



OPEN ACCESS

EDITED BY

Changlin Liu,
Chinese Academy of Agricultural Sciences
(CAAS), China

REVIEWED BY

Sang He,
Chinese Academy of Agricultural Sciences,
China
Mostofa Uddin Helal,
Shanxi Agricultural University, China

*CORRESPONDENCE

Leah Nandudu
✉ nanduduleah@gmail.com

RECEIVED 23 December 2023

ACCEPTED 07 March 2024

PUBLISHED 18 March 2024

CITATION

Nandudu L, Strock C, Ogonna A, Kawuki R
and Jannink J-L (2024) Genetic analysis
of cassava brown streak disease root
necrosis using image analysis and
genome-wide association studies.
Front. Plant Sci. 15:1360729.
doi: 10.3389/fpls.2024.1360729

COPYRIGHT

© 2024 Nandudu, Strock, Ogonna, Kawuki
and Jannink. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Genetic analysis of cassava brown streak disease root necrosis using image analysis and genome-wide association studies

Leah Nandudu^{1,2*}, Christopher Strock¹, Alex Ogonna¹,
Robert Kawuki² and Jean-Luc Jannink^{1,3}

¹School of Integrative Plant Sciences, Section of Plant Breeding and Genetics, Cornell University, Ithaca, NY, United States, ²Root Crops Department, National Crops Resources Research Institute (NaCRRI), Kampala, Uganda, ³US Department of Agriculture, Agricultural Research Service (USDA-ARS), Ithaca, NY, United States

Cassava brown streak disease (CBSD) poses a substantial threat to food security. To address this challenge, we used PlantCV to extract CBSD root necrosis image traits from 320 clones, with an aim of identifying genomic regions through genome-wide association studies (GWAS) and candidate genes. Results revealed strong correlations among certain root necrosis image traits, such as necrotic area fraction and necrotic width fraction, as well as between the convex hull area of root necrosis and the percentage of necrosis. Low correlations were observed between CBSD scores obtained from the 1-5 scoring method and all root necrosis traits. Broad-sense heritability estimates of root necrosis image traits ranged from low to moderate, with the highest estimate of 0.42 observed for the percentage of necrosis, while narrow-sense heritability consistently remained low, ranging from 0.03 to 0.22. Leveraging data from 30,750 SNPs obtained through DArT genotyping, eight SNPs on chromosomes 1, 7, and 11 were identified and associated with both the ellipse eccentricity of root necrosis and the percentage of necrosis through GWAS. Candidate gene analysis in the 172.2kb region on the chromosome 1 revealed 24 potential genes with diverse functions, including ubiquitin-protein ligase, DNA-binding transcription factors, and RNA metabolism protein, among others. Despite our initial expectation that image analysis objectivity would yield better heritability estimates and stronger genomic associations than the 1-5 scoring method, the results were unexpectedly lower. Further research is needed to comprehensively understand the genetic basis of these traits and their relevance to cassava breeding and disease management.

KEYWORDS

root necrosis, image analysis, PlantCV, genome-wide association studies (GWAS), cassava brown streak disease (CBSD)

Introduction

Cassava (*Manihot esculenta* Crantz) is a vital tropical food crop, ranking third in caloric contribution after maize and rice. Its large starchy roots and edible leaves serve as a primary food source for more than 800 million people, primarily in sub-Saharan Africa (Nassar & Ortiz, 2010).

Over the past 90 years, cassava production has faced escalating threats from biotic and abiotic stresses (Tomlinson et al., 2018) exacerbated by climate change (Jarvis et al., 2012). Cassava brown streak disease (CBSD), ranked among the seven most serious threats to world food security (Pennisi, 2010), stands as a significant biotic challenge to cassava production, particularly in East, Central and Southern Africa, with the potential to cause losses of up to 100% in susceptible varieties (Hillocks et al., 2000; Hillocks & Jennings, 2003; Kaweesi et al., 2014). CBSD is caused by a positive sense single-stranded RNA virus belonging to the genus *Ipomovirus* and family *Potyviridae* (Winter et al., 2010; Walker et al., 2022). It is caused by two distinct viruses: cassava brown streak virus (CBSV) and Uganda cassava brown streak virus (UCBSV), both collectively referred to as cassava brown streak viruses (CBSVs). Both viruses are transmitted by whiteflies (*Bemisa tabaci*) in a semi-persistent manner (Maruthi et al., 2005; Mero et al., 2021) and through the movement of infected stem cuttings by farmers, which adds to the complexity of controlling the disease. CBSD symptoms are characterized by the initial emergence of leaf chlorosis along secondary vein margins, which later develop into blotches. Subsequently, brown streaks develop on the stem, accompanied by radial root constrictions and root necrosis (Hillocks & Jennings, 2003; Alicai et al., 2007; Kaweesi et al., 2014). The most destructive symptom is root necrosis, as it makes the cassava tubers unfit for consumption by both humans and animals. Therefore, it is essential to thoroughly investigate CBSD root necrosis to uncover the genetic markers or genetic mechanisms that can be leveraged in breeding for resistance. This can be comprehended by adopting root necrosis image capture and analysis to complement the 1-5 visual severity scores. This incorporation of root necrosis imaging could introduce fresh opportunities for enriching the qualitative 1-5 scoring approach with more precise quantitative analyses.

High-throughput phenotyping based on images holds enormous promise in unraveling the genetic basis of root necrosis. Image analysis enables the dissection of the genetic architecture of root necrosis by harnessing valuable information from variation in root necrosis expression patterns, which are often challenging to characterize with the naked eye (Demidchik et al., 2020). The effectiveness of image analysis for crop phenotyping has been showcased in a variety of species, including beans (Kumar et al., 2013; Sun et al., 2016), rice (Wu et al., 2019), potatoes (Cassells et al., 1999; Siebring et al., 2019). Its success extends to the study of diverse traits, ranging from plant architecture to chlorophyll content (Yadav et al., 2010; Dutta Gupta et al., 2013; Walter et al., 2015). Image analysis in plants has also been expanded to multiple pathosystems, enabling the detection and quantification of plant diseases (Sun et al., 2014; Elliott et al., 2022; Palma et al., 2022; Shi et al., 2023). Success of image analysis relies on a diverse range of image analysis tools, including Image J/Fuji (Schindelin

et al., 2012, 2015, 2017), Brushlets (Meyer & Coifman, 1997), PlantCV (Gehan et al., 2017), and many others. Plant Computer Vision (PlantCV) is an open-source, high-throughput image analysis tool that can be used to create customized workflows for segmenting and measuring quantitative characteristics of plants from images (Gehan et al., 2017). PlantCV Version 2.1 is a Python-based package comprised of modular functions, which has remarkable flexibility, usability, and functionality in processing images from multiple platforms including red, green, and blue (RGB), Near-infrared (NIR), PSII fluorescence, thermal, and hyperspectral sensors. PlantCV has found application in the phenotypic characterization of a diversity of plant species including pennycress, *Arabidopsis*, wheat, teff, rice, common bean, quinoa, and more (Castillo et al., 2022; Knapp et al., 2022; Griffiths et al., 2023; Pierz et al., 2023), underscoring its significance as a valuable tool for quantitative classification of plant phenotypes.

Substantial advancements in next-generation sequencing (NGS) and statistical methodologies have opened new avenues for plant breeders to use state-of-the-art tools and methods to adopt more efficient strategies when developing improved cassava varieties (Wolfe et al., 2016). CBSD root necrosis is a trait in cassava for which the genetic architecture can be explored by leveraging DNA markers spread throughout the entire cassava genome. Integration of DNA markers has gained widespread acceptance in genome-wide association studies (GWAS) and has proven invaluable in revealing the extensive allelic diversity present within natural populations (Wang et al., 2020; Xiao et al., 2022). As a result, GWAS has enabled the identification of genomic regions or QTLs (Quantitative Trait Loci) associated with desirable traits that have been leveraged in implementing marker-assisted selection (MAS) and/or genomic selection (Rabbi et al., 2022). Quantitative root necrosis assessment using image analysis introduces novel traits that provide precise measurements and capture spatial characteristics overlooked by the traditional 1-5 scoring method, which relies on human observation to quantify root necrosis. To implement quantitative root necrosis assessments, Diversity Arrays Technology (DArT) markers, which detect thousands of genetic loci through microarray technology were used to conduct GWAS for root necrosis. These markers rely on hybridization of genomic DNA fragments to microarrays with probes for multiple loci, enabling detection of genetic polymorphisms such as SNPs and indels (Deres and Feyissa, 2023). Given that this marks the first attempt to investigate the genetic architecture of CBSD root necrosis using imaging. There is a necessity to identify genomic regions and alleles linked to these traits, potentially resulting in the discovery of novel genetic loci associated with CBSD resistance. This is because for more than 9 decades, CBSD evaluations have primarily relied on a qualitative 1-5 visual scoring scale (Tomlinson et al., 2018) which hinders the precise quantitative assessment of CBSD root necrosis, thus restricting the potential to reveal functional genomic insights. These insights can play a crucial role in utilizing marker-assisted selection or genomic selection to expedite genetic gains in breeding for CBSD root necrosis resistance.

In this study, we used data from PlantCV and CBSD severity scores to achieve four main objectives: (1) characterizing root necrosis image traits through broad-sense and narrow-sense

heritability, genotype by trait (GT), and genotype plus genotype-by-environment interaction (GGE) biplots, (2) determining phenotypic and genetic correlations, (3) identifying genomic regions through univariate and multivariate GWAS analyses, and (4) categorizing candidate genes associated with root necrosis image traits and CBSD severity in the Cycle 2 population of genomic selection.

Materials and methods

Plant material and field conditions

The Cycle (C2) population of genomic selection, which consisted of 471 cassava clones used in this study, was developed at the National Crops Resources Research Institute (NaCRRI), Uganda (Nandudu et al., 2023). This population was developed through a series of selection and hybridization cycles involving clones selected based on genomic estimated breeding values (GEBVs) derived from the initial cycle (C0) and the first cycle (C1) populations. A comprehensive selection process identified 95 clones from the C1 population, which were then crossed, resulting in the creation of a population of 6,570 seedlings. From this pool, 471 cassava clones for the C2 population were selected. This C2 population was then evaluated in two clonal evaluation trials (CETs) that were planted in two locations (Namulonge and Serere) in 2019/2020 and 2020/2021. CETs were set up using an augmented incomplete block design, with three check varieties (UG110017, TME204 and Mkumba) planted in each block. Each plot consisted of ten plants, arranged in a single row with a spacing of 1m *1m. Additionally, spreader rows of a CBSD susceptible clone

(TME 204) were planted around the experimental field and between blocks to enhance disease pressure.

Root necrosis image acquisition

During the harvest, five cassava roots were randomly sampled from each plant within a plot. The selection of only five roots was done to ensure consistency across all plots, considering that the number of cassava roots per plant varied. Subsequently, three images were captured by slicing each root at proximal, middle, and distal points as illustrated by (Ishmael and Emmanuel, 2020) using a sharp knife. In the field, RGB images were captured using Android tablets equipped with the Field Book App (Rife, 2016). This approach was employed to ensure the integrity of both the images and the associated data. 51,194 images (Figure 1) were collected from 471 clones. Note that we expected more images, but losses occurred due to natural calamities and poor sprouting of stem cuttings, particularly at the Serere experimental site. All images have been made accessible on Cassava base (Teclé et al., 2014; Morales et al., 2022) in Joint Photographic Experts Group (JPEG) formats.

Root necrosis image processing

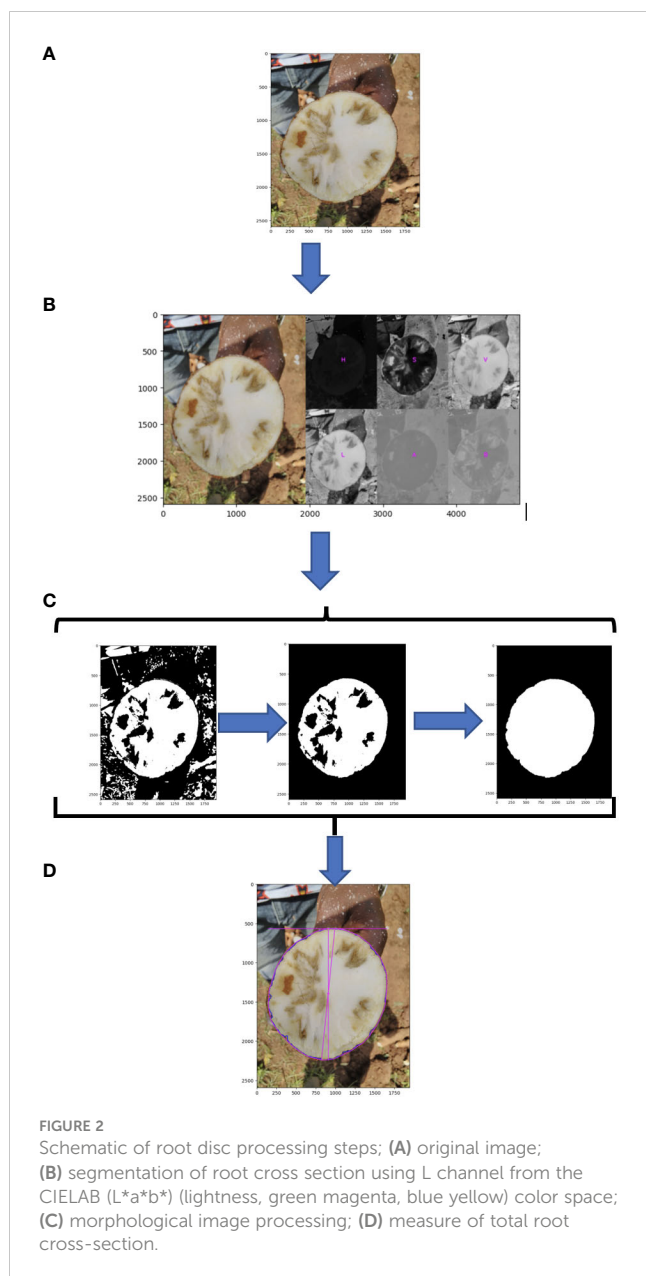
Root necrosis images underwent processing to measure root necrotic lesion traits from the root discs.

Original image

Root necrosis images (Figure 2A and Supplementary Figure 1A) were labeled with an identification code containing the clone's



FIGURE 1
Examples of root necrosis manifestation as caused by cassava brown streak viruses in the C2 population.



name, plot number, plant number, root number, and image position on the root.

Segmentation of root disc cross-section and root necrotic lesions

The original RGB image was transformed to HSV (Hue, saturation, value) color space to segregate the root disc from the background. To isolate necrotic lesions from the root disc, the RGB image was converted using the CIELAB (L*a*b*) (lightness, green magenta, blue, yellow) color space (Figure 2B). The L channel was then isolated, and the background was removed using the selection created in the root disc segmentation step. A default lower threshold of 160 and an upper threshold of 255 were used to analyze the L channel for the identification and selection of necrotic lesions. Subsequently, morphological image processing operations were employed in Python to eliminate background noise and fill holes

in both the root cross-section and necrotic lesions, leaving them both as single and solid objects (Figure 2C, Supplementary Figure 1B, and source code).

Segmented image processing

Using PlantCV, various shape properties of both root cross-section and necrotic area were measured including: (1) total root cross-sectional area (Figure 2D), (2) convex hull area of the cross-section, (3) cross-section perimeter, (4) cross-section width, (5) cross-section height, (6) longest path across the centroid of the cross-section, (7) cross-section eccentricity, (8) cross-section major axis, and (9) cross-section minor axis. For the root cross section, we refer to these shape properties as “root1” to “root9”. Shape properties of the total necrotic regions were also measured including: (1) total necrotic area, (2) convex hull necrotic area, (3) necrotic perimeter, (4) necrotic width, (5) necrotic height, (6) necrotic longest path across the centroid, (7) necrotic ellipse eccentricity, (8) necrotic ellipse major axis, and (9) necrotic ellipse minor axis. For the necrotic area, we refer to these shape properties as “necro1” to “necro9”. All quantified traits were documented in pixels and subsequently normalized to those of the root disc, as outlined in Table 1. Only normalized traits were employed for subsequent analyses.

Phenotyping CBSD using the 1-5 scoring method

Disease severity at 3, 6 and 12 MAP were measured with the 1-5 visual scoring scale (Legg and Thresh, 1998). CBSD foliar severities were determined based on symptom expression on the leaves and stems for each plant in a plot. A score of 1 denoted no symptoms; 2 slight foliar chlorotic leaf mottling with no stem lesions; 3 foliar chlorotic leaf mottling and blotches with mild stem lesions, but no die back; 4 foliar chlorotic leaf mottling and blotches with pronounced stem lesions, but no die back; and 5 defoliation with stem lesions and dieback. Root severity scores were ascertained by visually evaluating the ratio of necrotic lesions to the total area of cross-sectioned root discs obtained from five roots per plant within a plot. In the root severity scoring, a score of 1 indicated absence of necrosis, 2 represented less than 5% necrotic lesions; 3 denoted 6 – 10% necrotic lesions; 4 signified 11 – 25% necrotic lesions and mild root constriction; while 5 denoted >25% necrotic lesions with severe root constriction.

DARtseq genotyping

Of the 471 clones in the CET, 320 were chosen at random for genotyping. For each clone, two young top leaves were collected, folded, and punched using a 5mm hand puncher before being placed in 96-well plates. Subsequently, DNA extraction, Genotyping-by-Sequencing, and SNP calling were performed for each sample using the DARtseq genotyping platform (<https://www.diversityarrays.com/technology-and-resources/dartseq/>).

TABLE 1 Characterization and phenotypic variation of CBSD root necrosis traits in the C2 population.

	Trait Name	Acronyms	Trait description	Mean and standard deviation	Variance	Broad-sense heritability	Narrow-sense heritability
1	Solidity of necrosis	SN	Ratio of area of root necrosis to the convex hull area of root necrosis. SN = necro1/necro2.	0.29 ± 0.28	0.08	0.16	0.03
2	Convex hull area of root necrosis	CHAN	Area of the smallest convex polygon containing necrotic points. CHAN = necro2/root1.	3.03 ± 0.27	0.08	0.14	0.22
3	Ellipse eccentricity of root necrosis	EEN	Ratio of the distance of the focus from the center of the ellipse of the necrotic lesion and the distance of one end of the ellipse from the center. EEN = necro7.	0.79 ± 0.17	0.03	0.27	0.09
4	Percentage of necrosis	NECRO	Percentage area of the total root cross section that is necrotic. NECRO = necro1/root1.	0.69 ± 1.28	1.65	0.42	0.15
5	Necrotic area fraction	NAF	Ratio of the entire region of the root cross section affected by necrosis to Area of the smallest convex polygon containing necrotic points. NAF = necro2/root1.	-2.15 ± 1.80	3.25	0.23	0.17
6	Necrotic width fraction	NWF	Ratio of the widest part of the entire necrotic area to the width of the widest part of the root cross-section. NWF = necro4/root4.	-0.77 ± 0.98	0.95	0.21	0.13

28,434 markers were identified through DArTseq, and these were combined with an additional set of GBS-imputed genotypic data for the same lines (Marnin Wolfe, unpublished data) resulting in a dataset containing 51,860 SNPs. The integration of both marker datasets improved SNP coverage. To enhance the association power and address potential sequencing errors, an additional filtering step was applied to the combined marker dataset. Genotypes with more than 10% missing data and SNPs with over 5% missing data or minor allele frequency below 5% were removed. After filtering, a total of 30,750 SNP markers were obtained and used for downstream analyses.

Statistical analyses

Descriptive statistical measures describing root necrosis image traits

The root necrosis image data displayed significant skewness and non-normal distributions. To address this, most traits underwent natural logarithm transformation, except for solidity of necrosis, and necrotic ellipse eccentricity, which remained unchanged. Also, to visualize the relationships among clones in the C2 population and evaluate their performance across the two environments, the dataset containing root necrosis image traits underwent analysis using both genotype by trait (GT) and genotype plus genotype-versus-environment interaction (GGE) biplot analyses. GT and GGE analyses were executed via the *gtb* and *gge* functions respectively within the R package *metan* (Olivoto and Lúcio, 2020). These functions used the first two principal components (PC) extracted through singular value decomposition (SVD) to partition the standardized genotype by trait and genotype by environment tables, yielding eigenvalues for genotypes, traits,

environments, and singular values. In the biplot, the cosine of the angle between the vectors of two traits approximated the Pearson correlation with an angle less than 90° indicating a positive correlation, greater than 90° indicating a negative correlation and 90° indicating a zero correlation (Ma et al., 2004; Yan and Frégeau-Reid, 2018). Descriptive statistics of mean, standard deviation and variance were also calculated using base functions in R.

Phenotypic and genetic correlations

Phenotypic and genetic correlations were calculated for both CBSD root necrosis image traits and severity scores from the 1-5 visual scoring scale. These were computed using log-transformed phenotypic values and deregressed Best linear unbiased predictions (BLUPs). Correlation analyses were performed using the *cor* function in the R package (R Development Core Team, 2016), and visualization of the correlation matrices was done using the 'corrplot' R package (Wei and Simko, 2017).

Broad-sense and narrow-sense heritability

A standard linear mixed effects model was used to partition variance components using the *lme4* package in R (R Development Core Team, 2016) to determine broad-sense heritability.

A mixed linear model for the observed root necrosis image traits was fitted.

$$y = Xb + Z_c c + Z_b B + Z_p p + Z_n n + Z_r r + \epsilon$$

where y represents the vector of phenotypic data related to root necrosis image traits, while X denotes an $n \times q$ known design matrix encompassing fixed effects for location, year, checks, and image number. The vector b represents the coefficients associated with these fixed effects.

Z_c , Z_b , Z_p , Z_n and Z_r are known design matrices of random effects clone ID, block effects, plot ID, plant ID, and root ID with distributions of $c \sim N(0, I\sigma_c^2)$, $B \sim N(0, I\sigma_B^2)$, $p \sim N(0, I\sigma_p^2)$, $n \sim N(0, I\sigma_n^2)$, and $r \sim N(0, I\sigma_r^2)$. ϵ is the residual with a distribution of $\epsilon \sim N(0, I\sigma_\epsilon^2)$.

Broad-sense heritability was determined using $H^2 = \sigma_c^2 / [\sigma_c^2 + \sigma_p^2/n_p + \sigma_N^2/n_N + \sigma_R^2/n_R + \sigma_\epsilon^2/n_I]$ Where σ_c^2 is the genotypic variance, σ_p^2 plot variance, σ_N^2 variance of plants within a plot, σ_R^2 variance of roots within a plant, and σ_ϵ^2 is the residual variance. The values n_p , n_N , n_R and n_I are the harmonic means the number of plots per clone, number of plants per clone, number of roots per clone, and number of images per clone. Best linear unbiased predictions (BLUPs) were obtained from the mixed linear model, and these were deregressed using the following equation:

$$\text{deregressed BLUP} = \frac{BLUP}{1 - \frac{PEV}{\sigma_c^2}}$$

Where PEV was the prediction error variance of the BLUP and σ_c^2 variances of the genotypes (Garrick et al., 2009). Deregressed BLUPs were used for downstream analyses.

Narrow sense heritability was estimated using the function `emmreml` in the *EMMREML* package (Akdemir & Okeke, 2015) in R.

Genome-wide association studies

Deregressed BLUPs of the 320 clones with genomic data for both root necrosis image traits and CBSD severity scores from the 1-5 scoring method (de los Campos et al., 2013) were used to perform univariate and multivariate GWAS using GEMMA version 0.98.4 with default parameters (Zhou, 2012; Zhou and Stephens, 2014). Univariate GWAS analysis was based on model; $Y = W\alpha + x\beta + U + \epsilon$, $U \sim MVN_n(0, \lambda\tau^{-1}K)$, $\epsilon \sim MVN_n(0, \tau^{-1}I_n)$ where Y is an n -vector of CBSD root necrosis image traits and severity scores from the 1-5 scoring method; W is an $n \times c$ matrix of principal components as covariates including a column of 1s; α is a c -vector of the corresponding coefficients including the intercept; x is an n -vector of marker genotypes; β is the effect size of the marker; U is an n -vector of random effects; ϵ is an n -vector of errors; τ^{-1} is the variance of the residual errors; λ is the ratio between the random effect and error variances; K is a known $n \times n$ relatedness matrix and I_n is an $n \times n$ identity matrix. Using the `prcomp` function in R, principal components (PCs) were determined, and four (PCs) these were used to account for population structure.

Multivariate GWAS analysis was based on the model $Y = WA + x\beta + U + \epsilon$, $U \sim MN_{n \times d}(0, K, V_g)$, $\epsilon \sim MN_{n \times d}(0, I_{n \times n}, V_e)$ where Y is an $n \times d$ matrix of d CBSD root necrosis image traits phenotypes for n individuals; W is an $n \times c$ matrix of principal components as covariates including a column of 1s; A is a c by d matrix of the corresponding coefficients including the intercept; x is an n -vector of marker genotypes; β is a d vector of marker effect sizes for the d phenotypes; U is an n by d matrix of random effects; ϵ is an n by d matrix of errors; K is a known n by n relatedness matrix, $I_{n \times n}$ is a n by n identity matrix, V_g is a d by d symmetric matrix of genetic variance components, V_e is a d by d symmetric matrix of

environmental variance components and $MN_{n \times d}(0, V_1, V_2)$ denotes the $n \times d$ matrix normal distribution with mean 0, row covariance matrix V_1 (n by n), and column covariance matrix V_2 (d by d). Given the extensive number of traits in this study and our aim to enhance the likelihood of identifying shared candidate genes, we performed multivariate GWAS only on traits exhibiting high phenotypic and genotypic correlations. These trait combinations included: (1) convex hull area of root necrosis, percentage of necrosis, necrotic area fraction and necrotic width fraction, (2) necrotic area fraction and necrotic width fraction, (3) ellipse eccentricity of root necrosis, percentage of necrosis, necrotic area fraction and necrotic width fraction, (4) percentage of necrosis, necrotic area fraction and necrotic width fraction, (5) percentage of necrosis, necrotic area fraction and CBSD severity at 12 months after planting (CBSDS12). Despite the absence of strong phenotypic correlations, we conducted multivariate GWAS analyses for the necrosis percentage and CBSD severity at 3 months after planting (CBSDS3) as well as at 6 months after planting (CBSDS6). Additionally, a multivariate GWAS analysis was performed for the necrosis percentage and CBSD severity at 12 months after planting (CBSDS12). Visualization of Manhattan and quantile-quantile plots was carried out using the “*qqman*” R package (Turner and Turner, 2021). To account for multiple testing, the significance threshold was set at a corrected p -value of 0.05 divided by the number of markers on each chromosome, as described by Li and Ji (2005).

Candidate gene analysis

Significant SNP markers linked to both root necrosis image traits and CBSD severity scores were used to determine genomic regions that were characterized for candidate genes. Gene positions were established using *M. esculenta* genome version 6, and any genes that overlapped with these significant genomic regions were classified as candidate genes. BEDTools were employed to detect potential genes regions where the significant SNPs were identified (Quinlan & Hall, 2010). Identified genes were characterized for gene ontology including molecular and biological functioning using PANTHER version 17.0 (Mi et al., 2019) and *M. esculenta* genome version 6 gene ontology database in Phytozome (Goodstein et al., 2012).

Results

Characterization of CBSD root necrosis image traits in the C2 population

To visualize the relationships among genotypes, traits, and environments associated to both root necrosis image traits and the 1-5 visual severity scores, GT and GGE biplots were generated. Specifically, the convex hull area of root necrosis, necrotic area fraction, necrotic width fraction, and the percentage of necrosis exhibited strong positive correlations among themselves. Conversely, all these root necrosis image traits exhibited negative

correlations with both the ellipse eccentricity of root necrosis and the solidity of necrosis. Interestingly, the root necrosis image traits, including convex hull area of root necrosis, necrotic area fraction, necrotic width fraction, and the percentage of necrosis, demonstrated weak but positive correlations with CBSDs3, CBSDs6, and CBSDs12, as shown in Figure 3. In contrast, ellipse eccentricity of root necrosis and the solidity of necrosis exhibited strong negative correlations. The GGE biplots revealed a positive correlation between datasets from the Namulonge experimental site for both years and the Serere experimental site in the 2020/2021 season. The dataset from the Serere experimental site in the 2019/2020 season was not correlated with all the other experimental sites. Moreover, within the C2 population, most clones displayed stability across different environments, with only a few exceptions. Notably, outliers such as UG16F002P004 and UG16F291P132 were observed in the Namulonge experimental site for both years, while clones UG16F07P001, UG16F295P041, UG16FF294P095, UG16F296P025, UG16F082P001, and UG16F319P029 exhibited unique characteristics in the Serere 2019/2020 experimental site.

Broad-sense heritability estimates for root necrosis image traits consistently spanned from low to moderate values (ranging between 0.14 and 0.42). Among these traits, the percentage of necrosis displayed the highest heritability estimate at 0.42, whereas the solidity of necrosis and the convex hull area of root necrosis exhibited the lowest broad-sense heritability estimates, measuring 0.16 and 0.14, respectively (Table 1). Regarding the CBSD severity scores derived from the 1-5 scoring method, the broad-sense heritability estimates stood at 0.73, 0.74, and 0.71 for CBSDs3, CBSDs6, and CBSDs12, respectively. Narrow-sense heritability estimates consistently remained low, ranging from 0.03 to 0.22. Among these estimates, the convex hull area of root necrosis exhibited the highest value at 0.22, while the solidity of necrosis and ellipse eccentricity of root necrosis displayed the lowest narrow-sense heritability estimates, measuring 0.03 and 0.09, respectively (Table 1). The percentage of necrosis was estimated

at 0.15, whereas necrotic area fraction and necrotic width fraction had estimates of 0.17 and 0.13, respectively. In the context of CBSD severity scores obtained from the 1-5 scoring method, the narrow-sense heritability estimates were 0.34, 0.42, and 0.50 for CBSDs3, CBSDs6, and CBSDs12, respectively.

Phenotypic and genetic correlations between root necrosis image traits and severity scores from 1-5 scoring method

The analysis revealed several significant correlations among root necrosis image traits, as shown in Figure 4. Notably, the necrotic area fraction and necrotic width fraction exhibited the strongest correlation at 0.85. Additionally, the convex hull area of root necrosis correlated positively with the percentage of necrosis ($r = 0.53$), necrotic area fraction ($r = 0.50$), and necrotic width fraction ($r = 0.40$). A strong correlation was also found between the percentage of necrosis and both necrotic area fraction ($r = 0.79$) and necrotic width fraction ($r = 0.55$). Low but positive phenotypic correlations were observed between the solidity of necrosis and the ellipse eccentricity of root necrosis ($r = 0.07$). In contrast, negative correlations ranged from -0.25 to -0.81 among solidity of necrosis, convex hull area of root necrosis, percentage of necrosis, necrotic area fraction, and necrotic width fraction. Moreover, ellipse eccentricity of root necrosis showed negative correlations with convex hull area of root necrosis, percentage of necrosis, and necrotic area fraction. The correlation coefficients between root necrosis image traits and severity scores from the 1-5 scoring method obtained at 3, 6, and 12 months after planting (MAP) were relatively low, ranging from -0.08 to 0.31. Notably, positive but modest correlation coefficients (ranging from -0.08 to 0.26) were observed between CBSDs3 and CBSDs6 and all root necrosis image traits. In the case of CBSDs12, correlation coefficients with root necrosis traits ranged from -0.10 to 0.31, with the percentage of

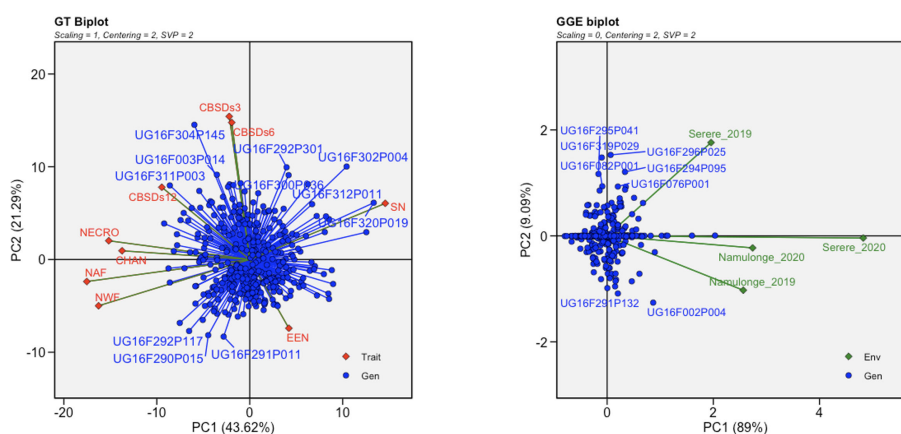


FIGURE 3

Genotype by trait (GT) and genotypic main effect plus genotype by environment interaction (GGE) biplots based on transformed root necrosis image data. GT trait codes are SD, Solidity of necrosis; CHAN, Convex hull area of root necrosis; EEN, Ellipse eccentricity of root necrosis; NECRO, Percentage of necrosis; NAF, Necrotic area fraction; NWF, Necrotic width fraction; CBSDs3, CBSD foliar severity at 3 MAP; CBSDs6, CBSD foliar severity at 6; CBSDs12, CBSD root severity at 12 MAP while GGE biplot shows the location-year scaled.

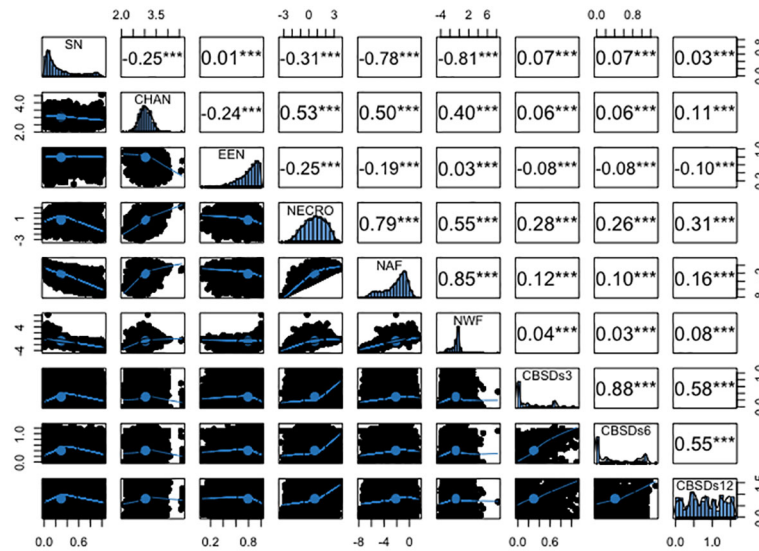


FIGURE 4
Phenotypic correlation coefficients among root necrosis imaging traits and visual scores from the 1-5 scoring method. Frequency distribution histograms are shown on the diagonal. SD, Solidity of necrosis; CHAN, Convex hull area of root necrosis; EEN, Ellipse eccentricity of root necrosis; NECRO, Percentage of necrosis; NAF, Necrotic area fraction; NWF, Necrotic width fraction; CBSDs3, CBSD foliar severity at 3 MAP; CBSDs6, CBSD foliar severity at 6MAP; CBSDs12, CBSD root severity at 12 MAP. Significant thresholds: *, P<0.05; **, P<0.01; ***, P<0.001.

necrosis demonstrating the highest correlation coefficient of 0.31. Interestingly, ellipse eccentricity of root necrosis consistently exhibited negative correlation coefficients with all the CBSD severity scores. Furthermore, correlations were evident within the CBSD severity scores themselves, including a strong correlation of $r = 0.88$ between CBSDs3 and CBSDs6. Additionally, moderately positive correlations ($r = 0.58$ and 0.55) were identified between

CBSDs12 and CBSDs3, as well as between CBSDs12 and CBSDs6, respectively.

The genotypic correlations mirrored the pattern observed in the phenotypic correlations, as illustrated in Figure 5. Notably, a high genetic correlation emerged between necrotic area fraction and necrotic width fraction ($r = 0.85$), along with a similarly strong association between the percentage of necrosis and necrotic area

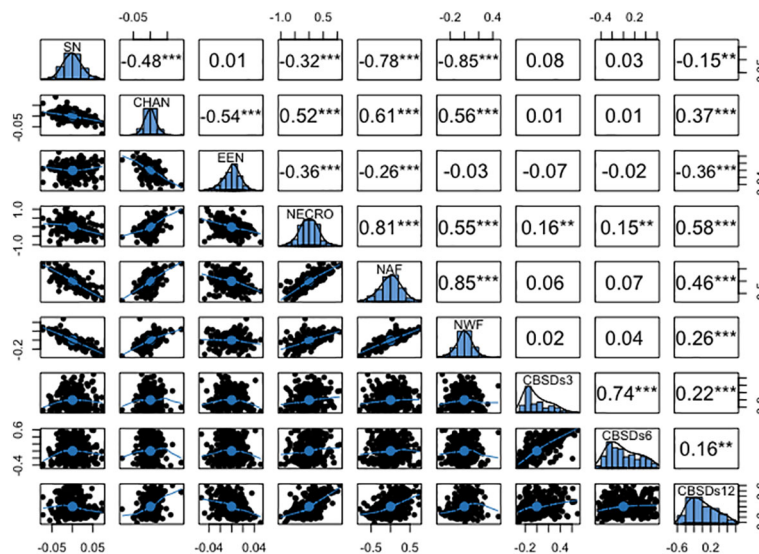


FIGURE 5
Genetic correlation coefficients among root necrosis imaging traits and visual scores from the 1-5 scoring method. Frequency distribution histograms are shown on the diagonal. SD, Solidity of necrosis; CHAN, Convex hull area of root necrosis; EEN, Ellipse eccentricity of root necrosis; NECRO, Percentage of necrosis; NAF, Necrotic area fraction; NWF, Necrotic width fraction; CBSDs3, CBSD foliar severity at 3 MAP; CBSDs6, CBSD foliar severity at 6 MAP; CBSDs12, CBSD root severity at 12 MAP. Significant thresholds: *, P<0.05; **, P<0.01; ***, P<0.001.

fraction ($r = 0.81$). Moderate yet positive genetic correlations were observed between the convex area of root necrosis and the percentage of necrosis ($r = 0.52$), necrotic area fraction ($r = 0.61$), and necrotic width fraction ($r = 0.56$). Conversely, negative genetic correlations ranging from -0.03 to -0.85 were consistently observed between the solidity of necrosis and the eccentricity of root necrosis, as well as with all other root necrosis traits. In terms of genetic correlations between root necrosis traits and CBSD severity scores on the 1-5 scale, they spanned from low to moderate ($r = 0.01 - 0.58$). The percentage of necrosis ($r = 0.58$), necrotic area fraction ($r = 0.46$), and the convex hull area of root necrosis (0.37) exhibited the most substantial correlation coefficients with CBSDs12. However, correlation coefficients between CBSDs3 and CBSDs6 and most root necrosis traits failed to reach significance at the p -value threshold of 0.001. Correlations were also evident among the CBSD severity scores themselves, including a strong correlation of $r = 0.74$ between CBSDs3 and CBSDs6. Additionally, moderately positive correlations ($r = 0.22$ and 0.16) were identified between CBSDs12 and CBSDs3, as well as between CBSDs12 and CBSDs6, respectively.

GWAS of CBSD traits in the C2 population

A comprehensive analysis was conducted on 320 clones within the C2 population, using both univariate and multivariate GWAS methodologies. In the univariate GWAS analysis, a total of eight SNPs were discovered across chromosomes 1, 7, and 11. These SNPs were determined to be linked with both the ellipse eccentricity of root necrosis and the percentage of root necrosis (Table 2, Figure 6, Supplementary Figure 2). Specifically, two SNPs on chromosomes 7 and 11 were linked to the ellipse eccentricity of root necrosis, while six SNPs on chromosome 1 were associated with the percentage of necrosis. Each significant marker accounted for a relatively small proportion of phenotypic variance, ranging from 0.01 to 0.41. Particularly, SNP markers S1_29063012 and S1_29103386 associated to the percentage of root necrosis were found to explain the highest variance among them (Table 2). No significant associations were observed with the other root necrosis image traits (Supplementary Figures 3 and 4). Concerning the CBSD severity scores, one SNP marker located on chromosome 1

was found to be associated with CBSDs12, whereas no SNPs displayed associations with CBSDs3 and CBSDs6 (Supplementary Figures 5 and 6). In the multivariate GWAS analysis, no significant associations were detected for various trait combinations, except for ellipse eccentricity of root necrosis, percentage of root necrosis, necrotic area fraction, and necrotic width fraction, where one SNP on chromosome 11 was identified (Supplementary Figures 7-10). This SNP, named S11_31660720, was also identified as associated with the ellipse eccentricity of root necrosis in the univariate GWAS analysis.

Candidate gene identification

Since most of the identified SNPs were located on chromosome 1, a region spanning 172.2kb was examined, leading to the identification of 24 potential candidate genes (Supplementary Table 1). These genes comprised a broad spectrum of functional annotations, including roles such as ubiquitin-protein ligase, ribosomal protein, transmembrane signal receptor, DNA-binding transcription factor, and RNA metabolism protein, among others. Molecular characterizations of these candidate genes suggest their involvement in catalytic activity (GO:0003824), binding (GO:0005488), and transcription regulator activity (GO:0140110) (Supplementary Table 2). Biological classifications of these potential candidate genes underscore their involvement in a wide array of processes, such as cellular processes (GO:0009987), metabolic processes (GO:0008152), biological regulation (GO:0065007), and localization (GO:0051179) (Supplementary Table 3).

Discussion

Using PlantCV to extract root necrosis image traits and the 1-5 visual scoring method for foliar and root severities, we conducted a comprehensive characterization of cassava brown streak disease. Our aim was to identify genomic regions associated with CBSD root necrosis image and severity traits through univariate and multivariate GWAS analyses. Additionally, we investigated the functions of annotated genes associated with these traits. Key traits of interest included solidity of necrosis, convex hull area of

TABLE 2 Significant SNPs associated to Ellipse eccentricity of root necrosis and Percentage of necrosis and the phenotypic variance explained.

Trait	SNP	Chromosome	Position	P-value	PVE (%)
Ellipse eccentricity of root necrosis	S7_718516	7	718516	5.14	0.07
	S11_31660720	11	31660720	5.22	0.01
Percentage of necrosis	S1_28931143	1	28931143	5.32	0.40
	S1_28972056	1	28972056	5.32	0.40
	S1_28977687	1	28977687	5.32	0.40
	S1_29058977	1	29058977	5.80	0.32
	S1_29063012	1	29063012	5.38	0.41
	S1_29103386	1	29103386	5.80	0.41

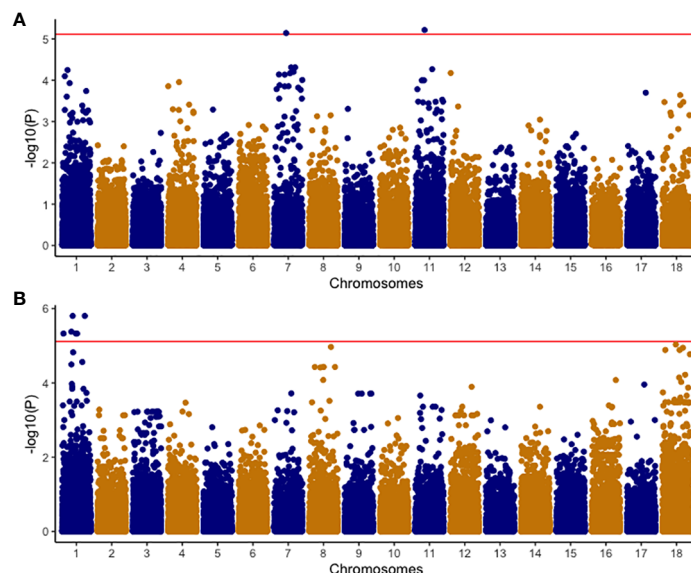


FIGURE 6

Manhattan plots of $-\log_{10}(P)$ showing chromosomal positions of SNP markers in univariate GWAS associated with (A) EEN; Ellipse eccentricity of root necrosis, and (B) NECRO; Percentage of necrosis. The red line represents the significant threshold $-\log_{10}(P)$ value of 5.1 which was determined by using the effective number of independent tests on each chromosome to modify the Bonferroni correction method.

root necrosis, ellipse eccentricity of root necrosis, percentage of necrosis, necrotic area fraction, and necrotic width fraction. Broad-sense and narrow-sense heritability estimates are key in providing insights into the genetic gain achievable through selection (Holland et al., 2002; Schmidt et al., 2019). Results showed that broad sense heritability estimates were low to moderate for all the root necrosis image traits while narrow sense heritability estimates were consistently low and ranged from 0.03 to 0.22. The low to moderate broad-sense heritability estimates provide evidence supporting the involvement of diverse levels of genetic factors in the manifestation of these root necrosis image traits while the narrow sense heritability estimates also provide additional evidence supporting the involvement of diverse levels of additive genetic factors. These heritability estimates indicate some genetic influence, especially for root necrosis traits with moderate estimates. We suggest that random factors may affect the spatial distribution of necrotic lesions, leading to lower heritability estimates. Additionally, we excluded images without root necrosis from our analysis, but these images were given a score of one on a 1-5 scale. This could mean that the set of root necrosis images may not fully represent all the phenotypes, potentially resulting in underestimated heritability for certain image traits compared to the true genetic variation. Broad sense and narrow sense heritability estimates for root necrosis severity scores were similar with those previously reported (Okul Valentor et al., 2018; Ozimati et al., 2019, 2021; Nandudu et al., 2023).

Phenotypic and genetic correlations were analyzed to understand the genetic basis of root necrosis image traits and CBSD severities. We observed varied phenotypic and genetic correlations among root necrosis image traits, including significant positive correlations. Traits such as percentage of necrosis, necrotic width fraction, and necrotic area fraction,

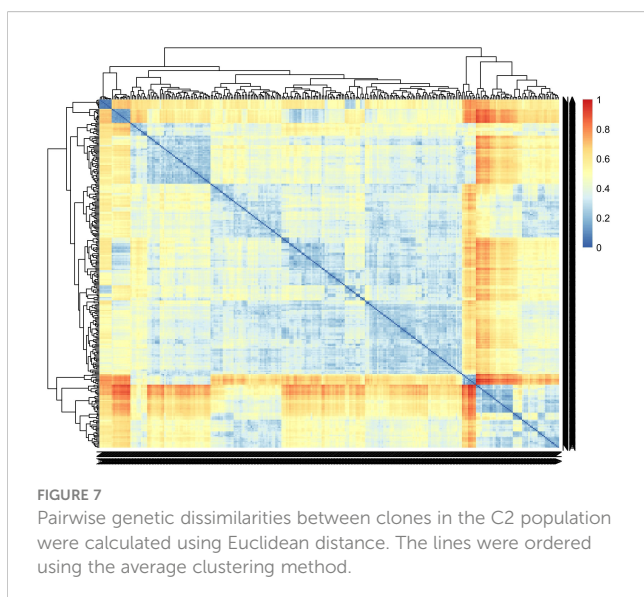
exhibited both high phenotypic and genotypic correlations (Figure 4). This means that these traits can be effective mutual predictors for indirect selection when quantifying root necrosis, resulting in cost efficiency, and reduced computational time. Nonetheless, for breeders to effectively utilize highly correlated root necrosis traits in breeding against root necrosis, it is essential to incorporate them into selection indexes (León et al., 2021). This would involve assigning appropriate weights to these traits based on their economic significance and heritability. Such indexes would enable breeders to prioritize and concentrate their efforts on the root necrosis traits that have the most substantial influence on overall crop enhancement. The high positive correlations between root necrosis image traits also signify the existence of common genetic mechanisms (Vattikuti et al., 2012; Sodini et al., 2018; van Rheenen et al., 2019). Correlations between root necrosis traits can be affected by various factors and one influencing factor is pleiotropism and linkage disequilibrium (Walsh & Blows, 2009; van Rheenen et al., 2019). Pleiotropism occurs when a single gene has an impact on multiple traits, leading to correlations between these traits. Another factor is linkage disequilibrium, a phenomenon described by Walsh and Blows (2009). In linkage disequilibrium, certain genetic variants tend to be inherited together, creating dependencies between traits. Both pleiotropism and linkage disequilibrium contribute to the complex relationships observed among different traits, highlighting the intricate nature of genetic interactions and their effects on trait correlations.

Incorporating genomic data into the analysis pipeline constitutes a pivotal strategy for unraveling the genetic foundations of CBSD root necrosis traits, even in scenarios characterized by low to moderate heritability estimates (Shao et al., 1991). We conducted a comprehensive analysis, encompassing both univariate and multivariate GWAS

approaches. The univariate GWAS found eight SNP markers on chromosomes 1, 7, and 11, linked to both the ellipse eccentricity of root necrosis and the percentage of root necrosis. The haplotype view (Supplementary Figure 11) highlighted the 6 SNPs identified in chromosome 1 to be in a high LD block region (Supplementary Figure 11A). When we zoomed into the high LD block region where the 6 SNPs clustered, three of the SNPs were seen to be in high LD with correlations ranging from 0.9 to 0.97 (Supplementary Figure 11B). On average, four SNPs captured the variation explained by the significant SNPs identified on chromosome 1. Despite being significant, these SNPs only accounted for small proportions of phenotypic variance. These small proportions of variance could arise due to the complexity of CBSD root necrosis image traits' genetic architecture which may arise from the influence of numerous rare variants with minor effects. Furthermore, standard GWAS methods may overlook low-frequency polymorphisms, which could also contribute to CBSD root necrosis image traits. Consequently, the significant SNPs identified in this study might only explain small proportions of phenotypic variance. The limited extent of phenotypic variability also suggests that the underlying causal variants linked to both root necrosis ellipse eccentricity and the percentage of root necrosis, despite their rarity, might reside at substantial distances from the significant SNPs pinpointed in this GWAS (Orozco et al., 2010; Wray et al., 2011). Consequently, the actual effect size could potentially be much more substantial than what is suggested by the significant SNPs identified in this study. To comprehend the impact of multiple rare variants, increasing the population size with diverse individuals is pivotal in enhancing statistical power for their detection. A larger sample size with diverse individuals ensures a more comprehensive representation of rare alleles, facilitating the identification of associations. In our study, the C2 population exhibited moderate genetic diversity (Figure 7). However, our sample size was significantly reduced by excluding root images lacking necrosis, potentially limiting our ability to detect multiple rare variants. Also, to determine causal SNPs linked to CBSD root

necrosis, it is advised to establish a sizable biparental cassava population with diverse root necrosis levels for precise genetic mapping (Nzuki et al., 2017; Ferguson et al., 2019). This population should be evaluated in various environments to enhance statistical power. Fine mapping on regions of interest to identify causal variants associated with synthetic associations should be carried out. Also, integrating traditional QTL mapping with association mapping for a thorough understanding of the genetic architecture of CBSD root necrosis traits is key. We highlight that the outcomes of our study relating to root necrosis traits and CBSD severity scores align with earlier GWAS and QTL analyses focused on CBSD severities utilizing the 1-5 scoring method. These genomic regions include (1) chromosome 1 (Nandudu et al., 2023), (2) chromosome 11 (Kayondo et al., 2018), (3) chromosome 11 (Nzuki et al., 2017), chromosome 11 (Masumba et al., 2017), and chromosome 11 (Kawuki et al., 2016). This implies that certain genes or closely linked genes on these chromosomes might play a role in contributing root necrosis image traits and CBSD severity. Notably, such genes may also originate from prominent founders like Namikonga and Kiroba, which have been extensively employed in CBSD resistance breeding across multiple cassava breeding programs. Ultimately, the persistence of these genomic regions across various studies further emphasizes their significance for future research endeavors aimed at comprehending and enhancing resistance against cassava brown streak disease.

A total of 24 potential candidate genes were identified in the 172.2kb region on chromosome 1, with diverse range of functions crucial in the context of CBSD. These functions include ubiquitin-protein ligase activity, potentially involved in degrading viral components; ribosomal proteins, essential for viral replication and translation; transmembrane signal receptors, aiding in the recognition of viral molecules; DNA-binding transcription factors, governing the expression of defense-related genes; and RNA metabolism proteins, influencing viral RNA stability and processing. Molecular characterizations revealed various gene ontologies associated with these candidate genes. Catalytic activity (GO:0003824) indicates their roles in enzymatic reactions that can either facilitate or hinder CBSD viral replication and spread within cassava plants. The identification of transporter activity (GO:0005215) suggests that these genes may be involved in transporting CBSD viral molecules across cellular membranes, potentially increasing the severity of infection. Transcription regulator activity (GO:0140110) highlights the influence of these genes on gene expression regulation during CBSD viral responses. The biological classification of these potential candidate genes also highlighted their involvement in various critical processes critical for plant-virus interactions. These encompass cellular processes (GO:0009987), which are fundamental for viral entry and replication within the host. Metabolic processes (GO:0008152) underscore their roles in managing energy allocation and resources during infection, while biological regulation (GO:0065007) fine-tunes the plant's responses to viral challenges. Processes related to localization (GO:0051179) are essential for precisely targeting viral components within the cell. Gene ontologies catalytic activity (GO:0003824), transporter activity



(GO:0005215), structural molecule activity (GO:0005198) and molecular, cellular processes (GO:0009987), and metabolic processes (GO:0008152) were previously reported (Amuge et al., 2017) and their presence played a crucial role in facilitating susceptibility to CBSD. Therefore, our study sheds light on the complex nature of CBSD, its genetic architecture, and several candidate genes involved in root necrosis and disease severity. These findings have important implications for understanding CBSD and developing strategies for resistance breeding, which is critical for food security in East and Central Africa.

Conclusions

Using PlantCV for root necrosis traits and a 1-5 scoring method for foliar and root severities in CBSD, we aimed to identify genomic regions associated with CBSD root necrosis through univariate and multivariate GWAS analyses. Key traits included solidity, convex hull area, ellipse eccentricity, percentage of necrosis, necrotic area fraction, and necrotic width fraction. Broad-sense heritability was low to moderate, indicating diverse genetic factors, while narrow-sense heritability was consistently low.

Phenotypic and genetic correlations were analyzed, revealing significant positive correlations among root necrosis traits, especially percentage of necrosis, necrotic width fraction, and necrotic area fraction. Integration of genomic data identified 8 significant SNPs associated with ellipse eccentricity and percentage of root necrosis, suggesting potential epistatic interactions with rare variants. A total of 24 candidate genes were found, involved in crucial processes like viral degradation, replication, transcription regulation, and defense-related gene expression. Gene ontologies highlighted their roles in catalytic activity, transporter activity, structural molecule activity, cellular and metabolic processes. These genes also played roles in cellular entry, energy allocation, regulatory responses, localization, homeostasis, reproduction, developmental and multicellular organismal processes, detoxification, and growth. Some of these ontologies were previously linked to CBSD susceptibility.

In conclusion our study has uncovered potential candidate genes linked to CBSD, shedding light on its genetic makeup and the intricacies of the disease. Despite our initial anticipation of enhanced heritability estimates and strong genomic associations through image analysis objectivity compared to the 1-5 scoring method, the results yielded unexpectedly lower outcomes than those achieved with the 1-5 scoring method. These findings hold considerable importance for comprehending CBSD and devising resistance breeding strategies, ultimately contributing to food security in East and Central Africa.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Author contributions

LN: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. CS: Data curation, Formal Analysis, Methodology, Software, Writing – original draft, Writing – review & editing. AO: Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. KR: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. JJ: Conceptualization, Formal Analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the NEXTGEN Cassava project, through a grant to Cornell University by the Bill & Melinda Gates Foundation (Grant INV-007637 <http://www.gatesfoundation.org>) and the UK's Foreign, Commonwealth & Development Office (FCDO).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2024.1360729/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Schematic of root necrosis processing steps; (A) root necrosis image after processing the root disc characteristics; (B) segmentation of root necrosis using L channel from the CIELAB (L*a*b*) (lightness, green-magenta, blue-yellow) color space; (C) morphological image processing; (D) measure of root necrosis.

SUPPLEMENTARY FIGURE 2

Quantile-quantile plots of: (A) EEN; Ellipse eccentricity of root necrosis, and (B) NECRO; Percentage of necrosis.

SUPPLEMENTARY FIGURE 3

Manhattan plots of $-\log_{10}(P)$ showing chromosomal positions of SNP markers in univariate GWAS associated with (A) SD; Solidity of necrosis, (B) CHAN; Convex hull area of root necrosis, (C) NAF; Necrotic area fraction, and (D) NWF; Necrotic width fraction. The red line represents the significant threshold $-\log_{10}(P)$ value of 5.1 which was determined by using the effective number of independent tests on each chromosome to modify the Bonferroni correction method.

SUPPLEMENTARY FIGURE 4

Quantile-quantile plots of: (A) SD; Solidity of necrosis, (B) CHAN; Convex hull area of root necrosis, (C) NAF; Necrotic area fraction, and (D) NWF; Necrotic width fraction.

SUPPLEMENTARY FIGURE 5

Manhattan plots of $-\log_{10}(P)$ showing chromosomal positions of SNP markers in univariate GWAS associated with (A) CBSDs3; CBSD foliar severity at 3 MAP, (B) CBSDs6; CBSD foliar severity at 6 MAP and CBSDs12; CBSD root severity at 12 MAP. The red line represents the significant threshold $-\log_{10}(P)$ value of 5.112404 which was determined by using the effective number of independent tests on each chromosome to modify the Bonferroni correction method.

SUPPLEMENTARY FIGURE 6

Quantile-quantile plots of (A) CBSDs3; CBSD foliar severity at 3 MAP, (B) CBSDs6; CBSD foliar severity at 6 MAP and CBSDs12; CBSD root severity at 12 MAP.

SUPPLEMENTARY FIGURE 7

Manhattan plots of $-\log_{10}(P)$ showing chromosomal positions of SNP markers in multivariate GWAS associated with (A) convex hull area of root necrosis, percentage of necrosis, necrotic area fraction and necrotic width fraction, (B) necrotic area fraction and necrotic width fraction, (C) ellipse eccentricity of root necrosis, percentage of necrosis, necrotic area fraction

and necrotic width fraction, (D) percentage of necrosis, necrotic area fraction and necrotic width fraction, (E) percentage of necrosis, necrotic area fraction and CBSD severity at 12 months after planting (CBSDs12). The red line represents the significant threshold $-\log_{10}(P)$ value of 5.112404 which was determined by using the effective number of independent tests on each chromosome to modify the Bonferroni correction method.

SUPPLEMENTARY FIGURE 8

Quantile-quantile plots of (A) convex hull area of root necrosis, percentage of necrosis, necrotic area fraction and necrotic width fraction, (B) necrotic area fraction and necrotic width fraction, (C) ellipse eccentricity of root necrosis, percentage of necrosis, necrotic area fraction and necrotic width fraction, (D) percentage of necrosis, necrotic area fraction and necrotic width fraction, (E) percentage of necrosis, necrotic area fraction and CBSD severity at 12 months after planting (CBSDs12).

SUPPLEMENTARY FIGURE 9

Manhattan plots of $-\log_{10}(P)$ showing chromosomal positions of SNP markers in multivariate GWAS associated with (A) percentage of necrosis, CBSD foliar severity at 3 MAP, and CBSD foliar severity at 6 MAP and (B) percentage of necrosis and CBSD root severity at 12 MAP. The red line represents the significant threshold $-\log_{10}(P)$ value of 5.112404 which was determined by using the effective number of independent tests on each chromosome to modify the Bonferroni correction method.

SUPPLEMENTARY FIGURE 10

Quantile-quantile plots of (A) percentage of necrosis, CBSD foliar severity at 3 MAP, and CBSD foliar severity at 6 MAP and (B) percentage of necrosis and CBSD root severity at 12 MAP.

SUPPLEMENTARY FIGURE 11

(A) Haplotype view chromosome 1 where the six SNPs associated to the percentage of necrosis were identified. (Red color intensity indicates the intensity of r^2 , i.e., higher color intensity means higher r^2). (B) Haplotype view of the 172.kb region on the chromosome 1 where the six SNPs associated to the percentage of necrosis were identified. (Red color intensity indicates the intensity of r^2 , i.e., higher color intensity means higher r^2).

References

- Akdemir, D., and Okeke, U. G. (2015). "EMMREML: fitting mixed models with known covariance structures," in *R package version 3.1*. Available at: <https://Cran.R-Project.Org/Package=EMMREML>.
- Alicai, T., Omongo, C. A., Maruthi, M. N., Hillocks, R. J., Baguma, Y., Kawuki, R., et al. (2007). Re-emergence of cassava brown streak disease in Uganda. *Plant Dis.* 91, 24–29. doi: 10.1094/PD-91-0024
- Amuge, T., Berger, D. K., Katari, M. S., Myburg, A. A., Goldman, S. L., and Ferguson, M. E. (2017). A time series transcriptome analysis of cassava (*Manihot esculenta* Crantz) varieties challenged with Ugandan cassava brown streak virus. *Sci. Rep.* 7, 1–21. doi: 10.1038/s41598-017-09617-z
- Cassells, A. C., Kowalski, B., Fitzgerald, D. M., and Murphy, G. A. (1999). The use of image analysis to study developmental variation in micropropagated potato (*Solanum tuberosum* L.) plants. *Potato Res.* 42, 541–548. doi: 10.1007/BF02358170
- Castillo, S. E., Tovar, J. C., Shamin, A., Gutierrez, J., Pearson, P., and Gehan, M. A. (2022). A protocol for Chenopodium quinoa pollen germination. *Plant Methods* 18, 1–11. doi: 10.1186/s13007-022-00900-3
- de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., and Calus, M. P. L. (2013). Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* 193, 327–345. doi: 10.1534/genetics.112.143313
- Demidchik, V. V., Shashko, A. Y., Bandarenka, U. Y., Smolikova, G. N., Przhevalskaya, D. A., Charnysh, M. A., et al. (2020). Plant phenomics: fundamental bases, software and hardware platforms, and machine learning. *Russian J. Plant Physiol.* 67, 397–412. doi: 10.1134/S1021443720030061
- Deres, D., and Feyissa, T. (2023). Concepts and applications of diversity array technology (DARt) markers for crop improvement. *J. Crop Improv.* 37, 913–933. doi: 10.1080/15427528.2022.2159908
- Dutta Gupta, S., Ibaraki, Y., and Pattanayak, A. K. (2013). Development of a digital image analysis method for real-time estimation of chlorophyll content in micropropagated potato plants. *Plant Biotechnol. Rep.* 7, 91–97. doi: 10.1007/s11816-012-0240-5
- Elliott, K., Berry, J. C., Kim, H., and Bart, R. S. (2022). A comparison of ImageJ and machine learning based image analysis methods to measure cassava bacterial blight disease severity. *Plant Methods* 18, 1–12. doi: 10.1186/s13007-022-00906-x
- Ferguson, M. E., Shah, T., Kulakow, P., and Ceballos, H. (2019). A global overview of cassava genetic diversity. *PLoS One* 14, 1–16. doi: 10.1371/journal.pone.0224763
- Garrick, D. J., Taylor, J. F., and Fernando, R. L. (2009). Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Select. Evol.* 41, 1–8. doi: 10.1186/1297-9686-41-55
- Gehan, M. A., Fahlgren, N., Abbasi, A., Berry, J. C., Callen, S. T., Chavez, L., et al. (2017). PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ* 2017, 1–23. doi: 10.7717/peerj.4088
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* 40, 1178–1186. doi: 10.1093/nar/gkr944
- Griffiths, M., Liu, A. E., Gunn, S. L., Mutan, N. M., Morales, E. Y., and Topp, C. N. (2023). A temporal analysis and response to nitrate availability of 3D root system architecture in diverse pennycress (*Thlaspi arvense* L.) accessions. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1145389
- Hillocks, R. J., and Jennings, D. L. (2003). Cassava brown streak disease: A review of present knowledge and research needs. *Int. J. Pest Manage.* 49, 225–234. doi: 10.1080/0967087031000101061
- Hillocks, R. J., Thresh, J. M., Mosaic, C., and Streak, C. B. (2000). Cassava mosaic and cassava brown streak virus diseases in Africa: a comparative guide to symptoms and aetiologies. *Roots* 7, 1–8.
- Holland, J. B., Nyquist, W. E., and Cervantes-Martinez, C. T. (2002). "Estimating and interpreting heritability for plant breeding: an update," in *Plant Breeding Reviews* (Wiley), 9–112. doi: 10.1002/9780470650202.ch2
- Ishmael, N. A., and Emmanuel, Y. H. B. (2020). Evaluation of a motorized cassava peeler with four lining materials. *Afr. J. Agric. Res.* 16, 1342–1354. doi: 10.5897/AJAR

- Jarvis, A., Ramirez-Villegas, J., Campo, B. V. H., and Navarro-Racines, C. (2012). Is cassava the answer to African climate change adaptation? *Trop. Plant Biol.* 5, 9–29. doi: 10.1007/s12042-012-9096-7
- Kaweesi, T., Kawuki, R., Kyaligonza, V., Baguma, Y., Tusiime, G., and Ferguson, M. E. (2014). Field evaluation of selected cassava genotypes for cassava brown streak disease based on symptom expression and virus load. *Virol. J.* 11. doi: 10.1186/s12985-014-0216-x
- Kawuki, R. S., Kaweesi, T., Esuma, W., Pariyo, A., Kayondo, I. S., Ozimati, A., et al. (2016). Eleven years of breeding efforts to combat cassava brown streak disease. *Breed. Sci.* 66, 560–571. doi: 10.1270/jsbbs.16005
- Kayondo, S. I., Del Carpio, D. P., Lozano, R., Ozimati, A., Wolfe, M., Baguma, Y., et al. (2018). Genome-wide association mapping and genomic prediction for CBSD resistance in *Manihot esculenta*. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-19696-1
- Knapp, A., Stefani, J., Katz, E., and Bloom, A. J. (2022). Novel method for the quantification of rosette area from images of *Arabidopsis* seedlings grown on agar plates. *Appl. Plant Sci.* 10, 1–7. doi: 10.1002/aps3.11504
- Kumar, M., Bora, G., and Lin, D. (2013). Image processing technique to estimate geometric parameters and volume of selected dry beans. *J. Food Meas. Charact.* 7, 81–89. doi: 10.1007/s11694-013-9142-7
- Legg, J. P., and Thresh, J. M. (1998). Cassava virus diseases in Africa. *May 2014*, 517–552.
- León, R., Rosero, A., García, J. L., Morelo, J., Orozco, A., Silva, G., et al. (2021). Multi-trait selection indices for identifying new cassava varieties adapted to the caribbean region of Colombia. *Agronomy* 11. doi: 10.3390/agronomy11091694
- Li, J., and Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95, 221–227. doi: 10.1038/sj.hdy.6800717
- Ma, B. L., Yan, W., Dwyer, L. M., Frégeau-Reid, J., Voldeng, H. D., Dion, Y., et al. (2004). Interactions on spring wheat yield. *Agron. J.* 96, 169–180. doi: 10.2134/agronj2004.1690
- Maruthi, M. N., Hillocks, R. J., Mtunda, K., Raya, M. D., Muhanna, M., Kiozia, H., et al. (2005). Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). *J. Phytopathol.* 153, 307–312. doi: 10.1111/j.1439-0434.2005.00974.x
- Masumba, E. A., Kapinga, F., Mkamillo, G., Salum, K., Kulembeka, H., Rounsley, S., et al. (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, 2069–2090. doi: 10.1007/s00122-017-2943-z
- Mero, H. R., Lyantagaye, S. L., and Bongcam-Rudloff, E. (2021). Why has permanent control of cassava brown streak disease in Sub-Saharan Africa remained a dream since the 1930s? *Infect. Genet. Evol.* 94, 105001. doi: 10.1016/j.meegid.2021.105001
- Meyer, F. G., and Coifman, R. R. (1997). Brushlets: A tool for directional image analysis and image compression. *Appl. Comput. Harmon. Anal.* 4, 147–187. doi: 10.1006/acha.1997.0208
- Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X., et al. (2019). Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* 14, 703–721. doi: 10.1038/s41596-019-0128-8
- Morales, N., Ogonna, A. C., Ellerbrock, B. J., Bauchet, G. J., Tantikanjana, T., Tecle, I. Y., et al. (2022). Breedbase: A digital ecosystem for modern plant breeding. *G3: Genes Genomes Genet.* 12. doi: 10.1093/g3journal/jkac078
- Nandudu, L., Kawuki, R., Ogonna, A., Kanaabi, M., and Jannink, J. L. (2023). Genetic dissection of cassava brown streak disease in a genomic selection population. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1099409
- Nassar, N., and Ortiz, R. (2010). Breeding cassava to feed the poor. *Sci. Am.* 302, 78–84. doi: 10.1038/scientificamerican0510-78
- Nzuki, I., Katari, M. S., Bredeson, J. V., Masumba, E., Kapinga, F., Salum, K., et al. (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.* 8. doi: 10.3389/fpls.2017.01168
- Okul Valentor, A., Ochwo-Ssemakula, M., Kaweesi, T., Ozimati, A., Mrema, E., Mwale, E. S., et al. (2018). Plot based heritability estimates and categorization of cassava genotype response to cassava brown streak disease. *Crop Prot.* 108, 39–46. doi: 10.1016/j.cropro.2018.02.008
- Olivoto, T., and Lúcio, A. D. C. (2020). metan: An R package for multi-environment trial analysis. *Methods Ecol. Evol.* 11, 783–789. doi: 10.1111/2041-210X.13384
- Orozco, G., Barrett, J. C., and Zeggini, E. (2010). Synthetic associations in the context of genome-wide association scan signals. *Hum. Mol. Genet.* 19. doi: 10.1093/hmg/ddq368
- Ozimati, A. A., Esuma, W., Alicai, T., Jannink, J. L., Egesi, C., and Kawuki, R. (2021). Outlook of cassava brown streak disease assessment: perspectives of the screening methods of breeders and pathologists. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.648436
- Ozimati, A., Kawuki, R., Esuma, W., Kayondo, S. I., Pariyo, A., Wolfe, M., et al. (2019). Genetic variation and trait correlations in an East African Cassava breeding population for genomic selection. *Crop Sci.* 59, 460–473. doi: 10.2135/cropsci2018.01.0060
- Palma, D., Blanchini, F., and Montessoro, P. L. (2022). A system-theoretic approach for image-based infectious plant disease severity estimation. *PLoS One* 17, 1–24. doi: 10.1371/journal.pone.0272002
- Pennisi, E. (2010). Armed and dangerous. *Science* 327, 804–805. doi: 10.1126/science.327.5967.804
- Pierz, L. D., Heslinga, D. R., Buell, C. R., and Haus, M. J. (2023). An image-based technique for automated root disease severity assessment using PlantCV. *Appl. Plant Sci.* 11, 1–9. doi: 10.1002/aps3.11507
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- R Development Core Team (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available at: <http://www.R-project.org>.
- Rabbi, I. Y., Kayondo, S. I., Bauchet, G., Yusuf, M., Aghogho, C. I., Ogunpaimo, K., et al. (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, 195–213. doi: 10.1007/s11103-020-01038-3
- Rife, T. (2016). *Field Book 3.0 Manual*.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019
- Schindelin, J., Rueden, C. T., Hiner, M. C., and Eliceiri, K. W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* 82, 518–529. doi: 10.1002/mrd.22489
- Schindelin, J., Rueden, C. T., Hiner, M. C., and Eliceiri, K. W. (2017). The ImageJ ecosystem: An open platform for biomedical image analysis. *analysis* 82, 518–529. doi: 10.1002/mrd.22489.The
- Schmidt, P., Hartung, J., Bennewitz, J., and Hans-Peter, P. (2019). Heritability in plant breeding on a genotype-difference basis. *Genetics* 212, 991–1008. doi: 10.1534/genetics.119.302134
- Shao, H., Burrage, L. C., Sinasac, D. S., Hill, A. E., Ernest, S. R., O'Brien, W., et al. (1991). "Genetic architecture of complex traits: Large phenotypic effects and pervasive epistasis," in *Sciences of the USA*, vol. Vol. 105. .
- Shi, T., Liu, Y., Zheng, X., Hu, K., Huang, H., Liu, H., et al. (2023). Recent advances in plant disease severity assessment using convolutional neural networks. *Sci. Rep.* 13, 1–13. doi: 10.1038/s41598-023-29230-7
- Siebring, J., Valente, J., Franceschini, M. H. D., Kamp, J., and Kooistra, L. (2019). Object-based image analysis applied to low altitude aerial imagery for potato plant trait retrieval and pathogen detection. *Sens. (Switzerland)* 19. doi: 10.3390/s19245477
- Sodini, S. M., Kemper, K. E., Wray, N. R., and Trzaskowski, M. (2018). Comparison of genotypic and phenotypic correlations: Cheverud's conjecture in humans. *Genetics* 209, 941–948. doi: 10.1534/genetics.117.300630
- Sun, J., Jiang, S., Mao, H., Wu, X., and Li, Q. (2016). Classification of black beans using visible and near infrared hyperspectral imaging. *Int. J. Food Prop.* 19, 1687–1695. doi: 10.1080/10942912.2015.1055760
- Sun, H., Wei, J., Zhang, J., and Yang, W. (2014). A comparison of disease severity measurements using image analysis and visual estimates using a category scale for genetic analysis of resistance to bacterial spot in tomato. *Eur. J. Plant Pathol.* 139, 125–136. doi: 10.1007/s10658-013-0371-8
- Tecle, I. Y., Edwards, J. D., Menda, N., Egesi, C., Rabbi, I. Y., Kulakow, P., et al. (2014). solGS: A web-based tool for genomic selection. *BMC Bioinf.* 15, 1–9. doi: 10.1186/s12859-014-0398-7
- Tomlinson, K. R., Bailey, A. M., Alicai, T., Seal, S., and Foster, G. D. (2018). Cassava brown streak disease: historical timeline, current knowledge and future prospects. *Mol. Plant Pathol.* 19, 1282–1294. doi: 10.1111/mpp.12613
- Turner, A. S., and Turner, M. S. (2021). *Package 'qqman'*. 1–5.
- van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H., and Wray, N. R. (2019). Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* 20, 567–581. doi: 10.1038/s41576-019-0137-z
- Vattikuti, S., Guo, J., and Chow, C. C. (2012). Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. *PLoS Genet.* 8. doi: 10.1371/journal.pgen.1002040
- Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Alfenas-Zerbini, P., et al. (2022). Recent changes to virus taxonomy ratified by the International Committee on Taxonomy of Viruses, (2022). *Arch. Virol.* 167, 2429–2440. doi: 10.1007/s00705-022-05516-5
- Walsh, B., and Blows, M. W. (2009). Abundant genetic variation + strong selection = multivariate genetic constraints: A geometric view of adaptation. *Annu. Rev. Ecol. Evol. Syst.* 40, 41–59. doi: 10.1146/annurev.ecolsys.110308.120232
- Walter, A., Liebisch, F., and Hund, A. (2015). Plant phenotyping: From bean weighing to image analysis. *Plant Methods* 11, 1–11. doi: 10.1186/s13007-015-0056-8
- Wang, Q., Tang, J., Han, B., and Huang, X. (2020). Advances in genome-wide association studies of complex traits in rice. *Theor. Appl. Genet.* 133, 1415–1425. doi: 10.1007/s00122-019-03473-3
- Wei, T., and Simko, V. (2017) *Corrplot. R Package, v. 0.84, 18*. Available online at: <https://cran.r-project.org/web/packages/corrplot/corrplot.pdf>.
- Winter, S., Koerber, M., Stein, B., Pietruszka, A., Paape, M., and Butgereit, A. (2010). Analysis of cassava brown streak viruses reveals the presence of distinct virus

species causing cassava brown streak disease in East Africa. *J. Gen. Virol.* 91, 1365–1372. doi: 10.1099/vir.0.014688-0

Wolfe, M. D., Rabbi, I. Y., Egesi, C., Hamblin, M., Kawuki, R., Kulakow, P., et al. (2016). Genome-wide association and prediction reveals genetic architecture of cassava mosaic disease resistance and prospects for rapid genetic improvement. *The Plant Genome* 92 (2), plantgenome2015-11.

Wray, N. R., Purcell, S. M., and Visscher, P. M. (2011). Synthetic associations created by rare variants do not explain most GWAS results. *PLoS Biol.* 9. doi: 10.1371/journal.pbio.1000579

Wu, W., Liu, T., Zhou, P., Yang, T., Li, C., Zhong, X., et al. (2019). Image analysis-based recognition and quantification of grain number per panicle in rice. *Plant Methods* 15, 1–14. doi: 10.1186/s13007-019-0510-0

Xiao, Q., Bai, X., Zhang, C., and He, Y. (2022). Advanced high-throughput plant phenotyping techniques for genome-wide association studies: A review. *J. Adv. Res.* 35, 215–230. doi: 10.1016/j.jare.2021.05.002

Yadav, S. P., Ibaraki, Y., and Gupta, S. D. (2010). Estimation of the chlorophyll content of micropropagated potato plants using RGB based image analysis. *Plant Cell Tissue Organ Cult.* 100, 183–188. doi: 10.1007/s11240-009-9635-6

Yan, W., and Frégeau-Reid, J. (2018). Genotype by Yield*Trait (GYT) biplot: A novel approach for genotype selection based on multiple traits. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-26688-8

Zhou, X. (2012). *GEMMA User Manual V0.91*. 1–12.

Zhou, X., and Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* 11, 407–409. doi: 10.1038/nmeth.2848