

Unraveling candidate genomic regions responsible for delayed post-harvest deterioration in Cassava (*Manihot esculenta* Crantz)

Diana C Solarte Certuche ¹, Gabriel M de Freitas ¹, Jean-Luc Jannink ², Cinara Fernanda Garcia Morales ³, Tamires Sousa Cerqueira ³, Bruna Santos de Santana ⁴, Eder Jorge de Oliveira³, Antônio Augusto F Garcia ¹

São Paulo University, Genetics Department, Piracicaba, São Paulo, Brazil. ²USDA-ARS, Robert W. Holley Center, Ithaca, NY, United States of America. ³Embrapa Mandioca e Fruticultura, EMBRAPA, Genetic Resources and Cultivar Development Center, Cruz das Almas, Bahia, Brazil. ⁴Federal University of Recôncavo da Bahia, CCAAB, Cruz das Almas, Bahia, Brazil.

Abstract

Post-harvest physiological deterioration (PPD) represents a significant challenge of cassava production and commercialization. This multifaceted biological process involves a series of mechanisms, including enzymatic stress responses, alterations in gene expression, protein synthesis, accumulation of secondary metabolites, and ultimately, programmed cell death. These changes render the storage roots unpalatable and unmarketable. Therefore, unraveling the genetic architecture of PPD and exploring the interactions of associated genes during its early and late stages is essential for the crop production. We used modern genetic resources to unravel the genetic basis of PPD, based on a genome-wide association study (GWAS), utilizing a combination of different models, including BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway), SUPER (Settlement of MLM Under Progressively Exclusive Relationship), and MLMM (Multi-locus mixed models). The phenotyping dataset spanned five years and included evaluations from 42 different trials, evaluating the Embrapa (Brazilian Agricultural Research Corporation) germplasm along with a population derivative from a genomic selection cycle. We utilized a genotype dataset comprising 26,000 high-quality SNPs (single nucleotide polymorphisms). Our findings indicated four significant genetic variants located on chromosomes 2, 5, and 13, which together explain 35.83 % of the phenotypic variation. These variants are associated with genes that are closely linked to the pathways activated during the early and late symptoms of PPD. The identification of these three key genes provides valuable insights into the genetic architecture of PPD and lays a strong foundation for molecular breeding, supporting the efforts to identify cassava genotypes with enhanced PPD tolerance, the identified genomic regions may be incorporated into genomic selection models, thereby enhancing marker-assisted selection (MAS) and improving breeding strategies for long shelf life and high-quality agronomic cassava cultivars for the cassava community.

Keywords: Cassava, GWAS, Post-harvest physiological deterioration

Introduction

Cassava (*Manihot esculenta* Crantz) has become a globally significant crop, with global production increasing by approximately 16–17% over the past decade. By 2024, global production exceeded 315 million tonnes, with Nigeria remaining the largest producer (~62.69 million tonnes) and Brazil ranking fifth (~18.51 million tons) (FAO, 2024). Cassava's agronomic resilience, including broad edaphic adaptability and notable drought tolerance, makes it a critical crop for both smallholder and commercial farming systems (Borku et al., 2025). Historically, cassava has underpinned food security in many developing countries, particularly in Africa, where it is the third most important staple food. At the same time, the demand for cassava-derived products has expanded across an increasingly diverse set of industries (food, beverages, confectionery, textiles, paper-making, pharmaceuticals, cosmetics, chemicals, and bioethanol) (Mohidin et al., 2023). This increase in demand has sparked significant scientific interest, as evidenced by the 21,136 related publications focused on cassava in the Scopus database (Otekunrin, 2024).

Despite its agronomic advantages, cassava breeding and commercialization face several persistent constraints embedded in the base germplasm: low yield potential in some clones, susceptibility to major pests and diseases, low micronutrient content in many cultivars, and elevated cyanogenic potential in others. Among these limitations, post-harvest physiological deterioration (PPD) of storage roots poses a particularly acute challenge for transport, storage and marketability (Devi et al., 2022). PPD is a complex, wound-triggered physiological syndrome that initiates in root parenchyma tissue after harvest. Mechanical damage during harvest activates a rapid enzymatic stress response that propagates signaling from wounded zones to distal tissues and triggers defense-related gene expression (Saravanan et al., 2016). The organoleptic and nutritional consequences of PPD, principally vascular parenchyma discoloration and loss of starch and other quality attributes, significantly reduce commercial value and consumer acceptance (Devi et al., 2022). Global production losses due to PPD have been estimated at ~19 %, with losses in regions having poor limited transport and storage infrastructure (e.g., many areas in Africa) reaching ~29 % (Saravanan et al., 2016). Extending the shelf life of cassava roots during storage is a common goal. This not only helps address food security challenges but also enhances marketing opportunities and facilitates the large-scale conversion of cassava into a commercial crop.

At the cellular and biochemical levels, PPD progresses through well-documented stages. Early responses include extensive reprogramming of gene expression, altered protein synthesis and accumulation of secondary metabolites (Iyer et al., 2010). An oxidative burst, marked by production of superoxide radicals and other reactive oxygen species (ROS), is a hallmark early event; ROS together with hydrogen peroxide modulate transcriptional networks and secondary metabolism (Lebot et al., 2023; Reilly et al., 2007). These redox dynamics are accompanied by increased respiration and conversion of starch to soluble sugars, producing a measurable decline in starch content during storage (Zainuddin et al., 2018). Concomitantly, volatile metabolites (cyanohydrins, aldehydes, ketones, alcohols) and ethylene are produced (Shi Liu et al., 2017), and Ca²⁺ signaling appears to contribute to the induction of oxidative bursts (Salcedo, 2011). Oxidation of hydroxycoumarins such as scopoletin by ROS and peroxidases yields colored complexes that explain the characteristic blue–dark discoloration observed in affected roots (Shi Liu et al., 2017). Recent evidence further indicates that phenolics and carotenoids can modulate

damage to metabolic pathways and thereby mitigate symptom severity in early PPD stages (Che et al., 2024). Later stages progress to programmed cell death, secondary microbial colonization, and tissue softening, culminating in severe quality losses (Zainuddin et al., 2018).

Overall, PPD is a complex phenomenon that results from interactions between different pathways, influencing whether a genotype shows tolerance or susceptibility to PPD. Given this substantial economic and food-security impact, extensive biochemical, molecular and genetic studies have sought to elucidate PPD mechanisms and identify tolerant germplasm (Buschmann et al., 2000; Iyer et al., 2010; Djabou et al., 2017; Lebot et al., 2023; Che et al., 2024; Cortes et al., 2002; Reilly et al., 2007; Ma et al., 2023; Ayalde et al., 2024). Breeding work has also explored phenotypic sources of tolerance and estimated genetic parameters for PPD (Kawano & Rojanaridpiched, 1983; Aina et al., 2007 ; Akinwale et al., 2010 ; Tumuhimbise et al., 2015; Venturini et al., 2016). Although significant knowledge has been gained regarding PPD in cassava, achieving long shelf-life genotypes remains a major challenge for breeding programs and stakeholders across the cassava value chain. PPD is notably difficult to breed for due to the strong influence of environmental conditions on trait expression. Developing longer shelf-life genotypes involves navigating a complex genetic structure, where numerous small genetic effects contribute to the polygenic nature of PPD symptom expression. Additionally, there is a negative correlation between PPD and key traits for cassava, such as dry matter content and carotenoids. Over the past few decades, considerable efforts have been made to clarify this intricate genetic structure, including the identification of several quantitative trait loci (QTL) in mapping populations in African germplasm. However, these QTL analyses have not accounted for a significant portion of the phenotypic variation observed. Despite extensive research, to our knowledge there have been no successful applications of marker-assisted breeding strategies to improve postharvest traits related to PPD reported to date.

Markers identified in association with PPD could be utilized as covariates in genomic selection (GS) and marker-assisted selection (MAS) as routine practices in the breeding pipeline. This would provide a viable solution to obtain long shelf life post-harvest accessions while also addressing other important agronomic traits in different cassava varieties, such as dry matter content, cyanogenic compounds, carotenoids, and resistance to major diseases (Mbinda & Mukami, 2022).

In this study, we build upon previous efforts to understand the genetic architecture of PPD by integrating multi-environment and multi-year phenotyping with an appropriate association mapping strategy. This approach enables us to better control for environmental interactions affecting PPD expression. As a result, we provide a comprehensive catalog of trait-locus combinations and candidate genes that can explain the observed variation. Building on prior statistical analyses of Brazilian germplasm that revealed substantial variability and environmental effects on PPD expression (Venturini et al., 2016), the present study investigates the genetic architecture underlying PPD in a broad Brazilian germplasm collection. We extend Venturini et al.'s work by including a new population derived from recent genomic selection cycles in the Embrapa cassava breeding program. Using genome-wide association studies (GWAS) implemented with multiple statistical models to account for population structure and kinship, we analyzed 812 cassava clones genotyped by Genotyping-by-Sequencing (GBS) and scored with 27,047 SNPs. These markers obtained could be utilized as covariates in (GS) and (MAS) as steps

routine practices in the breeding pipeline. This would provide a viable solution to obtain long shelf-life post-harvest accessions.

Material and Methods

Plant material and field trials

The study evaluated a panel of 812 cassava clones derived from the Embrapa cassava breeding program (Supplementary Table 1). Phenotypic data were collected across 42 field trials conducted in multiple years and locations throughout the state of Bahia, Brazil (Supplementary Table S2). The trials represent historical and ongoing breeding evaluations that include germplasm collections and populations derived from four genomic selection cycles (C0, C1, C2, and C3). Experimental designs varied by year and location as follows: in 2018 germplasm and the C0 population were evaluated at a single location using an augmented block design; in 2020 evaluations were expanded to 10 locations using a combination of augmented block and randomized complete block designs (RCBD); in 2021 the panel was assessed across 13 locations using an augmented block design, RCBD and a completely randomized design (CRD); 2022 trials included two designs across 10 locations; and 2023 trials covered eight locations. Plot-level management, planting density and standard crop husbandry followed local Embrapa protocols for cassava (Filho, 2013).

Post-harvest physiological deterioration (PPD) phenotyping

PPD was assessed using two complementary protocols to quantify root deterioration: (i) a conventional visual scoring of the proportion of damaged root tissue on a 0–100 scale (Venturini et al., 2016), and (ii) an image-based measure obtained from digital photographs processed in *ImageJ software v.2.14.0* (Schneider et al., 2012), employing automated area-detection routines and machine-learning plugins to estimate the percentage of affected tissue (scaled 0–100). For both methods, freshly harvested roots were cleaned and sanitized, then each root was cut into six longitudinal slices. Slices were stored for 10 days on open shelves beneath a metal-roofed shelter that allowed free air circulation on all sides. After the storage period, slices were photographed following a standardized imaging protocol (fixed camera-to-sample distance, consistent lighting, inclusion of a scale bar and color reference), and the resulting images were evaluated using both the visual scoring and ImageJ pipelines.

DNA extraction and SNP genotyping

DNA extraction and SNP genotyping were conducted by collaborators from Embrapa as part of a collaborative effort, following the protocol described below. Genomic DNA was isolated from young, freeze-dried leaf tissue using the cetyltrimethylammonium bromide (CTAB) extraction method by Doyle & Doyle, 1987. Genome-wide genotyping was performed using GBS following standard protocols; raw sequence reads were aligned to the cassava reference genome (v6) and variant calling was performed with the *Genome Analysis Toolkit (GATK v. 4.4)* pipeline (McKenna et al., 2010). Initial quality control removed markers with minor allele frequency (MAF) < 0.01 and sites with >60 % missing genotype calls, producing an initial high-quality set of 27,047

biallelic SNPs for downstream filtering and analyses. All genotype matrices were quality-checked and formatted for population-structure, kinship and association analyses.

Phenotypic data analysis

Phenotypic data used in this study were provided by collaborators from Embrapa and were collected according to the Embrapa's protocol. Phenotypic traits were analyzed using a linear mixed model using the package *ASReml-R v.4.1* (Butler et al., 2023) implementing a two-stage mixed-model framework adapted for multi-environment trial data (Endelman, 2023). In the first stage, environment-specific linear mixed models were fitted to plot-level PPD observations to adjust for design effects and to obtain genotype adjusted means Best Linear Unbiased Estimator (BLUES) within each environment. The general form of the per-environment statistical model was:

$$y_j = X_j\beta_j + Z_{g(j)}\mu_{g(j)} + Z_{c(j)}\mu_{c(j)} + \varepsilon_j \quad (1)$$

where, y_j represents the data vector of PPD evaluation plot in the location-year combination that represents the environment. X_j is an incidence matrix relating observations in environment j , β_j is a vector that represents the fixed effects specific for environment j^{th} , $Z_{g(j)}$ is an incidence matrix for clones, $\mu_{g(j)}$ is a vector that contains the random effects of the clone in environment j^{th} , $Z_{c(j)}$ incidence matrix for additional random design- related effects in environment j^{th} , such as row and columns, and blocks, $\mu_{c(j)}$ vector of random effects between environments in the combination location-year, and, ε_j vector of plot error associated with the environment j^{th} . Assumed that $\varepsilon_j \sim \mathcal{N}(\mathbf{0}, \mathbf{g}_j)$, $\mu_{g(j)} \sim \mathcal{N}(\mathbf{0}, \mathbf{g}_j)$, $\mu_{c(j)} \sim \mathcal{N}(\mathbf{0}, \mathbf{c}_j)$.

In the second stage, environment-level genotype estimates were combined to obtain single genotype genetic values. De-regressed BLUPs (Best Linear Unbiased Prediction) (or weighted BLUPs) were computed to account for heterogeneous error variances among environments. The second stage working model can be expressed as:

$$y = X\delta + Z\mu + \varepsilon \quad (2)$$

where y is the vector of de-regressed genotype values, X is the design matrix relating genotype level observations to fixed effects across environments, δ contains fixed effects (environmental, markers and inbreeding), Z is the incidence matrix relating observations to genotypes, μ represents the vector of random genetic effects structured by a marker-based relationship matrix, and ε is the vector of residuals.

The resulting BLUPs values obtained in this analysis were used to the GWAS analysis model in the next steps.

Broad sense heritability for the trait was estimated on an entry mean basis using the equation

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{ge}^2}{e} + \frac{\sigma_e^2}{er}}$$

Where σ_g^2 is the genotypic variance, σ_{ge}^2 is the genotype by environment interactions variance, σ_e^2 error variance, and respectively r and e denote the number of replications and the number of environments included in the analysis, respectively.

Population structure, kinship and linkage disequilibrium

Population structure was evaluated using multiple complementary approaches. Bayesian clustering was performed with *STRUCTURE* v2.3.4 (Pritchard et al., 2000) with multiple independent runs across $K=1$ to $K=6$; runs included appropriate burn-in and Markov Chain Monte Carlo (MCMC) sampling and were replicated to ensure convergence. The most likely number of clusters (ΔK) was determined with *STRUCTURE HARVESTER* v.0.6.9.94 (Earl & vonHoldt, 2012) using the method of Evanno et al. (2005). Principal component analysis (PCA) was performed in *TASSEL software* v. 5.0 (Bradbury et al., 2007). Pairwise linkage disequilibrium (LD) between markers was computed using *PLINK* v. 2.0 (Purcell et al., 2007) and LD decay and local r^2 patterns were evaluated to define candidate regions, considering SNP pairs with $r^2 \geq 0.8$ as being in strong LD. Kinship among individuals was estimated using a marker-based genomic relationship matrix was calculated following the method of Theo H. E. VanRaden (2008), based on centered genotype data and allele frequencies, where higher values indicate greater genetic similarity. Visualization of structure and clustering results was generated with construct (Bradburd et al., 2016) and standard R plotting tools.

Genome-wide association study (GWAS)

The association analyses used BLUPs derived from 812 clones evaluated across multiple location–year combinations as the response variable. Genotype data represented in 27,044 high-quality, biallelic SNPs. GWAS was performed in *GAPIT* v3 (Wang & Zhang, 2021) using several complementary multi-locus and mixed-model methods to improve power and reduce false positives: *i*) SUPER (Settlement of MLM Under Progressively Exclusive Relationship) (Wang et al., 2014), *ii*) BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway), which employs linkage disequilibrium to streamline marker selection (Huang et al., 2019), and *iii*) MLM (Multi-locus Mixed Model), which adjusts for cofactors associated with markers using forward and backward stepwise regression in mixed mode (Segura et al., 2012). All models incorporated marker-based kinship and population structure covariates as appropriate. Multiple testing correction was performed with a Bonferroni procedure ($\alpha = 0.05$). Results were visualized with Manhattan plots (nominal visualization threshold set at $-\log_{10}(\text{p-value}) = 5.5$) and quantile–quantile (Q–Q) plots to assess test statistic inflation and model fit. LD was estimated by pairwise correlation coefficients (r^2) to define candidate regions, and loci with $-\log_{10}(p) > 3.6$ ($p < 2.51 \times 10^{-4}$) were highlighted for further inspection. Multiple testing was addressed using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995), and SNPs with q -values below 0.05 were considered significant.

248

249 Genetic interactions analyses

250

251 To understand the genetic action of the candidate genes mapped in the GWAS we use a model
252 based in Cockerham model (Kao & Zeng, 2002 & Cockerham, 1954) for detecting epistatic two-
253 way bi-allelic interactions; denoting locus 1 with A and a alleles and locus 2 with B and b alleles.
254 By doing this, we were able to test if the two markers interact each other. The statistical model
255 was:

256

$$257 y_j = \mu + \alpha_1 S_{1j} + \delta_1 S_{1j} + \alpha_2 S_{2j} + \delta_2 S_{2j} + I_{12j} + Gu' + \varepsilon_j$$

258

259 Where y_i is the phenotypic value of clone i , μ is the overall mean, α_1 and α_2 are the additive effect
260 of allele substitution at locus 1 and locus 2, respectively, δ_1 and δ_2 denote the corresponding
261 dominance deviations, S_{1j} and S_{2j} represent the additive genotype coding for individual j at loci
262 1 and 2, respectively (coded as 0, 1, and 2 according to allele dosage). The terms αS_{1j} and αS_{2j}
263 capture the linear effects of allele dosage at each locus, whereas δS_{1j} and δS_{2j} account for
264 dominance deviations from the midpoint of the homozygotes., I_{12j} represents the epistatic
265 interaction effects between loci 1 and loci 2 for individual j (additive x additive, additive x
266 dominance, dominance x additive, dominance x dominance interactions), The kinship structure
267 among individuals was modeled using the matrix Gu' and ε_i is the residual error. A bonferroni
268 threshold ($\alpha = 0.05$) was used to correct the multiple comparisons. The proportion of phenotypic
269 variance explained by the additive loci identified through GWAS, as well as by both additive and
270 epistatic loci, was calculated as the r^2 from a simple linear regression of these loci on mean trait
271 performance.

272 Candidate gene identification, functional annotation and haplotype block analysis

273

274 Candidate genes within a conservative window (± 1 kb from each focal SNP based on LD results)
275 were extracted and annotated using Gene Ontology (GO), InterPro and Pfam descriptors (El-
276 Gebali et al., 2019). Homology and additional functional annotation were retrieved from National
277 Centre for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/>) and European
278 Molecular Biology Laboratory European Bioinformatics Institute (EMBL-EBI)
279 (<https://www.ebi.ac.uk/>) resources to support putative biological roles. Where appropriate, gene
280 annotations were cross-referenced against published studies implicating pathways relevant to PPD.
281 Candidate loci were prioritized based on statistical strength of association, biological plausibility
282 and local LD patterns. Haplotype block analysis was conducted using *HaploView* v.4.2 (Barrett et
283 al., 2005), and gene structure visualizations were generated with *SnapGene* v.7.0. Haplotype
284 blocks were defined based on linkage disequilibrium (LD) decay patterns inferred from the
285 genome-wide association study (GWAS).

Results

Phenotypic variation and heritability of PPD

We evaluated phenotypic variation for PPD across multiple location–year combinations. Visual assessments of PPD ranged from 0.0 % to 94.73 % (mean = 31.78 %), while ImageJ-derived measurements ranged from 0.0 % to 99.92 % (mean = 30.62 %) with a coefficient of variation of 64.5 %. The Q–Q plots compare the observed residual distributions with the expected normal distributions for each trial or environment included in the PPD analyses. In most environments, the residuals are closely aligned with the reference line, suggesting that the residuals approximate a normal distribution and that the fitted models effectively capture the data's variation (Supplementary Figure. S1). Broad-sense heritability estimated across environments varied between 0.22 and 0.58, with an overall mean of 0.48 (Supplementary Table S2), indicating a substantial genetic component to PPD expression alongside a large environmental contribution. Variance decomposition attributed 48 % of the phenotypic variance to genotypes $\times$ environments interaction ($V_{g \times e}$) and 12 % to genotype main effects (V_G), underscoring the importance of multi-environment evaluation in selection. The two scoring methods were highly concordant (correlation $r = 0.97$, Supplementary Figure. S2), supporting the adoption of the faster, visual assessments pipeline as a scalable alternative to ImageJ for routine breeding evaluations; until fully automated ImageJ workflows are implemented, visual scoring can continue as a practical interim solution.

SNP distribution and linkage disequilibrium

After SNP calling and quality control, 27,047 high-quality biallelic SNPs were retained and mapped across the 18 cassava chromosomes (reported mapped span = 35 Mb). SNP marker distribution was heterogeneous across chromosomes: chromosome 1 carried the largest number of markers ($n = 572$) while chromosome 14 contained the fewest ($n = 259$) (Figure. 1a). Minor allele frequencies ranged from 0.05 to 0.50 with a mean MAF of 0.24 (Figure. 1b), and approximately 75 % of markers had allele frequencies exceeding 5 % in the evaluated panel, ensuring reasonable informativeness for association mapping.

We characterized linkage disequilibrium (LD) by computing pairwise r^2 across each chromosome to assess mapping resolution. Mean genome-wide r^2 was 0.19, with a peak LD near $r^2 = 0.40$ that decayed to $r^2 \approx 0.18$ over short physical distances, on the order of ~ 1 kb in the present panel (Figure. 1c). This rapid LD decay indicates fine local resolution for association mapping and guided the window sizes used to define candidate regions around significant markers.

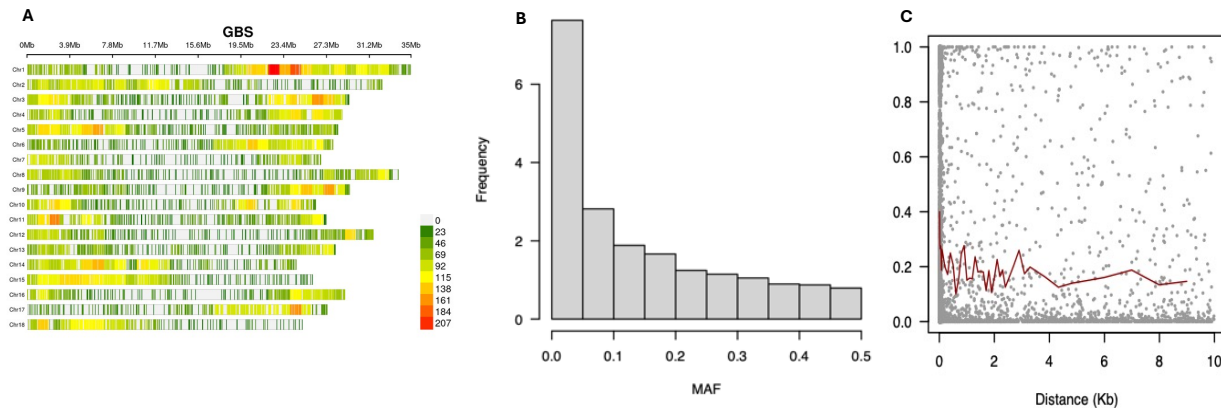


Figure 1. Overview of genotypic data: **(A)** SNP density distribution across the 18 chromosomes of cassava in the evaluated population. **(B)** Minor allele frequencies of the SNPs identified in the studied clones. **(C)** Linkage disequilibrium (LD) decay: This illustrates the squared correlation (r^2) between markers and their physical distance in kilobases (kb).

Population structure and relatedness

Population structure analyses were performed on the post-QC dataset used for structure inference (≈ 812 sample entries and $\sim 26k$ SNPs after the stricter QC used for these analyses). Principal component analysis (PCA) indicated modest stratification: the first two principal components (PC1 + PC2) together explained 14.23 % of the total genetic variance (Figure. 2a), and inclusion of PC3 increased explained variance to ~ 17.69 % (Supplementary Figure. S3). Bayesian clustering and hierarchical analyses revealed two clear subgroups within the panel (Figure. 2b) that correspond broadly to historical germplasm and populations derived from recent genomic-selection cycles. Identity-by-state and kinship matrices showed generally low pairwise relatedness among most clones, consistent with a relatively diverse association panel and favorable conditions for detecting marker–trait associations with limited confounding by close kinship (Figure. 2b).

Genomic regions associated with post-harvest deterioration

We conducted GWAS using three complementary statistical models: MLMM, BLINK, and SUPER. The quantile–quantile (Q–Q) plots indicated minimal genomic inflation or deflation, with observed distributions closely matching expectations except for the significant associations that clearly deviated, confirming their statistical robustness (Figure. 3c, 3d). Across the three models, we identified 10 significant SNPs distributed on chromosomes 2, 5, and 13 (Supplementary Table S1; Figure. 3a, 3b). Notably, the markers located on chromosome 2 were consistently detected by all three models (MLMM, SUPER, BLINK) and were validated using both phenotyping approaches (visual scoring and ImageJ analysis). Two SNPs in this region, S2_11347402 (allele A/T) and S2_11347477 (allele T/G), together explained 8.34 % of the phenotypic variance (PVE) (Supplementary Table S1; Figure. 3a, 3b). This consistency across models and phenotyping methods highlights chromosome 2 as a key genomic region influencing PPD expression.

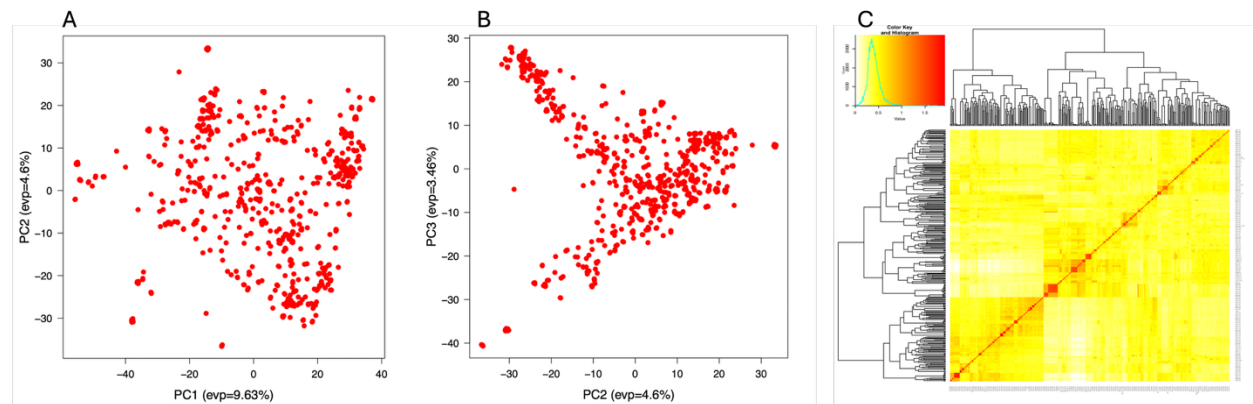


Figure 2. Population structure assessed by principal component analysis (PCA). **(A)** Scatter plot of the first and second principal components (PCs). **(B)** Scatter plot of the second and third principal components **(C)** Heatmap of the pairwise kinship matrix. Yellow and red represent weak and strong kinship relationships between genotypes, respectively. The clustering dendrogram is displayed outside the matrix.

The good agreement between the two phenotyping methods (visual and ImageJ) explains the consistency of the associations across analyses. On chromosome 5, a significant SNP (S5_8348709, allele A/G) was detected using the BLINK model with ImageJ data, explaining 23.87% of the phenotypic variance (PVE). On chromosome 13, the SNP S13_1760296 (alleles C/A) was identified by the BLINK model with ImageJ data, explaining 3.62 % of the PVE (Supplementary Table S1; Figure. 3a). In total the four significant SNPs detected on chromosomes 2, 5, and 13 accounted for 35.83 % of the total phenotypic variance in PPD observed across the population. Comprehensive details of the peak SNPs associated with all loci and their candidate genes are provided in the supplementary material (Supplementary Table S1).

Prediction of candidate genes

Functional annotation of significant loci was performed using the cassava reference genome v6.1 available on the Phytozome platform (Goodstein et al., 2012). This analysis enabled the prediction of causal genes underlying loci significantly associated with PPD expression. In total, three candidate genes were identified across the significant SNPs detected by the different GWAS models. On chromosome 2, SNPs S2_11347402 and S2_11347477 were located within the coding region of the gene *Manes.02G152300*, which encodes an aquaporin transporter (PTHR19139//PTHR19139:SF145) (Figure. 4a). Aquaporins are transmembrane proteins that mediate the transport of water and small molecules, including glycerol, CO₂, hydrogen peroxide, and lactic acid, across biological membranes (Kaplan et al., 2018). Given their established roles in water and ROS transport, aquaporins are highly relevant to wound responses and early-stage PPD progression.

On chromosome 5, SNP S5_8348709 was located within the coding region of *Manes.05G094000*, which encodes a serine/threonine protein phosphatase (PTHR11668//PTHR1668:SF257) (Figure. 4b). These enzymes are central regulators of plant defensive responses to biotic and abiotic stress,

often through the induction of oxidative stress and the generation of reactive oxygen species (ROS) that damage cellular structures. In chromosome 13, SNP S13_1760296 was positioned near *Manes.13G017800*, which encodes a leucine-rich repeat (LRR) protein kinase (PTHR27008//PTHR27008:SF4) (Figure. 4c). LRR kinases are essential in interpreting extracellular signals and initiating downstream transduction cascades.

Further inspection of allelic variants highlights the functional consequences of these SNPs. In *Manes.02G152300*, S2_11347477 results in a nonsynonymous substitution from leucine (Leu) to tryptophan (Trp), while S2_11347402 is located in the coding region. In *Manes.05G094000*, the SNP introduces an amino acid substitution from tyrosine (Tyr) to histidine (His), which could significantly affect protein folding and interactions, potentially explaining phenotypic differences between tolerant and susceptible clones. Finally, in *Manes.13G017800*, variation in the intergenic region likely influences transcriptional regulation of kinase activity, modulating the balance between ROS production and hypersensitive responses.

Haplotype-based GWAS analysis

To characterize local haplotype structure around trait-associated loci, we examined four flanking SNPs upstream and downstream of each peak GWAS marker on chromosomes 2, 5 and 13, and defined haplotype blocks from local LD patterns. Mean haplotype-block sizes were ~107 kb on chromosome 2, ~47 kb on chromosome 5 and ~202 kb on chromosome 13 (Figure. 5), reflecting chromosome-specific differences in local linkage disequilibrium and recombination. Block boundaries and local r^2 profiles informed the genomic intervals used for candidate-gene inspection. Haplotype composition was evaluated in a targeted subset of 19 clones selected to span the phenotypic range (low, medium and high PPD). Among the low-PPD group, 9 accessions carried one or two copies of the chromosome-2 alleles (S2_11347402 / S2_11347477) but carried zero copies of the variant alleles at the chromosome-5 and chromosome-13 loci. In contrast, clones with medium and high PPD typically carried fewer or no copies of the protective chromosome-2 alleles while frequently carrying two copies of the variant alleles on chromosomes 5 and 13. Boxplot summaries (Figure. 5) show a clear stepwise increase in PPD severity associated with presence of the derived alleles on chromosomes 5 and 13, supporting a contributory role of these loci to symptom severity. Collectively, these haplotype patterns are consistent with a multi-locus genetic architecture in which alleles on chromosomes 5 and 13 increase susceptibility, whereas alleles on chromosome 2 are associated with reduced PPD. These results complement the single-marker GWAS by (i) refining the physical intervals that harbor candidate genes, (ii) revealing haplotypic backgrounds that differentiate tolerant and susceptible clones, and (iii) suggesting that favorable allele combinations at chromosome-2 loci may partially mitigate the adverse effects of susceptibility alleles on chromosomes 5 and 13.

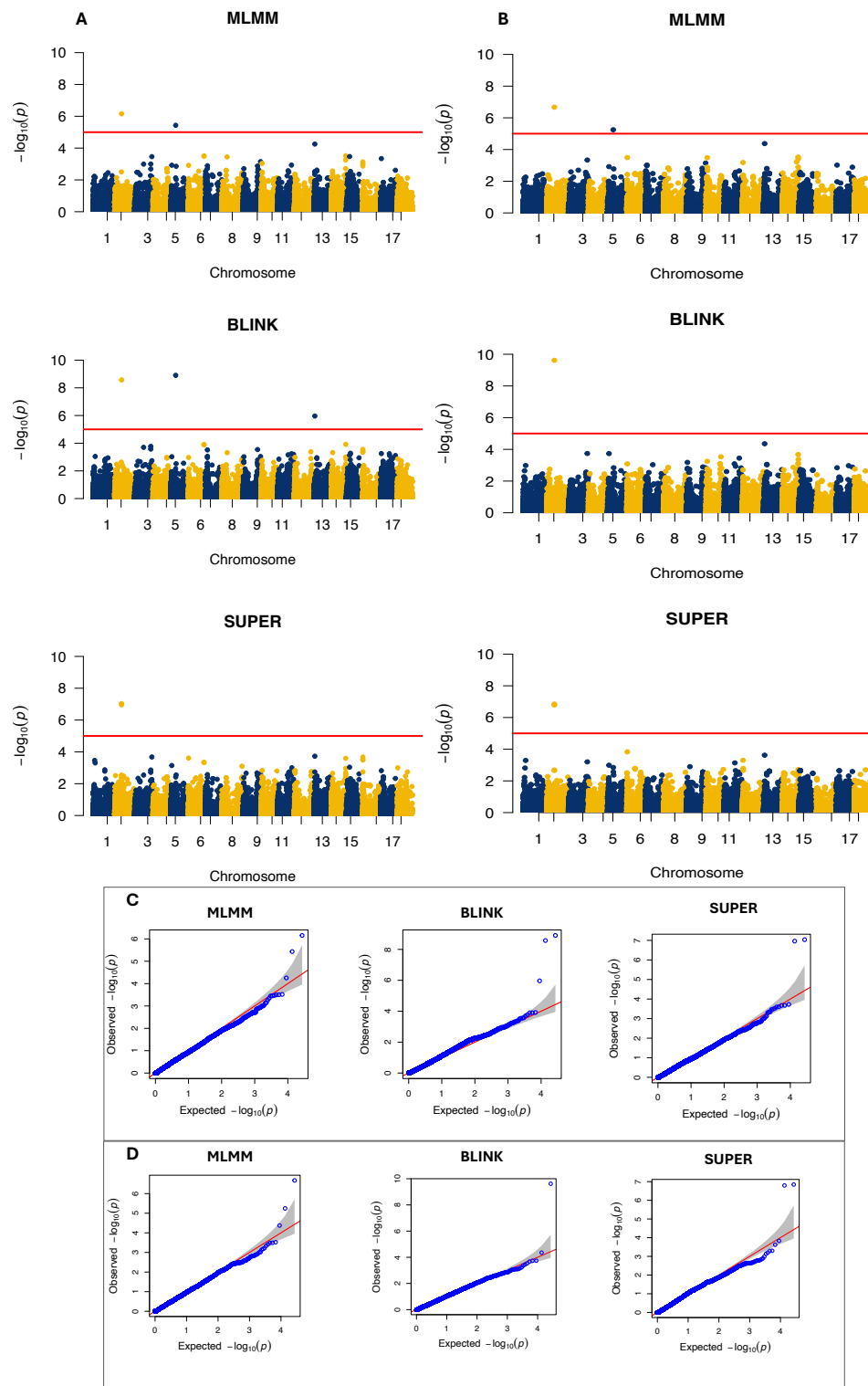


Figure 3. Results of genome-wide association studies (GWAS) using different statistical models. (A) Manhattan plots for PPD scored with ImageJ data employing MLMM, BLINK, and SUPER models. (B) Manhattan plots for PPD scored with visual data utilizing MLMM, BLINK, and

SUPER models. (C) Quantile–quantile (Q–Q) plots for the models using ImageJ data.
(D) Q–Q plots for the models using visual data.

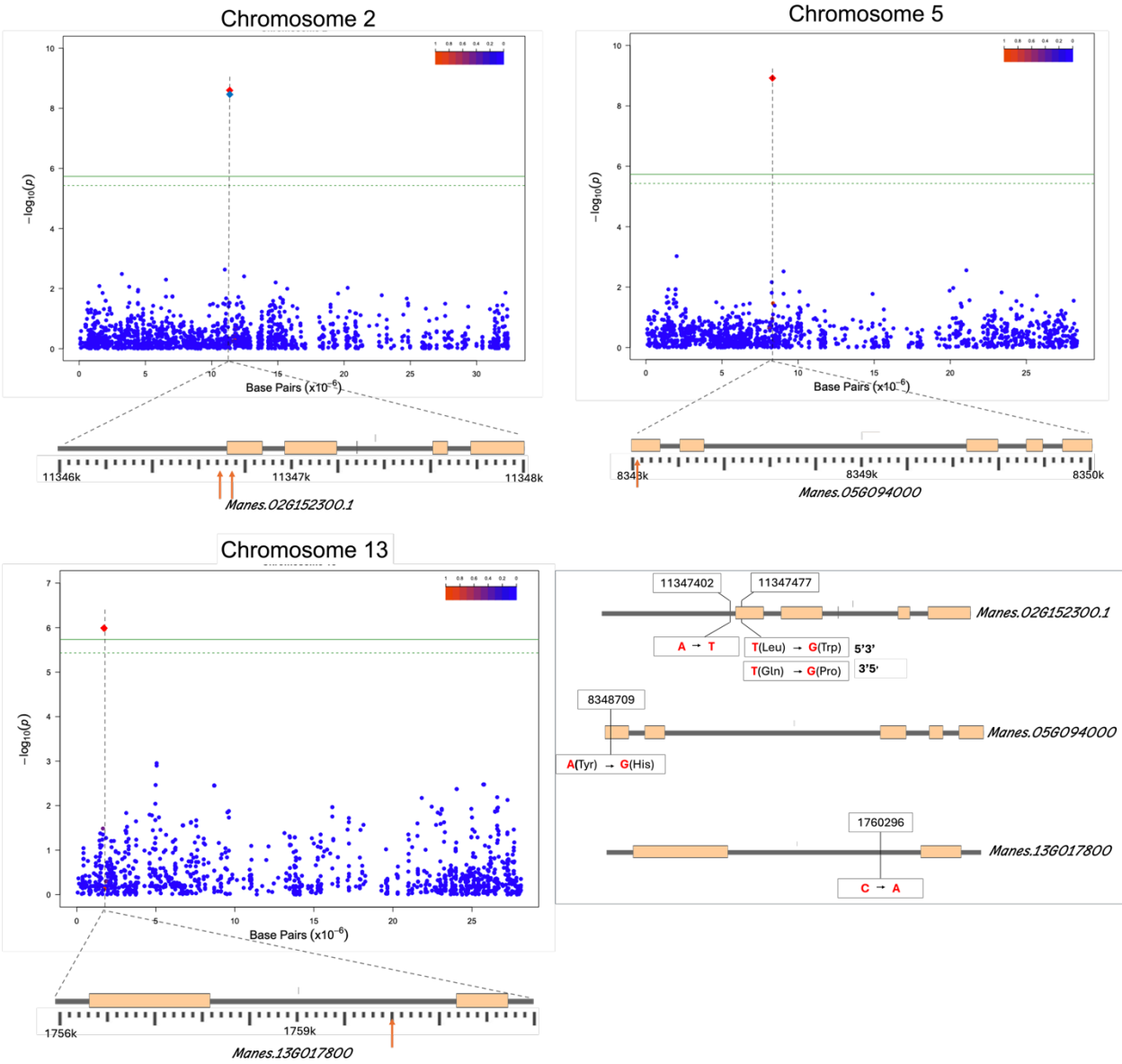


Figure 4. Candidate genes and local LD around peak SNPs. (A–C) Regional association plots for the three chromosomes showing $-\log_{10}(p)$ of GWAS signals (top), pairwise LD (r^2 ; color scale) and annotated genes in the interval (bottom). The lead SNP position is indicated by a dashed vertical line; the significance threshold is shown. Orange pointers mark SNP positions mapped to gene models. (D) Haplotype maps for candidate genes: coding regions (gray), coding sequences (bisque), mutated bases highlighted in red, and amino-acid changes (in parentheses) where applicable.

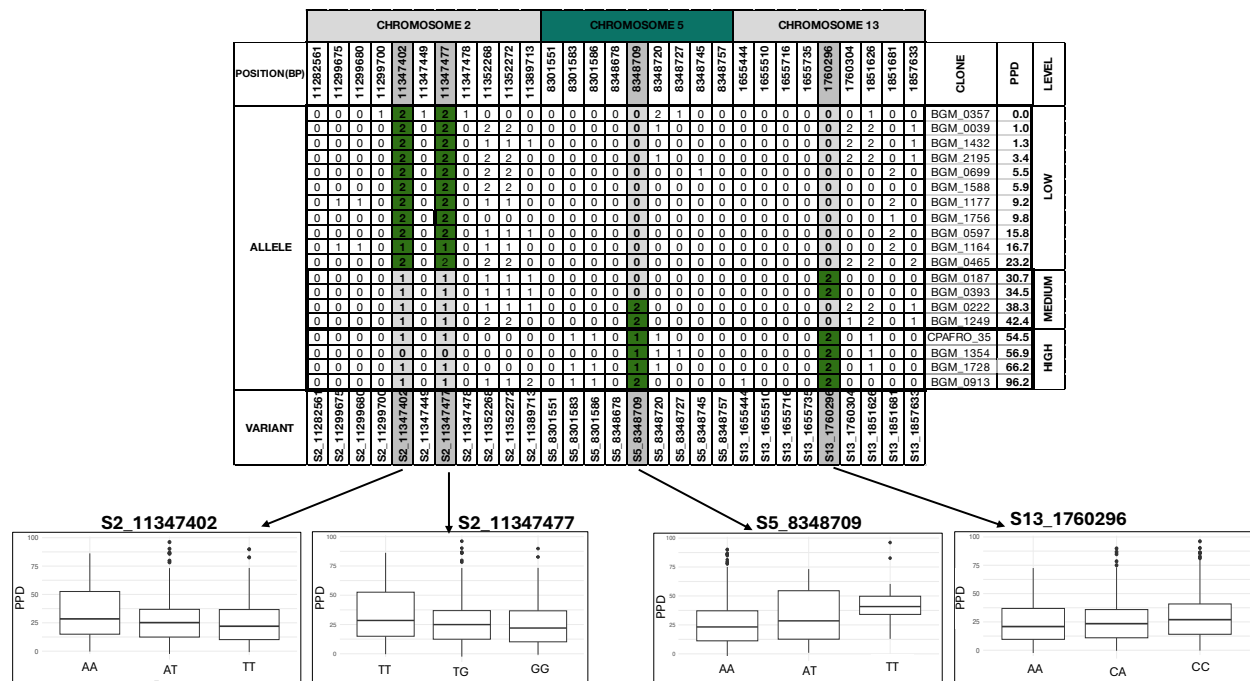


Figure 5. Haplotype analysis of candidate genes involved in PPD expression across the cassava genome. At the top, the SNP analysis is presented for chromosomes two, five, and thirteen, with the corresponding positions of the variants indicated in gray. On the right, the clones are categorized into three groups based on the severity of PPD symptoms: low, medium, and high. The numbers zero, one, and two represent the number of SNP copies present in each genotype. In the bottom section, a boxplot illustrates the levels of PPD symptoms associated with each genotypic allele combination, showing data for those with zero, one or two copies of the reference SNPs.

Statistical inference of epistatic interactions among candidate genes

The analysis of genetic interactions among candidate genes provides deeper insights into the genetic basis of complex traits such as PPD and facilitates the identification of favorable allelic combinations for pyramiding in breeding strategies. In this study, we focused on alleles located on chromosomes 2, 5, and 13, as revealed by GWAS, and evaluated their potential interactions through both haplotype and epistasis analyses. Haplotype analysis suggested that specific allele combinations can attenuate the impact of metabolic pathways underlying PPD expression in cassava roots. To quantify these effects, we performed an epistasis analysis using four SNPs, arranged into pairs. The results revealed clear interaction patterns: 1) The allele combination S2_11347477_GG $\times$ S5_8348709_AA was associated with significantly reduced PPD expression, consistent with an additive $\times$ additive interaction, with a estimated value of -43.25 ; 2) In contrast, the combination S5_8348709_AT $\times$ S13_1760296_CC was linked to markedly higher PPD levels, showing a dominance $\times$ additive effect, with value of 52.25 increasing of PPD expression in the genotypes with this allele combination related with the population mean; 3) Additionally, a significant additive $\times$ dominance interaction was detected between S2_11347477_GG $\times$ S13_1760296_CA, which reduced PPD expression by -26.25 compared with the population mean.

These results (Supplementary Table S3; Figure. 6) highlight the importance of epistatic interactions in shaping PPD expression, emphasizing that tolerance is not solely determined by individual loci but also by the interaction among alleles across different genomic regions.

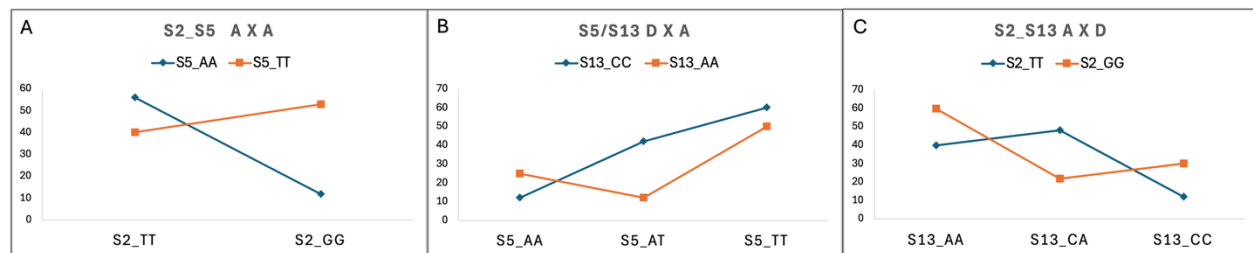


Figure 6. Epistasis plots for SNP pairs. **(A)** Additive × additive interaction between S2_11347477 and S5_8348709; the x-axis represent the genotypes at SNP S2_11347477 (TT/GG), the colored lines correspond to genotypes classes at SNP S5_8348709 (AA blue and TT orange) and the y-axis represents the phenotypic value of the PPD; **(B)** dominance × additive interaction between S5_8348709 and S13_1760296; the x-axis represent the genotypes at SNP S5_8348709 (AA/AT/TT), the colored lines correspond to genotypes classes at SNP S13_1760296 (CC blue and AA orange) and the y-axis represents the phenotypic value of the PPD **(C)** additive × dominance interaction between S2_11347477 and S13_1760296. the x-axis represents the genotypes at SNP S13_1760296 (AA/CA/CC), the colored lines correspond to genotypes classes at SNP S2_11347477 (TT Blue and GG orange) and the y-axis represents the phenotypic value of the PPD.

Discussion

Genomic approaches and the challenge of studying PPD

Genomics-driven approaches have become indispensable to modern crop improvement because they accelerate the development of varieties that combine favorable traits with stability across environments. For complex physiological traits such as PPD in cassava, understanding the underlying genetic architecture is a prerequisite for efficient selection. PPD is an intrinsically complex phenotype: it is triggered by harvest-induced wounding and involves coordinated responses across stress signaling, ROS dynamics, water and solute flux, and multiple metabolic pathways that together determine whether a genotype will be highly susceptible or comparatively tolerant.

Our findings sit within a growing literature that has characterized biochemical, transcriptomic and genetic contributors to PPD, including metabolomic studies (Lebot et al., 2023), transgenic experiments (Ma, et al., 2022), surveys of Brazilian germplasm (Venturini et al., 2016), and transcriptome profiling (Reilly et al., 2007). Integrating these layers (genome, transcriptome, and metabolome) strengthens candidate gene prioritization because it links statistical association to plausible biological mechanisms. A major strength of this study is the use of multi-environment,

multi-year phenotyping, which increases power to detect reproducible associations and helps to partition genetic signal from environmentally driven variation.

Panel diversity, structure, and heritability

The evaluation panel showed wide phenotypic variation for PPD (0–94 %), indicating segregation of alleles with contrasting effects. The panel’s average MAF = 24 % provides reasonable power for association detection while highlighting that some alleles may be moderately rare and therefore could have limited transferability across germplasm sets. Principal component analysis separated samples into two clusters but did not indicate strong stratification that would confound GWAS. Kinship estimates were low, consistent with considerable diversity in the panel, a favorable scenario for association mapping. Heritability estimates varied among environments with a mean near 48 %, which aligns well with Venturini et al. (2016), who reported ~52 %. An intermediate heritability implies that genetic improvement is attainable but will be moderated by environmental influence and G×E; thus, breeding strategies must account for both genetic effects and environmental robustness.

GWAS strategy and functional annotation

To maximize detection and to reduce false positives, we applied three complementary models: MLM, BLINK and SUPER. Each model balances control of structure/kinship and computational efficiency differently; convergence of signals across models increases confidence in putative loci. MLM was particularly sensitive to signals on chromosome 2, BLINK captured associations across chromosomes 2, 5 and 13, and SUPER both confirmed chromosome 2 signals and revealed additional local variants. All three models consistently identified S2_11347402 as a high-confidence SNP associated with PPD.

This study identifies three key biological and biochemical processes that occur after the harvest process and are associated with the candidate genes proposed for the early stages of PPD expression. The first process involves aquaporins, which facilitate transmembrane water and small-molecule movement and are known to respond rapidly to wounding and osmotic changes. Reilly et al. (2007) reported early upregulation of upregulation of Plasma membrane Intrinsic Protein PIP1 and PIP2 family members during PPD; Zainuddin et al. (2018) also highlight the importance of water/ion transport in the early wound response. We provided two SNPs on chromosome 2 (S2_11347402 and S2_11347477) jointly explain ~8.9 % of phenotypic variance. They map to a region that includes *Manes.02G152300*, annotated as an aquaporin (PIP-type) gene. We detected a coding-frame mutation co-located with S2_11347402 that results in a notable amino acid change between leucine (Leu) and tryptophan (Trp) (Figure 4d). The physicochemical change (hydrophobic aliphatic residue to bulkier aromatic residue) could alter channel conformation, gating dynamics, or interactions with regulatory partners, consistent with the observation that clones homozygous for the alternative allele show improved tolerance (Figure 5). These lines of evidence support a functional role for this aquaporin in modulating early PPD dynamics.

In following biochemical process related to PPD, Serine/threonine phosphatases are central regulators of signal transduction, cellular metabolism, and both biotic and abiotic stress responses (Máthé et al., 2019) the context of PPD, these enzymes influence oxidative stress pathways and the generation of reactive oxygen species (ROS), which underlie the characteristic darkened tissue observed in advanced PPD (García et al., 2013).

In biochemical processes associated with PPD, serine/threonine phosphatases act as central regulators of signal transduction, cellular metabolism, and responses to both biotic and abiotic stresses (Máthé et al., 2019). In this context, these enzymes are thought to modulate oxidative stress pathways and the production of reactive oxygen species (ROS), which contribute to the characteristic tissue darkening observed during advanced stages of PPD (García et al., 2013). They also participate in signaling cascades involving phytohormones such as jasmonic acid. Pharmacological studies using phosphatase inhibitors (e.g., okadaic acid) further underscore the role of phosphatase activity in modulating stress responses. (Choudhury et al., 2017) our findings reveal in chromosome 5, the SNP S5_8348709 explains approximately 23.87 % of the phenotypic variation in our panel. This variant maps to *Manes.05G094000*, which encodes a serine/threonine phosphatase. In our dataset, Figure 4d shows a non-synonymous substitution at this locus of tyrosine (Tyr) to histidine (His), a change in residue chemistry that may affect local folding, charge distribution, or protein–protein interactions. Figure 5 demonstrates that clones carrying the alternative allele display a marked increase in PPD. Together, these lines of evidence support a plausible functional role for this phosphatase in modulating PPD progression.

One of the late symptoms observed in the expression of PPD is the appearance of blue-dark marks in the tissues, which result from hypersensitivity and cell death. In this process of activation leucine-rich repeat receptor-like kinase LRR-RLKs play a crucial role in the membrane-localized receptors that mediate extracellular signal recognition and downstream signal transduction, and they play central roles in responses to both biotic and abiotic stress. of these receptors are frequently associated with triggered immunity (TI), a response that includes rapid ROS accumulation and a cascade of signaling events that may culminate in hypersensitive responses and programmed cell death (Ma, Hu, et al., 2022). Such immune-related responses are consistent with the late stages of PPD, when tissue integrity is severely compromised and necrotic regions become prevalent (Djabou et al., 2017). In our findings we report on chromosome 13 explains approximately 3.62 % of the phenotypic variation, the SNP S13_1760296. This variant lies in *Manes.13G017800*, which encodes a (LRR-RLK). This association is reflected in Figure 5, where clones carrying the alternative allele show higher PPD scores. Thus, one plausible hypothesis is that the variant at S13_1760296 predisposes roots to an exaggerated ROS-mediated hypersensitivity response after harvest, accelerating tissue collapse.

Understanding epistatic interactions is essential for analyzing the interactions between candidate genes and optimizing the pyramiding of genes in future genomic selection processes. Our epistasis tests revealed significant interactions between loci on chromosomes 2 and 5; when favorable alleles are combined at these loci, the average PPD score is reduced by 43.25 units (on the phenotyping scale used). In this study, we provide evidence of a significant interaction between the transcript products of genes *Manes.02G152300* and *Manes.05G094000*. The first gene encodes for aquaporins, which have been identified as essential regulators in molecular transport and play a crucial role in the plant's stress response. In the context of PPD, this means that if a plant

possesses a favorable allele combination in its genotype, it will exhibit improved mechanisms to control the PPD reaction during the initial stages or hours. This finding is linked to the second gene, responsible for transcribing antioxidant molecules. When a plant possesses a favorable combination of alleles for certain genes, it can produce a regulated amount of molecules involved in the oxidative process. With effective transport, this enables the creation of a complex metabolic pathway, allowing tissue cells to mount a slower but more effective stress response after damage occurs during harvesting or transportation. Importantly, our statistical analysis provides strong evidence of this epistatic interaction, which can be demonstrated in the phenotypic expression of genotypes that harbor the genetic combinations for both genes. As a result, these plants display long-shelf life and low symptoms of PPD. This information provides an excellent opportunity to utilize the most favorable allele combinations across different loci, aiming for optimal effects in the pathways associated with PPD's early and late symptoms. This approach will be valuable for developing the new genomic selection population in the coming years.

Together, the statistical and functional evidence point to four SNPs distributed across three chromosomes that jointly explain 35.83 % of the phenotypic variance. Each SNP maps within or near genes with clear roles in the physiological processes triggered by harvest wounding and mechanical injury. From these data we can reasonably hypothesize that there is a genetic interaction among the various regions where the SNPs are located, as well as interactions between the resulting products of the metabolic pathways during both the early and late phases of post-harvest physiological deterioration.

Perspectives for cassava breeding

The present study provides a comprehensive catalog of loci that significantly affect PPD expression. By utilizing historical data from multi-environment evaluations and employing two assessment methods, visual analysis and image, along with high-density genome-wide SNP data from GBS, we achieved a higher mapping resolution. As a result, we can identify candidate genes closely related to the PPD biological processes that occur in the roots after harvesting. We also identified SNP variants that are advantageous for extending cassava's shelf life. Furthermore, this study provides specific allele combinations in the haplotype. If this information could be converted into allele-specific high-throughput SNP markers, it would enable fast and efficient screening and identification of individuals in the segregating population that carry the favorable alleles during the genomic selection process. Additionally, our study offers substantial insights into the epistatic interactions between genes on chromosomes 2 and 5. This information enhances our understanding of the strategies to be used in breeding programs. By recognizing the strong relationship between the transcripts of *Manes.02G152300* and *Manes.05G094000*, we can utilize this information to select favorable alleles in the segregating population, facilitating the pyramiding of these genes in future generations.

Nevertheless, this information is derived from a rigorous statistical analysis and data sourced from a GenBank on genes and protein information. Our findings have the potential to advance future functional studies that enhance our understanding of the interactions among candidate genes. Furthermore, to explore how these genes' transcriptional products and their corresponding proteins interact to regulate the processes affecting root tissue during the hours following harvest and under transportation stress. Additionally, it would be interesting to develop CRISPR-Cas9

methodologies to silence gene expression. This could be validated with biological evidence demonstrating that these genes play a crucial role in the expression of PPD. The studies mentioned, along with other related research, would provide valuable insights into the pathways involved in PPD expression. This information would be essential for the cassava community, aiding in the management of breeding processes through molecular techniques, biochemical approaches, genomic selection, and technology transfer to producers.

Concluding remarks

In this study, we demonstrate the value of integrating a historical, multi-year phenotypic dataset with high-density SNP genotyping and complementary GWAS models to identify genomic regions associated with delayed PPD in cassava. This combined methodological framework increases power to detect reproducible associations and accelerates the discovery of candidate variants that can be translated into breeding tools for complex, environmentally sensitive traits. We identified four SNPs across three chromosomes that together explain 35.83 % of the phenotypic variance for PPD in the Brazilian germplasm panel. These loci are associated with three biologically plausible candidate genes: an aquaporin (*Manes.02G152300*), a serine/threonine phosphatase (*Manes.05G094000*), and an LRR-RLK (*Manes.13G017800*). Each gene is linked to processes relevant to early (water and solute transport, membrane integrity) and late (ROS-mediated signaling, hypersensitivity) stages of PPD. Significant epistatic interactions were also detected, particularly between loci on chromosomes 2 and 5.

While these results provide substantial insights into the genetic architecture of PPD, it is important to recognize additional factors that may influence trait expression. Heritability and environmental effects remain critical considerations, as they directly affect the stability of PPD-related traits across diverse production environments. Likewise, the role of heterozygosity within the population warrants attention, since it may shape both allele effects and the efficiency of selection strategies aimed at reducing PPD. Even with these complexities, the loci and candidate genes identified here hold strong potential for integration into MAS and GS pipelines, offering practical tools for breeding programs. We anticipate that these findings will provide a valuable resource for the cassava research community and, in the long term, contribute to enhancing the economic viability and sustainability of cassava production for farmers, processors, and retailers alike.

Conflict of interest: The authors have no relevant conflict of interest to declare.

Funding: DCSC were supported by fellowships from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - grant 141175/2023-0) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES - grant 88887.916369/2023-00). AAFG received a scholarship from produtive cientists (grant CNPq 313269/2021-1). This research was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil (grant numbers 310980/2021-6 and 402422/2023-6). This work was also partially funded by the UK's Foreign, Commonwealth & Development Office (FCDO) and the Bill & Melinda Gates Foundation (grant number INV-007637). The funder provided support in the form of fellowship and funds for the research, but did not have any additional role in the study design, data collection and analysis, decision to publish, nor preparation of the manuscript.

References

- Aina, O. O., Dixon, A. G. O., & Akinrinde, E. A. (2007). Genetic Variability in Cassava as it Influences Storage Root Yield in Nigueria. *Journal of Biological Sciences*, 7(5), 765–770.
- Akinwale, M. G., Aladesanwa, R. D., Akinyele, B. O., Dixon, A. G. O., & Odiyi, A. C. (2010). Inheritance of β -carotene in cassava (*Manihot esculenta crantz*). *International Journal of Genetics and Molecular Biology*, 2(10), 198–201.
- Amelework, A. B., & Bairu, M. W. (2022). Advances in Genetic Analysis and Breeding of Cassava (*Manihot esculenta Crantz*): A Review. *Plants*, 11(12), 1617. <https://doi.org/10.3390/plants11121617>
- Ayalde, D. G., Londoño, J. C. G., Mosquera, A. Q., Melendez, J. L. L., Gimode, W., Tran, T., Zhang, X., & Selvaraj, M. G. (2024). AI-powered detection and quantification of post-harvest physiological deterioration (PPD) in cassava using YOLO foundation models and K-means clustering. *Plant Methods*, 20(1), 178. <https://doi.org/10.1186/s13007-024-01309-w>
- Bakare, M. A., Kayondo, S. I., Kulakow, P., Rabbi, I. Y., & Jannink, J. (2024). Evaluating breeding for broad versus narrow adaptation for cassava in Nigeria using stochastic simulation. *Crop Science*, csc2.21170. <https://doi.org/10.1002/csc2.21170>
- Banerjee, S., Zeng, L., Schunkert, H., & Söding, J. (2018). Bayesian multiple logistic regression for case-control GWAS. *PLOS Genetics*, 14(12), e1007856. <https://doi.org/10.1371/journal.pgen.1007856>

Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2), 263–265. <https://doi.org/10.1093/bioinformatics/bth457>

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>

Bender, K. W., & Zipfel, C. (2023). Paradigms of receptor kinase signaling in plants. *Biochemical Journal*, 480(12), 835–854. <https://doi.org/10.1042/BCJ20220372>

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>

Bernardo, R. (2020). *Breeding for quantitative traits in plants* (Third edition). Stemma Press.

Beyene, G., Solomon, F. R., Chauhan, R. D., Gaitán-Solis, E., Narayanan, N., Gehan, J., Siritunga, D., Stevens, R. L., Jifon, J., Van Eck, J., Linsler, E., Gehan, M., Ilyas, M., Fregene, M., Sayre, R. T., Anderson, P., Taylor, N. J., & Cahoon, E. B. (2018). Provitamin A biofortification of cassava enhances shelf life but reduces dry matter content of storage roots due to altered carbon partitioning into starch. *Plant Biotechnology Journal*, 16(6), 1186–1200. <https://doi.org/10.1111/pbi.12862>

Bhat, J. A., Yu, D., Bohra, A., Ganie, S. A., & Varshney, R. K. (2021). Features and applications of haplotypes in crop breeding. *Communications Biology*, 4(1), 1266. <https://doi.org/10.1038/s42003-021-02782-y>

- 729 Borku, A. W., Tora, T. T., & Masha, M. (2025). Cassava in focus: A comprehensive literature
730 review, its production, processing landscape, and multi-dimensional benefits to
731 society. *Food Chemistry Advances*, 7, 100945.
732 <https://doi.org/10.1016/j.focha.2025.100945>
- 733 Bradburd, G. S., Ralph, P. L., & Coop, G. M. (2016). A Spatial Framework for Understanding
734 Population Structure and Admixture. *PLOS Genetics*, 12(1), e1005703.
735 <https://doi.org/10.1371/journal.pgen.1005703>
- 736 Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007).
737 TASSEL: Software for association mapping of complex traits in diverse samples.
738 *Bioinformatics*, 23(19), 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- 739 Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edsinger-Gonzales, E.,
740 Grimwood, J., Schmutz, J., Rabbi, I. Y., Egesi, C., Nauluvula, P., Lebot, V., Ndunguru,
741 J., Mkamilo, G., Bart, R. S., Setter, T. L., Gleadow, R. M., Kulakow, P., Ferguson, M. E.,
742 ... Rokhsar, D. S. (2016). Sequencing wild and cultivated cassava and related species
743 reveals extensive interspecific hybridization and genetic diversity. *Nature*
744 *Biotechnology*, 34(5), 562–570. <https://doi.org/10.1038/nbt.3535>
- 745 Briollais, L., Dobra, A., Liu, J., Friedlander, M., Ozcelik, H., & Massam, H. (2016). A Bayesian
746 graphical model for genome-wide association studies (GWAS). *The Annals of Applied*
747 *Statistics*, 10(2). <https://doi.org/10.1214/16-AOAS909>
- 748 Buschmann, H., Tohme, J., & Beeching, J. R. (2000). Biochemistry of Post-harvest
749 Deterioration in Cassava Root Tubers. *Deutscher Tropentag 2000*, 1.

- Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., & Thompson, R. (2023). *ASReml estimates variance components under a general linear.*
- Cantor, R. M., Lange, K., & Sinsheimer, J. S. (2010). Prioritizing GWAS Results: A Review of Statistical Methods and Recommendations for Their Application. *The American Journal of Human Genetics*, 86(1), 6–22. <https://doi.org/10.1016/j.ajhg.2009.11.017>
- Ceballos, H., Kulakow, P., & Hershey, C. (2012). Cassava Breeding: Current Status, Bottlenecks and the Potential of Biotechnology Tools. *Tropical Plant Biology*, 5(1), 73–87. <https://doi.org/10.1007/s12042-012-9094-9>
- Che, Y., Ding, Z., Shen, C., Fernie, A. R., Tang, X., Yao, Y., Liu, J., Wang, Y., Li, R., & Guo, J. (2024). Physiological and Microstructure Analysis Reveals the Mechanism by Which Formic Acid Delays Postharvest Physiological Deterioration of Cassava. *Antioxidants*, 13(10), 1245. <https://doi.org/10.3390/antiox13101245>
- Choudhury, F. K., Rivero, R. M., Blumwald, E., & Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. *The Plant Journal*, 90(5), 856–867. <https://doi.org/10.1111/tpj.13299>
- Cockerham, C. C. (1954). An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics*, 39, 859–882.
- Cortes, D. F., Reilly, K., Okogbenin, E., Beeching, J. R., & Tohme, J. (2002). Mapping wound-response genes involved in post-harvest physiological deterioration (PPD) of cassava (*Manihot esculenta* Crantz).

- Dankwa, K. O., Gimode, W., & Olanmi, B. (2025). A review of global cassava (*Manihot esculenta* Crantz) production trends, post-harvest physiological deterioration (PPD) challenge, and control strategies. *Discover Food*, 5(1), 307. <https://doi.org/10.1007/s44187-025-00629-7>
- De Carvalho, R. R. B., Bandeira E Sousa, M., De Oliveira, L. A., & De Oliveira, E. J. (2022). Phenotypic diversity and selection in biofortified cassava germplasm for yield and quality root traits. *Euphytica*, 218(12), 173. <https://doi.org/10.1007/s10681-022-03125-6>
- De Vetten, N., Wolters, A.-M., Raemakers, K., Van Der Meer, I., Ter Stege, R., Heeres, E., Heeres, P., & Visser, R. (2003). A transformation method for obtaining marker-free plants of a cross-pollinating and vegetatively propagated crop. *Nature Biotechnology*, 21(4), 439–442. <https://doi.org/10.1038/nbt801>
- Delgado, L. F., Moreta, D. E., Morante, N., Lenis, J. I., Aparicio, J. S., Londoño, L. F., Salazar, S. M., Tran, T., Ospina, M. A., Melendez, J. L. L., Alzate, J. L. M., Vargas, H. C., Duran, L. P., Alpala, E. A. R., & Zhang, X. (2024). Assessing realized genetic gains in biofortified cassava breeding for over a decade: Enhanced nutritional value and agronomic performance. *Crop Science*, 64(6), 3242–3258. <https://doi.org/10.1002/csc2.21369>
- Devi, B., Kumar, M. N., Chutia, M., & Bhattacharyya, N. (2022). Abiotic and biotic stress challenges of Cassava in changing climate and strategies to overcome: A review. *Scientia Horticulturae*, 305, 111432. <https://doi.org/10.1016/j.scienta.2022.111432>

Difabachew, Y. F., Frisch, M., Langstroff, A. L., Stahl, A., Wittkop, B., Snowdon, R. J., Koch, M., Kirchhoff, M., Cselényi, L., Wolf, M., Förster, J., Weber, S., Okoye, U. J., & Zenke-Philippi, C. (2023). Genomic prediction with haplotype blocks in wheat. *Frontiers in Plant Science*, 14, 1168547. <https://doi.org/10.3389/fpls.2023.1168547>

Ding, Z., Fu, L., Wang, B., Ye, J., Ou, W., Yan, Y., Li, M., Zeng, L., Dong, X., Tie, W., Ye, X., Yang, J., Xie, Z., Wang, Y., Guo, J., Chen, S., Xiao, X., Wan, Z., An, F., ... Hu, W. (2023). Metabolic GWAS-based dissection of genetic basis underlying nutrient quality variation and domestication of cassava storage root. *Genome Biology*, 24(1), 289. <https://doi.org/10.1186/s13059-023-03137-y>

Djabou, A. S. M., Carvalho, L. J. C. B., Li, Q. X., Niemenak, N., & Chen, S. (2017). Cassava postharvest physiological deterioration: A complex phenomenon involving calcium signaling, reactive oxygen species and programmed cell death. *Acta Physiologiae Plantarum*, 39(4). <https://doi.org/10.1007/s11738-017-2382-0>

Doyle, J. J., & Doyle, J. L. (1987). A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue. *Phytochemical Bulletin*, 19, 11–15.

Drapal, M., Ovalle Rivera, T. M., Becerra Lopez-Lavalle, L. A., & Fraser, P. D. (2024). Carotenoid composition and sequestration in cassava (*Manihot esculentum* Crantz) roots. *PLOS ONE*, 19(11), e0312517. <https://doi.org/10.1371/journal.pone.0312517>

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M.,
Richardson, L. J., Salazar, G. A., Smart, A., Sonnhammer, E. L. L., Hirsh, L., Paladin,
L., Piovesan, D., Tosatto, S. C. E., & Finn, R. D. (2019). The Pfam protein families
database in 2019. *Nucleic Acids Research*, 47(D1), D427–D432.
<https://doi.org/10.1093/nar/gky995>

Endelman, J. B. (2023). Fully efficient, two-stage analysis of multi-environment trials with
directional dominance and multi-trait genomic selection. *Theoretical and Applied
Genetics*, 136(4), 65. <https://doi.org/10.1007/s00122-023-04298-x>

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals
using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–
2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>

Ferguson, M. E., Shah, T., Kulakow, P., & Ceballos, H. (2019). A global overview of cassava
genetic diversity. *PLOS ONE*, 14(11), e0224763.
<https://doi.org/10.1371/journal.pone.0224763>

Filho, J. R. F. (2013). Cultivo, processamento e uso da mandioca. *Embrapa Mandioca e
Fruticultura*, 1.

Food and Agriculture Organization of the United Nations (FAO). *FAOSTAT Statistical
Database, Statistical Division. Rome. 2024. (n.d.)*.

Forkum, A. T., Wung, A. E., Kelese, M. T., Ndum, C. M., Lontum, A., Kamga, E. B., Nsaikila, M.
N., & Okwen, P. M. (2025). Safety of cassava and cassava-based products: A
systematic review. *Frontiers in Sustainable Food Systems*, 9, 1497609.
<https://doi.org/10.3389/fsufs.2025.1497609>

- García, J. A., Sánchez, T., Ceballos, H., & Alonso, L. (2013). Non-destructive sampling procedure for biochemical or gene expression studies on post-harvest physiological deterioration of cassava roots. *Postharvest Biology and Technology*, 86, 529–535. <https://doi.org/10.1016/j.postharvbio.2013.06.026>
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1), D1178–D1186. <https://doi.org/10.1093/nar/gkr944>
- Hamazaki, K., & Iwata, H. (2020). RAINBOW: Haplotype-based genome-wide association study using a novel SNP-set method. *PLOS Computational Biology*, 16(2), e1007663. <https://doi.org/10.1371/journal.pcbi.1007663>
- He, J., Meng, S., Zhao, T., Xing, G., Yang, S., Li, Y., Guan, R., Lu, J., Wang, Y., Xia, Q., Yang, B., & Gai, J. (2017). An innovative procedure of genome-wide association analysis fits studies on germplasm population and plant breeding. *Theoretical and Applied Genetics*, 130(11), 2327–2343. <https://doi.org/10.1007/s00122-017-2962-9>
- Hill, W. G., & Robertson, A. (1968). Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, 38(6), 226–231. <https://doi.org/10.1007/BF01245622>
- Hu, W., Kong, H., Guo, Y., Zhang, Y., Ding, Z., Tie, W., Yan, Y., Huang, Q., Peng, M., Shi, H., & Guo, A. (2016). Comparative Physiological and Transcriptomic Analyses Reveal the Actions of Melatonin in the Delay of Postharvest Physiological Deterioration of Cassava. *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.00736>

856 Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2019). BLINK: A package for the
857 next level of genome-wide association studies with both individuals and markers in
858 the millions. *GigaScience*, 8(2), giy154. <https://doi.org/10.1093/gigascience/giy154>

859 International Cassava Genetic Map Consortium (ICGMC). (2015). High-Resolution Linkage
860 Map and Chromosome-Scale Genome Assembly for Cassava (*Manihot esculenta*
861 Crantz) from 10 Populations. *G3 Genes|Genomes|Genetics*, 5(1), 133–144.
862 <https://doi.org/10.1534/g3.114.015008>

863 Ismail Rabbi, Rabbi, I. Y., Lovina I. Udoh, Udoh, L. I., Marnin D. Wolfe, Wolfe, M. D., Elizabeth
864 Parkes, Parkes, E., Melaku Gedil, Gedil, M., A. G. O. Dixon, Dixon, A. G. O., Punna
865 Ramu, Ramu, P., Jean-Luc Jannink, Jannink, J.-L., Peter Kulakow, & Kulakow, P. (2017).
866 Genome-Wide Association Mapping of Correlated Traits in Cassava: Dry Matter and
867 Total Carotenoid Content. *The Plant Genome*, 10(3).
868 <https://doi.org/10.3835/plantgenome2016.09.0094>

869 Iyer, S., Mattinson, D. S., & Fellman, J. K. (2010). Study of the Early Events Leading to Cassava
870 Root Postharvest Deterioration. *Tropical Plant Biology*, 3(3), 151–165.
871 <https://doi.org/10.1007/s12042-010-9052-3>

872 Jaldhani, V., Suman, K., Rao, D. S., Travis, T., Rao, I. S., Rathod, S., Voleti, S. R.,
873 Subrahmanyam, D., Rao, P. R., Barbadikar, K. M., Mangrauthia, S. K., Sundaram, R.
874 M., Price, A. H., & Neeraja, C. N. (2024). Superior haplotypes to enhance grain filling
875 in rice (*Oryza sativa* L.) identified through GWAS. *Discover Plants*, 1(1), 70.
876 <https://doi.org/10.1007/s44372-024-00069-0>

877 Jaramillo, A. M., Londoño, L. F., Orozco, J. C., Patiño, G., Belalcazar, J., Davrieux, F., & Talsma,
878 E. F. (2018). A comparison study of five different methods to measure carotenoids in
879 biofortified yellow cassava (*Manihot esculenta*). *PLOS ONE*, 13(12), e0209702.
880 <https://doi.org/10.1371/journal.pone.0209702>

881 Jia, M., Ma, R., Liu, C., Pan, X., Shen, W., & Tian, Y. (2025). Insight into the relationship
882 between cell wall, intracellular starch structure and physicochemical properties of
883 cassava cells from cassava varieties. *International Journal of Biological*
884 *Macromolecules*, 288, 138728. <https://doi.org/10.1016/j.ijbiomac.2024.138728>

885 Kao, C.-H., & Zeng, Z.-B. (2002). Modeling Epistasis of Quantitative Trait Loci Using
886 Cockerham's Model. *Genetics*, 160(3), 1243–1261.
887 <https://doi.org/10.1093/genetics/160.3.1243>

888 Kapilan, R., Vaziri, M., & Zwiazek, J. J. (2018). Regulation of aquaporins in plants under stress.
889 *Biological Research*, 51(1), 4. <https://doi.org/10.1186/s40659-018-0152-0>

890 Kawano, K., Fukuda, W. M. G., & Cenpukdee, U. (1987). Genetic and Environmental Effects
891 on Dry Matter Content of Cassava Root¹. *Crop Science*, 27(1), 69–74.
892 <https://doi.org/10.2135/cropsci1987.0011183X002700010018x>

893 Kawano, K., & Rojanaridpiched, C. (1983). Genetic study on post-harvest root deterioration
894 in cassava. *Kasetsart Journal*, 17(2), 14–26.

895 Kim, D.-S., & Iida, F. (2023). Nutritional composition of Cassava (*Manihot esculenta*) and its
896 application to elder-friendly food based on enzyme treatment. *International Journal*
897 *of Food Properties*, 26(1), 1311–1323.
898 <https://doi.org/10.1080/10942912.2023.2213410>

899 Kim Reilly, Reilly, K., Diana M. Bernal, Bernal, D., Diego Fernando Marmolejo Cortes, Cortes,
900 D. F., D. Cortes, R. Gómez-Vásquez, Gómez-Vásquez, R., Gómez-Vásquez, R., Joe
901 Tohmé, Tohme, J., John R. Beeching, & Beeching, J. R. (2007). Towards identifying the
902 full set of genes expressed during cassava post-harvest physiological deterioration.
903 *Plant Molecular Biology*, 64(1), 187–203. <https://doi.org/10.1007/s11103-007-9144-0>

904 Lebot, V., Lawac, F., Muñoz-Cuervo, I., Mercier, P.-E., & Legendre, L. (2023). Metabolite
905 fingerprinting of cassava (*Manihot esculenta* Crantz) landraces assessed for post-
906 harvest physiological deterioration (PPD). *Food Chemistry*, 421, 136217.
907 <https://doi.org/10.1016/j.foodchem.2023.136217>

908 Lin, J., & Musunuru, K. (2018). From Genotype to Phenotype: A Primer on the Functional
909 Follow-Up of Genome-Wide Association Studies in Cardiovascular Disease.
910 *Circulation: Genomic and Precision Medicine*, 11(2), e001946.
911 <https://doi.org/10.1161/CIRCGEN.117.001946>

912 Lin, Y.-C., Mayer, M., Valle Torres, D., Pook, T., Hölker, A. C., Presterl, T., Ouzunova, M., &
913 Schön, C.-C. (2024). Genomic prediction within and across maize landrace derived
914 populations using haplotypes. *Frontiers in Plant Science*, 15, 1351466.
915 <https://doi.org/10.3389/fpls.2024.1351466>

916 Luo, X., An, F., Xue, J., Zhu, W., Wei, Z., Ou, W., Li, K., Chen, S., & Cai, J. (2023). Integrative
917 analysis of metabolome and transcriptome reveals the mechanism of color
918 formation in cassava (*Manihot esculenta* Crantz) leaves. *Frontiers in Plant Science*,
919 14, 1181257. <https://doi.org/10.3389/fpls.2023.1181257>

- Ma, Q., Hu, Z., Mao, Z., Mei, Y., Feng, S., & Shi, K. (2022). The novel leucine-rich repeat receptor-like kinase MRK1 regulates resistance to multiple stresses in tomato. *Horticulture Research*, 9, uhab088. <https://doi.org/10.1093/hr/uhab088>
- Ma, Q., Xu, J., Feng, Y., Wu, X., Lu, X., & Zhang, P. (2022). Knockdown of p-Coumaroyl Shikimate/Quinate 3'-Hydroxylase Delays the Occurrence of Post-Harvest Physiological Deterioration in Cassava Storage Roots. *International Journal of Molecular Sciences* 2022, Vol. 23, Page 9231, 23(16), 9231. <https://doi.org/10.3390/IJMS23169231>
- Ma, Y., Xue, M., Zhang, X., & Chen, S. (2023). Genome-wide analysis of the metallothionein gene family in cassava reveals its role in response to physiological stress through the regulation of reactive oxygen species. *BMC Plant Biology*, 23(1). <https://doi.org/10.1186/s12870-023-04174-2>
- Malheiros, G. (2026). Brazil cassava starch exports hit record in 2025, eye EU growth—DatamarNews. *DatamarNews*, 2020604.
- Marcela Tonini Venturini, Venturini, M. T., Leandro Ribeiro dos Santos, Santos, L. R. dos, Carlos Ivan Aguilar Vildoso, Vildoso, C. I. A., Vinícius Silva dos Santos, da Silva Santos, V., E. J. de Oliveira, & de Oliveira, E. J. (2016). Variation in cassava germplasm for tolerance to post-harvest physiological deterioration. *Genetics and Molecular Research*, 15(2). <https://doi.org/10.4238/gmr.15027818>
- Máthé, C., Garda, T., Freytag, C., & M-Hamvas, M. (2019). The Role of Serine-Threonine Protein Phosphatase PP2A in Plant Oxidative Stress Signaling—Facts and

Hypotheses. *International Journal of Molecular Sciences*, 20(12), 3028.

<https://doi.org/10.3390/ijms20123028>

Mbinda, W., & Mukami, A. (2022). Breeding for postharvest physiological deterioration in cassava: Problems and strategies. *CABI Agriculture and Bioscience* 2022 3:1, 3(1), 1–18. <https://doi.org/10.1186/S43170-022-00097-4>

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>

Mohidin, S. R. N. S. P., Moshawih, S., Hermansyah, A., Asmuni, M. I., Shafqat, N., & Ming, L. C. (2023). Cassava (*Manihot esculenta* Crantz): A Systematic Review for the Pharmacological Activities, Traditional Uses, Nutritional Values, and Phytochemistry. *Journal of Evidence-Based Integrative Medicine*, 28, 2515690X231206227. <https://doi.org/10.1177/2515690X231206227>

Ohmura, T., Ueda, T., Hashimoto, Y., & Imoto, T. (2001). Tolerance of point substitution of methionine for isoleucine in hen egg white lysozyme. *Protein Engineering, Design and Selection*, 14(6), 421–425. <https://doi.org/10.1093/protein/14.6.421>

Olayide, P., Alexandersson, E., Tzfadia, O., Lenman, M., Gisel, A., & Stavolone, L. (2023). Transcriptome and metabolome profiling identify factors potentially involved in pro-vitamin A accumulation in cassava landraces. *Plant Physiology and Biochemistry*, 199, 107713. <https://doi.org/10.1016/j.plaphy.2023.107713>

962 Olayinka, A. F., Dzidzienyo, D. K., Mbanjo, E. G. N., Offei, S. K., Tongoona, P. B., Danquah, E.
963 Y., Egesi, C., & Rabbi, I. Y. (2025). Genome-wide association study identifies novel
964 genes for plant architecture and yield traits in cassava (*Manihot esculenta* Crantz).
965 *Frontiers in Plant Science*, 16, 1660789. <https://doi.org/10.3389/fpls.2025.1660789>

966 Ospina, J. A., Lopez-Alvarez, D., Gimode, W., Wenzl, P., & Carvajal-Yepes, M. (2024).
967 Genome-wide association study of cassava brown streak disease resistance in
968 cassava germplasm conserved in South America. *Scientific Reports*, 14(1), 23141.
969 <https://doi.org/10.1038/s41598-024-74161-6>

970 Otekunrin, O. A. (2024). Cassava (*Manihot esculenta* Crantz): A global scientific footprint—
971 production, trade, and bibliometric insights. *Discover Agriculture*, 2(1), 94.
972 <https://doi.org/10.1007/s44279-024-00121-3>

973 Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using
974 Multilocus Genotype Data. *Genetics*, 155(2), 945–959.
975 <https://doi.org/10.1093/genetics/155.2.945>

976 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J.,
977 Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for
978 Whole-Genome Association and Population-Based Linkage Analyses. *The American*
979 *Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>

980 Qi, M., Chen, Z., & Li, H. (2025). Synergistic Effects of Extrusion on the Structural, Functional
981 and Safety Improvement of Cassava Starch: A Review. *Plant Foods for Human*
982 *Nutrition*, 80(4), 182. <https://doi.org/10.1007/s11130-025-01419-3>

983 Rabbi, I. Y., Udoh, L. I., Wolfe, M., Parkes, E. Y., Gedil, M. A., Dixon, A., Ramu, P., Jannink, J.-
984 L., & Kulakow, P. (2017). Genome-Wide Association Mapping of Correlated Traits in
985 Cassava: Dry Matter and Total Carotenoid Content. *The Plant Genome*, 10(3),
986 plantgenome2016.09.0094. <https://doi.org/10.3835/PLANTGENOME2016.09.0094>
987 Rogers, A. R., & Huff, C. (2009). Linkage Disequilibrium Between Loci With Unknown Phase.
988 *Genetics*, 182(3), 839–844. <https://doi.org/10.1534/genetics.108.093153>
989 Saba, M., Junaid, J. A., Mustafa, S. E., Javaid, A., Raja, S., & Saleem, H. (2024). Cassava: A
990 potential candidate for biofortification exploration. In *Biofortification of Grain and*
991 *Vegetable Crops* (pp. 255–265). Elsevier. [https://doi.org/10.1016/B978-0-323-91735-](https://doi.org/10.1016/B978-0-323-91735-3.00014-5)
992 [3.00014-5](https://doi.org/10.1016/B978-0-323-91735-3.00014-5)
993 Sampaio Filho, J. S., De Souza Campos, M., & De Oliveira, E. J. (2024). Stability and genetic
994 parameters for cassava yield attributes in the tropical humid region of Brazil.
995 *Euphytica*, 220(8), 127. <https://doi.org/10.1007/s10681-024-03384-5>
996 Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., & Kilian, A.
997 (2011). Diversity Arrays Technology (DArT) and next-generation sequencing
998 combined: Genome-wide, high throughput, highly informative genotyping for
999 molecular breeding of Eucalyptus. *BMC Proceedings*, 5(S7), P54, 1753-6561-5-S7-
1000 P54. <https://doi.org/10.1186/1753-6561-5-S7-P54>
1001 Santos, C. S. D., Sousa, M. B., Brito, A. C., De Oliveira, L. A., Carvalho, C. W. P., & De Oliveira,
1002 E. J. (2022). Genome-wide association study of cassava starch paste properties.
1003 *PLOS ONE*, 17(1), e0262888. <https://doi.org/10.1371/journal.pone.0262888>

1004 São Julião, S. M., De Oliveira Gonçalves, L., Zago, L., & Santana, I. (2026). What cassava
1005 means to Brazilians: A contemporary analysis of cultural and nutritional narratives
1006 about the root and its derived products. *International Journal of Gastronomy and*
1007 *Food Science*, 43, 101388. <https://doi.org/10.1016/j.ijgfs.2025.101388>

1008 Saravanan, R., Ravi, V., Stephen, R., Thajudhin, S., & George, J. (2016). Post-harvest
1009 physiological deterioration of cassava (*Manihot esculenta*)—A review. *The Indian*
1010 *Journal of Agricultural Sciences*, 86(11). <https://doi.org/10.56093/ijas.v86i11.62869>

1011 Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of
1012 image analysis. *Nature Methods*, 9(7), 671–675. <https://doi.org/10.1038/nmeth.2089>

1013 Segura, V., Vilhjálmsdóttir, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., & Nordborg, M. (2012).
1014 An efficient multi-locus mixed-model approach for genome-wide association studies
1015 in structured populations. *Nature Genetics*, 44(7), 825–830.
1016 <https://doi.org/10.1038/ng.2314>

1017 Shi Liu, Liu, S., Ima Mulyama Zainuddin, Zainuddin, I. M., Hervé Vanderschuren,
1018 Vanderschuren, H., James Doughty, Doughty, J., John R. Beeching, & Beeching, J. R.
1019 (2017). RNAi inhibition of feruloyl CoA 6'-hydroxylase reduces scopoletin
1020 biosynthesis and post-harvest physiological deterioration in cassava (*Manihot*
1021 *esculenta* Crantz) storage roots. *Plant Molecular Biology*, 94(1), 185–195.
1022 <https://doi.org/10.1007/s11103-017-0602-z>

1023 Silva, R. M. D., Fernandes, A. M., Leonel, M., Pelvine, R. A., Figueiredo, R. T. D., Rangel, M. A.
1024 S., Ringenberg, R., Oliveira, L. A. D., Santos, V. D. S., & Vieira, E. A. (2023).
1025 Measurement of Dry Matter and Starch in Modern Cassava Genotypes during Long

1026 Harvest Cycles. *Horticulturae*, 9(7), 733.
1027 <https://doi.org/10.3390/horticulturae9070733>

1028 Siriwan, W., Vannatim, N., Chaowongdee, S., Roytrakul, S., Charoenlappanit, S.,
1029 Pongpamorn, P., Paemanee, A., & Malichan, S. (2023). Integrated Proteomic and
1030 Metabolomic Analysis of Cassava cv. Kasetart 50 Infected with Sri Lankan Cassava
1031 Mosaic Virus. *Agronomy*, 13(3), 945. <https://doi.org/10.3390/agronomy13030945>

1032 Sivabharathi, R. C., Rajagopalan, V. R., Suresh, R., Sudha, M., Karthikeyan, G., Jayakanthan,
1033 M., & Raveendran, M. (2024). Haplotype-based breeding: A new insight in crop
1034 improvement. *Plant Science*, 346, 112129.
1035 <https://doi.org/10.1016/j.plantsci.2024.112129>

1036 Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redoña, E., Jannink, J.-L., & McCouch, S.
1037 (2016). Genome-wide prediction models that incorporate de novo GWAS are a
1038 powerful new tool for tropical rice improvement. *Heredity*, 116(4), 395–408.
1039 <https://doi.org/10.1038/hdy.2015.113>

1040 Sun, T., Rao, S., Zhou, X., & Li, L. (2022). Plant carotenoids: Recent advances and future
1041 perspectives. *Molecular Horticulture*, 2(1), 3. [https://doi.org/10.1186/s43897-022-](https://doi.org/10.1186/s43897-022-00023-2)
1042 [00023-2](https://doi.org/10.1186/s43897-022-00023-2)

1043 Sunvittayakul, P., Wonnapijit, P., Wannitikul, P., Phanthanong, P., Changwichukarn, K.,
1044 Suttangkakul, A., Utthiya, S., Phraemuang, A., Kongsil, P., Prommarit, K., Ceballos,
1045 H., Gomez, L. D., Kittipadaku, P., & Vuttipongchaikij, S. (2025). Genome-wide
1046 association studies unveils the genetic basis of cell wall composition and

1047 saccharification of cassava pulp. *Plant Physiology and Biochemistry*, 218, 109312.

1048 <https://doi.org/10.1016/j.plaphy.2024.109312>

1049 Tumuhimbise, R., Melis, R., & Shanahan, P. (2015). Genetic variation in cassava for

1050 postharvest physiological deterioration. *Archives of Agronomy and Soil Science*,

1051 61(9), 1333–1342. <https://doi.org/10.1080/03650340.2014.995641>

1052 Uchendu, K., Njoku, D. N., Paterne, A., Rabbi, I. Y., Dzidzienyo, D., Tongoona, P., Offei, S., &

1053 Egesi, C. (2021). Genome-Wide Association Study of Root Mealiness and Other

1054 Texture-Associated Traits in Cassava. *Frontiers in Plant Science*, 12, 770434.

1055 <https://doi.org/10.3389/fpls.2021.770434>

1056 Umar, A. B., Ramalingam, A. P., Sadia, B., Awan, F. S., Khan, F. A., Nasir, M., Bernardo, A., St.

1057 Amand, P., Bai, G., Prasad, P. V. V., & Perumal, R. (2025). Genome-wide association

1058 study and haplotype analyses reveal the genetic architecture of agronomic traits and

1059 sugars in sweet sorghum. *Frontiers in Genetics*, 16, 1611863.

1060 <https://doi.org/10.3389/fgene.2025.1611863>

1061 VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy*

1062 *Science*, 91(11), 4414–4423. <https://doi.org/10.3168/jds.2007-0980>

1063 Villwock, S. S. C., Parkes, E., Nkouaya, E. G. N., Rabbi, I., & Jannink, J. L. (2025). *Bivariate*

1064 *genome-wide association study reveals polygenic contributions to covariance*

1065 *between total carotenoid and dry matter contents in yellow-fleshed cassava (Manihot*

1066 *esculenta)*.

1067 Wang, D., Liu, Q., Xie, X., Zhang, J., Xiao, J., & Wang, W. (2025a). Genomic-Wide Association
1068 Markers and Candidate Genes for the High-Protein Trait in Storage Roots of Cassava
1069 (*Manihot esculenta*). *Plants*, 14(20), 3162. <https://doi.org/10.3390/plants14203162>

1070 Wang, D., Liu, Q., Xie, X., Zhang, J., Xiao, J., & Wang, W. (2025b). Genomic-Wide Association
1071 Markers and Candidate Genes for the High-Protein Trait in Storage Roots of Cassava
1072 (*Manihot esculenta*). *Plants*, 14(20), 3162. <https://doi.org/10.3390/plants14203162>

1073 Wang, J., & Zhang, Z. (2021). GAPIT Version 3: Boosting Power and Accuracy for Genomic
1074 Association and Prediction. *Genomics, Proteomics & Bioinformatics*, 19(4), 629–640.
1075 <https://doi.org/10.1016/j.gpb.2021.08.005>

1076 Wang, Q., Tian, F., Pan, Y., Buckler, E. S., & Zhang, Z. (2014). A SUPER Powerful Method for
1077 Genome Wide Association Study. *PLoS ONE*, 9(9), e107684.
1078 <https://doi.org/10.1371/journal.pone.0107684>

1079 Wang, X., Chang, L., Tong, Z., Wang, D., Yin, Q., Wang, D., Jin, X., Yang, Q., Wang, L., Sun, Y.,
1080 Huang, Q., Guo, A., & Peng, M. (2016). Proteomics Profiling Reveals Carbohydrate
1081 Metabolic Enzymes and 14-3-3 Proteins Play Important Roles for Starch
1082 Accumulation during Cassava Root Tuberization. *Scientific Reports*, 6(1), 19643.
1083 <https://doi.org/10.1038/srep19643>

1084 Wheatley, C. C., & Schwabe, W. W. (1985). Scopoletin Involvement in Post-Harvest
1085 Physiological Deterioration of Cassava Root (*Manihot esculenta* Crantz). *Journal of*
1086 *Experimental Botany*, 36(5), 783–791. <https://doi.org/10.1093/jxb/36.5.783>

1087 Xiao, L., Cao, S., Shang, X., Xie, X., Zeng, W., Lu, L., Kong, Q., & Yan, H. (2021). Metabolomic
1088 and transcriptomic profiling reveals distinct nutritional properties of cassavas with

1089 different flesh colors. *Food Chemistry: Molecular Sciences*, 2, 100016.

1090 <https://doi.org/10.1016/j.fochms.2021.100016>

1091 Xu, Y., Xu, C., & Xu, S. (2017). Prediction and association mapping of agronomic traits in

1092 maize using multiple omic data. *Heredity*, 119(3), 174–184.

1093 <https://doi.org/10.1038/hdy.2017.27>

1094 Yang, Q., Zhao, D., & Liu, Q. (2020). Connections Between Amino Acid Metabolisms in Plants:

1095 Lysine as an Example. *Frontiers in Plant Science*, 11, 928.

1096 <https://doi.org/10.3389/fpls.2020.00928>

1097 Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., McMullen, M. D.,

1098 Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A

1099 unified mixed-model method for association mapping that accounts for multiple

1100 levels of relatedness. *Nature Genetics*, 38(2), 203–208.

1101 <https://doi.org/10.1038/ng1702>

1102 Zainuddin, I. M., Fathoni, A., Sudarmonowati, E., Beeching, J. R., Grisse, W., &

1103 Vanderschuren, H. (2018). Cassava post-harvest physiological deterioration: From

1104 triggers to symptoms. *Postharvest Biology and Technology*, 142, 115–123.

1105 <https://doi.org/10.1016/j.postharvbio.2017.09.004>

1106 Zegeye, H., Rasheed, A., Makdis, F., Badebo, A., & Ogbonnaya, F. C. (2014). Genome-Wide

1107 Association Mapping for Seedling and Adult Plant Resistance to Stripe Rust in

1108 Synthetic Hexaploid Wheat. *PLoS ONE*, 9(8), e105593.

1109 <https://doi.org/10.1371/journal.pone.0105593>

1110 Zhang, K., Hu, B., Wang, W., Li, W.-X., & Ning, H. (2025). Integration of GWAS and
 1111 transcriptome and haplotype analyses to identify QTNs and candidate genes
 1112 controlling oil content in soybean seeds. *Scientific Reports*, 15(1), 16803.
 1113 <https://doi.org/10.1038/s41598-025-00869-8>

1114 Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu,
 1115 J., Arnett, D. K., Ordovas, J. M., & Buckler, E. S. (2010). Mixed linear model approach
 1116 adapted for genome-wide association studies. *Nature Genetics*, 42(4), 355–360.
 1117 <https://doi.org/10.1038/ng.546>

1118
 1119
 1120