



Genetic mapping and validation of QTL for whitefly resistance in cassava (*Manihot esculenta* Crantz)

Adriana Bohorquez-Chaux¹ · Luis Augusto Becerra Lopez-Lavalle² · Vianey Barrera-Enriquez³ · María Isabel Gómez-Jiménez¹ · Camilo E. Sanchez-Sarria¹ · Luis Fernando Delgado¹ · Xiaofei Zhang⁴ · Winnie Gimode¹

Received: 10 February 2025 / Accepted: 30 May 2025 / Published online: 24 June 2025
© The Author(s) 2025

Abstract

Key message QTL associated with whitefly resistance were identified in a cassava F₂ population and KASP markers applicable in selection for the trait were validated.

Abstract Whitefly species pose a major threat to cassava production in tropical regions causing direct plant damage and transmitting viruses that lead to devastating cassava diseases. *Aleurotrachelus socialis* whitefly is one of the pests that affect cassava in South America. Developing resistant cassava varieties is the most sustainable control strategy for managing whiteflies. This study aimed to map the quantitative trait loci (QTL) associated with resistance to *A. socialis* and develop molecular markers to facilitate marker-assisted selection. An F₂ cassava population ($N=183$) was generated by selfing a highly resistant F₁ derived from a cross between ECU72 (resistant) and COL2246 (susceptible) landraces. Phenotyping was performed using an efficient glasshouse screening method and high throughput image analysis of infested leaves (Nymphstar). We identified QTL on chromosomes 1, 2, 5, 6, 8, 9, and 14, with a stable and highly significant QTL on chromosome 8 (*MeF2Wfly8.1*), explaining 35.44% of the phenotypic variation. To enable efficient selection, high-throughput KASP markers were developed and validated across diverse genetic backgrounds. Three SNPs displayed the highest association with whitefly resistance, with Chr08_6483145 as the most effective marker for selection in diverse backgrounds. These markers are provided for improving the efficiency of whitefly resistance breeding in the global cassava community.

Introduction

Whiteflies are a major threat to cassava (*Manihot esculenta* Crantz) production, not only causing direct plant damage but also transmitting viruses responsible for cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) (Colvin et al. 2004; Maruthi et al. 2005). While these pests

pose a significant challenge in many cassava growing regions, they are particularly devastating in sub-Saharan Africa, where frequent super-abundant whitefly outbreaks exacerbate the spread (Otim-Nape et al. 2001; Alicai et al. 2007). In addition to significant yield losses caused by the spread of cassava diseases, whiteflies, especially in the Neotropics, contribute to direct damage by feeding on the leaf phloem, leading to symptoms such as chlorosis and premature leaf drop (Bellotti and Arias 2001; Bellotti et al. 2012). They also excrete sticky and sugary substances that coat the foliage and serve as a substrate for sooty mold (Bellotti and Arias 2001; Nelson 2008), resulting in a reduction in photosynthesis and subsequently in root yield, if the infestation is prolonged.

The most damaging whitefly species in Africa is *Bemisia tabaci*, which includes more than 40 morphologically indistinguishable putative whitefly species distributed worldwide (de Moya et al. 2019; Mugerwa et al. 2021). The damage caused by the *B. tabaci* species continues to increase, with the areas affected with CBSD also rapidly expanding

Communicated by Damaris Odeny.

✉ Winnie Gimode
w.gimode@cgiar.org

- ¹ Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT), Km 17, Recta Cali-Palmira, A. A. 6713 Cali, Colombia
- ² International Center for Agricultural Research in the Dry Areas (ICARDA), BP 6202, Rabat, Morocco
- ³ European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK
- ⁴ University of California Davis, Davis, CA 95616, USA

(MacFadyen et al. 2018). In the Americas, eleven whitefly species have been reported affecting cassava (Bellotti et al. 1999), with *Aleurotrachelus socialis* Bondar reported to cause significant yield losses of up to 79% (Vargas & Bellotti, 1981; Farias 1994; Bellotti et al. 1999) primarily through direct plant damage. Although the *B. tabaci* complex is well known for transmitting many plant viruses, *A. socialis* has not generally been considered a major virus vector. However, past research indicated *A. socialis* could transmit cassava virus X (Angel et al. 1987, 1989). Furthermore, it is hypothesized that *A. socialis* may act as an aerial vector for cassava frog skin disease (CFSD) in South America, recently established to be caused by torradoviruses (Jimenez et al. 2024), but its involvement in CFSD transmission is currently under investigation.

The control of whitefly populations in the fields relies largely on integrated pest management (IPM) measures, including cultural practices and chemical controls (Bellotti and Arias 2001; Carabalí et al. 2010). The best strategy, which is a cheaper and more environmentally sustainable option for management, would be to utilize cassava varieties with host resistance against the pest. Screening of the germplasm collection at the International Center for Tropical Agriculture (CIAT) has identified donor lines with whitefly resistance (Bellotti et al. 1987; Parsa et al. 2015; Bohorquez-Chaux et al. 2023; Atim et al. 2024), with some sources including ECU72 exhibiting broad resistance not only against *A. socialis*, but also against various species of *Bemisia tabaci* (Omongo et al. 2012; Atim et al. 2024).

Quantitative trait loci (QTL) associated with whitefly resistance have been described in other crops including tomato, soybean, melon, cotton and cabbage (Nombela and Muñoz 2010; Boissot et al. 2010; Xu et al. 2011; Broekgaarden et al. 2018; Aslam et al. 2023). However, no QTL for whitefly resistance has been previously reported in cassava. High-resolution trait mapping in cassava is a significant and ongoing challenge. This is largely due to cassava's inherent genetic complexity: it is highly heterozygous, predominantly clonally propagated and possesses complex trait inheritance. Nonetheless, metabolomic and transcriptomic studies have suggested potential mechanisms of resistance associated with *A. socialis* (Nye et al. 2023; Perez-Fons et al. 2023). To pinpoint the underlying genetic loci, genomics studies are crucial, building upon the insights from these omics approaches regarding resistance to *A. socialis*. Identification of loci linked to whitefly resistance would enable the development of molecular markers thereby increasing the efficiency of trait introgression into elite cassava.

The objective of this study was to identify QTL associated with *A. socialis* whitefly resistance in an F₂ cassava population derived from ECU72, and to develop

molecular markers linked to QTL to enable marker-assisted selection for the trait.

Materials and methods

Plant material

The AM1588 F₂ mapping population used in this study was developed by selfing an F₁ produced from a cross between the whitefly-resistant landrace ECU72 (Bellotti and Arias 2001; Parsa et al. 2015; Atim et al. 2024), and the whitefly-susceptible landrace COL2246. ECU72 which is mostly male sterile, is from Ecuador and was used as the female parent, while COL2246 is from Colombia. The crossing blocks were established at the International Center for Tropical Agriculture (CIAT), Palmira, Colombia, where hand pollinations were performed in the field.

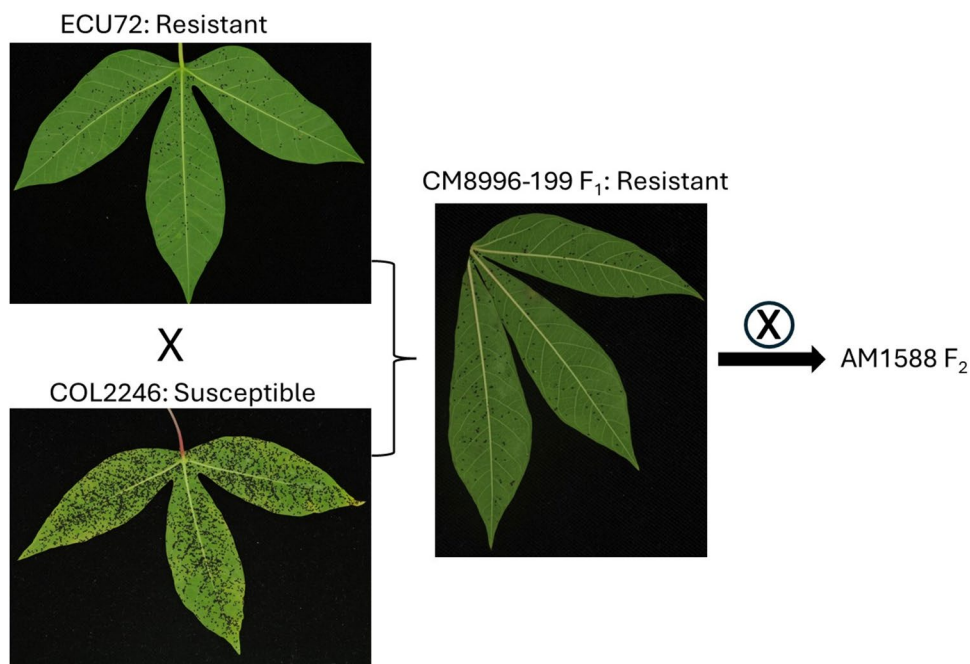
The F₁ individual selected for generating the F₂ mapping population, CM8996-199, exhibited high levels of resistance similar to its resistant parent, ECU72 (Fig. 1) and produced abundant female and male flowers. It was selfed in the 2018–2019 season, resulting in an F₂ progeny of 183 individuals. These progeny were used for mapping and identification of loci associated with whitefly resistance.

To validate the QTL identified in the AM1588 population, two sets of genetic backgrounds were used. These included other F₂ populations derived from the ECU72 background, as well as CIAT's genomic selection training population (GS), consisting of cassava accessions of various pedigrees. Specifically, three F₂ populations with ECU72 genetic background were used: GM12202 (a pseudo-F₂ population derived from CM8996-199 × CM8996-758 with 119 progeny) and two smaller F₂ populations, AM1620 (70 progeny) and AM1621 (24 progeny), both derived from GM8586 F₁ [ECU72 × TMS60444 (susceptible African line)]. Additionally, 673 accessions from various families of the GS population, cohort 2–2021, were also used for QTL validation (Table S1).

Mass rearing of *Aleurotrachelus socialis* (bondar) and phenotyping

A methodology for mass rearing *A. socialis*, developed for glasshouse assays and capable of producing about 7600 whitefly adults per plant was used (Bohorquez-Chaux et al. 2023). Phenotypic evaluations were conducted in a glasshouse over three years: 2020–2021 for F₂ populations, and in 2023 for the GS population using choice experiments. The plants were propagated in one of three methodologies, from in vitro (AM1588), stakes (AM1588 and GS) and micro-stakes (F₂ and pseudo-F₂ validation populations),

Fig. 1 Phenotypes of the two cassava genotypes ECU72 (resistant) and COL2246 (susceptible), and their F_1 (resistant) 34 days post infestation with *A. socialis*. CM8996-199 was selfed for the AM1588 mapping population development



F_2 sexual seeds were planted to obtain mother plants. The mother plants from AM1588 were subsequently used to obtain meristems for in vitro propagation and to produce stakes in field conditions. This allowed us to evaluate potential differences in the phenotyping experiments based on the propagation method. Micro-stakes were obtained from mother plants from the other F_2 for QTL validation. The infestation was made using the methodology described by Bohorquez-Chaux et al. (2023).

An incomplete randomized block design was used separately for each propagation method and population. For the evaluation of AM1588, respective resistant and susceptible checks [ECU72 as resistant, CM8996-199 F_1 which is the parent of the population, as well as an infestation/susceptible check (COL1468)] were included in each batch of plants. Large white mesh cages (18 m length $\times$ 3 m width $\times$ 3 m height) located in a glasshouse were used for the evaluation. Each week three batches of 100 plants (total = 300 plants) with at least five completely opened leaves were infested at the same time. Since the number of genotypes to be evaluated exceeded the capacity of each cage, the number of plants per genotype in every experimental batch was unequal. For AM1588, four plants per genotype were evaluated for each propagation method in eighteen and thirteen experimental batches for plants from in vitro and stakes, respectively. For validation, the F_2 populations (AM1620 F_2 , AM1621 F_2 and GM12202 pseudo- F_2) and GS population were evaluated using the same methodology described for AM1588 F_2 . Thirty-four (34) days after infestation, the two most infested leaves were collected (L1 and L2) and photographs taken in a photo

box. Analysis for nymph count (NC) and percentage of the leaf occupied by nymphs (%Area) were made using the Nymphstar automated nymph counting method (Bohorquez-Chaux et al. 2023). In total, 12 traits were calculated and used for downstream analyses, including NC and %Area for leaf 1, leaf 2 and BLUPs from both leaves, in the in vitro and stake datasets ($2 \times 3 \times 2$). These traits are coded as L1_NC, L2_NC, BLUPs_NC and L1_%Area, L2_%Area and BLUPs_%Area in subsequent text.

Statistical analysis

The statistical analyses were performed in SAS 9.4 software (SAS Institute Inc., Cary, NC) for Linux using the PROC GLIMMIX (generalized linear mixed models—GLMM) procedure. Correlation between the stake and in vitro datasets was assessed using Pearson correlation (r) analysis in the R stats package (R Core Team 2023). Analysis of variance (ANOVA) were performed to measure the effect of genotypes, leaves (L1 and L2), and propagation methods (stake or in vitro) on the NC and %Area. The statistical model applied for each analysis is defined as follows:

$$y = \mu + G + \text{rep} + L + \varepsilon$$

where y represents either the NC or the %Area, μ denotes the overall mean effect, G captures the effect of different cassava genotypes, rep accounts for the different experimental batches, L represents the effect of leaf position (i.e., L1 (younger leaf) and L2), and ε is the residual effect. All effects were treated as random. Best Linear Unbiased

Predictions (BLUPs) for each accession were calculated for the different experiments and used for QTL identification. The phenotypic distributions of scores from the NC and %Area, for stakes and in vitro, were tested for deviations from normality with Shapiro–Wilk tests (Shapiro and Wilk 1965).

Broad sense heritability (H^2) (Cullis et al. 2006) was calculated using the following equation:

$$H^2 = 1 - V_{BLUP}/2\sigma_g^2$$

where V_{BLUP} is the mean–variance difference of two AM1588 F_2 individuals based on BLUPs and σ_g^2 is the genetic variance of these genotypes. A linear mixed-effects model was fitted to estimate the variance components using the R package lme4 (function lmer) (Bates et al. 2015).

Genotyping, SNP analysis and linkage map construction

Total genomic DNA was extracted from fresh young leaf tissues harvested from the AM1588 F_2 family and parents, according to Doyle and Hortorium (1991). RAD sequencing (Davey and Blaxter 2010) was performed on 183 genotypes of the AM1588 F_2 family at BGI (China). Data were aligned to the *Manihot esculenta* v6.1 genome (https://phytozome-next.jgi.doe.gov/info/Mesculenta_v6_1) using BWA-MEM algorithm (Li and Durbin 2010). Single nucleotide polymorphisms (SNPs) were called and filtered using GATK (Depristo et al. 2011), retaining only biallelic SNPs. To ensure high-quality SNPs, additional filters were applied including no more than 10% missingness, $\geq 5\%$ MAF, quality ≥ 30 , as well as a depth ≥ 6 , using VCFtools (Danecek et al. 2011). Finally, a custom Python script was used to remove non-informative SNPs (e.g., identical loci) and those with high segregation distortion (chi-squared test $> 5\%$). The genetic map was constructed using JoinMap v5 (Van Ooijen J., 2006). Since the AM1588 F_2 family does not come from two fully homozygous diploid parents (ECU72 and COL2246), heterozygous SNPs were encoded as hkhk in a .loc file to detect the linkage phases. The results of the phases were then extracted from each linkage group and encoded in an ABH format following this guideline: Initial encoding HH, HK, KK; final encoding to loci with phase {00} A, H, B; final encoding to loci with phase {11} B, H, A. Identical loci and individuals were excluded and linkage groups determined with the independence LOD criterion. The F_2 map was constructed using regression mapping and distance between markers was calculated using the Kosambi mapping function (Kosambi, 1943).

QTL mapping, candidate gene identification and marker development

QTL mapping was performed using datasets from stakes (6 variables) and in vitro (6 variables) of the AM1588 F_2 population. This was determined for L1_NC, L2_NC, BLUPs_NC, L1_%Area, L2_%Area and BLUPs_%Area (total of 12 datasets/traits). Mapping was performed using composite interval mapping (CIM) (Zeng 1994) in WinQTLCart 2.5 (Wang et al. 2007) with threshold values calculated through permutation tests (1,000 permutations, $\alpha = 0.05$) (Churchill and Doerge 1994). CIM analysis was performed with a window size of 10 cM using the standard model (Model 6) with a walk speed of 1 cM and 5 marker cofactors determined by forward and backward regression. Candidate genes within the 2-LOD interval of significant QTL were identified using the *Manihot esculenta* V6 genome (https://phytozome-next.jgi.doe.gov/info/Mesculenta_v6_1).

As a complement, to identify the most significant SNPs associated with whitefly resistance and to compare these results with QTL mapping, a genome-wide association analysis (GWAS) was also performed on the F_2 population using BLUPs for the 12 traits. For the GWAS, a final marker set of 247,117 was used following filtering (missingness ≤ 0.1 , maf ≥ 0.05). Four models were used for the GWAS on the GAPIT R package (Wang and Zhang 2021): the General Linear Model (GLM), Fixed and Random Model Circulating Probability Unification (FarmCPU), Bayesian-Information and Linkage-disequilibrium Iteratively nested Keyway (BLINK) and the Multi-Locus Mixed Model (MLMM) (Price et al. 2006; Segura et al. 2012; Liu et al. 2016; Huang et al. 2019). We incorporated three principal components (PCs) derived from the genotype matrix, along with a kinship matrix to ensure model robustness. We applied a Bonferroni threshold using the formula α/n where α is 0.05 and n the number of SNPs, and we also applied a false discovery rate (FDR) correction using the Benjamini–Hochberg method to control false positives in our models (Bonferroni 1936; Benjamini and Hochberg 1995).

Kompetitive allele specific PCR (KASP) assays for significant SNPs were developed by Intertek Group plc, Australia. Three markers were selected from the peak region of the consistent QTL identified from linkage mapping, and the other thirty-three SNPs obtained from the GWAS (total of 36 markers) for assay development. The allele frequencies and marker quality of the significant SNPs were calculated using a metric that calculates the false positive rate (FPR) and false negative rate (FNR) (Platten et al. 2019; Mbanjo et al. 2024).

Marker validation

Significant markers were validated on the GM12202 pseudo- F_2 , AM1620 F_2 and AM1621 F_2 as well as on the GS training populations described. Leaf punches from the samples were sent to Intertek (Australia) for Low-Density SNP Genotyping (LD SG) using thirty-six KASP markers. These samples included 673 samples from different genetic backgrounds of the GS population as well as a total of 365 samples from ECU72 genetic background, developed for whitefly resistance mapping/validation {AM1588 F_2 (152), AM1620 F_2 (70), AM1621 F_2 (24), GM12202 pseudo- F_2 (119)}. For analysis, ten whitefly-resistant and susceptible checks were also included as haplotype information for these SNPs was available from previous sequencing. These checks included: ECU72, COL2246, TMS60444, CM8996-199, CM8996-758, GM8586-103, GM8586-64, PER368, PER415 and PER608. Resistant (R), intermediate (I) and susceptible (S) categories were determined for the ECU72 genetic background and the genomic selection training population. This was done by performing a K-means analysis based on the BLUPs for NC and %Area using the base R package 'stats' (R Core Team 2023). For each population, we assessed the performance of the K-means analysis when the number of cluster (k) varied from 2 to 10. A quality measure of clusters was estimated based on the sum of squares in which the between-cluster sum of squares was divided by the total within-cluster sum of squares, multiplied by 100. This score ranges from 0 – 100%, in which a higher value indicates a better capture of the variance of clusters, however, a value close to 100% could indicate an overfitting of clustering. Finally, using an Elbow method, the optimal number of clusters per population was determined.

Results

Whitefly resistance in the F_2 mapping population

The phenotypic distribution for whitefly resistance following infestation confirmed the quantitative nature of the trait. The distributions slightly deviated from a normal distribution according to the Shapiro–Wilk test for normality ($P=0.002$) (Fig. 2).

The summary statistics for the 12 traits are presented in Table 1. A significant ($P<0.0001$) positive correlation of $r=0.54$ and $r=0.52$ was observed between the datasets from stakes and in vitro, for NC and %Area, respectively (Fig. S1). ANOVA revealed significant differences ($P<0.0001$) among genotypes, replicates and leaves (Table S2) with a calculated broad-sense heritability (H^2) of 0.58 and 0.56 for in vitro BLUPs_NC and BLUPs_%Area, and 0.73 and 0.74 for stake BLUPs_NC and BLUPs_%Area, respectively.

Genotyping, SNP analysis and map construction

A total of 390,234 SNPs were obtained from the RAD sequencing. After filtering, a genetic map was constructed using 2,017 SNPs on 18 chromosomes, with a total length of 2658.2 cM and an average of 112 markers per chromosome. Chromosome 18 was significantly shorter (66 cM) compared to the other chromosomes, despite relatively high density of markers (101) with shorter inter-locus distance, while chromosome 1 was the longest (164.1 cM) (Fig. 3; Table S3).

QTL identification

For the 12 different traits mapped in the AM1588 F_2 population, QTL were identified in chromosomes 1, 2, 5, 6, 8, 9 and 14 with R^2 ranging from 0.35%–35.44% (Fig. 4; Table 2). Among these, the QTL on chromosome 8 was the most consistent and with high LOD across all twelve traits (maximum LOD = 14.02), with QTL on chromosomes 2, 5, and 14 also consistent across at least 6 traits. These 4 consistent QTL (2, 5, 8 and 14) also had the highest R^2 values. Table 1 represents the various QTL for the different traits with their respective LOD scores. Several minor QTL were identified; therefore, the focus was on the consistently overlapping QTL (2, 5, 8 and 14) across traits.

For chromosome 8 that had overlapping QTL (*MeF2Wfly8.1*) across all traits (both from in vitro and stakes), the R^2 range was 8.4–20.18% for plants from in vitro and 24.13–35.44% for stakes. For chromosome 2, two separate QTL were obtained only from stakes, and none from in vitro data. *MeF2Wfly2.1* overlapped for all six traits from stake data ($R^2=8.49$ –21.8%), while *MeF2Wfly2.2* ($R^2=6.58$ –7.71%) was identified in the same position for two traits (leaf 2 and the total BLUPs for nymph count). For chromosome 5 QTL (*MeF2Wfly5.1*) which were also only obtained from stake data, the locus for all 6 traits overlapped. The R^2 range was 1.42–15.61%.

Although the QTL on chromosome 14 had peaks at different positions, there was overlap among traits (both in vitro and stakes) and the R^2 range was 0.35–11.72% for in vitro and 7.98–14.46% for stakes (Table 2). Overall, the loci from the stake data explained higher phenotypic variation. This is consistent with the broad sense heritability obtained in these experiments, where the 2021 (stakes) assessments had higher H^2 (0.73 NC; 0.74% Area) compared to the moderate H^2 in the 2020 (in vitro) assessments (0.58 NC; 0.56%Area).

GWAS analysis and marker validation

Thirty-three SNPs were identified as significant across the four different models. Only 3 SNPs were consistently significant across all models (Chr08_6483145,

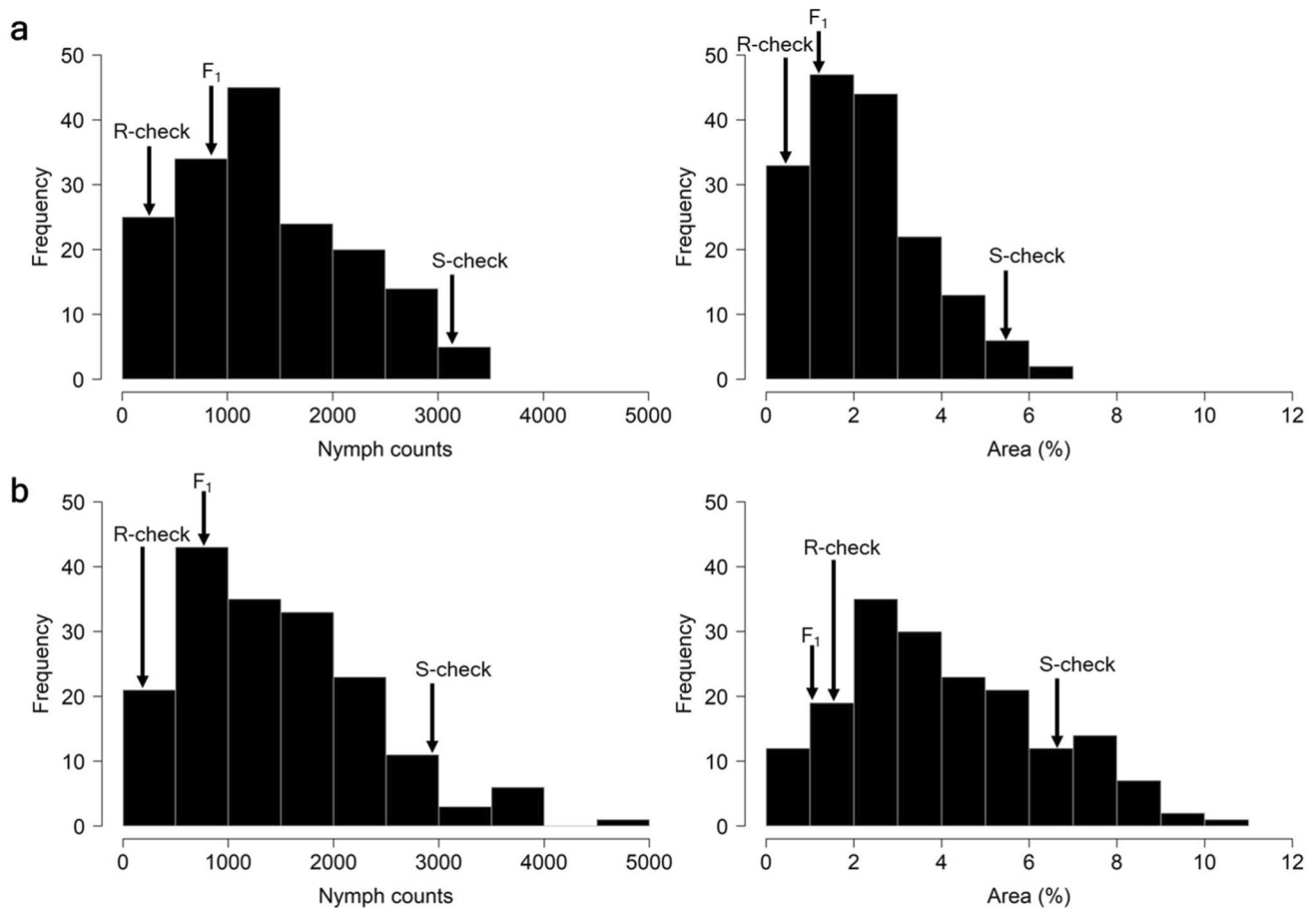


Fig. 2 Frequency distribution of AM1588 F₂ cassava mapping population using **a** stakes and **b** in vitro for nymph counts and percentage of leaf area affected following whitefly infestation ($N=183$). Arrows

indicate the checks (R =resistant, S =susceptible, F_1 =parent of AM1588 F₂ mapping population)

Table 1 Summary statistics of phenotypic data associated with the 12 traits

Plant material	Trait	Checks			AM1588 F ₂	
		CM8996-199 (F ₁)	ECU72	COL1468	Mean	Range
Stake	BLUPs_%Area	1.30	0.51	5.40	2.29	0.10–6.82
Stake	BLUPs_NC	743.13	402.30	3219.89	1381.50	61.17–3332.63
Stake	L1_%Area	1.22	0.42	6.00	2.41	0.45–6.40
Stake	L1_NC	776.50	209.42	3290.15	1436.45	382.61–3176.30
Stake	L2_%Area	1.28	0.51	4.93	2.15	0.45–5.80
Stake	L2_NC	706.73	301.60	3126.00	1326.31	390.36–3074.73
In vitro	BLUPs_%Area	1.04	1.60	6.10	4.05	0.0422–10.39
In vitro	BLUPs_NC	613.50	605.60	3010.37	1542.00	2.25–4805.33
In vitro	L1_%Area	1.36	1.90	7.17	3.70	0.01–12.2
In vitro	L2_%Area	0.69	1.07	5.00	2.25	0.04–8.97
In vitro	L1_NC	847.50	578.80	3248.07	1840.02	771.97–3425.26
In vitro	L2_NC	378.50	330.85	2510.49	1231.30	469.16–2742.46

L1 = leaf 1, L2 = leaf 2, NC = nymph count

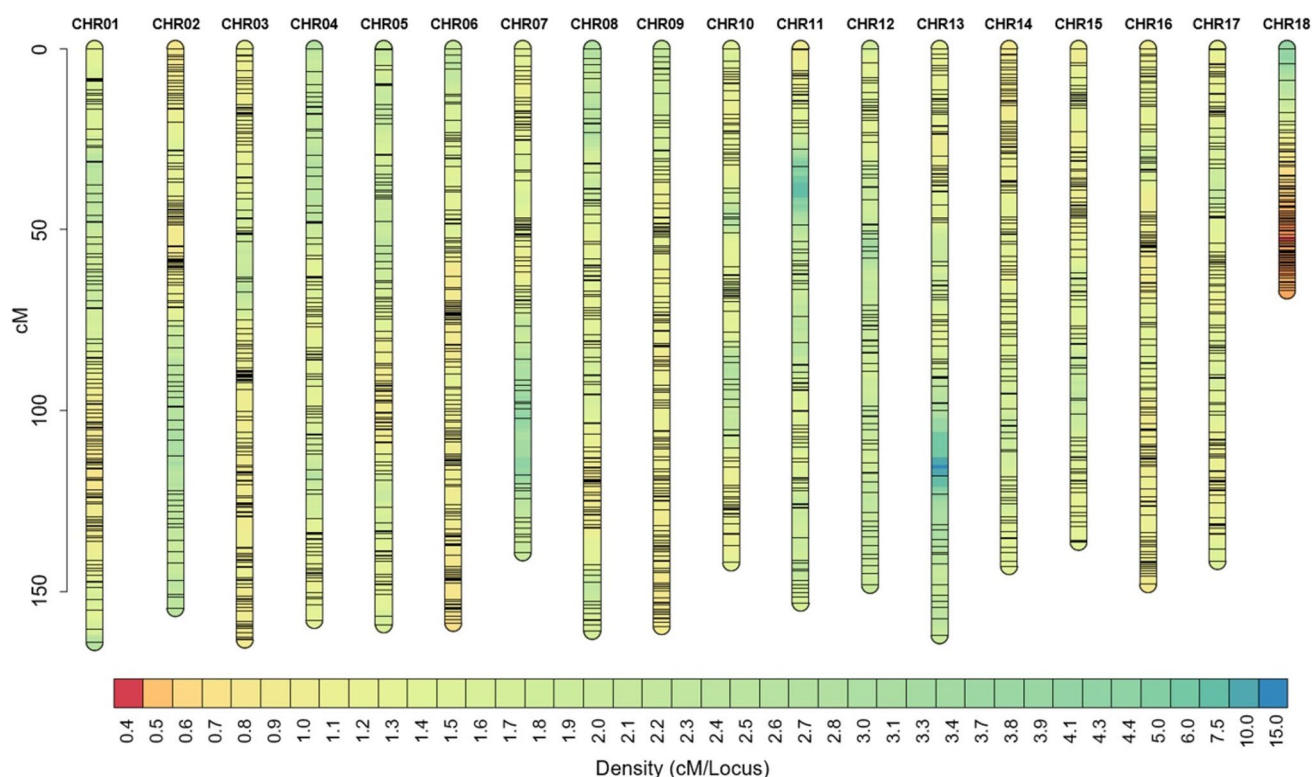


Fig. 3 Genetic map of the AM1588 F₂ population, using 2,017 SNP markers. The horizontal bars on each chromosome represent mapped SNPs, and the scale bar to the left indicates the chromosome length in

cM. The bottom bar with the color gradient represents the intra-locus density, denoting distance between markers

Chr08_6512259, and Chr08_6512325), while the rest of the SNPs were identified using only the GLM and/or FarmCPU models (Fig. S2, Tables 3 and S4). The 3 SNPs were within the chromosome 8 peak region from the QTL mapping results. The most significant marker (Chr08_6483145) had an R^2 value of 25% in the mapping population. The 36 markers (33 from GWAS and 3 in the peak region from linkage mapping) converted to KASP assays were tested for applicability in marker-assisted selection, of which 12 did not amplify correctly and therefore were not informative in the analyses. In the ECU72 background (F₂ mapping and validation populations), markers 1–18 (Table 3) indicated additive segregation patterns, with heterozygotes and genotypes carrying the homozygous unfavorable allele, exhibiting intermediate and susceptible resistance levels to whitefly, respectively. Genotypes with homozygous favorable allele were highly resistant. In contrast, for SNP markers 19–24 (Table 3), there was a dominance segregation pattern with no differences in whitefly resistance levels between genotypes carrying the homozygous favorable allele and heterozygotes, while the genotypes with homozygous unfavorable allele were significantly different, exhibiting susceptibility to whitefly.

Validation of significant SNPs in the ECU72 genetic background and GS training population

For the ECU72 genetic background populations ($N=365$; AM1588 F₂, AM1620 F₂, AM1621 F₂, and GM12202 pseudo-F₂), we observed a clustering quality of 80.58% when $k=3$. Therefore, the 365 genotypes were classified in three clusters: 171 as resistant (R), 142 as intermediate (I), and 52 as susceptible (S). Of the 6 SNPs that showed a dominant segregation pattern, 4 (Chr08_5641063, Chr08_3623233, Chr08_3607706, and Chr08_6618913) had acceptable percentages of FPR and FNR, and the other 2 had very high percentages of either FPR or FNR (Table 2). Of the 18 SNPs that showed an additive segregation pattern, 14 (Chr08_6483145, Chr08_6512259, Chr08_6512325, Chr08_6512307, Chr08_6512329, Chr08_6640496, Chr08_6640586, Chr08_6825255, Chr08_3607807, Chr08_3778097, Chr08_7985587, Chr08_3607718, Chr08_3623304, Chr08_5725346) had acceptable percentages of FPR and FNR, and the other 4 had very high percentages of either FPR or FNR (Table 3).

For the GS training population, a 78.80% quality measure was calculated when $k=3$. The 673 genotypes were therefore classified into three clusters: 228 as

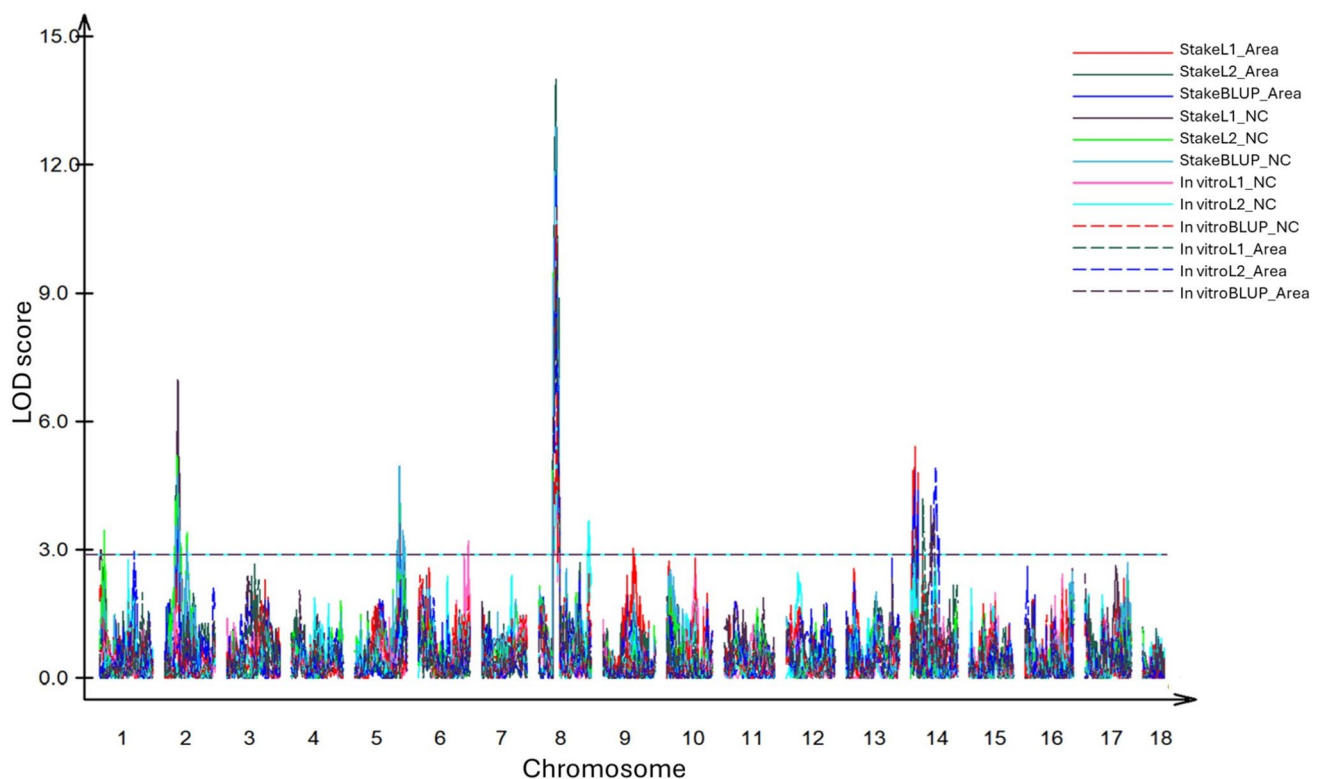


Fig. 4 QTL associated with *Aleurotrachelus socialis* whitefly resistance in the AM1588 F₂ cassava population ($N=183$)

resistant (R), 310 as intermediate (I), and 135 as susceptible (S).

Based on the validation across the ECU72 background and GS training population, 3 SNPs (Chr08_6483145, Chr08_6512325, and Chr08_6640496), which displayed an additive segregation pattern were the most promising as markers for selection, with Chr08_6483145 best applicable in other genetic backgrounds (Fig. 5). 69.8% of the genotypes evaluated had at least one copy of the favorable allele (C) for the best marker (Chr08_6483145) and were in the R (resistant) or I (intermediate) category. For the other significant markers, Chr08_6512325 (C) and Chr08_6640496 (G), 52.6% and 84.8% of the evaluated genotypes had at least one copy of the favorable allele.

Candidate gene annotation and functional analysis

Potential candidate genes within the various QTL were identified (Table S5). Table 4 summarizes genes within the loci that have been previously described as involved in host resistance against whiteflies or other insects. *MeF2Wfly8.1* harbors 115 annotated genes. Worth highlighting, are genes encoding functions related to resistance against sucking insects, such as those involved in cuticular wax formation, cell wall modification, lignin biosynthesis, antibiosis and association with insect

effectors. These genes include MYB106, lectin receptor kinases (LecRLKs), peroxidase 53, strictosidine synthase-like 2 (SSL2), patatin-like phospholipase and trichome birefringence-related gene among others (Table 4). *MeF2Wfly14.1* and *MeF2Wfly14.2* contain 522 and 400 genes, respectively, with key genes including MYB domain proteins and other genes involved in lignin biosynthesis, trichome development, and immune response against phloem-feeding insects, such as aphids and whiteflies (Table 3). The genes in *MeF2Wfly5.1* are 204 and include brassinosteroid insensitive 1 (BRI1), phytoalexin 2 and MYB103 associated with defense mechanisms against insects. The QTL on chromosome 2, *MeF2Wfly2.1* and *MeF2Wfly2.2* contain 426 and 299 genes, respectively. These genes include MYB domain proteins, ABA-deficient 4, terpenoid synthase and HXXXD-type acyl-transferase family protein, which are involved in cuticular wax biosynthesis, trichome development, response to aphids and whiteflies, and secondary metabolism (Table 4).

Table 2 Quantitative trait loci (QTL) associated with *A. socialis* whitefly in the AM1588 F₂ cassava population and the corresponding 2-LOD support interval for different traits

Trait	Plant source	QTL name	Chr	Peak (cM)	LOD ^a	Additive ^b	Dominant ^c	2-LOD interval (cM) ^d	Right flanking marker (Mb)	Left flanking marker (Mb)	R ² (%) ^e
L1_NC	Stake	MeF2WFly1.1	1	5.01	3.02	473.00	148.70	0–8.3	C01_4531869	C01_6648455	3.60
L2_NC	Stake	MeF2WFly1.2	1	14.41	3.47	344.40	177.28	12.9–15.2	C01_7051287	C01_7477364	2.25
L2_%Area	In vitro	MeF2WFly1.3	1	105.31	2.98	–0.56	–0.53	104.7–107	C01_26237311	C01_26584494	1.02
L1_%Area	Stake	MeF2WFly2.1	2	40.01	2.97	–0.45	0.40	30–47.8	C02_2390274	C02_5496285	8.49
L2_NC	Stake	MeF2WFly2.1	2	38.01	5.24	–123.92	489.43	36.1–45.5	C02_3110867	C02_4327514	9.22
BLUPs_NC	Stake	MeF2WFly2.1	2	39.01	4.88	–146.97	483.28	36.6–48	C02_3132983	C02_5496285	9.33
BLUPs_%Area	Stake	MeF2WFly2.1	2	34.21	3.49	–0.46	0.33	33.9–36	C02_2709930	C02_3110867	9.75
L2_%Area	Stake	MeF2WFly2.1	2	34.21	4.51	–0.46	0.43	30.9–43.6	C02_2390274	C02_4167810	12.45
L1_NC	Stake	MeF2WFly2.1	2	40.01	6.98	–353.10	487.97	30.2–47.4	C02_2390274	C02_5496285	21.80
BLUPs_NC	Stake	MeF2WFly2.2	2	68.61	3.17	–96.02	397.38	65.6–76.2	C02_9733910	C02_12321797	6.58
L2_NC	Stake	MeF2WFly2.2	2	68.61	3.4	–113.20	397.01	65.5–71.5	C02_9733910	C02_12321797	7.71
BLUPs_%Area	Stake	MeF2WFly5.1	5	136.41	2.91	–0.02	–0.69	134.5–140.2	C05_25698421	C05_26500914	1.42
L1_%Area	Stake	MeF2WFly5.1	5	136.41	3.32	0.01	–0.85	133.7–140.2	C05_25679518	C05_26500914	1.91
L2_NC	Stake	MeF2WFly5.1	5	135.41	4.74	–34.97	–449.19	134.5–145.3	C05_25698421	C05_26941212	2.05
L1_NC	Stake	MeF2WFly5.1	5	150.81	3.76	122.62	–335.38	134.1–155.9	C05_25698421	C05_28201769	5.12
L2_%Area	Stake	MeF2WFly5.1	5	146.01	2.57	0.38	–0.48	133.5–148	C05_25679518	C05_27076022	10.54
BLUPs_NC	Stake	MeF2WFly5.1	5	147.01	4.98	234.19	–365.54	135.1–148	C05_25896546	C05_27076022	15.61
L1_NC	In vitro	MeF2WFly6.1	6	152.31	3.22	335.78	–259.01	139.3–154.3	C06_25713854	C06_27159371	10.02
L1_NC	In vitro	MeF2WFly8.1	8	54.01	5.64	505.94	181.92	48.4–61.5	C08_5317824	C08_7997296	8.40
L2_NC	In vitro	MeF2WFly8.1	8	54.01	6.17	451.50	77.82	42.5–53.9	C08_6640496	C08_8097664	11.75
BLUPs_NC	In vitro	MeF2WFly8.1	8	54.01	6.71	475.13	99.46	42.5–61.5	C08_5317824	C08_7997296	12.13
L1_%Area	In vitro	MeF2WFly8.1	8	54.01	10.01	1.70	0.31	48.1–61.5	C08_5317824	C08_7997265	18.17
L2_%Area	In vitro	MeF2WFly8.1	8	54.01	10.18	1.34	0.14	41.9–63	C08_5317824	C08_7380199	18.43
BLUPs_%Area	In vitro	MeF2WFly8.1	8	54.01	11.12	1.48	0.18	53.5–61.5	C08_6640496	C08_7380199	20.18
L1_NC	Stake	MeF2WFly8.1	8	52.11	9.62	531.94	–113.48	43.5–62.6	C08_5725346	C08_7380199	24.13
L1_%Area	Stake	MeF2WFly8.1	8	52.11	10.6	1.04	–0.19	43.7–62	C08_5725346	C08_7380199	27.35
BLUPs_%Area	Stake	MeF2WFly8.1	8	52.11	12.16	0.99	–0.23	42.5–60	C08_5250567	C08_7307520	31.06
BLUPs_NC	Stake	MeF2WFly8.1	8	52.11	13.03	588.37	–68.51	43.5–62.6	C08_5250567	C08_7380199	32.46
L2_NC	Stake	MeF2WFly8.1	8	52.11	13.01	570.68	–77.69	53.3–63.9	C08_5250567	C08_6618913	33.02
L2_%Area	Stake	MeF2WFly8.1	8	52.11	14.02	0.99	–0.22	43.5–61.5	C08_5250567	C08_8097606	35.44
L2_NC	In vitro	MeF2WFly8.2	8	152.41	3.7	194.73	–457.65	147.9–153.7	C08_31471311	C08_32428244	9.87
L1_%Area	Stake	MeF2WFly9.1	9	92.91	3.04	–0.49	0.07	87.7–94.5	C09_15940829	C09_18730270	6.22
BLUPs_NC	Stake	MeF2WFly14.1	14	13.01	3.28	165.29	–316.79	12.2–15.7	C14_2692787	C14_4488144	7.98
L1_NC	Stake	MeF2WFly14.1	14	13.01	3.55	178.84	–360.52	12.2–15.7	C14_2692787	C14_4488144	8.33
L2_%Area	Stake	MeF2WFly14.1	14	13.01	3.39	0.32	–0.47	11.5–24.2	C14_2505816	C14_5118096	8.50
BLUPs_%Area	Stake	MeF2WFly14.1	14	13.01	4.89	0.42	–0.59	12.2–14.9	C14_2692787	C14_3068061	12.80

Table 2 (continued)

Trait	Plant source	QTL name	Chr	Peak (cM)	LOD ^a	Additive ^b	Dominant ^c	2-LOD interval (cM) ^d	Right flanking marker (Mb)	Left flanking marker (Mb)	R ² (%) ^e
L1_%Area	Stake	<i>MeF2Wfly/4.1</i>	14	13.01	4.81	0.48	-0.74	1–23	C14_1021098	C14_4844595	14.46
L2_NC	In vitro	<i>MeF2Wfly/4.2</i>	14	75.91	2.79	317.86	-74.39	74.2–78.7	C14_11842045	C14_9988785	7.35
BLUPs_%Area	In vitro	<i>MeF2Wfly/4.2</i>	14	72.41	4.03	0.87	-0.20	54.4–72.9	C14_8089554	C14_11400601	9.20
L2_%Area	In vitro	<i>MeF2Wfly/4.2</i>	14	75.91	4.97	0.93	-0.25	63.9–87.4	C14_8860012	C14_14868789	11.72
L1_%Area	In vitro	<i>MeF2Wfly/4.3</i>	14	37.01	4.21	0.70	1.10	35.1–37.5	C14_6986571	C14_6985846	0.35
BLUPs_NC	In vitro	<i>MeF2Wfly/4.3</i>	14	37.01	2.91	117.12	421.18	36.7–37.5	C14_6986750	C14_6985846	0.40

^aLogarithm of odds ratios at the position of the peak^bAdditive effect of QTL^cDominance effect of QTL^dThe QTL interval on genetic map^ePercent of phenotypic variance explained by the QTL

L1 = leaf 1, L2 = leaf 2, NC = nymph count

Discussion

Phenotyping and genetic control of whitefly resistance

The high-throughput and semi-automated screening method for assessing whitefly resistance responses—Nymphstar (Bohorquez-Chaux et al. 2023) enabled efficient phenotyping of the populations used in this study. The phenotypic distribution of the evaluated traits were slightly skewed in the AM1588 F₂ mapping population, with transgressive segregation observed in the direction of resistance. The correlation between the two datasets of plants derived from in vitro (2020) and those from stakes (2021) was $r=0.54$ for nymph counts, and $r=0.52$ for the percentage of leaf area occupied by nymphs. Broad sense heritability estimates (0.57–0.75) suggested a significant proportion of whitefly resistance is genetically controlled.

The first linkage map of cassava was published by Akano et al., (2002), using RFLPs and SSRs, followed by SNP-based maps (Masumba et al. 2017; Nzuki et al. 2017; Garcia-Oliveira et al. 2020). These maps were generated from bi-parental families with contrasting parents for different traits including CMD, CBSD and cassava green mite (CGM). This study represents the first report of a linkage map from a cassava population generated from selfing (CM8996-199×CM8996-199=AM1588 F₂, 183 individuals) and the first genetic mapping for resistance to *A. socialis* whitefly in cassava. Using high throughput phenotyping and genetic mapping, we identified QTL associated with nymph count and percentage of leaf area occupied by nymphs, on chromosomes 1, 2, 5, 6, 8, 9 and 14, in cassava.

Whitefly resistance in cassava appears to be quantitatively controlled, as evidenced by multiple QTL identified in this population, indicating that the trait is controlled by several genes. The stable QTL on chromosome 8 (*MeF2Wfly8.1*), which explained up to 35.44% of the phenotypic variance holds the highest potential for introgression into elite cassava lines due to its consistent effect across traits. Interestingly, *MeF2Wfly8.1* co-localizes with the QTL underlying resistance to CGM in genetically diverse cassava panels (Ezenwaka et al. 2018; Rabbi et al. 2022) and may be a good target for multi-pest resistance breeding in cassava. Among the detected loci, we focused our attention on QTL on chromosomes 2, 5, 8 and 14, as they were stable across multiple traits, and explained relatively high proportion of phenotypic variation.

Cassava defense mechanisms against whiteflies

Whiteflies and other phloem-feeding insects manipulate plant physiology to their advantage, altering host metabolism, reducing photosynthetic efficiency, and inducing

Table 3 Frequencies of the 24 markers obtained using GWAS and QTL mapping. FPR and FNR are included for each marker, and segregation patterns are shown

No	Marker	Approach	Favorable allele	Unfavorable allele	% Hom favorable allele	% Het	% Hom unfavorable allele	FPR (%)	FNR (%)	Segregation pattern
1	Chr08_6483145 (snpME00572)	GWAS and within <i>MeF2WFLy8.1</i>	C	G	24.5	45.3	30.1	18	25	Additive
2	Chr08_6512259	GWAS and within <i>MeF2WFLy8.1</i>	T	C	22.5	31.5	44.2	14.7	55.7	Additive
3	Chr08_6512325 (snpME00575)	GWAS and within <i>MeF2WFLy8.1</i>	C	T	24.8	27.8	43.1	28	36.5	Additive
4	Chr08_6512307	GWAS and within <i>MeF2WFLy8.1</i>	G	A	14	22.6	63.8	23.3	38.4	Additive
5	Chr08_6512329	GWAS and within <i>MeF2WFLy8.1</i>	C	T	16.8	23.6	45	16	40.3	Additive
6	Chr08_6640496 (snpME00577)	GWAS and within <i>MeF2WFLy8.1</i>	G	A	45.9	38.9	15.3	39	9.6	Additive
7	Chr08_6640586	GWAS and within <i>MeF2WFLy8.1</i>	T	G	22.2	46.2	31	14.3	42.3	Additive
8	Chr08_6825255	GWAS and within <i>MeF2WFLy8.1</i>	A	G	8.4	48.9	37.2	3.2	26.9	Additive
9	Chr08_3607807*	Only GWAS	C	A	16.3	42.8	40.8	7	32.7	Additive
10	Chr08_3778097*	Only GWAS	C	T	17	49.5	33.5	17.4	26.9	Additive
11	Chr08_7985587	GWAS and within <i>MeF2WFLy8.1</i>	G	A	24.4	48.9	25.9	15.9	28.8	Additive
12	Chr08_4043469*	Only GWAS	A	T	28.3	7.7	17.1	58	20.6	Additive
13	Chr08_7617732	GWAS and within <i>MeF2WFLy8.1</i>	C	T	75.5	17.9	6.5	68.8	2	Additive
14	Chr08_3607718*	Only GWAS	T	C	24.2	47.6	27.5	16.2	23	Additive
15	Chr08_3623304*	Only GWAS	T	C	31.3	24.1	36.9	17.6	17.3	Additive
16	Chr08_5627630	GWAS and within <i>MeF2WFLy8.1</i>	G	C	71.7	20.9	6.7	86	13.3	Additive
17	Chr08_5411780	GWAS and within <i>MeF2WFLy8.1</i>	G	A	8.5	30.4	60.4	2.6	67.3	Additive
18	Chr08_5725346**	Within <i>MeF2WFLy8.1</i>	T	A	38.9	41.6	19.1	43.2	11.5	Additive
19	Chr08_5641063	GWAS and within <i>MeF2WFLy8.1</i>	T	A	36.2	4.9	54.4	36	25	Dominant
20	Chr08_3623233*	Only GWAS	C	T	42.2	24	32.5	19.8	21.1	Dominant
21	Chr08_3607706*	Only GWAS	G	C	47	24.5	25.6	27	25	Dominant
22	Chr08_4078044*	Only GWAS	G	T	66.2	26.4	6.2	52.3	8	Dominant
23	Chr08_2600888	Only GWAS	A	G	6.7	27.1	63.3	13.3	67.3	Dominant
24	Chr08_6618913**	Within <i>MeF2WFLy8.1</i>	C	T	19	29.5	45.5	13.7	38	Dominant

* and ** represent markers that were identified using only association or linkage mapping, respectively. All other markers were identified using GWAS and were also within the QTL interval. snpME00572, snpME00575 and snpME00577 are Intertek IDs for the respective markers. Hom = homozygous, Het = heterozygous

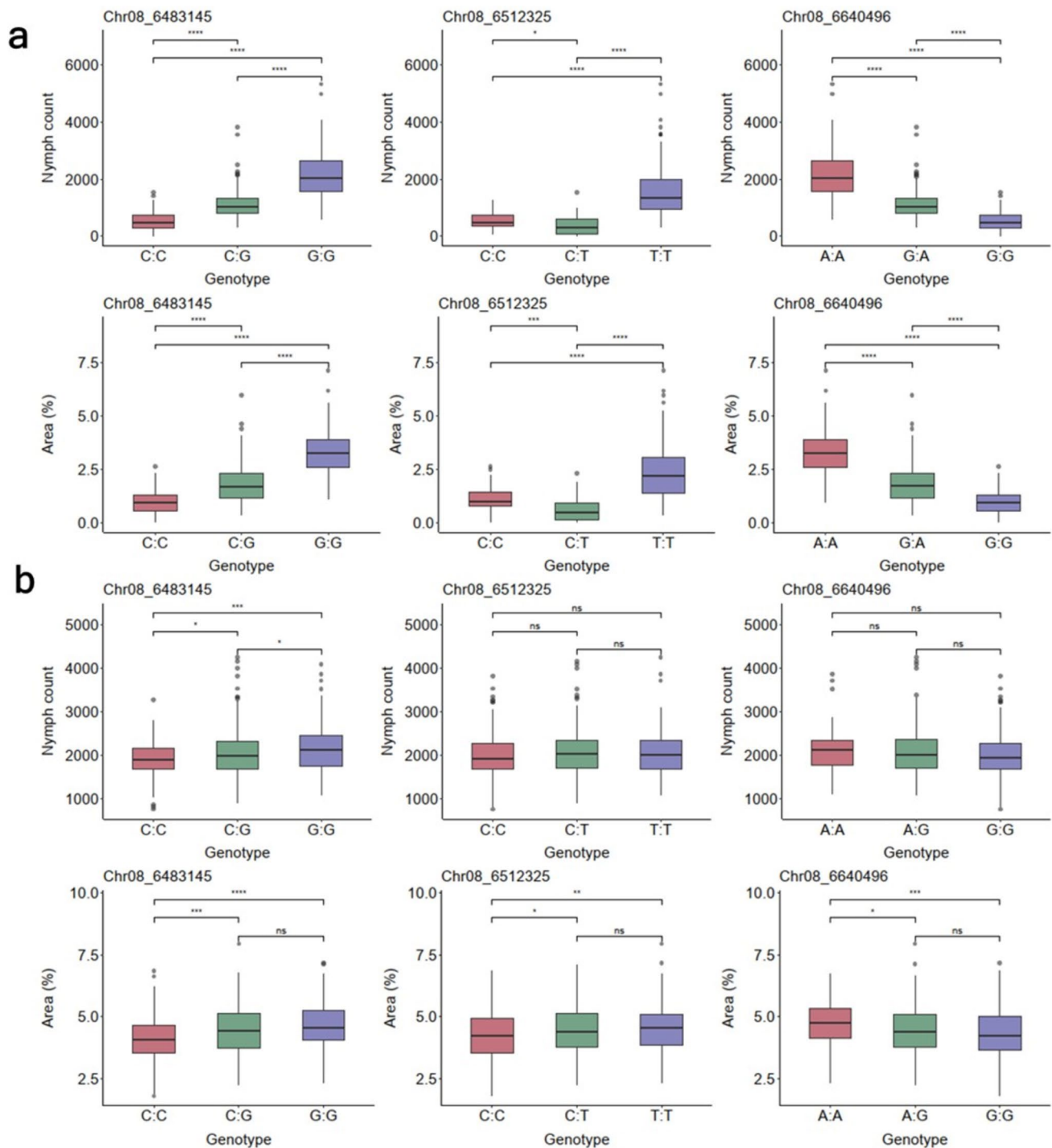


Fig. 5 Boxplots displaying the performance of three KASP markers for whitefly resistance in **a**) ECU72 background and **b**) GS training population, for nymph count and percentage leaf area affected. The

asterisks indicate levels of statistical significance. *, **, ***, **** significant at $p \leq 0.05$, $p \leq 0.01$, $p \leq 0.001$, $p \leq 0.0001$, respectively; ns = non-significant ($p > 0.05$)

structural modifications in plant tissues (Thompson and Goggin 2006). In response, cassava has evolved multiple resistance mechanisms, broadly classified into antibiosis, antixenosis, and hormonal regulation, each mediated by specific genetic pathways.

Antibiosis

One of the key mechanisms identified in this study is antibiosis, whereby plants produce secondary metabolites and other compounds that disrupt insect growth and

Table 4 Potential candidate genes associated with whitefly resistance in the identified loci

Plant source	Trait	Approach	SNP	Chr	Gene ID	Gene description	Gene homolog	References	Biological function, crop and insect species
Stake	BLUPs_NC & L1_NC	<i>MeF2WFLy2.1</i>	Chr02_5122509	2	Manes.02G068600	Abscissic acid (ABA)-deficient 4	ABA4	Guo et al. (2020), Nye et al. (2023)	<i>Aloutrachelus socialis</i> , <i>Bemisia tabaci</i> , <i>Acyrtosiphon pisum</i> , tomato, cassava
Stake	BLUPs_NC & L2_NC	<i>MeF2WFLy2.2</i>	Chr02_10698625	2	Manes.02G144500	HXXXD-type acyltransferase family protein	BAHD acyltransferases	Xu et al. (2023), Nye et al. (2023)	Acylation reactions of secondary metabolites, several crops
Stake	BLUPs_NC, L1_%Area, L1_NC & L2_NC	<i>MeF2WFLy2.1</i>	Chr02_3551813	2	Manes.02G046100	myb domain protein 30	ATMYB30, MYB30	Zhang et al. (2019), Nye et al. (2023)	Cuticular wax biosynthesis, <i>Malus domestica</i>
Stake	BLUPs_NC, L1_%Area, L1_NC & L2_NC	<i>MeF2WFLy2.1</i>	Chr02_3191070	2	Manes.02G041300	myb domain protein 5	ATMYB5, MYB5	Li et al. (2009)	Trichome morphogenesis, <i>Arabidopsis</i>
Stake	BLUPs_%Area & L2_%Area	<i>MeF2WFLy2.1</i>	Chr02_3110867	2	Manes.02G040100	Pentatricopeptide repeat (PPR) superfamily protein		Simon et al. (2020)	<i>Phyllocolpa</i> sp, <i>Populus trichocarpa</i>
Stake	BLUPs_NC, L1_%Area, L1_NC & L2_NC	<i>MeF2WFLy2.1</i>	Chr02_2530211	2	Manes.02G032800	sterol-4alpha-methyl oxidase 1-1	ATSMO1, ATSMO1-1, SMO1-1	Behmer et al. (2013)	<i>Mycus persicae</i> , <i>Nicotiana tabacum</i> , <i>Phaseolus vulgaris</i>
Stake	BLUPs_NC, L1_%Area, L1_NC & L2_NC	<i>MeF2WFLy2.1</i>	Chr02_2718591	2	Manes.02G035200	Terpenoid synthase superfamily protein		Boncan et al. (2020)	Responses to herbivory, several crops
Stake	BLUPs_NC & L1_NC	<i>MeF2WFLy2.1</i>	Chr02_4402369	2	Manes.02G059400	Thioredoxin superfamily protein		Sytykiewicz et al. (2020)	<i>Rhopalosiphum padi</i> & <i>Metopolophium dirhodum</i> , maize
Stake	BLUPs_NC, L2_%Area & L1_NC	<i>MeF2WFLy5.1</i>	Chr05_27223321	5	Manes.05G196700	PHYTOCYSTATIN 2	AtCYS2, CYS2	Martinez et al. (2016)	Hemiptera, Acari, and several crops
Stake	L1_NC	<i>MeF2WFLy5.1</i>	Chr05_27883494	5	Manes.05G203900	WRKY DNA-binding protein 2	ATWRKY2, WRKY2	Tang et al. (2021)	<i>Ostrinia furnacalis</i> , maize
Stake	L1_NC	<i>MeF2WFLy5.1</i>	Chr05_28074549	5	Manes.05G205700	myb domain protein 103	MYB103, MYB80, MS188	Ohman et al. (2013)	syringyl lignin biosynthesis, <i>Arabidopsis</i>

Table 4 (continued)

Plant source	Trait	Approach	SNP	Chr	Gene ID	Gene description	Gene homolog	References	Biological function, crop and insect species
Stake	BLUPs_NC, L2_%Area & L1_NC	<i>MeF2WFIy5.1</i>	Chr05_26523004	5	Manes.05G192200	Transducin/WD40 repeat-like superfamily protein		Guerriero et al. (2015)	Cell wall biosynthesis, <i>Arabidopsis</i>
Stake	L1_NC	<i>MeF2WFIy5.1</i>	Chr05_27867256	5	Manes.05G203800	BRASSINOSTEROID INSENSITIVE 1 (BRI1)	BRI1	Prince et al. (2014)	<i>Myzus persicae</i> , <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS	Chr08_4043469	8	Manes.08G043700	Peroxidase 53-Related		Singh et al. (2013), Nye et al. (2023)	Elicited by <i>Apis craccivora</i> & <i>B. tabaci</i> , <i>A. socialis</i> , Cotton, tomato, cowpea, cassava
Stake & In vitro	All traits	<i>MeF2WFIy8.1</i>	Chr08_5725346	8	Manes.08G055900	strictosidine synthase-like 2	SSL2	Gu et al. (2023)	Monoterpenes alkaloids
Stake & In vitro	All traits	GWAS	Chr08_3778097	8	Manes.08G041400	Dirigent Protein 20-Related	DIR proteins	Pei et al. (2023)	Lignan biosynthesis, <i>Phryna leptostachya</i>
Stake & In vitro	All traits	GWAS	Chr08_4078044	8	Manes.08G044000	Protein Trichome Birefringence-related	TBL	Sun et al. (2020)	Modify cell wall properties through O-acetylation, <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS	Chr08_3778097	8	Manes.08G041800	Protein STICHEL-Like 2	STICHEL	Ilgenfritz et al. (2003)	Trichome branch number, <i>Arabidopsis</i>
Stake & In vitro	All traits	<i>MeF2WFIy8.1</i>	Chr08_5363907	8	Manes.08G053200	Plant invertase/pectin methylesterase inhibitor superfamily	PME	Silva-Sanzana et al. (2019)	<i>Myzus persicae</i> , <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS	Chr08_6483145	8	Manes.08G058400	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein		Yang et al. (2012)	Synthesis of cutin & suberin, <i>Arabidopsis</i>
Stake & In vitro	All traits	<i>MeF2WFIy8.1</i>	Chr08_5725346	8	Manes.08G055700	transhyretin-like protein	TTL	Kyoung et al. (2004)	BRI1 substrate, <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS & <i>MeF2WFIy8.1</i>	Chr08_6640496, 6,640,586	8	Manes.08G059700	basic transcription factor 3	ATBTF3, BTF3	Wang et al. (2014)	Stress response, <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS	Chr08_6483145	8	Manes.08G058500	C2H2-like zinc finger protein	zinc finger protein ZAT3-like	Benyó et al. (2023)	Cell wall biogenesis, <i>Arabidopsis</i>
Stake & In vitro	All traits	<i>MeF2WFIy8.1</i>	Chr08_5725346	8	Manes.08G055800	calmodulin binding		Yadav et al. (2022)	<i>Spodoptera sp</i> , <i>Glycine max</i>

Table 4 (continued)

Plant source	Trait	Approach	SNP	Chr	Gene ID	Gene description	Gene homolog	References	Biological function, crop and insect species
Stake & In vitro	All traits	GWAS	Chr08_5411780	8	Manes.08G054200	IQ-domain 12	IQD12	Levy et al. (2005)	<i>Myzus persicae</i> , <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS	Chr08_6512259, 6,512,307, 6,512,325, 6,512,329	8	Manes.08G058900	Membrane trafficking VPS53 family protein	ATVPS53, HIT1, VPS53	Rodriguez et al. (2017)	<i>Myzus persicae</i> , <i>Arabidopsis</i> , <i>Solanum tuberosum</i>
Stake & In vitro	All traits	<i>MeF2WFFy</i> 8.1	Chr08_5725346	8	Manes.08G056000	Patatin-like phospholipase family protein	SDP1	Simiyu et al. (2023)	Regulates lignin biosynthesis, <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS	Chr08_6512259, 6,512,307, 6,512,325, 6,512,329	8	Manes.08G058800	Concanavalin A-like lectin protein kinase family protein		Tang et al. (2020); Sauvion et al. (2004)	<i>Psyllid</i> , <i>Acyrtosiphon pisum</i> , tomato, pea
Stake & In vitro	All traits	GWAS	Chr08_5627630, 5,641,063	8	Manes.08G055000	dehydration-induced protein (ERD15)	CID1, ERD15, LSR1	Kariola et al. (2006)	<i>Erwinia carotovora</i> , <i>Arabidopsis</i>
Stake & In vitro	All traits	GWAS	Chr08_6825255	8	Manes.08G060300	CDPK-related kinase 1	ATCBK3, ATCRK1, CRK1	Hettenhausen et al. (2016); Santamaria et al. (2018)	<i>Spodoptera exigua</i> , <i>Aphis glycines</i> , <i>Glycine max</i> .
Stake & In vitro	All traits	GWAS	Chr08_6483145	8	Manes.08G058000	myb domain protein 106	MYB106, MIXTA-like R2R3-MYB family	Wang et al. (2020); Oshima et al. (2013)	Cuticular wax formation, <i>Arabidopsis</i> , <i>Eustoma grandiflorum</i>
Stake	L1_%Area	<i>MeF2WFFy</i> 14.1	Chr14_1314046	14	Manes.14G014100	ETHYLENE-INSENSITIVE3-like 3	AEIL3, ATSLIM, EIL3, SLIM1	He et al. (2017)	Stress response, <i>Arabidopsis</i>
Stake	L1_%Area	<i>MeF2WFFy</i> 14.1	Chr14_2200538	14	Manes.14G026300	myb domain protein 19	AMYB19, MYB19	Wang et al. (2017)	<i>Macrosiphoniella sanborni</i> , <i>Chrysanthemum morifolium</i>
Stake	All traits	<i>MeF2WFFy</i> 14.1	Chr14_2863792	14	Manes.14G034400	myb domain protein 42	AMYB42, MYB42	Zhong et al. (2008); Geng et al. (2020)	Lignin biosynthesis during secondary cell wall formation, <i>Arabidopsis</i>
In vitro	BLUPs_%Area & L2_NC	<i>MeF2WFFy</i> 14.2	Chr14_9958663	14	Manes.14G115200	myb domain protein 82	AMYB82, MYB82	Liang et al. (2014)	Trichome development, <i>Arabidopsis</i>

Table 4 (continued)

Plant source	Trait	Approach	SNP	Chr	Gene ID	Gene description	Gene homolog	References	Biological function, crop and insect species
In vitro	BLUPs_ %Area & L2_NC	<i>MeF2WFly14.2</i>	Chr14_11400601	14	Manes.14G128300	plant natriuretic peptide A	PNP-A	Patané et al. (2022)	Effector of <i>Bemisia tabaci</i> , <i>Arabidopsis</i>
Stake	All traits	<i>MeF2WFly14.1</i>	Chr14_2863792	14	Manes.14G035800	RPM1 interacting protein 2	RIN2	Lee et al. (2015)	Immune responses, <i>Arabidopsis</i>
Stake	All traits	<i>MeF2WFly14.1</i>	Chr14_2863792	14	Manes.14G035200	SKU5 similar 2	SKS2	Chen et al. (2023)	Root cell wall formation, <i>Arabidopsis</i>
In vitro	L2_NC & L2_ %Area	<i>MeF2WFly14.2</i>	Chr14_11842022	14	Manes.14G133300	Tetratricopeptide repeat (TPR)-like superfamily protein		Zhou et al. (2021)	Disease resistance, tomato
In vitro	L2_NC & L2_ %Area	<i>MeF2WFly14.2</i>	Chr14_9821251	14	Manes.14G114400	xyloglucan endotransglycosylase 6		Divol et al. (2007); Nye et al. (2023)	<i>Aleurotrachelus socialis</i> , <i>Myzus persicae</i> , <i>Celery</i> , <i>Arabidopsis</i> , cassava

L1 = leaf 1, L2 = leaf 2, NC = nymph count

development, and form barriers that defend the plant from subsequent insect attack (Li et al. 2022). Plant lectins, such as ConA-like lectins, were identified within *MeF2WFly8.1* and play an essential role in plant immunity by disrupting the insect digestion and impairing development (Sauvion et al. 2004; Tang et al. 2020). Similarly, peroxidase 53 contributes to whitefly resistance by triggering the reactive oxygen species (ROS)-mediated localized cell death, creating a barrier against further infestation (Singh et al. 2013; Nye et al. 2023). Some enzyme inhibitors including phytoestrogen 2, block pest digestive proteases, reducing their ability to extract nutrients (Martinez et al. 2016). These findings reinforce the hypothesis that cassava uses protein-based defenses to reduce the survival of whitefly nymphs and limit population expansion.

At the biochemical level, jasmonic acid (JA) signaling pathway enhances plant defense responses. Patatin-like phospholipase proteins identified in *MeF2WFly8.1* have been reported to increase JA accumulation, deterring insect feeding (Canonne et al. 2011; Simiyu et al. 2023). In addition, flavonoids and phenolics regulated by HXXXD-type acyl-transferase proteins (Xu et al. 2023) identified within *MeF2WFly2.2*, also contribute to plant resistance by disrupting insect physiology. Volatile organic compounds (VOCs) synthesized by terpenoid synthases, identified within *MeF2WFly2.1*, may contribute to cassava's ability to deter whiteflies and in field conditions, attract signal predatory insects (Boncan et al. 2020). Furthermore, the presence of strictosidine synthase-like 2 (SSL2) gene in *MeF2WFly8.1* suggests that cassava's resistance mechanisms include monoterpene alkaloid biosynthesis, which provide chemical defense against herbivores and pathogens in other plant species like maize (Gu et al. 2023). Together, these metabolic processes highlight cassava's ability to actively manipulate its chemical environment to resist whitefly infestations.

Antixenosis

Beyond biochemical defenses involved in antibiosis, cassava also relies on structural barriers (antixenosis), preventing insect feeding and egg-laying (Bellotti and Arias 2001). These barriers include waxes, cell wall modifications, changes in cutin or suberin, and the formation of trichomes. Lignin biosynthesis is a well-documented resistance mechanism in plants against sap-sucking insects. Metabolomics and transcriptomic studies have shown that cell wall reinforcement and changes in lignin content contribute to resistance against pests such as *A. socialis* in cassava (Perez-Fons et al. 2019; Nye et al. 2023). Interestingly Ferguson et al. (2023) found evidence for involvement of the lignin pathway in resistance to CBSD. Notably, key genes responsible for lignin biosynthesis were

located within CBSD resistance QTL. This aligns with earlier findings by Amuge et al. (2017), who demonstrated that these lignin biosynthesis genes were upregulated in CBSD-resistant cassava when challenged with the Ugandan cassava brown streak virus. Consequently, genes involved in the lignin pathway hold significant promise as targets for breeding efforts aimed at developing cassava varieties resistant to both CBSD and whiteflies. In this study, MYB domain proteins involved in plant cuticle formation, trichome development and lignin biosynthesis which are crucial for plant defense were also identified in the QTL. Specific examples are MYB106 that regulates cuticular wax formation (Oshima and Mitsuda 2013; Wang et al. 2020), reported in loci associated with CGM severity (Ezenwaka et al. 2018; Rabbi et al. 2022) and MYB103 involved in lignin deposition and secondary cell wall reinforcement (Ohman et al. 2013), which is related to resistance against *A. socialis* whitefly in cassava (Nye et al. 2023). Notably, our lead SNP for whitefly resistance (Chr08_6483145) lies approximately 74 kb downstream of S8_6409580, the top marker for CGM resistance in Rabbi et al. (2022), suggesting that this region may harbor gene(s) conferring broad-spectrum resistance. Other MYB domains (19, 42 and 82) were identified, associated with lignin biosynthesis, secondary cell wall thickening, trichome development and aphid resistance (Zhong et al. 2008; Liang et al. 2014; Wang et al. 2017; Geng et al. 2020). Additional genes associated with plant defense against insects through cell wall structure modifications are described in Table 3 (Ilgenfritz et al. 2003; Divol et al. 2007; Yang et al. 2012; Behmer et al. 2013; Guerriero et al. 2015; Silva-Sanzana et al. 2019; Sun et al. 2020; Benyó et al. 2023; Nye et al. 2023; Pei et al. 2023).

Trichomes play a role in resistance to pests such as whiteflies, though their role varies by plant species. For example, in tomato and *Nicotiana*, glandular trichomes producing acyl sugars are associated with resistance to the whitefly *Bemisia argentifolii* (Liedl 1995), while cassava studies show no strong correlation between trichome density and whitefly resistance (Parsa et al. 2015; Pastório et al. 2023). Genes involved in trichome formation were identified in this study, however from observation of our populations, we did not note any correlation between genotypes that had visible trichomes, with high levels of whitefly resistance.

Hormonal signaling

Hormonal regulation is another critical component of whitefly resistance, orchestrating defense responses at a systemic level through abscisic acid (ABA), brassinosteroids (BRs), and ethylene (ET). ABA-deficient mutants lack sufficient callose deposition, making them more susceptible to aphids (*Acyrtosiphon pisum*) and whiteflies (*Bemisia tabaci*) (Guo et al. 2020; Nye et al. 2023). Brassinosteroid

signaling, mediated by brassinosteroid insensitive 1 (BRI1) and its co-receptor BRI1-associated kinase 1 (BAK1), supports plant growth and pattern-triggered immunity (PTI) restricting herbivore success, as seen in the green peach aphid's (*Myzus persicae*) interaction with *Arabidopsis* (Prince et al. 2014). Other hormone signaling genes identified include ethylene-insensitive3 (EIN3) transcription factors which enhance defense gene expression against herbivores through ethylene signaling (He et al. 2017). This pathway interacts with JA signaling, inducing secondary metabolite production and ROS accumulation. The synergistic regulation of ABA, BRs, and ET provides cassava with a multi-layered resistance system, reinforcing physical barriers, metabolic defenses, and signaling cascades that deter whitefly infestation and feeding. Overall, the QTL identified in the study harbor several key genes that potentially contribute to increased levels of resistance against whiteflies in cassava by enhancing various plant defense mechanisms.

Development and validation of KASP markers for breeding

KASP markers have been developed and implemented for multiple traits in cassava. These include markers for CMD, dry matter content, total carotenoid content, and hydrogen cyanide concentration (Esuma et al. 2022; Rabbi et al. 2022; Kanaabi et al. 2024; Mbanjo et al. 2024). These markers have proven invaluable for accelerating genetic gains and facilitating precision breeding through marker-assisted selection (MAS). In this study, we identified and validated KASP markers associated with whitefly (*A. socialis*) resistance in cassava. The study primarily focused on *A. socialis*, however, the whitefly resistance source (ECU72) demonstrates resistance not only against *A. socialis*, but also against various cryptic species of *Bemisia tabaci* (Omongo et al. 2012; Atim et al. 2024), with likely similar resistance mechanisms. The validation process involved individuals from multiple genetic backgrounds including progeny derived from ECU72 and other genetic backgrounds in the genomic selection training population of the cassava breeding program at CIAT. Three SNPs displayed the highest association with whitefly resistance, with Chr08_6483145 showing the strongest and most consistent association with whitefly resistance, demonstrating high potential for marker-based breeding applications across diverse genetic backgrounds. The utility of these SNPs for marker-assisted selection should be validated in segregating populations from additional genetic backgrounds in other cassava breeding programs, as they show great potential for the genetic enhancement of whitefly resistance in the global cassava community. As more genomic tools and resources

are refined for cassava, marker-assisted breeding will significantly improve genetic gains for vital agronomic and quality traits.

Conclusion

In this study, we employed linkage and association mapping to identify genetic loci conferring whitefly resistance in a cassava F_2 population. Among the identified QTL, *MeF2Wfly8.1* demonstrated stable expression and accounted for up to 35.44% of the observed phenotypic variance. This stability across various genetic backgrounds highlights its significant potential for enhancing whitefly resistance in cassava. Markers applicable in selection were developed and validated, with Chr08_6483145 emerging as the most significant. This marker shows great promise for use in marker-assisted selection (MAS) to improve whitefly resistance in different cassava genetic backgrounds.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00122-025-04949-1>.

Acknowledgements We thank Carmen Adriana Bolanos, Janneth Gutierrez, Luz Andrea Gomez, Juan Cuasquer and the CIAT Cassava Genetics, Entomology and Breeding teams for technical assistance. The authors also thank Paul Fraser, Laura Perez-Fons and Linda Walling for their insightful contributions during the planning and execution of this study.

Authors contributions ABC and LABL-L conceived the project idea, designed the research and secured the financial investment for this project. ABC and MIG conducted glasshouse experiments and phenotyping using images. VB, CS, LFD, ABC, and WG were involved in data analysis. XZ designed the field trials for the genomic selection population used for marker validation. ABC and WG wrote the first draft of the manuscript. All authors contributed to and approved the final manuscript.

Funding This research was funded by the Bill and Melinda Gates Foundation (BMGF) through the Natural Resources Institute (NRI), University of Greenwich, UK, under the African Cassava Whitefly Project Phase II (ACWP II) Project Number: INV-010435. We are also grateful for the genotyping support from the CGIAR Breeding Resources Initiative.

Data availability Data is provided in Electronic Supplementary Material. The leaf images used for phenotyping the mapping population are uploaded on CassavaBase (<https://cassavabase.org/>).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akano AO, Dixon AGO, Mba C, Barrera E, Fregene M (2002) Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theor Appl Genet* 105(4):521–525. <https://doi.org/10.1007/s00122-002-0891-7>
- Alicai T, Omongo CA, Maruthi MN, Hillocks RJ, Baguma Y, Kawuki R, Bua A, Otim-Nape GW, Colvin J (2007) Re-emergence of cassava brown streak disease in Uganda. *Plant Dis* 91(1):24–29. <https://doi.org/10.1094/PD-91-0024>
- Amuge T, Berger DK, Katari MS, Myburg AA, Goldman SL, Ferguson ME (2017) A time series transcriptome analysis of cassava (*Manihot esculenta* Crantz) varieties challenged with Ugandan cassava brown streak virus. *Sci Rep* 7:9747. <https://doi.org/10.1038/s41598-017-09617-z>
- Angel JC, Pineda LB, Nolt B, Velasco AC (1989) Whiteflies (Homoptera: Aleyrodidae) associated with the transmission of cassava viruses. *Fitopatologia Colombiana* 13(2):65–71
- Angel SJ, Nolt B & Pineda L, B (1987) Estudios sobre la transmisión por “moscas blancas” (homoptera: aleyrodidae) de virus asociados con el “cuero de sapo” en yuca (*manihot esculenta crantz*). Universidad Nacional de Colombia Sede Palmira
- Aslam MH, Khan N, Fatima M, Rashid MA, Davies SJ (2023) Strategic replacement of soybean meal with local cotton seed meal on growth performance, body composition, and metabolic health status indicators in the major South Asian carp *Catla catla* for aquaculture. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0296220>
- Atim J, Kalyebi A, Bohorquez-Chaux A, Lopez-Lavalle LAB, Omongo CA, Colvin J, Maruthi MN (2024) Identifying cassava genotypes resistant to the african cassava whitefly, *Bemisia tabaci* (Genadius). *Agriculture* 14(7):1016. <https://doi.org/10.3390/agriculture14071016>
- Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw*. <https://doi.org/10.18637/jss.v067.i01>
- Behmer ST, Olszewski N, Sebastiani J, Palka S, Sparacino G, Sciarrno E, Grebenok RJ (2013) Plant phloem sterol content: forms, putative functions, and implications for phloem-feeding insects. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2013.00370>
- Bellotti AC, Arias B (2001) Host plant resistance to whiteflies with emphasis on cassava as a case study. *Crop Prot* 20(9):813–823
- Bellotti A, Campo BVH, Hyman G (2012) Cassava production and pest management: present and potential threats in a changing environment. In *Tropical Plant Biol* 5(1):39–72. <https://doi.org/10.1007/s12042-011-9091-4>
- Bellotti AC, Smith L, & Lapointe SL (1999) Recent advances in cassava pest management. In *Annu. Rev. Entomol* (Vol. 44). www.annualreviews.org.
- Bellotti AC (1987) *Cassava breeding: a multidisciplinary review: Workshop: Papers*. Centro Internacional de Agricultura Tropical

- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc: Ser B (Methodol)* 57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Benyó D, Bató E, Faragó D, Rigó G, Domonkos I, Labhane N, Zsigmond L, Prasad M, Nagy I, Szabados L (2023) The zinc finger protein 3 of *Arabidopsis thaliana* regulates vegetative growth and root hair development. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2023.1221519>
- Bohorquez-Chaux A, Gómez-Jiménez MI, Leiva-Sandoval LF, Lopez-Lavalle LAB (2023) Nymphstar: an accurate high-throughput quantitative method for whitefly (*Aleurotrachelus socialis* Bondar) resistance phenotyping in cassava. *Plant Phenome J*. <https://doi.org/10.1002/ppj.2.20089>
- Boissot N, Thomas S, Sauvion N, Marchal C, Pavis C, Dogimont C (2010) Mapping and validation of QTLs for resistance to aphids and whiteflies in melon. *Theor Appl Genet* 121(1):9–20. <https://doi.org/10.1007/s00122-010-1287-8>
- Boncan DAT, Tsang SSK, Li C, Lee IHT, Lam HM, Chan TF, Hui JHL (2020) Terpenes and terpenoids in plants: Interactions with environment and insects. *Int J Mol Sci* 21(19):7382. <https://doi.org/10.3390/ijms21197382>
- Bonferroni C (1936) Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni Del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8:3–62. <https://doi.org/10.4135/9781412961288.n455>
- Broekgaarden C, Pelgrom KTB, Bucher J, Van Dam NM, Grosser K, Pieterse CMJ, Van Kaauwen M, Steenhuis G, Voorrips RE, De Vos M, Vosman B, Worrish A, Van Wees SCM (2018) Combining QTL mapping with transcriptome and metabolome profiling reveals a possible role for ABA signaling in resistance against the cabbage whitefly in cabbage. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0206103>
- Canonne J, Froidure-Nicolas S, Rivas S (2011) Phospholipases in action during plant defense signaling. *Plant Signal Behav* 6(1):13–18. <https://doi.org/10.4161/psb.6.1.14037>
- Carabalí A, Bellotti AC, Montoya-Lerma J, Fregene M (2010) Resistance to the whitefly, *Aleurotrachelus socialis*, in wild populations of cassava, *Manihot tristis*. *J Insect Sci* 10(1):170
- Chen C, Zhang Y, Cai J, Qiu Y, Li L, Gao C, Gao Y, Ke M, Wu S, Wei C, Chen J, Xu T, Friml J, Wang J, Li R, Chao D, Zhang B, Chen X, Gao Z (2023) Multi-copper oxidases SKU5 and SKS1 coordinate cell wall formation using apoplastic redox-based reactions in roots. *Plant Physiol* 192(3):2243–2260. <https://doi.org/10.1093/plphys/kiad207>
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138(3):963–971
- Colvin J, Omongo CA, Maruthi MN, Otim-Nape GW, Thresh JM (2004) Dual begomovirus infections and high *Bemisia tabaci* populations: two factors driving the spread of a cassava mosaic disease pandemic. *Plant Pathol* 53:577–584. <https://doi.org/10.1111/j.1365-3059.2004.01062.x>
- Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated data. *J Agric Biol Environ Stat* 11(4):381–393. <https://doi.org/10.1198/108571106X154443>
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Lunter GT, Sherry ST, McVean G, Durbin R (2011) The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Davey JL, Blaxter MW (2010) RADseq: next-generation population genetics. *Brief Funct Genomics* 9(5–6):416–423. <https://doi.org/10.1093/bfpg/elq031>
- de Moya RS, Brown JK, Sweet AD, Kimberly KK, Paredes-Montero JR, Waterhouse RM, Johnson KP (2019) Nuclear orthologs derived from whole genome sequencing indicate cryptic diversity in the *Bemisia tabaci* (Insecta: Aleyrodidae) complex of whiteflies. *Diversity* 11(9):151. <https://doi.org/10.3390/d11090151>
- Depristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491–501. <https://doi.org/10.1038/ng.806>
- Divol F, Vilaine F, Thibivilliers S, Kusiak C, Sauge MH, Dinant S (2007) Involvement of the xyloglucan endotransglycosylase/hydrolases encoded by celery XTH1 and *Arabidopsis* XTH33 in the phloem response to aphids. *Plant, Cell Environ* 30(2):187–201. <https://doi.org/10.1111/j.1365-3040.2006.01618.x>
- Doyle J, & Hortorium B (1991) DNA protocols for plants CTAB Total DNA isolation
- Esuma W, Eyoo O, Gwandu F, Mukasa S, Alicai T, Ozimati A, Nuwamanya E, Rabbi I, Kawuki R (2022) Validation of KASP markers associated with cassava mosaic disease resistance, storage root dry matter and provitamin A carotenoid contents in Ugandan cassava germplasm. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2022.1017275>
- Ezenwaka L, Del Carpio DP, Jannink JL, Rabbi I, Danquah E, Asante I, Danquah A, Blay E, Egesi C (2018) Genome-wide association study of resistance to cassava green mite pest and related traits in cassava. *Crop Sci* 58(5):1907–1918. <https://doi.org/10.2135/cropsci2018.01.0024>
- Farias A (1994) Flutuação populacional de *Aleurothrixus aepim* em mandioca, em São Miguel das Matas, Bahia. *Revista Brasileira Da Mandioca* 13:119–122
- Ferguson ME, Eyles RP, Garcia-Oliveira AL, Kapinga F, Masumba EA, Amuge T, Bredeson JV, Rokhsar DS, Lyons JB, Shah T, Rounsley S, Mkamilo G (2023) Candidate genes for field resistance to cassava brown streak disease revealed through the analysis of multiple data sources. *Front Plant Sci* 14:1270963. <https://doi.org/10.3389/fpls.2023.1270963>
- Garcia-Oliveira LKB, Kasele S, Kapinga F, Masumba E, Mkamilo G, Sicalwe C, Bredeson JV, Lyons JB, Hah T, Muranaka S, Katari MS, Ferguson ME (2020) Genetic analysis and qtl mapping for multiple biotic stress resistance in cassava. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0236674>
- Geng P, Zhang S, Liu J, Zhao C, Wu J, Cao Y, Fu C, Han X, He H, Zhao Q (2020) MYB20, MYB42, MYB43, and MYB85 regulate phenylalanine and lignin biosynthesis during secondary cell wall formation. *Plant Physiol* 182(3):1272–1283. <https://doi.org/10.1104/PP.19.01070>
- Gu L, Cao Y, Chen X, Wang H, Zhu B, Du X, Sun Y (2023) The genome-wide identification, characterization, and expression analysis of the strictosidine synthase-like family in maize (*Zea mays* L.). *Int J Mol Sci*. <https://doi.org/10.3390/ijms241914733>
- Guerriero G, Hausman JF, Ezcurra I (2015) Wd40-repeat proteins in plant cell wall formation: current evidence and research prospects. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2015.01112>
- Guo H, Sun Y, Yan H, Li C, Ge F (2020) O3-induced priming defense associated with the abscisic acid signaling pathway enhances plant resistance to *Bemisia tabaci*. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2020.00093>
- He X, Jiang J, Wang CQ, Dehesh K (2017) ORA59 and EIN3 interaction couples jasmonate-ethylene synergistic action to antagonistic salicylic acid regulation of PDF expression. *J Integr Plant Biol* 59(4):275–287. <https://doi.org/10.1111/jipb.12524>
- Hettenhausen C, Sun G, He Y, Zhuang H, Sun T, Qi J, Wu J (2016) Genome-wide identification of calcium-dependent protein kinases in soybean and analyses of their transcriptional responses to insect herbivory and drought stress. *Scientif Rep*. <https://doi.org/10.1038/srep18973>

- Huang M, Liu X, Zhou Y, Summers RM, Zhang Z (2019) BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. GigaScience. <https://doi.org/10.1093/gigascience/giy154>
- Ilgenfritz H, Bouyer D, Schnittger A, Mathur J, Kirik V, Schwab B, Chua NH, Jürgens G, Hülskamp M (2003) The arabidopsis STICHEL gene is a regulator of trichome branch number and encodes a novel protein. *Plant Physiol* 131(2):643–655. <https://doi.org/10.1104/pp.014209>
- Jimenez J, Caicedo S, Pardo JM, Gil-Ordoñez A, Alvarez-Quinto R, Molloy D, Cuellar WJ et al (2024) Single torradovirus infections explain the mysterious cassava frogskin disease in the Americas. *Sci Rep* 14:29648. <https://doi.org/10.1038/s41598-024-81142-2>
- Kanaabi M, Mukasa SB, Nuwamanya E, Iragaba P, Baguma JK, Nanyonjo AR, Wagaba H, Muhumuza N, Namakula FB, Wembabazi E, Ozimati A, Kayondo IS, Esuma W, Kawuki RS (2024) Validation of KASP markers associated with hydrogen cyanide in fresh cassava roots in uganda cassava germplasm. *Agronomy* 14(12):2765. <https://doi.org/10.3390/agronomy14122765>
- Kariola T, Brader G, Helenius E, Li J, Heino P, Palva ET (2006) Early responsive to dehydration 15, a negative regulator of abscisic acid responses in arabidopsis. *Plant Physiol* 142(4):1559–1573. <https://doi.org/10.1104/pp.106.086223>
- Kosambi DD (1943) The estimation of map distances from recombination values. *Ann Eugen* 12(1):172–175. <https://doi.org/10.1111/j.1469-1809.1943.tb02321.x>
- Kyoung HN, Li J (2004) The arabidopsis transthyretin-like protein is a potential substrate of Brassinosteroid-Insensitive 1. *Plant Cell* 16(9):2406–2417. <https://doi.org/10.1105/tpc.104.023903>
- Lee DH, Bourdais G, Yu G, Robatzek S, Coaker G (2015) Phosphorylation of the plant immune regulator RPM1-INTERACTING PROTEIN4 enhances plant plasma membrane h⁺-ATPase activity and inhibits flagellin-triggered immune responses in arabidopsis. *Plant Cell* 27(7):2042–2056. <https://doi.org/10.1105/tpc.114.132308>
- Levy M, Wang Q, Kaspi R, Parrella MP, Abel S (2005) Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *Plant J* 43(1):79–96. <https://doi.org/10.1111/j.1365-3113X.2005.02435.x>
- Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26(5):589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Li SF, Milliken ON, Pham H, Seyit R, Napoli R, Preston J, Koltunow AM, Parish RW (2009) The Arabidopsis MYB5 transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. *Plant Cell* 21(1):72–89. <https://doi.org/10.1105/tpc.108.063503>
- Li S, Liu S, Zhang Q, Cui M, Zhao M, Li N, Wang S, Wu R, Zhang L, Cao Y, Wang L (2022) The interaction of ABA and ROS in plant growth and stress resistance. *Front Plant Sci* 13:1050132. <https://doi.org/10.3389/fpls.2022.1050132>
- Liang G, He H, Li Y, Ai Q, Yu D (2014) MYB82 functions in regulation of trichome development in Arabidopsis. *J Exp Bot* 65(12):3215–3223. <https://doi.org/10.1093/jxb/eru179>
- Liedl B (1995) Acylsugars of wild tomato *Lycopersicon penellii* alters settling and reduce oviposition of *Bemisia argentifolii* (Homoptera: Aleyrodidae). *J Econ Entomol* 88:742
- Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet*. <https://doi.org/10.1371/journal.pgen.1005767>
- MacFadyen S, Paull C, Boykin LM, De Barro P, Maruthi MN, Otim M, Kalyebi A, Vassão DG, Sseruwagi P, Tay WT, Delatte H, Seguni Z, Colvin J, Omongo CA (2018) Cassava whitefly, *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) in East African farming landscapes: a review of the factors determining abundance. *Bull Entomol Res* 108(5):565–582. <https://doi.org/10.1017/S0007485318000032>
- Martinez M, Santamaria ME, Diaz-Mendoza M, Arnaiz A, Carrillo L, Ortego F (2016) Phytocystatins: defense proteins against phytophagous insects and acari. In *Int J Mol Sci* 17(10):1747. <https://doi.org/10.3390/ijms17101747>
- Maruthi MN, Hillocks RJ, Mtunda K, Raya MD, Muhanna M, Kiozia H, Rekha AR, Colvin J, Thresh JM (2005) Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). *J Phytopathol* 153(5):307–312. <https://doi.org/10.1111/j.1439-0434.2005.00974.x>
- Masumba EA, Kapinga F, Mkamilo G, Salum K, Kulembeka H, Rounsley S, Bredeson JV, Lyons JB, Rokhsar DS, Kanju E, Katari MS, Myburg AA, van der Merwe NA, Ferguson ME (2017) QTL associated with resistance to cassava brown streak and cassava mosaic diseases in a bi-parental cross of two Tanzanian farmer varieties, Namikonga and Albert. *Theor Appl Genet* 130(10):2069–2090. <https://doi.org/10.1007/s00122-017-2943-z>
- Mbanjo EGN, Ogungbesan A, Agbona A, Akpotuzor P, Toyinbo S, Iluebey P, Rabbi IY, Peteti P, Wages SA, Norton J, Zhang X, Bohórquez-Chaux A, Mushoriwa H, Egesi C, Kulakow P, Parkes E (2024) Validation of SNP markers for diversity analysis, quality control, and trait selection in a biofortified cassava population. *Plants*. <https://doi.org/10.3390/plants13162328>
- Mugerwa H, Wang HL, Sseruwagi P, Seal S, Colvin J (2021) Whole-genome single nucleotide polymorphism and mating compatibility studies reveal the presence of distinct species in sub-Saharan Africa *Bemisia tabaci* whiteflies. *Insect Sci* 28(6):1553–1566. <https://doi.org/10.1111/1744-7917.12881>
- Nelson S (2008). Sooty Mold. *Plant Disease*, 52, 1–6 <http://www.ctahr.hawaii.edu/freepubs>
- Nombela G, Muñoz M (2010) Host plant resistance for the management of *Bemisia tabaci*: a multi-crop survey with emphasis on tomato. *Bionomics and Management of a Global Pest*, In *Bemisia*. https://doi.org/10.1007/978-90-481-2460-2_14
- Nye DG, Irigoyen ML, Perez-Fons L, Bohorquez-Chaux A, Hur M, Medina-Yerena D, Lopez-Lavalle LAB, Fraser PD, Walling LL (2023) Integrative transcriptomics reveals association of abscisic acid and lignin pathways with cassava whitefly resistance. *BMC Plant Biol*. <https://doi.org/10.1186/s12870-023-04607-y>
- Nzuki I, Katari MS, Bredeson JV, Masumba E, Kapinga F, Salum K, Mkamilo GS, Shah T, Lyons JB, Rokhsar DS, Rounsley S, Myburg AA, Ferguson ME (2017) QTL mapping for pest and disease resistance in cassava and coincidence of some QTL with introgression regions derived from *Manihot Glaziovii*. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2017.01168>
- Ohman D, Demedts B, Kumar M, Gerber L, Gorzsa A, Goeminne G, Hedenström M, Ellis B, Boerjan W, Sundberg B (2013) MYB103 is required for FERULATE-5-HYDROXYLASE expression and syringyl lignin biosynthesis in Arabidopsis stems. *Plant J* 73(1):63–76. <https://doi.org/10.1111/tpj.12018>
- Omongo C, Kawuki R, Bellotti A, Alicai T, Baguma Y, Maruthi M, Bua A, Colvin J (2012) African cassava whitefly, *Bemisia tabaci*, Resistance in African and South American cassava genotypes. *J Integr Agric* 11(2):327–336. [https://doi.org/10.1016/S2095-3119\(12\)60017-3](https://doi.org/10.1016/S2095-3119(12)60017-3)
- Van Ooijen JW (2006). JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, Netherlands
- Oshima Y, Mitsuda N (2013) The MIXTA-like transcription factor MYB16 is a major regulator of cuticle formation in vegetative organs. *Plant Signal Behav*. <https://doi.org/10.4161/psb.26826>
- Otim-Nape GW, Alicai T, Thresh JM (2001) Changes in the incidence and severity of Cassava mosaic virus disease, varietal

- diversity and cassava production in Uganda. *Annals Appl Biol* 138(3):313–327
- Parsa S, Medina C, Rodríguez V (2015) Sources of pest resistance in cassava. *Crop Prot* 68:79–84. <https://doi.org/10.1016/j.cropro.2014.11.007>
- Pastório MA, Hoshino AT, Kitzberger CSG et al (2023) The Leaf color and trichome density influence the whitefly infestation in different cassava cultivars. *InSects*. <https://doi.org/10.3390/insec14010004>
- Patané JSL, Moreira LM, de Melo Teixeira M, Martins J, Setubal JC, Varani AM (2022) New insights into plant natriuretic peptide evolution: from the lysogenic conversion in *Xanthomonas* to the lateral transfer to the whitefly *Bemisia tabaci*. *Gene*. <https://doi.org/10.1016/j.gene.2022.146326>
- Pei Y, Cao W, Yu W, Peng C, Xu W, Zuo Y, Wu W, Hu Z (2023) Identification and functional characterization of the dirigent gene family in *Phryma leptostachya* and the contribution of PIDIR1 in lignan biosynthesis. *BMC Plant Biol*. <https://doi.org/10.1186/s12870-023-04297-6>
- Perez-Fons L, Bohorquez-Chaux A, Irigoyen ML, Garceau DC, Morel K, Boerjan W, Walling LL, Becerra Lopez-Lavalle LA, Fraser PD (2019) A metabolomics characterisation of natural variation in the resistance of cassava to whitefly. *BMC Plant Biol*. <https://doi.org/10.1186/s12870-019-2107-1>
- Perez-Fons L, Ovalle TM, Drapal M, Ospina MA, Gkanogiannis A, Bohorquez-Chaux A, Lopez-Lavalle LAB, Fraser PD (2023) Integrated genetic and metabolic characterization of Latin American cassava (*Manihot esculenta*) germplasm. *Plant Physiol* 192(4):2672–2686. <https://doi.org/10.1093/plphys/kiad269>
- Platten JD, Cobb JN, Zantua RE (2019) Criteria for evaluating molecular markers: comprehensive quality metrics to improve marker-assisted selection. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0210529>
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38(8):904–909. <https://doi.org/10.1038/ng1847>
- Prince DC, Drurey C, Zipfel C, Hogenhout SA (2014) The leucine-rich repeat receptor-like kinase BRASSINOSTEROID INSENSITIVE1-ASSOCIATED KINASE1 and the Cytochrome P450 PHYTOALEXIN DEFICIENT3 contribute to innate immunity to aphids in arabidopsis. *Plant Physiol* 164(4):2207–2219. <https://doi.org/10.1104/pp.114.235598>
- R Core Team (2023). *R Core Team (2023) R: A Language and Environment for Statistical Computing*. *R Foundation for Statistical Computing*, (4.3.0)
- Rabbi IY, Kayondo SI, Bauchet G, Yusuf M, Aghogho CI, Ogunpaimo K, Uwugiaren R, Smith IA, Peteti P, Agbona A, Parkes E, Lydia E, Wolfe M, Jannink JL, Egesi C, Kulakow P (2022) Genome-wide association analysis reveals new insights into the genetic architecture of defensive, agro-morphological and quality-related traits in cassava. *Plant Mol Biol* 109(3):195–213. <https://doi.org/10.1007/s11103-020-01038-3>
- Rodríguez PA, Escudero-Martínez C, Bos JIB (2017) An aphid effector targets trafficking protein VPS52 in a host-specific manner to promote virulence. *Plant Physiol* 173(3):1892–1903. <https://doi.org/10.1104/pp.16.01458>
- Santamaria ME, Arnaiz A, Gonzalez-Melendi P, Martinez M, Diaz I (2018) Plant perception and short-term responses to phytophagous insects and mites. *Int J Mol Sci* 19(5):1356. <https://doi.org/10.3390/ijms19051356>
- Sauvion N, Charles H, Febvay G, Rahbé Y (2004) Effects of jackbean lectin (ConA) on the feeding behaviour and kinetics of intoxication of the pea aphid. *Acyrtosiphon Pisum Entomol Exp Et Applicata* 110(1):31–44. <https://doi.org/10.1111/j.0013-8703.2004.00117.x>
- Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat Genet* 44(7):825–830. <https://doi.org/10.1038/ng.2314>
- Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52(3):591–611
- Silva-Sanzana C, Celiz-Balboa J, Garzo E, Marcus SE, Parra-Rojas JP, Rojas B, Olmedo P, Rubilar MA, Rios I, Chorbadian RA, Fereres A, Knox P, Saez-Aguayo S, Blanco-Herrera F (2019) Pectin methylesterases modulate plant homogalacturonan status in defenses against the aphid *myzus persicae*. *Plant Cell* 31(8):1913–1929. <https://doi.org/10.1105/tpc.19.00136>
- Simiyu DC, Jang JH, Lee OR (2023) A group III patatin-like phospholipase gene pPLAIIIδ regulates lignin biosynthesis and influences the rate of seed germination in *Arabidopsis thaliana*. *Front Plant Sci* 14:1212979. <https://doi.org/10.3389/fpls.2023.1212979>
- Simon SJ, Tschaplinski TJ, LeBoldus M et al (2020) Host plant genetic control of associated fungal and insect species in a *Populus* hybrid cross. *Ecol Evol* 10(11):5119–5134. <https://doi.org/10.1002/ece3.6266>
- Singh H, Dixit S, Verma PC, Singh PK (2013) Differential peroxidase activities in three different crops upon insect feeding. *Plant Signal Behav*. <https://doi.org/10.4161/psb.25615>
- Sun A, Yu B, Zhang Q, Peng Y, Yang J, Sun Y, Qin P, Jia T, Smeekens S, Teng S (2020) MYC2-activated trichome Birefringence-like37 acetylates cell walls and enhances herbivore resistance. *Plant Physiol* 184(2):1083–1096. <https://doi.org/10.1104/pp.20.00683>
- Sytykiewicz H, Łukasik I, Goławska S, Sprawka I, Goławski A, Sławianowska J, Kmiec K (2020) Expression of thioredoxin/thioredoxin reductase system genes in aphid-challenged maize seedlings. *Int J Mol Sci* 21(17):1–19. <https://doi.org/10.3390/ijms21176296>
- Tang XT, Ibanez F, Tamborindeguy C (2020) Concanavalin a toxicity towards potato psyllid and apoptosis induction in midgut cells. *InSects*. <https://doi.org/10.3390/insects11040243>
- Tang Y, Guo J, Zhang T, Bai S, He K, Wang Z (2021) Genome-wide analysis of WRKY gene family and the dynamic responses of key WRKY genes involved in *Ostrinia furnacalis* attack in *Zea mays*. *Int J Mol Sci*. <https://doi.org/10.3390/ijms222313045>
- Thompson GA, Goggin FL (2006) Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. *J Exp Bot* 57(4):755–766. <https://doi.org/10.1093/jxb/erj135>
- Vargas O, Bellotti AC (1981) Perdidas en rendimiento causadas por moscas blancas en el cultivo de la yuca. *Revista Colombiana De Entomol* 7(1–2):13–20. <https://doi.org/10.25100/socolen.v7i1-2.10319>
- Wang J, Zhang Z (2021) GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genom Proteom Bioinform* 19(4):629–640. <https://doi.org/10.1016/j.gpb.2021.08.005>
- Wang J, Wan X, Li H, Pfeiffer WH, Crouch J, Wan J (2007) Application of identified QTL-marker associations in rice quality improvement through a design-breeding approach. *Theor Appl Genet* 115(1):87–100. <https://doi.org/10.1007/s00122-007-0545-x>
- Wang W, Xu M, Wang Y, Jamil M (2014) Basal transcription factor 3 plays an important role in seed germination and seedling growth of rice. *BioMed Res Int* 2014(1):465739. <https://doi.org/10.1155/2014/465739>
- Wang Y, Sheng L, Zhang H, Du X, An C, Xia X, Chen F, Jiang J, Chen S (2017) CmMYB19 over-expression improves aphid tolerance in chrysanthemum by promoting lignin synthesis. *Int J Mol Sci*. <https://doi.org/10.3390/ijms18030619>
- Wang L, Xue W, Li X, Li J, Wu J, Xie L, Kawabata S, Li Y, Zhang Y (2020) EgMIXTA1, a MYB-type transcription factor, promotes cuticular wax formation in *eustoma grandiflorum* leaves. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2020.524947>

- Xu Y, Li HN, Li GJ, Wang X, Cheng LG, Zhang YM (2011) Mapping quantitative trait loci for seed size traits in soybean (*Glycine max* L. Merr.). *Theor Appl Genet* 122(3):581–594. <https://doi.org/10.1007/s00122-010-1471-x>
- Xu D, Wang Z, Zhuang W, Wang T, Xie Y (2023) Family characteristics, phylogenetic reconstruction, and potential applications of the plant BAHD acyltransferase family. *Front Plant Sci* 14:1218914. <https://doi.org/10.3389/fpls.2023.1218914>
- Yadav M, Pandey J, Chakraborty A, Hassan MI, Kundu JK, Roy A, Singh IK, Singh A (2022) A comprehensive analysis of calmodulin-like proteins of glycine max indicates their role in calcium signaling and plant defense against insect attack. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2022.817950>
- Yang W, Simpson JP, Li-Beisson Y, Beisson F, Pollard M, Ohlrogge JB (2012) A land-plant-specific glycerol-3-phosphate acyltransferase family in arabidopsis: substrate specificity, sn-2 preference, and evolution. *Plant Physiol* 160(2):638–652. <https://doi.org/10.1104/pp.112.201996>
- Zeng Z-B (1994) Precision mapping of quantitative trait loci. *Genetics* 136(4):1457–1468
- Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell* 20(10):2763–2782. <https://doi.org/10.1105/tpc.108.061325>
- Zhou XN, Zheng Y, Cai Z, Wang X, Liu Y, Yu A, Chen X, Liu J, Zhang Y, Wang A (2021) Identification and functional analysis of tomato TPR gene family. *Int J Mol Sci* 22(2):758. <https://doi.org/10.3390/ijms22020758>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.