerable variation. The number of lines developed from each of the nine bi-parental crosses varied from 2 to 47. Among the bi-parental crosses, the lowest mean MSV damage score was observed for lines developed from IITATZi1715/TZISTR1262 (score = 2.8), followed by lines derived from BBB/KU1409/SC55-KU1409)-S2-19-1-BBB-17-B-3-B-TZiSTR1262 (score = 2.9). Lines from these crosses are potential sources of resistance to MSV in Africa. Among the inbred lines, TZiSTR1100 had the lowest MSV symptom score of 1.9. Maize lines developed from SW 5 (S) C6-18-2-1-
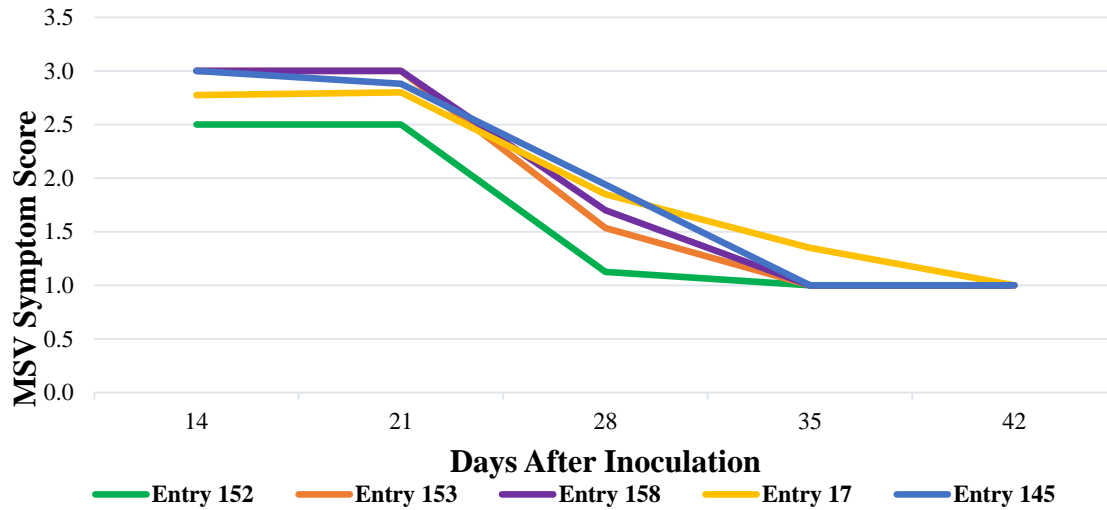
Figure 1. Trends in MSV symptom score of five maize lines with the lowest AUDPC value.
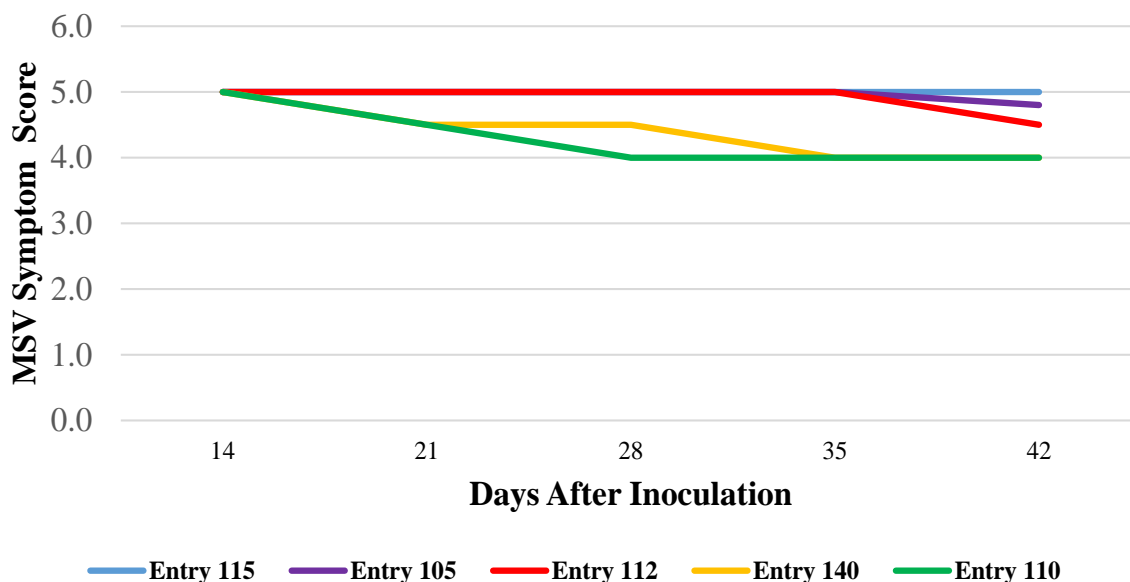
Figure 2. Trends in MSV symptom score of five maize lines with the highest AUDPC values.

B/TZiSTR1248 and KS 27 (S) C3-2-6-2-B-1/TZiSTR1262 bi-parental crosses had the highest average symptoms severity scores value (4.2), indicating highest susceptibility to MSV. Lines developed from IITATZi1715/TZiSTR1262 had better recovery resistance compared to the other maize lines. Amongst the 32 lines developed from IITATZi1715/TZiSTR1262 bi-parental crosses, 20 had symptoms severity score values less than or equal to 2. As resistant plants grew, resulting in an improvement in the level of resistance, the symptom gradually decreased from the lower to the upper leaf, especially on the resistant maize lines.

Maize streak virus detection through enzyme-linked immunosorbent assay

The virus titer determined from DAC-ELISA ranged from 0.3 to 8.5 ng/mg for the upper leaves and 0.4 to 12.1 ng/mg for the lower leaves. Thus, the MSV titer value for the lower leaves was in general higher than the titer value for the upper leaves. The mean concentration of the MSV negative control was 0.15 ng/mg. MSV was detected on the lower leaves in 90 of the 160 maize lines while the remaining 70 maize lines were negative for MSV, using the negative control as reference. The trend was different

for the upper leaves; MSV was not detected on a total of 134 lines while 26 lines were MSV positive.

All the 18 maize lines without the favorable allele of the SNP markers were positive for MSV in their lower leaves based on the ELISA test. In the upper leaves, however, 16 lines were positive while 2 were negative. For the 142 maize lines that carried the favorable allele, 72 were positive for MSV in the lower leaf while 70 were negative for MSV. In the upper leaf, the maize lines with favorable allele and which were positive for MSV reduced to 7 while the maize lines which were negative increased to 132. For both the lower and the upper leaves the association between MSV titer response and the SNP markers allele was significant *(P<0.01)* with chi-square value of 78.6 for the upper leaf and 13.6 for the lower leaf.

The 61 maize lines that had symptom scores considered resistant (1-2) on the MSV resistance scoring scale had the favorable allele and did not have MSV detected in their upper leaf by ELISA. Of this, 24 had MSV detected in their lower leaves while it was not detected in the upper leaves of the remaining 37 maize lines. All the lines in this group had higher MSV titer values in the lower than the upper leaves, except five lines which had the same value for leaves from the two positions. MSV was not detected in the upper and lower leaves of five lines as the titer values were low.

Of the 88 maize lines with MSV symptom scores considered as moderately resistant, 10 did not carry the favorable allele while the remaining 78 had favorable alleles. The ten lines were derived from KS27/TZISTR1205 (1 line) and SW5/TZISTR1248 (9 lines). Eight of the 10 lines had MSV detected in their lower and upper leaves while MSV was not detected in the upper leaves of two lines derived from SW/TZISTR1428. The MSV titer values of the lower leaves for the two lines were 4.7 and 4.2 while for the upper leaves the value were 2.3 and 3.1 respectively. These results suggest that the resistance in the two lines were different from that linked to the three markers. MSV was detected by ELISA in the lower leaves of 56 of the 88 maize lines; 18 of the 88 lines had MSV detected in their upper leaves.

Of the 11 lines that had MSV scores considered as susceptible, three carried the favorable SNP marker allele while the remaining eight lines did not. MSV was detected by ELISA on the lower leaves of the 11 lines with titer values ranging from 4.1 to 8.9. However, MSV was not detected in the upper leaves of the three lines with the favorable alleles; titer values for these ranged from 2.4 to 2.7. MSV was detected in the upper leaves of the remaining eight lines for which titer values ranged between 4.0 and 5.3. These results suggest the effectiveness of the marker allele in detecting resistance. However, the resistance does not include infection (of the lower leaves), but it did not allow the virus to spread to the upper leaves.

Chi-square analysis of the relationship between resistance/susceptible class (resistance, moderate resistance and susceptible) and ELISA MSV status showed significant *(P<0.01)* association for the lower (χ^2 =17.62, df=2) and upper leaves (χ^2 =38.75, df=2). Thus the association was stronger for the upper than the lower leaf. The percentage recovery of maize lines based on virus titer on upper leaves relative to the lower leaves varied considerably among the maize lines. The percentage recovery value ranged from -12.2% (no recovery) to 74.1% (high recovery). The negative percentage recovery value showed that the titer of the virus was higher on upper leaf compared to the lower leaf. Of the 160 tested lines, sixty-four (40%) had percent recovery greater than or equal to 50 % and ninety-six (60%) had below 50% recovery. Of the 64 high recovery resistance maize lines, only one did not carry the favorable allele for MSV resistance. The percentage recovery of phenotypically resistant lines ranged from 0 to 74.1%, for moderately resistant lines it ranged was from 0.0 to 70.6% while for the phenotypically susceptible lines range was from -12.2 to 49%.

There was significant association at six week after inoculation between the type of SNP allele carried by the maize lines and titer of the virus. Significant association was detected for both the lower and upper leaves. More maize lines than expected (in the calculation of the Chi-square value) without the allele of the SNP markers had MSV detected in their leaf tissue. The higher Chi-square value obtained for the upper leaf (78.63 vs 13.40) indicate the upper leaf at week 6 after inoculation to be more effective in assaying response to MSV by ELISA due to the resistance gene detected by the SNP markers. Of the 142 maize lines with the favorable allele of the marker, 132 (93.0%) were negative for MSV in their upper leaf, a much higher amount than the 70 (49.3%) that were negative for the virus in their lower leaves.

MSV symptoms assessment by image analysis

Significant differences *(P<0.01)* among the maize lines were observed on the percentage of disease severity on the sampled portion of the leaf. The mean disease severity on the fifth leaf ranged from 4.6 to 81.2% for the most resistant and susceptible maize lines respectively. In maize lines with the favorable allele for resistant, disease severity ranged from 4.6 to 66.4% while without the favorable allele it ranged from 29.8 to 81.2%. The mean disease severity for maize lines that had the favorable allele for MSV resistance and those that did not have favorable allele were 28.9 and 43.3% respectively. Entries 155,158 and 159 had the lowest disease severity values of 4.6, 6.7 and 10.7% respectively while entries 105, 140 and 117 had the highest disease severity values of 81.2, 66.4 and 63.3% respectively. The result obtained from image analysis significantly correlated with the symptom score with correlation coefficient of 0.67 and

ELASA test result 0.23 and 0.24 with the lower and upper leave test result respectively.

DISCUSSION

The level of resistance of the maize lines with the favorable allele indicated by the three SNP markers linked to MSV resistance varied with the source. These markers are linked to MSV resistance gene on Chromosomes 1. Molecular markers that are associated with MSV resistance in a range of genetic backgrounds could potentially enable pre-selection of genomic regions in tropical germplasm developed within and outside SSA. This can facilitate accelerated genetic gains.The SNP markers used in this study were reported to have an MSV reaction prediction efficiency of 0.94 (Nair et al., 2015). Inheritance of MSV resistance in maize is complex, involving at least three major genes and a number of minor genes. A major QTL (Msv1) for MSV resistance on Chromosome bin 1.05 had been reported in several studies on MSV resistance (Kyetere et al. 1999). Due to their high prediction efficiency, kompititive allele specific PCR (KASP) assay was developed for the three SNP markers that mapped to this chromosomes location.

In this study all the maize lines evaluated and the Pool 16 developed sever symptom which confirm the success of the artificial inoculation technique. MSV resistance score of the maize line evaluated ranged between highly resistant (score = 1) to highly susceptible (score = 5). Majority of the maize lines tested were moderately resistant to MSV following successful inoculation of the virus using C. triangular as vector. In similar studies Markham et al. (1984), reported that C. triangular had 60 to 100 % efficiency in transmitting MSV after acquiring the virus. In the present study, the inoculated plants develop symptoms 3 to 5 days after inoculation. These results are consistent with the report of Mesfin et al. (1995) that symptoms of MSV appear faster in younger maize plants, usually 3 to 5 days in one week old plants and 7 to 9 days in 9 week old plants. Viral symptoms observed in this study did not differ from the symptoms described in the literature (Asanzi et al., 1994; Bosque-Perez, 2000).

Among the maize lines with the favorable marker allele for resistant, 43% showed a significance recovery, 54.9% showed moderate recovery while 2.1% did not show significant recovery. As resistant plants grew, resulting in an improvement in the level of resistance, the symptom gradually decreases from the lower to the upper leaf, especially on the resistant maize lines. This pattern of resistance is referred to as mature plant resistance (Mesfin et al., 1995). Mature plant resistance has been reported for resistance to MSV and MLN in maize (Bosque-Perez, 2000; Sitta et al., 2017). The highly susceptible lines had less vigour compared to the resistant maize lines. Other author reported that affected

maize plants are shorter and has less vigour especially when the infection is early (Okoth et al., 1988).

MSV has been reported to infect all cell types of the host plant, with streak symptoms manifest only on inoculated leaves or on leaves produced after infection of the plant (Thottappilly et al., 1993). Bosque-Perez (2000) reported that the streak pattern is a result of the failure of chloroplasts to develop tissues surrounding the vascular bundles. The basal region of the maize leaf laminae expresses the symptoms of MSV disease first, and this gradually spreads towards the leaf apex.

The favorable allele of the SNP markers was significantly associated with resistance to MSV inferred from symptoms of the disease of the host plants. The three SNP markers were successful in identifying 139 of the 142 maize lines (97.9%) as resistant. The susceptibility of three lines carrying the favorable allele for resistance may be due to broken-down linkage between the SNP markers and the gene for resistance to MSV in the maize lines and/or the influence of different genetic backgrounds. However, the successful rate in identifying susceptible maize lines was 44.4% (8 out of 18). These results indicate that the SNP markers are useful for identifying many resistant maize lines, but are not able to identify all maize lines resistant to MSV. Reliance on the three SNP markers alone may result in discarding resistant maize lines that would otherwise be useful in breeding programmes.

In the present study, the ELISA serological viral detection method further confirmed the results obtained from phenotyping following artificial inoculation. The ELISA technique, in addition to offering insight on the nature of resistance in the genetic materials studied and also provided information on the distribution of MSV antigen in the leaves of infected plants. The titer values of the virus obtained from the upper and the lower leaves were significantly different among maize lines. The differences in virus titer between the susceptible and resistant maize lines were related with the symptom severity on the leaves. In this study, higher virus titer was associated with the more severe streak symptoms and vice versa. Peterschmitt et al. (1991) showed that virus concentration in leaves increases with increase in density of chlorotic streaks.

In this study, the MSV titer ranged from 0.3 to 12.1 ng/mg. A study conducted by Peterschmitt et al. (1991) reported MSV concentration value of 4.0 ng/mg on maize leaves 15 days after inoculation using indirect ELISA serological viral detection method. The titer of the leaf sample is a function of the level of resistance/ susceptibility of the maize line, the time (days/weeks) after infection when leaves were assayed and the position on the plant of leaf assayed. The alleles of the SNP markers strongly associated with the leaf virus titer response for the upper leaf. These results are in agreement with the observation and inference on the mature plant type resistance (based on symptom severity

score) to MSV earlier made for the lines used in the study.

In the present study, the digital imaging of the 5th leaves of the maize lines processed using the Leaf Doctor Software provided estimates of symptoms severity in agreement with the genotype information provided by the SNP markers, visual symptoms scores and ELISA result. The software has a potential for use in severity assessment of MSV disease. Sanjay and Shrikant (2011) reported that similar digital image processing achieved an accuracy of 98.6% in the estimation of brown spot disease severity in sugar cane *(Saccharum officinarum)* leaves. The approach has been credited with improvement in accuracy, precision, and reliability of estimates of plant disease severity over visual score (Bardsley and Ngugi, 2013). Assessment of disease severity visually has drawbacks which include rater fatigue, the decrease in accuracy and precision of rater estimates over time due to the repetitiveness of task and physically tiring nature of assessment task (Sanjay and Shrikant, 2011).

Martin and Rybicki (1998) used digital image processing to estimate disease severity of MSV and compared the results obtained with visual assessment, using commercial software package as well as an in-house customized package. They concluded that the commercial and customized software packages had approximately the same performance, and both computer-based methods achieved better accuracy and precision than the visual method.

CONCLUSION

The three SNP markers were useful in identifying maize lines with resistance to MSV. The MSV symptoms scores were higher on the older leaves and reduced over time on the upper leaves. The severity of symptoms displayed was dependent on the level of resistance of the maize lines. Resistant maize lines had reduced symptom on the upper leaves at six week after inoculation at which time a total of 139 out of 142 (97.9%) maize lines with the favorable marker allele for resistant to MSV showed some resistance based on symptoms scores on the upper leaves; 61 of these were resistant while the remaining 78 were moderately resistant. A total of 10 maize lines out of the 18 that did not have the favorable marker allele for resistance to MSV had moderate resistant to the virus. This result suggests that the existence of other resistance gene(s) not linked to the marker allele among the lines.

The three SNP markers have potential for use in marker assisted selection for the development of MSV resistant varieties. However, there is need to develop additional markers that can be used to identify other genes responsible for resistance in some of the lines used in the study.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGMENT

REFERENCES

Asanzi MC, Bosque-Perez NA, Nault LR, Gordon DT and Thottapilly G (1994). Biology of Cicadulina species (Homoptera: Cicadellidae) and transmission of maize streak virus. African Entomology 3:173-179.

Beyene YS, Mugo T, Tefera, T, Gethi JJ, Gakunga J, Ajanga S (2012). Yield stability of stem borer resistant maize hybrids evaluated in regional trials in East Africa. African Journal of Plant Science *6:77-83.*

Bosque-Perez NA (2000). Eight decades of maize streak virus research. Virus Research 71:107-121.

Bardsley SJ, Ngugi HK (2013). Reliability and accuracy of visual methods used to quantify severity of foliar bacterial spot symptoms on peach and nectarine. Plant Pathology 61:118-132.

Cairns JE, Sonder K, Zaidi PH, Verhulst N, Mahuku G, Babu R (2012). Maize production in a changing climate: Impacts, adaptation and mitigation strategies. Advances in Agronomy 114:1-58.

CIMMYT (1985). Managing trials and reporting data for CIMMYT International maize testing program. Mexico D. F. CIMMYT.

Efron Y, Kim SK, Fajemisin JM, Mareck JH, Tang CY, Dabrowski ZT, Rossel HW, Thottappilly G, Buddenhagen IW (1989). Breeding for resistance to maize streak virus: a multidisciplinary team approach. Plant Breeding 103:1-36.

Fajemisin JM, Efron Y, Kim SK, Khadr FH, Dabrowski ZT, Mareck JH, Bjarnason, M, Parlinson V, Everett LA, Diallo A (1986). Population and varietal development in maize for tropical Africa through resistance breeding approach. Istituto Agronomico per l'Oltremare 100(A2).

Gebrekidan B, Wafula BM, Njoroge K (1992). Agro ecological zoning in relation to maize research priorities in Kenya. IN KARI, Proceedings of a workshop. Review of the National Maize Research Program, KARIIISNAR Management Training Linkage Project. 1990:100.

Guthrie EJ (1978). Measurement of yield losses caused by maize streak disease. Plant Disease 62(10):839-841.

Kyetere DT, Ming R, McMullen MD, Pratt RC, Brewbaker J, Musket T (1999). Genetic analysis of tolerance to maize streak virus in maize. Genome 42:20-26.

Lagat M, Danson M, Kimani M, Kuria A (2008). Quantitative trait loci for resistance to maize streak virus in maize lines used in hybrid development. Africa Journal of Biotechnology 7:2573-2577.

Magenya OEV, Mueke J, Omwega C (2008). Significance and transmission of maize streak virus disease in Africa and options for management. African Journal of Biotechnology 7(25):4897-4910.

Markham PG, Pinner MS, Boulton MI (1984). The transmission of maize streak virus by leafhoppers, a new look at host adaptation. Entomology 57(4):431-432.

Martin DP, Rybicki EP (1998). Microcomputer-based quantification of maize streak virus symptoms in *Zea mays*. Pathology 88:422-427.

Mesfin T, Den Hollander J, Markham PG (1995). Feeding activities of *Cicadulina mbila* (Hemiptera: Cicadellidae) on different host-plants. Entomology 85:387-396.

Nair Sudha K, Raman Babu, Cosmos M, George M, Kassa S, Yoseph B, Biswanath D, Dan M, Lava KP, Michael O, Prasanna MB (2015). Fine mapping of *Msv1*, a major QTL for resistance to maize streak virus leads to development of production markers for breeding pipelines. Theoretical and Applied Genetics 128(9):1839-1854.

Njuguna JGM, Theuri JM, Ininda JM (1990). The status of maize streak disease in Kenya. In: Gebrekidan B (ed) Maize Improvement, Production and Protection in Eastern and Southern Africa pp. 218-223.

Okoth VAO, Dabrowski ZT, Thottappilly G, Van Emden HF (1988). Comparative analysis of some parameters affecting maize streak virus (MSV) transmission of various Cicadulina spp. populations. Journal Institute Science 8:445-46.

Peterschmitt M, Reynaud B, Sommermeyer G, Baudin P (1991). Characterization of maize streak virus isolates using monoclonal and polyclonal antibodies and by transmission to a few hosts. Plant Disease 75:27-32.

Peterschmitt M, Quiot JB, Reynaud B, Baudin P (1992). Detection of maize streak virus antigens over time in different parts of maize plants of a sensitive and a so-called tolerant cultivar by ELISA. Annal of Applied Biology 121:641-653.

Pratt RC, Gordon SG (2006). Breeding for resistance to maize foliar pathogens. Plant Breeding Reviews 27:119-174.

Redinbaugh MG, Jones MW, Gingery RE (2004). The genetics of virus resistance in Maize (Zea mays L.). Maydica 49:183-190.

Rossel HW, Thottapilly G (1985). Virus diseases of important food crops in tropical Africa. IITA, Ibadan, Nigeria.

Sanjay BP, Shrikant K (2011). Leaf disease severity measurement using image processing. International Journal of Engineering and Technology 3(5):297-301.

Sitta J, Nzuve FM, Olubayo FM, Mutinda C, Muiru WM, Miano DW, Muthomi JW, Leley PK (2017). Response of Assorted Maize Germplasm to the Maize Lethal Necrosis Disease in Kenya. Journal of Plant Studies 6(2):65.

Thottappilly G, Bosque-Perez NA, Rossel HW (1993). Viruses and virus diseases of maize in tropical Africa. Plant Pathology 42(4):494-509.

Vivek BS, Odongo O, Njuguna J, Imanywoha J, Bigirwa G, Diallo A (2010). Diallel analysis of grain yield and resistance to seven diseases of 12 African maize (Zea mays L.) inbred lines. Euphytica 172:329-340.

Wambugu FM, Wafula J (2000). Advances in Maize Streak Virus disease research in Eastern and Southern Africa: workshop report. KARI and ISAAA 8:35.

Welz HG, Schechert AA, Pernet KV, Pixley H, Geiger HH (1998). A gene for resistance to the maize streak virus in the African CIMMYT maize inbred line CML 202. Molecular Breeding 4:147-154.

Wilcoxson RD, Skovmand B, Latif AH (1975). Evaluation of wheat cultivars for ability to retard development of stem rust. Annuals of Applied Biology 80:275-281.

Related Journals:

African Journal of
Biochemistry Research

Journal of
Microbiology and Antimicrobials

African Journal of
Microbiology Research

Journal of
Biophysics and Structural Biology

International Journal of
Biotechnology and Molecular Biology Research

Journal of
Bioinformatics and Sequence Analysis

African Journal of
Biochemistry Research

Journal of
Microbiology and Antimicrobials

African Journal of
Microbiology Research