Virus Research xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

# High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding☆

Ismail Y. Rabbi[a,*], Martha T. Hamblin[b], P. Lava Kumar[a], Melaku A. Gedil[a], Andrew S. Ikpan[a], Jean-Luc Jannink[b,c], Peter A. Kulakow[a]

[a] International Institute for Tropical Agriculture (IITA), Ibadan, Nigeria
[b] Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA
[c] USDA-ARS, R.W. Holley Center for Agriculture and Health, Ithaca, NY, USA

ARTICLE INFO

*Article history:*
Available online xxx

*Keywords:*
Cassava mosaic disease
Breeding
Phenotyping
Monogenic resistance
Genotyping-by-sequencing
QTL

ABSTRACT

Cassava mosaic disease (CMD), caused by different species of cassava mosaic geminiviruses (CMGs), is the most important disease of cassava in Africa and the Indian sub-continent. The cultivated cassava species is protected from CMD by polygenic resistance introgressed from the wild species *Manihot glaziovii* and a dominant monogenic type of resistance, named *CMD2*, discovered in African landraces. The ability of the monogenic resistance to confer high levels of resistance in different genetic backgrounds has led recently to its extensive usage in breeding across Africa as well as pre-emptive breeding in Latin America. However, most of the landraces carrying the monogenic resistance are morphologically very similar and come from a geographically restricted area of West Africa, raising the possibility that the diversity of the single-gene resistance could be very limited, or even located at a single locus. Several mapping studies, employing bulk segregant analysis, in different genetic backgrounds have reported additional molecular markers linked to supposedly new resistance genes. However, it is not possible to tell if these are indeed new genes in the absence adequate genetic map framework or allelism tests. To address this important question, a high-density single nucleotide polymorphism (SNP) map of cassava was developed through genotyping-by-sequencing a bi-parental mapping population ($N = 180$) that segregates for the dominant monogenic resistance to CMD. Virus screening using PCR showed that CMD symptoms and presence of virus were strongly correlated ($r = 0.98$). Genome-wide scan and high-resolution composite interval mapping using 6756 SNPs uncovered a single locus with large effect ($R^2 = 0.74$). Projection of the previously published resistance-linked microsatellite markers showed that they co-occurred in the same chromosomal location surrounding the presently mapped resistance locus. Moreover, their relative distance to the mapped resistance locus correlated with the reported degree of linkage with the resistance phenotype. Cluster analysis of the landraces first shown to have this type of resistance revealed that they are very closely related, if not identical. These findings suggest that there is a single source of monogenic resistance in the crop's genepool tracing back to a common ancestral clone. In the absence of further resistance diversification, the long-term effectiveness of the single gene resistance is known to be precarious, given the potential to be overcome by CMGs due to their fast-paced evolutionary rate. However, combining the quantitative with the qualitative type of resistance may ensure that this resistance gene continues to offer protection to cassava, a crop that is depended upon by millions of people in Africa against the devastating onslaught of CMGs.



## 1. Introduction

### 1.1. Cassava mosaic disease

Cassava (*Manihot esculenta* Crantz, family Euphorbiaceae) is a starchy root crop that supplies carbohydrate energy to millions of people in the tropics (Ceballos et al., 2004) and it is being used increasingly as an industrial crop (Jansson et al., 2009). Though its remarkable ability to tolerate unfavourable conditions such as



* Corresponding author at: IITA, Headquarters & West Africa Hub, PMB 5320, Oyo Road, Ibadan 200001, Oyo State, Nigeria. Tel.: +234 2 7517472x2719; fax: +44 208 7113785.

0168-1702/$ – see front matter © 2013 The Authors. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.virusres.2013.12.028

drought and poor soils makes it a food security crop in many parts of sub-Saharan Africa (SSA), on-farm productivity of cassava has remained stagnant for many years due to several production constraints. Cassava mosaic disease (CMD), caused by several species of cassava mosaic geminiviruses (CMGs), is the most economically important constraint to cassava in SSA and the Indian sub-continent (Herrera-Campo et al., 2011). Though significant efforts have been expended on combating this disease, it still causes huge losses to production. The most striking example of the devastating potential of CMD to undermine food security in Africa is the severe pandemic that started as an epidemic in Uganda in the 1990s and led farmers to abandon the crop in many parts of the country (Otim-Nape and Thresh, 1998), and later spread to most countries in East and Central Africa (Legg and Fauquet, 2004). The pandemic is characterized by rapid spread through super-abundant *Bemicia tabaci* vectors (Legg and Ogwal, 1998) and is associated with a recombinant strain of the *East African cassava mosaic virus – Uganda* (EACMV-UG) along with *African cassava mosaic virus* (ACMV) belonging to the genus *Begomovirus*, within the *Geminiviridiae* family (Harrison et al., 1997).

Important control measures against CMD include rogueing of symptomatic plants, use of virus-free planting materials and deployment of resistant varieties. The first two options are not only labour intensive and difficult to implement but also require continuous and long-term intervention. Use of resistant varieties is the most effective solution in mitigating the negative effect of CMD in farmers' fields because this approach not only reduces yield losses due to the disease, but also reduces levels of the virus inoculum in the farming system particularly in varieties that suppress virus accumulation.

### 1.2. Sources of resistance to the disease

Currently deployed resistance against CMD in Africa is of two types: (i) quantitative resistance derived from *Manihot glaziovii*; and (ii) qualitative resistance conferred by a single resistance gene(s). The quantitative resistance was introgressed into cultivated cassava following an unsuccessful worldwide search for resistant clones in the 1930s (Nichols, 1947; Jennings, 1976). Genetic studies reveal that the polygenic resistance from *M. glaziovii* is recessive with a heritability of about 60% (Jennings, 1976). The second type of resistance, which is conditioned by a single-gene with a dominant effect, was discovered in the 1980s in landraces from Nigeria and other West African countries (Akano et al., 2002; Fregene et al., 2001). These landraces, which display near-immunity against nearly all species of CMGs, are currently maintained in the IITA germplasm collection referred to as the Tropical *Manihot esculenta* (TMe) series. Diversity studies using molecular markers have previously shown that most of the original landraces bearing this qualitative resistance to CMD are genetically very similar if not identical (Fregene et al., 2000; Lokko et al., 2006). This suggests that the genetic base of this type of resistance in the African cassava genepool may be narrow, or even just a single locus. In contrast, the relative ease with which the highly heritable monogenic resistance can be transferred between germplasm through simple crosses, has resulted in its extensive usage in breeding across Africa as well as pre-emptive breeding in Latin America (Okogbenin et al., 2007). The long-term stability of this single-gene type of resistance in diverse geographical regions with heterogeneous species and recombinants of CMGs is uncertain given the high evolutionary rate of geminiviruses (Duffy and Holmes, 2009).

Several genetic mapping studies have been conducted to find molecular markers linked to the qualitative resistance in the African germplasm. The first study identified two markers, a microsatellite (SSRY28) and an RFLP (GY1) that flank a single locus named *CMD2* at distances of 9 and 8 cM, respectively (Akano et al., 2002). Subsequent to the discovery of the *CMD2* locus, later studies have reported several additional markers that are linked to new resistance genes in other genetic backgrounds, including landraces and improved varieties derived from them (Lokko et al., 2005; Okogbenin et al., 2012). However, nearly all of these studies relied on the bulk segregant analysis (BSA) approach and/or very sparse maps for interval mapping analysis. The BSA approach provides little or no information regarding the chromosomal location of the identified markers, making it difficult to ascertain the number of unique loci/genes associated with a trait. When sparse maps are used, the confidence interval surrounding a QTL is usually large, making it difficult to determine the precise QTL location. For example, the *CMD2*-containing linkage group of Akano et al. (2002) had a total of five markers. Lokko et al. (2005) used a linkage map with just 45 markers, of which only three were in the linkage group containing the resistance locus.

The objective of this study was to provide a comprehensive framework for describing the breadth of the genetic base of the single-gene resistance to CMD in the African cassava germplasm. Firstly, a full-sib mapping population segregating for qualitative resistance to CMD was developed and phenotyped for three growing seasons. The population was genotyped using genotype-by-sequencing (GBS), and a dense genetic linkage map with more than 8000 single nucleotide polymorphism (SNPs) was constructed. Using this resource, a high-resolution genetic mapping of the CMD resistance locus was carried out. The markers previously reported to be linked to CMD resistance were then projected onto the newly generated genetic map. This revealed their genomic locations, and the spatial relationship between them and the mapped resistance locus from the present study. To confirm the relationship among the CMD resistant landraces, cluster analysis was carried out using genome-wide SNP markers.

## 2. Materials and methods

### 2.1. Mapping population development, phenotyping and genotyping

A full-sib mapping population segregating for dominant monogenic resistance to CMD was generated by crossing two non-inbred clones. Both parents are elite lines developed by IITA in Nigeria. The female parent, IITA-TMS-011412, is highly resistant to CMD and also rich in pro-vitamin A. Cloned in 1974, the male parent, IITA-TMS-4(2)1425, is an improved variety from a cross between a landrace from Nigeria (TME109, locally known as Oyarugbafunfun) and variety 58308, a hybrid derived directly from recombination of the *M. glaziovii* × *M. esculenta* triple-backcrosses (Jennings, 1976). Variety 58,308 was the main source of the quantitative resistance to CMD and produced some of the first generation Tropical Manihot Selection lines (see the discussion section for more background). IITA-TMS-4(2)1425 shows considerable susceptibility to CMD (Fig. 1).

The 180 F1 seeds produced were germinated in sterilized garden soil and transplanted one month after sowing. At maturity, the seedlings were cloned, regardless whether they were infected or not, and planted at Ibadan, Nigeria (7.40° North latitude, 3.90° East longitude) using a randomized complete block design. Each clone was planted in two replicated plots of five stands per plot with plant spacing of 1 m × 0.5 m for three 12-month long cropping seasons established in 2011, 2012 and 2013. Generation-to-generation propagation through cloning was based on use of 12-cm long stem cuttings in both the infected and non-infected F1s. A local landrace that is highly susceptible to CMD, TME117, was planted as spreader row every fifth plot and as border row surrounding the experimental field to facilitate whitefly-mediated inoculation of the F1 population. The Ibadan site is known for high CMD pressure and

Please cite this article in press as: Rabbi, I.Y., et al., High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. Virus Res. (2014), http://dx.doi.org/10.1016/j.virusres.2013.12.028