Theor Appl Genet (2002) 105:521–525
DOI 10.1007/s00122-002-0891-7

A.O. Akano · A.G.O. Dixon · C. Mba · E. Barrera
M. Fregene

Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease

Received: 9 October 2001 / Accepted: 12 December 2001 / Published online: 8 May 2002

Abstract Cassava mosaic disease (CMD) is the most-important disease of cassava (*Manihot esculenta*) in Africa, and is a potential threat to Latin American (LA) cassava production. Although this viral disease is still unknown in LA, its vector – the whitefly – has recently been found. The disease is best controlled through host-plant resistance, which was first found in third backcross derivatives of an interspecific cross between cassava and *Manihot glaziovii*, and is thought to be polygenic. Recently, high levels of resistance were also found in several Nigerian cassava landraces. Classical genetic analysis and molecular genetic-mapping of the landraces showed that a major dominant gene confers this resistance. Bulk segregant analysis (BSA) was used to quickly identify a simple sequence repeat (SSR) marker linked to the CMD-resistance gene. The marker, SSRY28, is located on linkage group R of the male-parent-derived molecular genetic map. The gene, designated as *CMD2*, is flanked by the SSR and RFLP marker GY1 at 9 and 8 cM, respectively. To our knowledge, this is the first report of qualitative virus resistance in cassava, and of molecular markers that tag CMD resistance in cassava. We discuss the use of markers linked to *CMD2* for marker-assisted breeding of CMD resistance in Latin America and for increasing the cost-effectiveness of resistance breeding in Africa.

Keywords Cassava mosaic disease (CMD) · Molecular markers · Bulk segregant analysis

Communicated by H.C. Becker

A.O. Akano · C. Mba · E. Barrera · M. Fregene (✉)
CIAT, A.A. 6713, Cali, Colombia
e-mail: M.Fregene@cgiar.org
Tel.: +57-2-4450073, Fax: +57-2-4450073

A.G.O. Dixon
IITA, PMB 5320, Ibadan, Nigeria

Introduction

The most-widespread cassava disease of economic importance in Africa is, undoubtedly, cassava mosaic disease (CMD). Epidemics are particularly ravaging, with root yield losses as high as 100% (Jennings 1994; Thresh et al. 1994). Even in the absence of a serious outbreak, yield losses of 20% to 90% are common in farm fields throughout sub-Saharan Africa (Hahn et al. 1980b; Muimba-Kankolongo and Phuti 1987). CMD is commonly found in the fields of smallholder cassava farmers who cannot always consistently follow good crop-sanitation practices, such as planting CMD-free cuttings and roguing diseased plants. Host-plant resistance to CMD, based on resistance originally obtained from a wild relative of cassava, *Manihot glaziovii* (Nichols 1947), is the best method of containing the disease. Recently, the International Institute of Tropical Agriculture found high levels of resistance in closely related Nigerian cassava landraces (IITA 1990).

Cassava mosaic disease is caused by at least four geminiviruses of the genus *Begomovirus* (Family Geminiviridae), and is transmitted by the whitefly *Bemisia tabaci* (Gennadius) biotype A. These viruses are the African cassava mosaic virus (ACMV), the East African cassava mosaic virus (EACMV), the Ugandan variant of the EACMV (EACMV-UgV), a hybrid virus of EACMV and ACMV, and the South African cassava mosaic virus (SACMV) (Swanson and Harrison 1994; Zhou et al. 1997). Outside Africa, a variant of the ACMV – the Indian cassava mosaic virus (ICMV) – causes a similar disease that is the most important cassava disease in India (Rajendran et al. 1993). The CMD begomoviruses are unknown in the Americas.

The whitefly vector does not colonize cassava in the New World, although, recently, a new biotype of *B. tabaci*, biotype B (also referred to as *Begomovirus argentifolia*), has become widespread in the Americas and has a wide host range, including cassava (Polston and Anderson 1997). This is a frightening prospect for Latin American (LA) cassava production, considering that most LA

germplasm of this root crop is highly susceptible to CMD (Okogbenin et al. 1998). Identifying and breeding host-plant resistance to CMD is therefore of strategic importance to all major cassava-producing regions.

Because, as for LA, breeding for CMD resistance must be done in the absence of the virus, and different sources of resistance must be pyramided, a project to genetically map CMD-resistance genes, using molecular markers, was set up. Approaches to the genetic mapping of cassava – a putative segmental allopolyploid (Magoon et al. 1969; Umanah and Hartmann 1973) – are similar to those used in mapping polyploid genomes (Ritter et al. 1991; Wu et al. 1992; Al Janabi et al. 1993; Da Silva et al. 1993). These approaches attempt to simplify the determination of allelism by analysing a special class of markers known as single-dose restriction fragments (SDRFs) (Wu et al. 1992). SDRFs are DNA markers that are present in one parent and absent in the other, and segregate in a 1:1 ratio (absence to presence) in the progeny. They represent the segregation equivalent of an allele at a heterozygous locus in a diploid or allopolyploid genome, or of a simplex allele in an autopolyploid.

Linkage analysis, using SDRFs in an F1 population, requires the presence of several unique alleles in either or both parents, and results in two separate linkage maps, based on male and female sources of markers. A molecular genetic map of cassava has been constructed using an F1 intraspecific cross at CIAT as a source of markers for gene tagging (Fregene et al. 1997).

In this paper, we report the results of genetic mapping of the dominant gene that controls the new source of CMD resistance.

Materials and methods

Planting materials and CMD resistance evaluation

The mapping population of CMD-resistant cassava was an F_1 progeny from a cross between a Nigerian landrace (TME 3) that represented the new source of CMD resistance and a susceptible improved line (TMS 30555). The progeny, comprising 158 individuals, was established in vitro from embryo axes as described by Akano et al. (1998), then subcloned. Six copies from a total of eight per genotype were transferred to a field with low CMD pressure in Abuja, Central Nigeria, in July 1997. Stakes, 15–20-cm long, were cut from these plants and planted in two sites with high CMD pressure: Onne (humid forest), and Ikenne (subhumid forest), both in June 1998. The experiment had a randomized complete block design, with rows comprising ten plants per genotype and three replicates.

CMD resistance was evaluated at 3 and 6 months after planting. Visual assessment of symptom intensity for each leaf on each plant was conducted according to a scale of 1 to 5, where 1 = no observable symptoms and 5 = very severe chlorosis and heavily reduced leaf area. Scores on symptoms were averaged across all leaves per genotype (Hahn et al. 1980a).

Genotypes were observed to be either entirely symptomless (scoring 1) or altogether diseased (scoring 4 or 5) in all three replicates and at both sites, suggesting qualitative inheritance. A chi-square test was therefore performed for a monogenic inheritance model of resistance. After 1 year, all plants were pruned and, 3 and 6 months later, assessed visually for symptoms. On re-growing, pruned CMD-infected plants generally had very severe symptoms. Again, genotypes were observed to be either entirely symptom-free or severely diseased.

Bulk segregant analysis

To quickly identify markers associated with the dominant CMD-resistance gene, bulk segregant analysis (BSA) was employed. One bulk of 40 susceptible and another of 40 resistant genotypes from the mapping progeny and a set of 186 SSR markers (Mba et al. 2000) were used. The optimal size of bulks for BSA depends on the type of population and the marker being screened (Michelmore et al. 1991). In the analysis of single-dose restriction fragment (SDRF) markers in F_1 populations derived from non-inbred parents, the probability of a bulk of *n* individuals having an SDRF and a second bulk of not having it when the locus is unlinked to a targeted gene is:

$$\text{Probability} = 2[1 - (1/2)^n](1/2)^n.$$

Forty individuals per bulk (i.e. $n = 40$) would therefore give a probability of 1.5×10^{-23} of an unlinked marker being polymorphic in the bulk. This is a very small chance, even with the large number of SSR markers being screened. Markers found to be polymorphic in the two parents and the two bulks were used to evaluate the 80 individuals of the bulks and another additional subset of 78 genotypes from the mapping population.

PCR analysis and electrophoresis of amplified products

Genomic DNAwas prepared from about 3 g of young fresh leaves, according to Dellaporta et al. (1983). The two bulks, along with the two parental clones, were evaluated, using PCR conditions as described by Mba et al. (2000). PCR amplification products were mixed with 4 µl of loading dye (98% formamide, 10 mM EDTA, pH = 8.0, bromophenol blue and xylene cyanol), heated for 2 min at 96 °C and chilled on ice for 3 min. Of this mixture, 4 µl were loaded onto a 6% denaturing polyacrylamide gel. Electrophoresis was in 1 x TBE at 40 V/cm for 2 h, and DNA was visualized by silver staining according to the manufacturer's manual for the PAGE gel silver-staining kit (Promega, Madison, Wis).

Genetic mapping of a CMD-resistance gene

Of the 186 SSR markers, the map positions of 80 are known (Mba et al. 2000). One of these, SSRY28, could distinguish between CMD resistant and susceptible genotypes. This marker is located on linkage group R of the male-parent-derived map of cassava (Fregene et al. 1997), and is flanked by GY1 (an RFLP marker) and Ai19b (a RAPD marker). These latter markers were evaluated in the bulks, parents and individuals of the CMD-mapping population to determine the precise location of the CMD-resistance gene. Southern-hybridization analysis with GY1 of the bulks and parents, using four restriction enzymes, namely *Eco*RI, *Hind*III, *Dra*I and *Hae*III, and RAPD analysis with Ai19b were performed as described by Fregene et al. (1997). The 158 genotypes of the mapping progeny were then analysed, using the restriction enzyme that gave polymorphisms in the bulks and parents.

To develop a linkage group around the CMD-resistance locus, markers analysed in the CMD mapping progeny and scored as SDRFs (Wu et al. 1992), and resistance scored in the progeny as a qualitative trait, were subjected to linkage analysis as described by Fregene et al. (1997). Testing for linkages and the calculation of distances were done with the computer package MapMaker 2.0 (Lander et al. 1987) on a G3 Macintosh computer. Thresholds for declaring linkage were a LOD score of 4.0 and a recombination fraction of 0.3. Map units (cM) were derived using the Kosambi function (Kosambi 1944). Maximum-likelihood orders of markers were verified by the 'ripple' function, and markers were said to belong to the framework map if the LOD value, as calculated by the 'ripple' command, was ≥ 2.0.

If the genetic model of a single dominant gene for CMD resistance is correct, the markers identified as co-segregating with the gene should explain most of the CMD resistance phenotypic variance. To test this, single-point marker analysis by simple linear regression was performed, using the JMP statistical package (SAS Institute 1995) on a G3 Macintosh computer. Association between a marker and CMD resistance was declared at $P < 0.001$.

Results

Evaluation and bulk segregant analysis of CMD resistance

Variation in response to CMD in the mapping population was qualitative, i.e. all ten plants of each resistant genotype in all three replicates and at both sites showed no visible symptoms, even when re-growing after pruning. In contrast, all plants of the susceptible genotypes were always heavily infected. The chi-square of the ratio of resistant to susceptible plants gave a value of 1.1, which is not significantly different from a 1:1 ratio at a probability level of 0.05. This fits the expected segregation ratio for a single dominant gene heterozygous in the CMD-resistant parent. The BSA revealed that an allele of the SSRY28 marker was present in the resistant parent and bulk, but was absent in the susceptible parent and bulk. The polymorphism was confirmed when individuals of the bulks were screened with the SSR marker (Fig. 1). The SSR marker was then analysed in all 158 F1 progeny. The gene was designated as *CMD2*. *CMD1*, described previously by Fregene (2000), controls currently deployed (polygenic) resistance to CMD.

Genetic mapping of a dominant CMD-resistance gene

The marker SSRY28 was found located on linkage group R (Fregene et al. 1997) of the male-parent-derived map

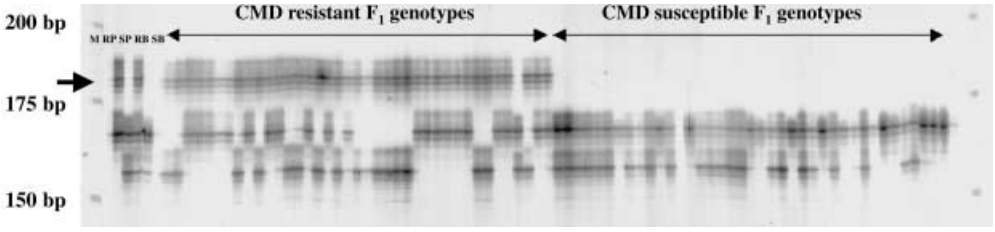

Fig. 1 Polyacrylamide gel image of SSR marker SSRY28 analysed in the CMD-resistant parent (*RP*), susceptible parent (*SP*), resistant bulk (*RB*), susceptible bulk (*SB*), and 40 resistant and 40 susceptible genotypes used as bulks for bulk segregant analysis (BSA). The *arrow* points to the SSR allele that is associated with CMD resistance. The molecular-weight marker (*M*) is a 25-bp ladder

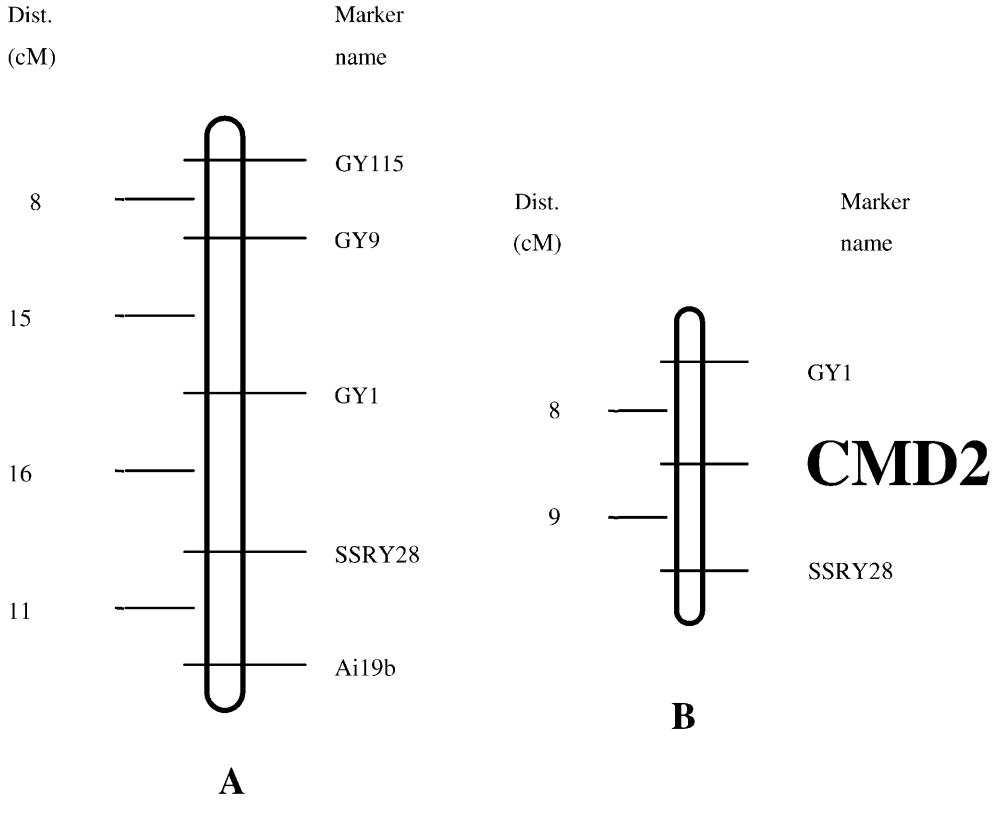

Fig. 2 Linkage group R of the CM 2177-2-derived map (**A**) showing markers GY1 and SSRY28 that flank the dominant gene *CMD2*, and a TME 3-derived map of the region around *CMD2* (**B**). Map distances shown on the left of each map are in Kosambi map units, expressed as centimorgans (cM)

of cassava, linked in coupling to marker GY1 (RFLP) and in repulsion to marker Ai19b (RAPD). RFLP analysis of the bulks and parents of the mapping population with GY1 revealed an RFLP allele unique in the CMD-resistant parent and bulk with the restriction enzymes *Eco*RI and *Hae*III. The 158 progeny were analysed with a *Hae*III/GY1 combination. The RAPD analysis, with primer Ai19b, of the CMD resistant and susceptible parents and bulks showed no polymorphism. Linkage analysis of SSRY28, GY1 and CMD resistance in the mapping progeny revealed that *CMD2* is located at a distance of 9 cM and 8 cM to GY1 and SSRY28 respectively (Fig. 2). The distance of 16 cM between GY1 and SSRY28 on the CM 2177-2-derived cassava map (Fregene et al. 1997) is comparable with the combined distances between GY1, *CMD2* and SSRY28 (17 cM), based on the TME 3-derived map. Results of single-marker analysis showed that SSRY28 and GY1 explain 68% and 70% ($P < 0.0001$) of the phenotypic variance of CMD resistance, which again confirms the hypothesis of single-gene inheritance of CMD resistance.

Discussion

A dominant gene for resistance to CMD has been found by conventional genetic analysis and molecular genetic mapping in a F1 cross between resistant and susceptible parents. To our knowledge, this is the first report of qualitative resistance to viruses in cassava. The single-dominant-gene nature of the new source of resistance makes it particularly useful in breeding for CMD resistance in the light of the common constraints to cassava breeding: heterozygosity and a long cropping cycle. The major gene nature also means that a genetic marker for marker-assisted selection (MAS) can easily be identified. MAS would thus become an invaluable tool for breeding CMD resistance in Latin America where the disease is not found, but where the presence of the vector makes it a threat.

Selecting for high levels of resistance with a marker may be more efficient than conventional breeding in Africa, where rapid deployment of high resistance into cassava gene pools is needed to protect cassava from the ravages of CMD. The advantage of MAS is that the breeder can, in early stages, eliminate CMD-susceptible genotypes. In the case of a heterozygous CMD-resistant donor parent, elimination would be 50%, reducing the costs of disease evaluation by half and increasing selection efficiency. The breeder can then concentrate on fewer genotypes at the seedling and crucial single-row trial stages where progenies are reduced by as much as 95%. Identification of markers for other traits in addition to CMD resistance can be used to chose parents more efficiently that combine the different traits

The gene designated as *CMD2* is different from the earlier found *CMD1*, which controls the currently deployed resistance (Fregene 2000). *CMD2* is located on linkage group R, whereas *CMD1* is on linkage group D of the cassava molecular map (Fregene et al. 1997). The action of the two genes is also different: *CMD2* is dominant, whereas *CMD1* appears recessive in that its effect is detected only in backcross progeny, and not in the F1.

The presence of two different sources of CMD resistance, and the markers in tight linkage with them, provides a means of combining multiple sources of resistance. The recessive nature of the older source of resistance, however, makes it less attractive, given cassava's out-crossing and heterozygous nature.

Acknowledgements We thank Drs. Melaku Gedil (IITA, Ibadan, Nigeria), Matthew Blair, Lee Calvert, John Miles (CIAT, Cali, Colombia) and two anonymous reviewers for helpful comments. This research was supported by a Rockefeller Foundation (RF) grant to CIAT and IITA. Author A.A. was also supported by an RF post-doctoral fellowship granted through CIAT.

References

Akano AO, Ng SYC, Dixon AGO, Thottappilly G (1998) Application of embryo culture in germinating an African cassava mosaic disease-resistant gene-mapping population. Trop Agric (Trinidad) 75:217–219

Al Janabi SM, Honeycutt RJ, McClelland M, Sobral BWS (1993) A genetic map of *Saccharum spontaneum* L. "SES 208". Genetics 134:1249–1260

Da Silva JAG, Sorrels ME, Burnquist WL, Tanksley SD (1993) RFLP linkage map and genome analysis of *Saccharum spontaneum*. Genome 36:782–791

Dellaporta SL, Wood J, Hicks JR (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1:19–21

Fregene M (2000) Marking progress: collaboration to improve cassava. In: Kinley D (ed) Synergies in science. CGIAR (Consultative Group on International Agricultural Research), Washington, D.C. pp 6–7

Fregene M, Angel F, Gómez R, Rodríguez F, Chavarriaga P, Roca W, Tohme J, Bonierbale M (1997) A molecular genetic map of cassava (*Manihot esculenta* Crantz). Theor Appl Genet 95:431–441

Hahn SK, Terry ER, Leuschner K (1980a) Breeding cassava for resistance to cassava mosaic disease. Euphytica 29:673–683

Hahn SK, Howland AK, Terry ER (1980b) Correlated resistance of cassava to mosaic and bacterial blight diseases. Euphytica 29:305–311

IITA (International Institute of Tropical Agriculture) (1990) In: Annual report 1990. Ibadan, Nigeria

Jennings DL (1994) Breeding for resistance to African cassava mosaic geminivirus in East Africa. Trop Sci 34:110–122

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Magoon ML, Krishnan R, Bai KV (1969) Morphology of the pachytene chromosomes and meiosis in *Manihot esculenta* Crantz. Cytologia 34:612–626

Mba REC, Stephenson P, Edwards K, Melzer S, Nkumbira J, Gullberg U, Apel K, Gale M, Tohme J, Fregene M (2000) Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards a SSR-based molecular genetic map of cassava. Theor Appl Genet 102: 21–31

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Muimba-Kankolongo A, Phuti K (1987) Relationship of cassava mosaic severity in planting material to mosaic development, growth and yield of cassava in Zaire. Exp Agric 23:221–225

Nichols RFW (1947) Breeding cassava for virus resistance. East Afr Agric J 12:184–194

Okogbenin E, Porto MCM, Dixon AGO (1998) Influence of planting season on incidence and severity of African cassava mosaic disease in the subhumic zone of Nigeria. In: Akoroda MO, Ekanayake IJ (eds) Root crops and poverty alleviation. International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria, pp 388–392

Polston JP, Anderson PK (1997) The emergence of whitefly-transmitted geminiviruses in tomato in the Western Hemisphere. Plant Dis 81:1358–1369

Rajendran PG, Nair SG, Easwari A, Vasudevan K, Sreekumari MT (1993) Cassava breeding, agronomy research and technology transfer in Asia. In: Howeler RH (ed) Proc 4th Regional Workshop, Trivandrum, Kerala, India, pp 84–96

Ritter E, Debener T, Barone A, Salamini F, Gebhardt C (1991) RFLP mapping on potato chromosomes of two genes controlling extreme resistance to potato virus X (PVX). Mol Gen Genet 227:81–85

SAS Institute (1995) JMP Users guide version 3; 4th edn. Cary, North Carolina

Swanson MM, Harrison BD (1994) Properties, relationships and distribution of cassava mosaic geminiviruses. Trop Sci 34:15–25

Thresh JM, Fargette D, Otim-Nape GW (1994) Effects of African cassava mosaic geminivirus on the yield of cassava. Trop Sci 34:26–42

Umanah EE, Hartmann RW (1973) Chromosome numbers and karyotypes of some *Manihot* species. J Am Soc Hortic Sci 98:272–274

Wu KK, Burnquist W, Sorrels ME, Tew TL, Moore PH, Tanksley SD (1992) The detection and estimation of linkage in polyploids using single dose restriction fragments. Theor Appl Genet 81:471–476

Zhou X, Liu Y, Calvert L, Muñoz C, Otim-Nape GW, Robinson DJ, Harrison BD (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. J Gen Virol 78: 2101–2111